

CHARLES UNIVERSITY

Faculty of Science

Developmental and Cell Biology



**New types of RNA non-canonical caps and their
degradation by Nudix enzymes**

Ph.D. Dissertation

Maria Bianca Mititelu

Supervisor: Ing. Hana Macíčková Cahová, Ph.D.

Institute of Organic Chemistry and Biochemistry of the CAS

Prague, 2024

Declaration

I declare that I wrote the thesis independently and that I cited all the information sources and literature. This work will not be presented to obtain another academic degree or equivalent.

Prague, 2024

Maria Bianca Mititelu

Acknowledgments

First of all, I would like to thank my supervisor, Hana Macíčková Cahová, not only for giving me the opportunity to pursue my PhD in her group, but also for her constant encouragement, support and belief in my abilities, even during challenging times. I appreciated all your advice, your kindness, and your ability to bring out the best in all of us.

A special thanks goes to Ondřej for being an important point of reference throughout almost my entire PhD. Thank you for your honesty, precision, and hard work in all aspects.

I would also like to thank my friends-colleagues from Cahová group: Jiří, Pavel, Kristinka, Anton, Zuzi, Klara and Matouš. Thank you all for your support, for the scientific discussions and for the lab life tricks. You all are amazing scientists, and I am more than grateful to have met you!

This journey would not have been the same without my girls, Ambri, Vale, and Flami. Everything with you was funnier, easier and better (Italians do it better!). We shared not only the scientific goals, future dreams, and continuous help, but also every disappointment, fear, and anxiety. Thank you from the bottom of my heart for always being by my side and for sharing with me these unforgettable years of our lives!

Special thanks to my dear friends, Luana, Elena, and Chiara, for their friendship that goes even across great distances.

Thanks to Kristián, my other half, for literally everything! I am so grateful to live this life with you by my side!

Infine, vorrei ringraziare la mia famiglia. Sono immensamente grata per il vostro sostegno incondizionato, per l'incoraggiamento e l'amore che mi date ogni giorno ed in ogni momento della mia vita. Grazie!

Table of Contents

Abstract (English)	8
Abstrakt (Cesky)	9
1 Introduction	10
1.1 RNA (Ribonucleic Acid)	10
1.1.1 Types of RNA	11
1.1.2 RNA modifications and epigenetics	13
1.1.3 Internal RNA modifications	13
1.1.3.1 N ⁶ -Methyladenosine (m ⁶ A)	14
1.1.3.2 N ¹ -Methyladenosine (m ¹ A):	14
1.1.3.3 5-Methylcytosine (m ⁵ C)	15
1.1.3.4 5-Hydroxymethylcytosine (hm ⁵ C):	15
1.1.3.5 Pseudouridine (Ψ)	15
1.1.3.6 2'-O-Methylation (2'-O-Me)	16
1.1.3.7 Inosine (I)	16
1.1.4 5' RNA modifications	16
1.1.4.1 m ⁷ G (N ⁷ -methylguanosine)	17
1.1.4.2 NAD (Nicotinamide adenine dinucleotide)	18
1.1.4.3 CoA (Coenzyme A)	20
1.1.4.4 FAD (Flavin adenine dinucleotide)	21
1.1.4.5 ADPR (Adenosine diphosphate ribose)	21
1.1.4.6 Np _n Ns (Dinucleoside polyphosphates)	22
1.2 Dinucleoside polyphosphates (Np_nNs) function in Eukaryotes	23
1.2.1 Cell Signaling	24
1.2.2 Immune response	25
1.2.3 Cellular stress response	26
1.3 mRNA decapping enzymes	27
1.3.1 Nudix Hydrolases	27
1.3.2 DXO (Decapping and exoribonuclease) protein	30
1.3.3 HIT (Histidine Triad Proteins)	31
1.3.4 ApaH-like phosphatases	31
2 Aims of the study	33
3 Materials and Methods	34
3.1 Cell culture, menadione treatment and RNA isolation	34
3.2 LC-MS analysis of small molecules and short RNA	35
3.2.1 Small molecules extraction and LC-MS analysis	35
3.2.2 LC-MS analysis of short RNA	35
3.3 Preparation of radiolabeled 35-mer RNA	35
3.3.1 In vitro transcription with T7 RNA polymerase (RNAP)	35
3.3.2 DNase I treatment	36

3.3.3	RNA 5'-polyphosphatase treatment	36
3.3.4	Terminator™ 5'-phosphate-dependent exonuclease treatment	36
3.4	Purification of recombinant enzymes	36
3.4.1	Purification of Plant NudIX enzymes (AtNUDT6, 7, 19 and 27)	36
3.4.2	Purification of mammalian enzymes	37
3.4.2.1	NUDT2 expression and purification	37
3.4.2.2	DXO expression and purification	37
3.4.2.3	NUDT12 expression and purification	38
3.4.2.4	RLIG1 human ligase expression and purification	38
3.5	Cleavage of capped RNA and small molecules by NudIX enzymes.....	39
3.5.1	Cleavage of capped RNA by <i>A. thaliana</i> NudIX enzymes AtNUDT6, 7, 19 and 27.....	39
3.5.2	Enzymatic assay and kinetics assay of <i>A. thaliana</i> NudIX enzymes AtNUDT6, 7, 19 and 27 on small molecules	39
3.5.3	NUDT2 activity on free Ap ₄ A and Ap ₄ A-RNA.....	40
3.5.4	NUDT12 and DXO activity on capped RNA and small molecules	40
3.5.5	Calculations of RNA decapping assay	40
3.5.6	Kinetic parameter calculations	40
3.5.7	In vitro RLIG1 ligase assay for production of Ap ₂ A-RNA.....	41
3.6	Capturing method for Ap₂A-RNA for Illumina RNA Sequencing (RNASeq)	41
3.6.1	T4 RNA ligase 2 truncated K227Q reaction for 35-mer capped RNA.....	41
3.6.2	mRNA Decapping enzyme (MDE), Terminator™ 5'-Phosphate-Dependent Exonuclease, <i>E. coli</i> Poly(A) Polymerase treatment for U1 capped RNA and short RNA from Expi293F cells.....	41
3.6.3	Final PCR for Illumina RNA sequencing	42
3.7	RT-qPCR and Western Blot analysis	42
3.7.1	NUDT12, DXO and RLIG1 enzyme expression in Expi293F cells.....	42
3.7.2	Ap ₂ A-U1, m ⁷ Gp ₃ A-U1, ppp-U1 as controls for qPCR, Sanger sequencing and RNA sequencing protocol	42
3.7.3	Western Blot	43
3.8	Statistical Analysis	43
4	Results	45
4.1	PART 1: Detection of dinucleoside polyphosphates as non-canonical RNA caps in eukaryotic cells	45
4.1.1	Diadenosine Tetraphosphate (Ap ₄ A) serves as a 5' RNA cap in mammalian cells	45
4.1.2	Diadenosine diphosphate (Ap ₂ A) serves as 5' RNA cap in mammalian cells	46
4.1.2.1	In vitro incorporation of Ap ₂ A into RNA by T7 RNA polymerase	46
4.1.2.2	Detection of Ap ₂ A capped RNAs using LC-MS in Expi293F cell line	47
4.1.2.3	Quantification of Ap ₂ A-RNA after internal oxidative stress.....	49
4.1.2.4	Detection and quantification of Ap ₂ N-RNAs in Expi293F cell line.....	51
4.1.2.5	Detection of free NAD, Ap ₄ A and Ap ₂ N in Expi293F cell line	53
4.1.2.6	Human RNA ligase 1 (RLIG1) forms Ap ₂ N-RNA in vitro using ATP and p-RNA.....	56
4.1.2.7	Mammalian NUDT12 and mouse DXO are decapping enzymes for Ap ₂ A-RNA in vitro	57
4.1.2.8	Identification and classification of Ap ₂ A-RNA using T4 RNA ligase 2, truncated K227Q ...	59

4.2	PART 2: Decapping enzymes for non-canonical RNA caps	65
4.2.1	<i>Arabidopsis Thaliana</i> Plant NudIX enzymes	65
4.2.1.1	AtNUDT6,7,19 and AtNUDT27 in vitro activity on capped RNA	66
4.2.1.2	AtNUDT6,7,19 and AtNUDT27 in vitro activity on small molecules	72
4.2.1.3	Kinetic parameters of AtNUDT6,7 and AtNUDT27 for the cleavage of free Ap ₄ A.....	73
4.2.1.4	Cleavage efficiency of free Ap ₄ A and Ap ₄ A-RNA by AtNUDT6, AtNUDT7, and AtNUDT27..	75
4.2.1.5	Inhibition of Ap ₄ A-RNA by AtNUDT6, 7 and 27 in presence of free Ap ₄ A	76
4.2.2	NUDT2 and DXO decapping activity on non-canonical capped RNAs	79
5	Discussion	83
5.1	Dinucleoside diphosphate Ap ₂ A as novel 5' RNA cap in mammalian cells	83
5.2	<i>Arabidopsis Thaliana</i> NudIXes have RNA decapping enzyme	87
5.3	Human NUDT2 and mouse DXO as decapping enzymes	89
6	Conclusions.....	91
7	References.....	92
8	Appendices.....	108
8.1	Abbreviations	108
8.2	Supplementary figures.....	111
8.3	Pubblications	119

Abstract (English)

Currently, more than 170 RNA modifications are known in both prokaryotes and eukaryotes, but the role of many of them is not well understood. The 5' termini of eukaryotic mRNAs and certain viral RNAs are distinguished by a specific structure known as the 7-methylguanosine (m^7G) cap. This cap plays crucial roles in transcription, stability, and translation of mRNAs. The recent discovery of various non-canonical RNA caps, such as Nicotinamide adenine dinucleotide (NAD), Coenzyme A (CoA), Flavin adenine dinucleotide (FAD), and Dinucleoside polyphosphates (Np_nNs) in different organisms, has prompted a reassessment of our understanding of RNA cap function and metabolism.

My PhD project aimed to discover new dinucleoside polyphosphates (Np_nNs) acting as 5' cap in mammalian cells and to investigate their degradation by certain mouse, human and plant Nudix enzymes. Using Liquid Chromatography–Mass Spectrometry (LC–MS) techniques, we identified four different Ap_2Ns acting as 5' cap in mammalian cells: Ap_2A , Ap_2C , Ap_2G and Ap_2U . Moreover, we observed an increase in the concentration of Ap_2A -RNA after exposing Expi293F cells to menadione, a molecule known to generate ROS-based cellular stress.

Furthermore, considering the similarity between Ap_2A -RNA and NAD-RNA, we identified two RNA decapping enzymes responsible for its degradation in vitro: mammalian DXO and NUDT12. We also explored possible biosynthetic pathways and searched for the free forms of these Ap_nNs in Expi293F cells. Since the free forms of these dinucleoside polyphosphates were not detected in cellular extracts, we focused on putative RNA capping enzymes, such as the newly discovered RLIG1 human ligase. We also established a sequencing protocol employing T4 RNA ligase 2, truncated K227Q, and identified the RNAs harboring this specific dinucleoside diphosphate structures at the 5' end.

Finally, we characterized mammalian NUDT2, mouse DXO, and four Nudix enzymes from plant *Arabidopsis Thaliana* (AtNUDT6,7,19 and AtNUDT27), showing their substrate specificity for both Np_nNs -capped RNAs and free Np_nNs molecules.

Overall, my thesis demonstrates that various Ap_nNs can be part of the 5' termini of eukaryotic RNAs and that specific Nudix enzymes are able to cleave these newly discovered caps.

Abstrakt (Cesky)

V současnosti je známo více než 170 modifikací RNA jak u prokaryot, tak u eukaryot, ale role mnohých z nich stále ještě není známá. 5' konce eukaryotických mRNA a určitých virových RNA jsou odlišeny specifickou strukturou známou jako 7-methylguanosinová (m^7G) RNA čepička. Tato čepička hraje klíčovou roli v transkripci, stabilitě a translaci mRNA. Nedávné objevy různých nekanonických RNA čepiček, jako je nikotinamid adenin dinukleotid (NAD), koenzym A (CoA), flavin adenin dinukleotid (FAD) a dinukleosidové polyfosfáty (Np_nNs) v různých organismech, vedly k přehodnocení našeho chápání funkce a metabolismu RNA čepiček.

Můj doktorský projekt si kladl za cíl objevit nové dinukleosidové polyfosfáty (Np_nNs), které působí jako 5' čepičky v buněčných liniích savců, a zkoumat jejich biodegradaci myšmi, lidskými a rostlinnými NudIX enzymy. Proto jsme pomocí techniky kapalinové chromatografie– hmotnostní spektrometrie (LC–MS) identifikovali čtyři různé Ap_nNs , které působí jako 5' víčka v buněčných liniích savců: Ap_2A , Ap_2C , Ap_2G a Ap_2U . Navíc jsme pozorovali nárůst koncentrace Ap_2A -RNA po vystavení buňkám Expi293F menadionu, molecule, která způsobuje buněčný stress.

Dále, vzhledem k podobnosti mezi Ap_2A -RNA a NAD-RNA, jsme identifikovali dva RNA odčepičkovací enzymy odpovědné za jejich biodegradaci in vitro: savčí DXO a NUDT12. Prozkoumali jsme také možné biosyntetické dráhy a hledali volné formy těchto Ap_nNs v buňkách Expi293F. Protože volné formy těchto dinukleosidových polyfosfátů nebyly detekovány v buněčných extraktech, zaměřili jsme se na potenciální RNA čepičkovací enzymy, jako je nově objevená lidská ligáza RLIG1. Zavedli jsme také protokol sekvenování pomocí T4 RNA ligázy 2, zkrácené K227Q, a identifikovali RNA nositele této specifické struktury dinukleosidového difosfátu na 5' konci.

Nakonec jsme studovali substrátovou specifitu savčí NUDT2, DXO a čtyři NudIX enzymy z rostlin *Arabidopsis thaliana* (AtNUDT6,7,19 a AtNUDT27). Jako jejich substráty jsme testovali Np_nNs -čepičkovanou RNA a volné Np_nN molekuly.

V mojí disertační práci jsem ukázala, že různé Ap_nNs mohou být součástí 5' konců eukaryotických RNA a že specifické NudIX enzymy jsou schopny štěpit tyto nově objevené RNA čepičky.

1 Introduction

1.1 RNA (Ribonucleic Acid)

RNA, or Ribonucleic Acid, is a molecule found in all living cells, fundamental to various biological processes. RNA serves as a bridge between the genetic information stored in DNA and the proteins that carry out the functions necessary for life. Similar to DNA, RNA is composed of nucleotides, which consist of a sugar (ribose), a phosphate group, and a nucleobase. The four nucleobases in RNA are adenine (A), cytosine (C), guanine (G), and uracil (U), which replaces thymine (T) found in DNA. Nucleotides in RNA and DNA molecules are joined in chains by the same chemical bond, called a phosphodiester bond (Figure 1). As its name suggests, a phosphodiester bond is composed of two individual ester bonds: one ester bond links a phosphorus atom to the oxygen atom on the 3' carbon of the upstream ribose sugar, and the other one links the same phosphorus atom to the oxygen atom attached to the 5' carbon atom of the downstream ribose sugar.

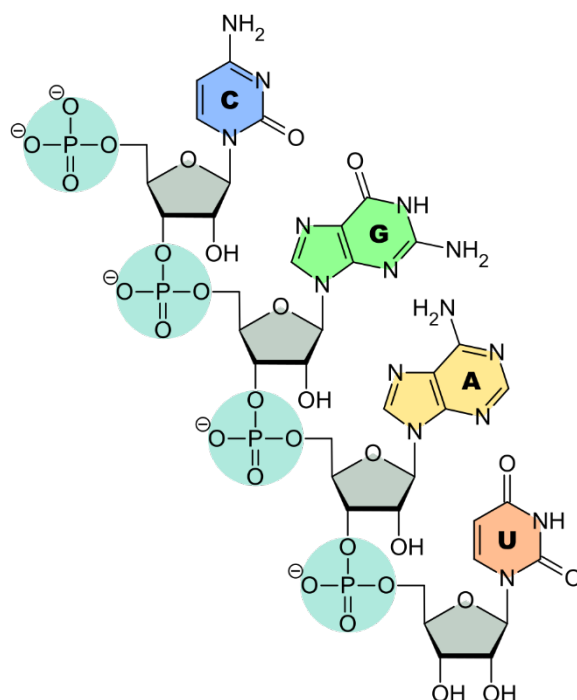


Figure 1. Structure of a fragment of RNA, including all four nucleobases, adenine (A), cytosine (C), guanine (G), and uracil (U), the sugar (ribose) and the phosphates groups; the negative charge on the phosphate chain is important for the interaction with enzymes and ions; Source: Wikimedia commons.

Because of the phosphodiester bonds that join nucleotides together in a polynucleotide chain, every RNA molecule starts with a 5' phosphate group and finishes with a 3'-OH group. This gives the RNA strand a directional polarity, denoted as 5'-3'. The phosphodiester bond plays a critical role in RNA biology considering that the cleavage of this bond in both RNA and DNA is an important reaction in cells.

RNA is a versatile molecule, involved in many cellular processes, including transcription (the synthesis of mRNA from DNA), translation (the synthesis of proteins from mRNA), and various regulatory mechanisms within the cell. In recent years, it has become increasingly evident that RNA plays additional biological roles in living organisms, including catalytic and gene-regulatory functions^{1,2}. Moreover, research into RNA has led to significant advancements in understanding genetics, evolution, and various diseases. It also holds promise for therapeutic applications in areas such as gene therapy and vaccine development.

1.1.1 Types of RNA

RNA can take various forms, each with specific functions (Figure 2). The three main types of RNA are messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). Messenger RNA carries genetic information from DNA, which is used as a template for protein synthesis. Ribosomal RNA is a structural component of ribosomes, cellular machinery responsible for protein synthesis, aiding in aligning mRNA and tRNA during protein assembly. Transfer RNA carries amino acids to the ribosome, where they are joined together to form a protein chain according to the information provided by mRNA.

Additionally, there are other types of RNA that can be broadly divided into coding (cRNA, i.e., mRNA) and noncoding RNA (ncRNA). Noncoding RNA includes housekeeping ncRNAs (rRNA and tRNA) and regulatory ncRNAs, further classified based on their size. Long ncRNAs (lncRNAs) are at least 200 nucleotides long, while small ncRNAs (sncRNAs) are shorter than 200 nucleotides. Small ncRNAs are further subdivided into small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), microRNA (miRNA), small-interfering RNA (siRNA), and PIWI-interacting RNA (piRNA) (Figure 2). They have distinct functions, including mRNA splicing, ribosomal RNA modification, and gene expression regulation.

SnRNAs primarily splice pre-messenger RNA (pre-mRNA) during RNA processing, forming complexes with proteins to create small nuclear ribonucleoproteins (snRNPs). These complexes recognize specific sequences at splice junctions in pre-mRNA and are typically found in the nucleus. SnoRNAs, found in the nucleolus, guide chemical modifications, such as methylation and pseudouridylation, on specific nucleotides of rRNA molecules, involved in processing and modifying ribosomal RNA.

MiRNAs, siRNAs and piRNAs are involved in post-transcriptional gene regulation through RNA interference (RNAi)³. miRNAs and siRNAs, typically 18-22 nucleotides long, bind to the 3' untranslated region (UTR) of target mRNA molecules. This binding can lead to mRNA degradation or translational repression, reducing protein production. They are found in the cytoplasm and processed by the DICER enzyme and incorporated into the RISC complex. Notably, siRNAs originate from exogenous double-stranded RNAs, while miRNAs derive from endogenous noncoding RNAs (single-stranded). SiRNAs typically match their target mRNA perfectly, while miRNAs can repress translation of different mRNAs due to imperfect pairing.

Another category of repressive small RNAs is PIWI-interacting RNA (piRNA), approximately 26 to 41 nucleotides long. PiRNAs regulate transposon (jumping genes) expression by preventing their transcription in germ cells (sperm and eggs). Many piRNAs exhibit complementarity to various transposons, enabling selective targeting and regulation of these transposable elements.

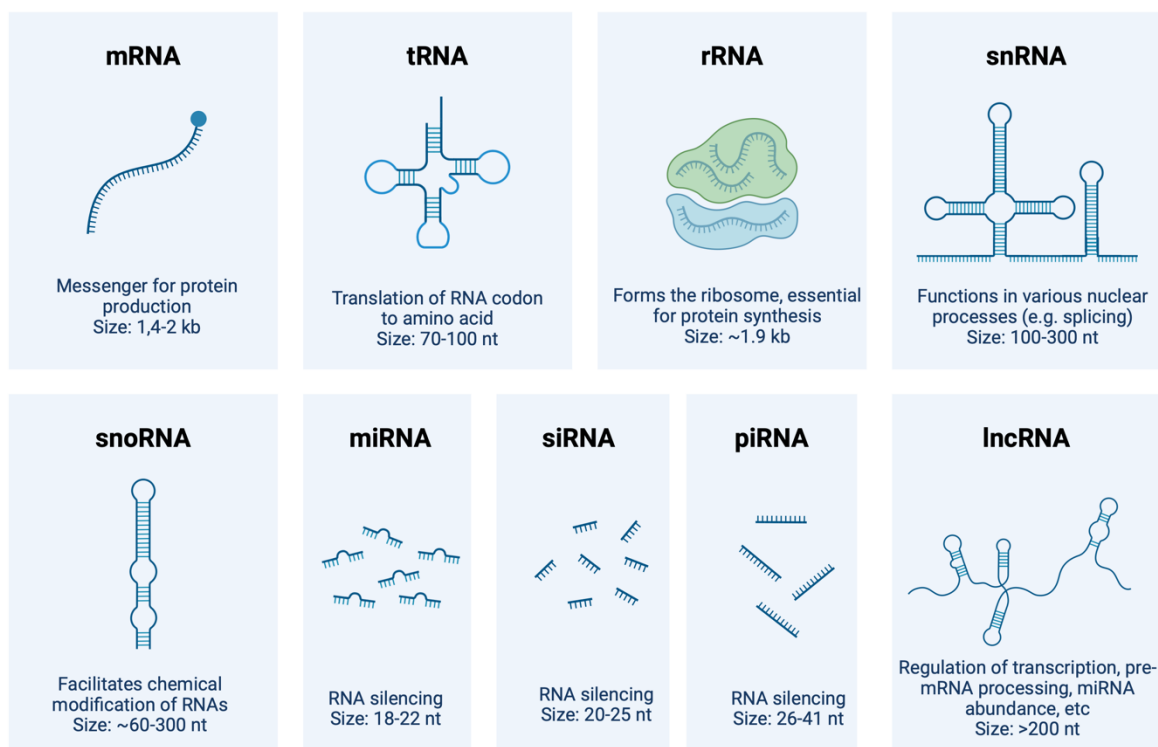


Figure 2. The structure and roles of the principal kinds of RNA and their mean size in nucleobases.

Understanding the functions of RNA, combined with the ongoing development of sequencing technologies, may lead us to the usage of RNA as therapeutic target using innovative strategies involving RNA regulation and function.

1.1.2 RNA modifications and epigenetics

The field of RNA modifications has rapidly expanded in recent years, with over 170 RNA modifications currently known across different organisms⁴. These modifications entail chemical alterations within the nucleotides of RNA molecules and can play critical roles in various cellular processes, including gene expression, RNA stability, and protein synthesis. The dynamic nature of RNA modifications has led to their consideration as a form of “RNA epigenetics”, adding another layer to biological regulation. While some modifications occur at internal positions, there is limited diversity at the RNA termini, specifically 5’ and 3’ ends^{5,6}.

Understanding RNA modifications is crucial for deciphering their roles in regulating gene expression, RNA processing, translation, and RNA-protein interactions. Moreover, dysregulation of RNA modifications has been associated with various diseases, including cancer, neurological disorders, and metabolic diseases⁷⁻⁹. Recent advancements in sequencing technologies, such as high-throughput sequencing and mass spectrometry, have empowered researchers to map RNA modifications transcriptome-wide and explore their functional significance¹⁰.

Additionally, the development of new biochemical and bioinformatics approaches has facilitated the identification and characterization of novel RNA modifications and their associated modifying enzymes¹¹. This rapidly expanding field offers exciting opportunities to uncover new layers of RNA regulation and potentially develop novel therapeutics targeting RNA-modifying enzymes or modified RNAs themselves. Understanding the roles and regulation of these internal RNA modifications represents a rapidly evolving area of molecular biology with implications for diverse fields, including cancer research, neurobiology, and virology.

1.1.3 Internal RNA modifications

Some of the most well-known modifications of RNA occur internally, including N⁶-Methyladenosine (m⁶A), 5-Methylcytosine (m⁵C), Pseudouridine (Ψ), 2’-O-Methylation (2’-O-Me), Inosine (I), N¹-Methyladenosine (m¹A), 5-Hydroxymethylcytosine (hm⁵C) (Figure 3). Detecting these modifications can provide insights into cellular regulation processes such as transcriptional modulation of gene expression, immune response, or epigenetic alterations. Additionally, these modifications can influence various cellular processes, including stress responses, differentiation, and disease states.

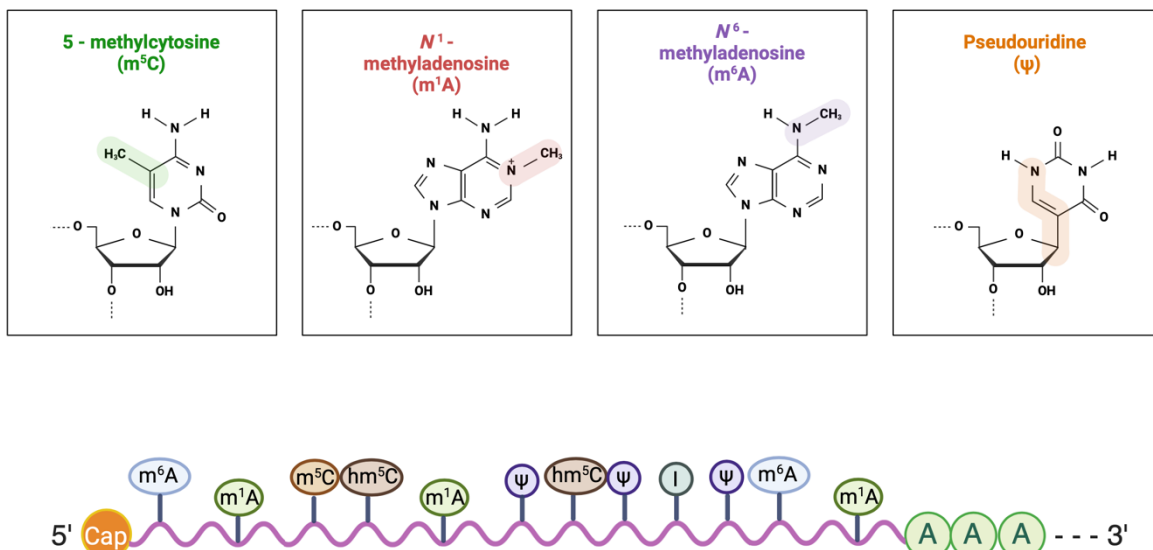


Figure 3. Examples of the structures of the most studied internal RNA modifications: 5-Methylcytosine, *N*¹-Methyladenosine, *N*⁶-Methyladenosine, Pseudouridine.

1.1.3.1 *N*⁶-Methyladenosine (*m*⁶A)

This is one of the most abundant and extensively studied RNA modifications. It involves the addition of a methyl group to the nitrogen at position 6 of adenine. The *m*⁶A modification is installed by *m*⁶A methyltransferase enzymes, called writers, and can be erased by *m*⁶A demethylases enzymes, called erasers. This dynamic regulation determines the abundance of *m*⁶A modifications and plays a decisive role in its distribution¹².

Recent discoveries have shown that dysregulation of *m*⁶A RNA modification is crucial in various physiological and pathological contexts, particularly in the initiation and progression of different human cancer types¹³. Although numerous findings related to the functional roles of this modification have been reported, ongoing studies aim to elucidate its full functional significance.

1.1.3.2 *N*¹-Methyladenosine (*m*¹A):

*N*¹-Methyladenosine is another prevalent internal RNA modification commonly found in tRNA¹⁴, rRNA¹⁵, mRNA¹⁶ and in some mitochondrial (mt) transcripts¹⁷. This modification can disrupt Watson-Crick base pairing, influencing mRNA structure and consequently its translation efficiency. In tRNA, *m*¹A modifications have been observed at positions 9, 14, and

58, helping with the stabilization of the tertiary structure of tRNAs and thus being indispensable for tRNA stability¹⁸.

1.1.3.3 5-Methylcytosine (m⁵C)

5-Methylcytosine (m⁵C) is a prevalent RNA modification characterized by the addition of a methyl group to carbon 5 in cytosine. This modification is widespread in various types of RNA, including messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), and non-coding RNA (ncRNA). The addition of the methyl group to cytosine is catalyzed by RNA methyltransferases, enzymes responsible for installing m⁵C modifications onto RNA molecules. These modifications can occur at specific sites within RNA sequences and play crucial roles in regulating RNA stability, splicing, translation, and interactions with RNA-binding proteins.

Research indicates that m⁵C modifications are dynamically regulated and reversible, suggesting a dynamic regulatory mechanism in RNA metabolism. Dysregulation of m⁵C modification has been implicated in various diseases, including cancer, neurological disorders, and metabolic diseases¹⁹. Recent sequencing technologies and bioinformatics tools have facilitated the genome-wide mapping of m⁵C modifications and the exploration of their functional significance in different biological contexts²⁰. Moreover, ongoing studies aim to elucidate the mechanisms underlying m⁵C deposition and removal, as well as to develop strategies for modulating m⁵C levels for therapeutic purposes.

1.1.3.4 5-Hydroxymethylcytosine (hm⁵C):

Similar to its counterpart in DNA, 5-hydroxymethylcytosine is a modification in RNA resulting from the addition of a hydroxymethyl group to position 5 of cytosine²¹. This modification is rarer compared to others and it is still not well understood. Although it has been identified in various RNA species, ongoing research is needed to fully explore its functional significance²².

1.1.3.5 Pseudouridine (Ψ)

Pseudouridine (Ψ) is one of the most prevalent naturally occurring modifications found in RNA^{23–25}. It is formed by the isomerization of uridine, which converts the N-glycosidic bond between the uracil base and the ribose sugar to a C-glycosidic bond. This modification is catalyzed by enzymes called pseudouridine synthases and is present in various types of RNA, including ribosomal RNA (rRNA)²⁶, transfer RNA (tRNA), messenger RNA (mRNA)^{27,28}, and small nuclear RNA (snRNA)^{29,30}. Pseudouridine modifications can significantly impact RNA structure, stability, and function.

For example, pseudouridine modifications within rRNA contribute to the proper folding and function of the ribosome during protein synthesis³¹. In tRNA, pseudouridine modifications can influence stability and accuracy during translation. In the case of U1 snRNA,

pseudouridylation could potentially affect its ability to interact with other spliceosomal components and participate in the splicing process³². The exact effects of pseudouridylation on U1 snRNA would depend on the specific sites of modification and the cellular context.

Further understanding of the functions and regulatory mechanisms of pseudouridine may lead to the development of novel diagnostic and therapeutic strategies for treating diseases associated with dysregulated RNA modifications.

1.1.3.6 2'-O-Methylation (2'-O-Me)

In the world of RNA, the 2'-O-methylation is one of the most common modifications. This modification involves the addition of a methyl group to the 2' oxygen of the ribose ring in the nucleotide, and it can significantly impact RNA structure, stability, and interactions, and various cellular processes. 2'-O-Me is present in several types of RNA, including ribosomal RNA (rRNA)³³, transfer RNA (tRNA)³⁴, small nuclear RNA (snRNA) such as U2 snRNA³⁵⁻³⁹, and in various eukaryotic mRNAs³⁵⁻³⁸. This modification contributes to RNA stability and helps prevent degradation by ribonucleases.

1.1.3.7 Inosine (I)

Inosine (I) is one of the most extensive forms of RNA modification, created by the deamination of adenosine catalyzed by enzymes known as adenosine deaminases acting on RNA (ADARs)⁴⁰⁻⁴². It pairs with cytosine (similar to guanine) during RNA transcription. Since inosine is interpreted as guanine by the translation machinery it can diversify the coding potential of a transcript. Inosine is particularly abundant in transfer RNA (tRNA)⁴³ and messenger RNA (mRNA)⁴⁴, although it can also be found in other types of RNA.

In tRNA, inosine plays a crucial role in expanding the decoding capacity of the genetic code because it can base pair with adenine, cytosine, or uracil, allowing a single tRNA molecule to recognize multiple codons. This flexibility contributes to the accuracy and efficiency of protein synthesis⁴⁵. In mRNA, this modification can occur within coding sequences, leading to amino acid substitutions in the encoded protein, or within non-coding regions, affecting splicing, RNA stability, and translation efficiency.

1.1.4 5' RNA modifications

Other types of modifications can indeed occur at the 5' end of the RNA. In eukaryotic RNAs, the textbook cap structure typically consists of a methylated base, known as m⁷G (N⁷-methylguanosine). On the other hand, in prokaryotes, the m⁷G cap is absent, and the 5' end is characterized by a triphosphate group^{46,47}.

However, in the last decade, numerous studies have reported the presence of alternative molecules at the 5' end of the RNAs, using liquid chromatography–mass spectrometry (LC–MS) techniques.

1.1.4.1 m⁷G (N⁷-methylguanosine)

All eukaryotic mRNA is equipped with a cap structure at the 5' end, known as m⁷G (N⁷-methylguanosine), which forms a reverse 5' to 5' triphosphate linkage with the first nucleotide of the RNA (Figure 4). The m⁷G cap is the initial modification made to RNA, occurring co-transcriptionally in the nucleus as soon as the first 25–30 nucleotides are incorporated into the nascent transcript. This capping process involves three enzymatic reactions: RNA triphosphatase (TPase), RNA guanylyltransferase (GTase), and guanine-N⁷ methyltransferase (guanine-N⁷ MTase). Each of these enzymatic activities plays a crucial role in converting the 5' triphosphate of nascent RNA into the cap structure.

The mRNA cap serves several crucial functions. It acts as a protective group for exonuclease cleavage of 5' to 3' nucleic acids and it is essential for the recruitment of protein factors involved in pre-mRNA splicing, polyadenylation, and nuclear export. Additionally, it serves as an anchor for recruiting factors that initiate protein synthesis and facilitates the 5' to 3' looping of mRNA during translation^{48,49}.

The m⁷G cap, also known as “cap-0” typically undergoes further modifications in higher eukaryotes. “Cap-1” includes an additional methyl group added to the 2'-O position of the first transcribed nucleotide adjacent to the m⁷G cap (Figure 4). This methylation is catalyzed by an enzyme called cap-specific mRNA (nucleoside-2'-O-)-methyltransferase 1 (CMTR1). Cap-1 modification enhances translation efficiency and may also contribute to mRNA stability. Moreover, it has been shown to be important in evading the cellular innate immune response in vivo⁵⁰.

“Cap-2” represents another modified form of the m⁷G cap, containing an additional methyl group at the 2'-O position of the second transcribed nucleotide adjacent to the m⁷G cap, catalyzed by the same enzyme CNMT2 (Figure 4). Cap-2 modification is less prevalent than cap-1 and is found in specific subsets of mRNA molecules⁵¹. Although its precise functional significance remains unclear, it is believed to play a role in fine-tuning translational regulation.

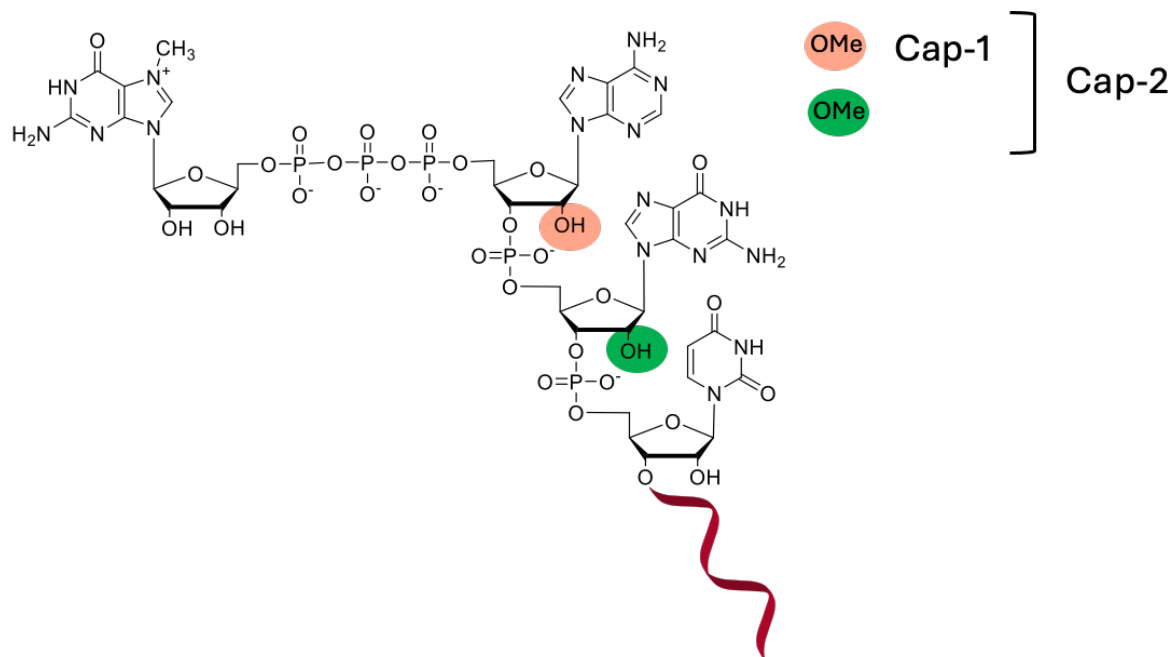


Figure 4. Structure of m^7G (N^7 -methylguanosine) cap, also called cap-0. The addition of a methyl group to the 2'-O position of the first transcribed nucleotide forms cap-1; a second addition of a methyl group to the 2'-O position of the second transcribed nucleotide forms cap-2.

1.1.4.2 NAD (Nicotinamide adenine dinucleotide)

NAD is a coenzyme present in all living cells, plays an essential role in various metabolic pathways, primarily by transportation high-energy electrons to the electron transport chain in aerobic respiration. Beyond its primary function, NAD serves as a coenzyme in numerous other reactions in the body, including immune response⁵², circadian rhythm regulation⁵³, nitrogen fixation^{54,55}, DNA repair⁵⁶, and more. In 2009, a pioneering study developed a general LC-MS technique capable of examining any small molecule-RNA conjugate, regardless of its chemical structure⁵⁷. They digested the total RNA into nucleotides using NucleaseP1 being able to selectively enrich any small molecules covalently bound to RNA mononucleotides, including 5' caps. Surprisingly, they discovered NAD as an inherent component of RNA in *Escherichia Coli* and *Streptomyces venezuelae* (Figure 5).

Later on, in 2015, in order to understand NAD-transcripts, a sequencing technique for specifically labeling, enriching, and sequencing NAD-capped RNA was published⁵⁸. This method relies on the labile N-glycosidic bond of NAD: using ADP-ribosylcyclase (ADPRC) catalysis, it replaces the nicotinamide of NAD (but not NADH) with an alkynyl alcohol in a glycosyl transfer reaction. The resulting alkyne is then reacted with biotin-azide in a copper-catalyzed cycloaddition, facilitating enrichment of 5'-biotinylated transcripts on streptavidin beads for subsequent cDNA library preparation.

Using this technique, they observed enrichment of NAD caps in a subset of small RNAs (sRNAs) and fragments of mRNAs. Subsequently, several other next-generation sequencing (NGS) techniques have emerged for the analysis and quantification of NAD-capped RNAs^{59,60}. Furthermore, the NAD capSeq analysis of budding yeast⁶¹, human⁶², and plant⁶³ transcripts revealed that NAD capped RNA is present not only in prokaryotes but also eukaryotes.

In human cells, NAD-capped mRNA undergo normal posttranscriptional processing, including splicing and polyadenylation. They can vary in length and function⁶⁴. On the other hand, in *Saccharomyces cerevisiae*, the distribution of introns and exons reads in NAD captureSeq datasets suggest that splicing is not obligatory⁶¹, which seems contradictory considering that NAD is added to mRNA transcripts before splicing. Another modified NAD captureSeq revealed that nuclear NAD-mRNAs in yeast may exhibit alternative transcription start sites, resulting in shortened or extended 5' UTRs compared to m⁷G-capped mRNAs⁶⁰. In contrast, mitochondrial NAD-capped RNAs (mtRNAs) have their own distinct NAD cap landscape. Mitochondrial RNAs (mtRNAs), including COX2, ATP9, 15S rRNA, and 21S rRNA, can bear NAD caps. These NAD-capped mtRNAs are distinguishable from the typical 5'-uncapped, leaderless transcripts and are transcribed by mitochondrial RNA polymerase⁶¹. Additionally, regulatory RNAs, like small nucleolar RNAs (snoRNAs) and small Cajal body RNAs (scaRNAs), can be decorated with an NAD cap. These NAD-capped regulatory RNAs are derived from introns and may acquire the NAD cap at a later processing stage, although they can sometimes yield false positive hits⁶².

In summary, NAD-RNAs encompass various types, including mRNAs, mitochondrial RNAs, and regulatory RNAs, each with its own distinct characteristics and functions within the cell. Although many techniques were developed in order to detect NAD-RNAs in the cells, there are very little information about the biological role of this types of RNAs.

In bacteria, studies have shown that NAD-capped RNAs are more stable compared to non-capped transcripts, suggesting a role in modulating RNA stability. This stability may influence various cellular processes, including gene expression and response to environmental stimuli⁶⁵. Moreover, NAD capping may also have an impact on RNA metabolism and processing, including splicing, export, and translation. The presence of NAD at the 5' end of RNA molecules may influence their interactions with RNA-binding proteins and other factors involved in RNA processing pathways⁶¹. One another important function of these transcripts can be the regulation of stress responses since under certain stress conditions, such as oxidative stress or nutrient deprivation, the production and accumulation of NAD-capped RNAs may increase, suggesting a potential role in cellular adaptation to stress⁶⁶.

In the past years, it has been implied that NAD-capped RNA may have a potential role also in disease. The dysregulation of NAD metabolism has been implicated in various diseases, including cancer and neurodegenerative disorders suggesting that this may contribute to disease pathogenesis and progression⁶⁷, suggesting that NAD-RNA could also be implicated in these processes.

Recently, it has also been shown that the presence of NAD-RNA can be important during the infection of *E. Coli* by the T4 bacteriophage. This phage is able to attach NAD capped RNAs to proteins, using its own adenosine diphosphate (ADP)-ribosyltransferase ModB. When this protein is inactivated, the lytic cycle of the bacteriophage is slowed down⁶⁸.

Clearly, in order to fully elucidate the biological roles of NAD-capped RNA further research is needed. As this field continues to evolve, more insights into the functions and mechanisms of NAD-capped RNA are expected to emerge.

1.1.4.3 CoA (Coenzyme A)

Together with NAD, CoA is another important cofactor that can also serve as a capping molecule for RNA, adding another layer of complexity to our understanding of RNA biology (Figure 5). In 2009, not only CoA but also its derivatives, 3'-dephospho-CoA (dpCoA), succinyl-CoA, acetyl-CoA, and methylmalonyl-CoA were identified for the first time using LC-MS as 5' caps on RNA in *E. coli* and *S. venezuelae*⁶⁹. The abundance of these transcripts in vivo, however, was significantly lower compared to NAD-RNAs, making it even more difficult to be identified. Another study revealed the identification and quantification of dpCoA caps in the small RNA fraction of the human pathogen *S. aureus*, using an enzymatic assay⁷⁰. The presence of dpCoA as 5' RNA cap in eukaryotes was described for the first time in 2023, together with the first sequencing method⁷¹. This novel method described for dpCoA capped RNA is called dpCoA tagSeq and it implies the usage of dpCoA's reactive thiol group that can be used for chemical conjugation using maleimides⁷¹. This method led to the identification of more than 40 genes that code for dpCoA in mouse livers. Finding a better sequencing method for CoA-RNA transcripts will help us understand the biological role and the real function of these RNAs.

Another relevant topic is the biosynthesis of CoA-RNA. It is important to mention that 3'-dephospho-CoA can be accepted as a NCIN by RNA polymerase in vitro⁷². However, some studies revealed that dpCoA can also be added post-transcriptionally: phosphopantetheine adenylyltransferase (PPAT), an enzyme involved in the CoA biosynthetic pathway, can utilize 5' triphosphate RNA transcripts as acceptor substrates and catalyze the transfer of 4'-phosphopantetheine post-transcriptionally, resulting in the formation of CoA-RNA⁷³. Nevertheless, PPAT prefers ATP as a substrate, making it a potential inhibitor for the formation of CoA-RNA considering the abundance of ATP in the cells. For this reason, the incorporation by T7 RNA polymerase is improbable for CoA-linked RNAs due to the low

intracellular concentration of the necessary initiator nucleotide, 3'-dephospho-CoA (dpCoA)⁷⁴.

It was shown that bacterial NudC is able to hydrolyze dpCoA and dpCoA-RNA in vitro. Interestingly, many other homologues of NudC such as Npy1 enzyme in *S. cerevisiae* and NUDT19 in *Arabidopsis Thaliana* were able to cleave dpCoA-capped RNAs in vitro⁷⁵. More details about NudIX enzymes and their ability to cleave CoA-RNA substrates will be described in a later chapter.

1.1.4.4 FAD (Flavin adenine dinucleotide)

Considering the fact that FAD is also accepted as a NCIN by *E.coli* RNA polymerase⁷², latest research confirmed the presence of FAD-RNA (Figure 5) in both prokaryotic and eukaryotic cells^{76,77}. The same studies showed that using a method developed for RNA cap screening (CapQuant) combined with LC-MS technique, it is possible to detect FAD-RNA not only in prokaryotes and eukaryotes but even in the RNA genome of the dengue virus. Later on, it was possible to quantify FAD-capped transcripts in the short RNA fraction of HEK293T cells using an enzyme-based colorimetric method called FAD capQ⁷⁸.

Interestingly, in the past years, one potential role of FAD-RNA came from studies on Hepatitis C virus (HCV). It is well known that HCV requires FAD for his replication in the cells, since it needs an adenosine to start the replication of his positive strand into the negative one^{79,80}. In this case, using a selective NudIX enzyme from *Arabidopsis Thaliana* that specifically cleave FAD-RNA (AtNUDT23), it was possible to adapt CapZyme seq and identify that up to 75% of the HCV transcripts are capped at the 5' end with FAD.

Moreover, transfecting cells with FAD-RNA showed that this type of RNA can escape the nucleic acid sensor Retinoic acid-inducible gene I (RIG-I). This shows a possible function of FAD-RNA that can help the virus to protect himself from innate immune recognition. So far, there are no more other functions attributed to this type of transcripts in mammalian cells⁸¹.

1.1.4.5 ADPR (Adenosine diphosphate ribose)

ADP-ribosylation is a post-translational modification process where the ADP-ribose group from nicotinamide adenine dinucleotide (NAD) is added to different molecules such as proteins, nucleic acids, and small molecules. This modification can regulate several biological processes such as DNA repair, transcription, chromatin remodeling, host-virus interactions, cellular stress response and many other⁸². In the past years it was discovered that also the 5' phosphorylated ends of RNA can be ADP-ribosylated (Figure 5) and the best studied group of ADP-ribosyl transferases are Poly(ADP-ribose) polymerases (PARPs) and tRNA 2'-phosphotransferase 1 (TRPT1)⁸³. In higher organisms, homologues of TRPT1 and human PARP10, PARP11 and PARP15 can also ADP-ribosylate phosphorylated ends of RNA. However, the majority of PARPs transfer only a single ADPr moiety, resulting in mono ADP-

ribosylation (MARylation)⁸⁴. This process can be reversible and mediated by several human hydrolases as well as by some viral and bacterial enzymes. Moreover, several viruses can code for their own hydrolases capable of decapping ADPr-RNA, such as Venezuelan equine encephalitis virus (VEEV) and severe acute respiratory syndrome (SARS)⁸⁵, making this modification an interesting research topic for the role of these types of RNAs.

1.1.4.6 Np_nNs (Dinucleoside polyphosphates)

Recently it was discovered that dinucleoside polyphosphates molecules can also act as 5' RNA caps (Figure 5). Using liquid chromatography–mass spectrometry (LC–MS) technique the first nine new Np_nNs caps were identified in the short fraction of *E. Coli*: Ap₃A, Ap₃G, Ap₅A, mAp₃A, m⁷Gp₄Gm, mAp₅G, 2mAp₅G, mAp₄G, mAp₅A⁸⁶. The RNA was isolated from cells, digested by Nuclease P1 into mononucleotides and eventually separated and analysed by LC-MS. Moreover, it was also shown that some Np_nNs can be incorporated by in vitro transcription into RNA by various RNA polymerases as non-canonical initiating nucleotides (NCINs).

Even though the main function of Np_nN-RNAs is still unknown, their degradation is well described. Two *E. coli* enzymes, RNA 5'-pyrophosphohydrolase (RppH) and bis(5'-nucleosyl)-tetraphosphatase (ApaH) were able to cleave the Np_nN caps in RNA. Considering that in bacteria, during late stationary phase, the concentration of methylated Np_nNs caps increased significantly in comparison with exponential phase of growth, this may indicate that bacteria use RNA capping as reaction to stress conditions. Moreover, during the same phase, the Np_nN-RNAs can also be methylated by unknown methyltransferase and this methylation is protecting the RNA from the cleavage by RppH. ApaH enzyme, on the other hand, is still able to cleave the methylated caps. This indicates that the methylation gives higher stability to the RNA during the stationary phase, when the cells lack nutrients and energy⁸⁶.

The process of how Np_nNs are integrated into RNA in living organisms is not entirely clear. Currently, no specific capping enzyme responsible for incorporating Np_nNs into RNA has been identified. There is a hypothesis suggesting that LysRS (Lysyl-tRNA synthetase) might play a role in this process, potentially acting as a capping enzyme. However, this proposed mechanism has only been demonstrated in laboratory settings so far, not in living organisms⁸⁷.

It has also been observed that the incorporation of Np_nNs into RNA is influenced by the nucleotides present in the promoter region. Specifically, RNA polymerase incorporates any dinucleoside polyphosphate once it is available in situ and the amount of produced capped RNA depends on the length of the polyphosphate bridge and the -1, +1, and +2 positions in the DNA template. Interestingly, the highest yield of Np_nN-RNA occurs when templates with Adenine at the -1 and G at +1 positions are used⁸⁸. In order to have a clearer picture of the

biological function of these RNAs, more research involving their metabolism and biosynthesis should be address. The development of new sequencing techniques involving specific decapping enzymes would allow us to understand the types of RNAs bearing this modification.

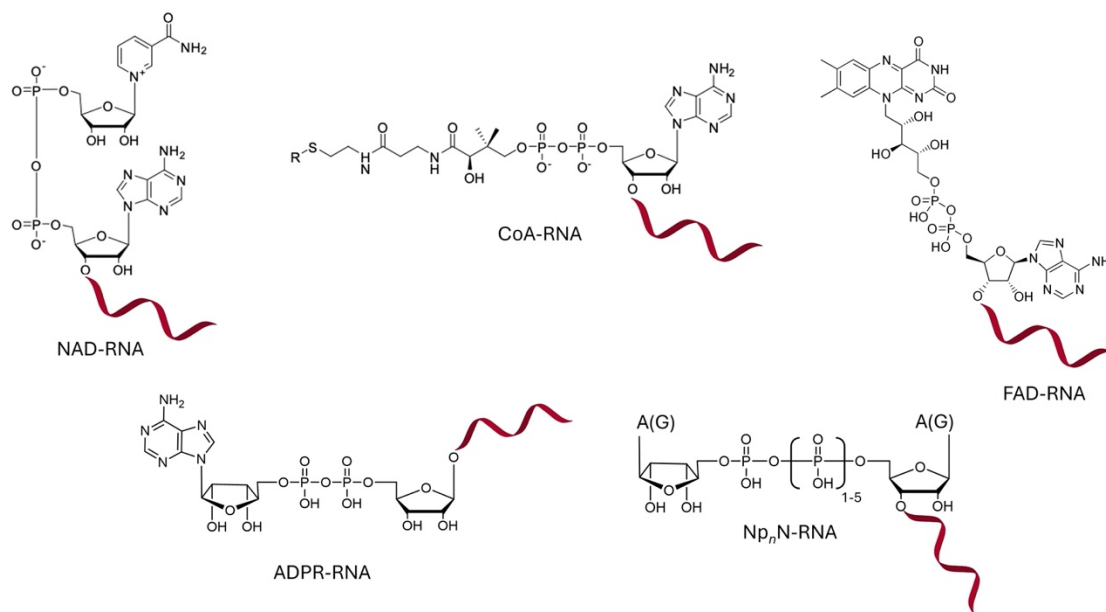


Figure 5. Structure of the most studied 5' RNA cap modifications: Nicotinamide adenine dinucleotide (NAD), Coenzyme A (CoA), Flavin adenine dinucleotide (FAD), Adenosine diphosphate ribose (ADPR), and Dinucleoside polyphosphates (Np_nNs).

1.2 Dinucleoside polyphosphates (Np_nNs) function in Eukaryotes

As free molecules, dinucleoside polyphosphates (Np_nNs) consist of two nucleosides linked together by one or more phosphate groups (Figure 6). The abbreviation “N” stands for the 5' nucleoside and “n” is the number of phosphate residues in the polyphosphate chain that links the two nucleosides via 5' end. These molecules have a negative charge on the phosphates chain, which is important for the interaction with enzymes and ions, such as Mg²⁺ or Zn²⁺.

Np_nNs were discovered more than 50 years ago and over the years and over the last decade they received considerable attention because of their multiple biological and pharmacological activities. They have been identified in bacteria⁸⁹, mammalian cells⁹⁰, yeast⁹¹ and even archaea⁹². They can be synthesized by some aminoacyl-tRNA synthetase⁹³⁻⁹⁶, firefly luciferase⁹⁷ and certain transferases^{98,99}. Notably, they have been found to serve

various functions in eukaryotic organisms, although their roles are not as extensively studied as those in prokaryotes¹⁰⁰. They were used to be called “alarmones” because they can act as signaling molecules, modulating cellular processes by interacting with specific receptors or enzymes¹⁰¹. Diadenosine tri- and tetraphosphates (Ap_3A , Ap_4A) were the first dinucleoside polyphosphates to be identified in human platelets^{102–104}. Diadenosine pentaphosphate (Ap_5A) and diadenosine hexaphosphate (Ap_6A) have been described as vasoconstrictive substances isolated from human platelets¹⁰⁵. Moreover, diadenosine heptaphosphate (Ap_7A) has been isolated from human platelets and has been postulated to play a role in the control of vascular tone¹⁰⁶. Furthermore, dinucleoside polyphosphates containing adenosine and guanosine (Ap_{3-6}G) or containing two guanosines (Gp_{3-6}G) were also identified in human platelets¹⁰⁷. It was reported that Ap_{5-6}G have a vasoconstrictive effect, whereas Gp_{3-6}G do not affect vascular tone in the isolated perfused rat kidney. Both Ap_{3-6}Gs and Gp_{3-6}Gs are growth-stimulating factors of VSMCs¹⁰⁷.

Dinucleoside polyphosphates are also implicated in mediating responses to stress, such as oxidative stress, and may play roles in apoptosis (programmed cell death) and cell proliferation. They can have both intracellular and extracellular function, acting as ligands and modulators on purine receptors and their concentration may vary depending on organism and actual cell status^{108,109}.

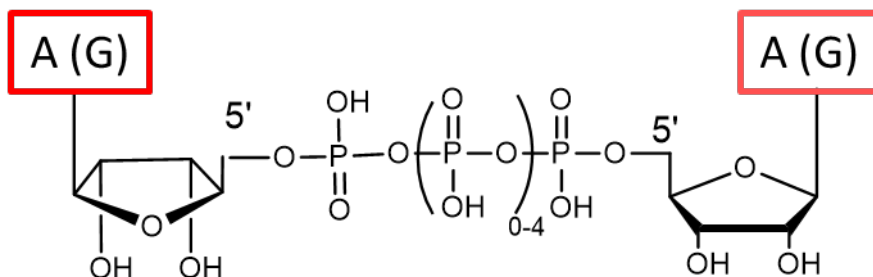


Figure 6. The general structure of dinucleoside polyphosphates. The red rectangle represents either Adenine (A) or Guanine (G). The phosphate chain length can be between 2 and 7 phosphates long.

1.2.1 Cell Signaling

Research on dinucleoside polyphosphates and their role in cell signaling has been a topic of interest in biochemistry and molecular biology for many years. More recently, some Np_nNs , such as diadenosine polyphosphates (Ap_nAs) have emerged as putative low-molecular extra and intracellular signals implicated in the most essential cellular functions, including neurotransmission, vascular tone regulation, and immune response modulation¹¹⁰.

Different tissues or cells, such as heart¹¹¹, hippocampus¹¹², sperm¹¹³, hepatocytes¹¹⁴, neutrophils¹¹⁵, pancreatic cells¹¹⁶, etc. may respond differently to Ap_nAs intracellular concentration indicating that this molecule's effect is widespread and not restricted to any specialized cell type or tissue.

Some Np_nNs have been shown to modulate ion channel activity^{117,118}. For instance, diadenosine tetraphosphate (Ap₄A) has been reported to regulate certain ion channels, such as the ATP-sensitive potassium (KATP) channels in ventricular myocytes as well as regulating KATPs in glucose stimulated Pancreatic β-Cell^{117,119,120}, which are involved in regulating cellular metabolism and membrane excitability.

In liver cells, Np_nNs can interact with calcium channels¹²¹. They can activate the release channels of Ca²⁺ ryanodine receptors in several cell types, increasing their sensitivity to Ca²⁺¹²².

Certain Np_nNs, like Ap_nAs, have been found in the brain and are thought to be involved in neurotransmission regulation. They can be released in the secretory granules of chromaffin cells and bind to the extracellular P2Y purinergic receptors^{123,124}. In these sites in the hippocampus, they can cause a block in the excitation of synapses and stop the signal transmission¹²⁵. Although they are involved in the modulation of synaptic transmission, neuronal excitability, and other aspects of neuronal function, the exact mechanisms are not fully understood.

Today, it is well established that dinucleoside polyphosphates ranging from shorter ones such as Ap₂A or Ap₂G in the myocardium or the placenta, to the longer chained ones such as Gp₆G or Ap₇A in platelets are agonists of specific purinergic receptors and affect vasoconstriction, smooth muscle cell proliferation, and more^{126,127}.

1.2.2 Immune response

Dinucleoside polyphosphates are also present in a number of immune cells. T-lymphocytes, for example, have been shown to increase their intracellular concentration of dinucleoside polyphosphates upon Interleukin-2 (IL-2) stimulation, before new DNA is synthesized¹²⁸. The cellular potential for DNA synthesis decreases upon the removal of IL-2. At the same time, the concentration of Np_nNs also decreases pointing to a role as a second messenger or an initiator of DNA replication^{129,130}. This was further proven, when the molecule showed a high affinity for binding to DNA polymerase alpha¹³¹. Immune cells from the myeloid lineage have also respond to dinucleoside polyphosphates. Induced monocytes and granulocytes increase their production of reactive oxygen species (ROS) upon stimulation by dinucleosides with 4 and 5 phosphate groups which shows a pro-inflammatory role of Np_nNs¹³².

Dendritic cells (DCs) are one of the types of antigens presenting cells (APCs) possessing MHC II. They are highly mobile cells able to bring the captured antigen to the lymph nodes and present it to the native lymphocytes in order to elicit an adaptive immune response¹³³. Dinucleoside polyphosphates can also affect the antigen presentation on the major histocompatibility complex class II (MHC). Béguez-Chédiak-Higashi syndrome (CHS) results in an impaired immune response due to a worsened presentation of peptides on the MHC class II, and this is accompanied by a decrease in intracellular Np_nNs concentration¹³⁴.

When the amount of Np_nNs , specifically diadenosine tetraphosphate Ap_4A , is increased by knocking out the enzyme responsible for the hydrolysis of Ap_4A , the cells exhibit increased motility as well as an enhancement of cross-presentation of antigens. The signaling cascade from LysRS through Ap_4A and Hint1 to MITF is a process that was previously observed only in mast cells, and it has been implicated as a key component of this dendritic cell process¹³⁵.

1.2.3 Cellular stress response

It is known that cells under stress (elevated temperatures, ethanol, cadmium, etc.) accumulate various Np_nNs , molecules that function as alarmones. However, no clear metabolite or molecular target of these alarmones has been experimentally demonstrated. In prokaryotes, the levels of Np_nNs inside the bacteria increases significantly under stress conditions from μM to mM range, according to their function. The same behavior was also observed in eukaryotic cells for Ap_4A , the most studied dinucleoside polyphosphate.

In bacteria, the concentration of Ap_4A is slightly higher than in eukaryotes¹³⁶. However, Ap_4A levels have been found to increase in response to various stressors, such as oxidative stress (H_2O_2 , ACDQ or other types of quinones, heavy metal cadmium, menadione, etc.), heat shock, and DNA damage, suggesting a potential role in cellular adaptation to stress conditions¹³⁷. Moreover, exposure to antibiotic kanamycin is also connected to increased levels of Ap_4A ¹³⁸. This could lead to some different mechanisms of Ap_4A signalization in both organisms.

Overall, while the functions of Np_nNs in eukaryotes are less well-understood compared to their roles in prokaryotes, accumulating evidence suggests that they are involved in various cellular processes, including cell signaling, neurotransmission, ion channel regulation, RNA modification, and stress response¹¹⁰. Further research is needed to fully elucidate the molecular mechanisms and physiological significance in cellular machinery and signaling pathways of Np_nNs in eukaryotic organisms.

1.3 mRNA decapping enzymes

The degradation of RNA is a crucial cellular process because it enables cells to regulate information for various cellular processes, including gene expression and translation. mRNA decapping enzymes catalyze the removal of the 5' cap structure, typically a 7-methylguanosine (m⁷G) cap, from mRNA molecules. This process is referred to as mRNA decapping, and it is an essential step in the degradation of mRNA^{139,140}. The removal of the cap structure by decapping enzymes marks the mRNA for degradation by exonucleases. This process is a key mechanism for controlling mRNA turnover and thus regulating gene expression. The main four classes of RNA decapping enzymes are: NudiX enzymes¹⁴¹, DXO family^{142,143}, HIT (histidine triad) family^{144,145} and ApaH-like phosphatases¹⁴⁶.

1.3.1 NudiX Hydrolases

NudiX hydrolases can cleave the 5' end cap of RNA, making them involved in cellular metabolism, homeostasis, and RNA processing. As their name suggests, they catalyze hydrolysis reactions, specifically cleaving nucleoside diphosphates (Np_nNs) linked to X (representing any moiety). All proteins in the NudiX family contain the conserved NudiX catalytic domain (NudiX box or motif) consisting of 23 amino acids GX₅EX₅[UA]XREX₂EEXGU, where U represents isoleucine, leucine or valine, and X is any amino acid¹⁴⁷ (Figure 7). This motif may vary to some extent among family members.

It was previously thought that NudiX hydrolases had a “housekeeping” function as their main substrates were oxidized forms of GTP, thus preventing its incorporation into DNA or RNA. However, these enzymes are also able to cleave (deoxy)nucleosides triphosphates and their oxidized forms, nucleotide sugars and alcohols, Np_nNs, and RNA cap structures. Despite ongoing studies on their substrate specificity, the primary function of these enzymes still requires confirmation.

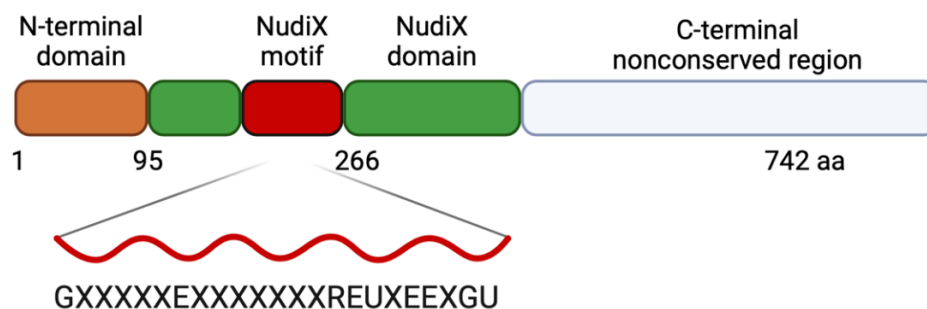


Figure 7. Schematic representation of NudiX enzymes; in orange is represented the N-terminal domain, in green the NudiX domain and in red the NudiX motif (consensus

sequence $GX_5EX_7REUXEEXGU$, where U represents isoleucine, leucine or valine, and X is any amino acid).

NudiX enzymes are characterized by different substrate specificity and intercellular localization. Currently, there are 24 genes encoding for NudiX hydrolases that have been characterized in mammalian cells¹⁴⁸ with specific cleavage activities. For example, NUDT1, NUDT5 and NUDT15 cleave oxidized nucleotides, preventing their incorporation into DNA and RNA, thereby maintaining genome stability, and promoting cell resistance to oxidative stress^{149–152} (Table 1).

On the other hand, some NudiX enzymes can also cleave capped RNAs. Dcp2 (NUDT20) was the first eukaryotic decapping enzyme identified, initially in yeast, and for a long time, it was thought to be the only one^{153–156}. Recent studies have shown that NUDT2, NUDT3, NUDT12, NUDT15, NUDT16, NUDT17 and NUDT19 can cleave both methylated and unmethylated capped RNAs^{141,157,158}. Along with them, NUDT20 also removes the 5'-cap (7-methylguanosine) from RNA, thereby regulating mRNA and small nuclear RNA stability, as well as gene expression^{159,160} (Table 1).

In light of the discovery of non-canonical RNA caps like NAD, CoA and Np_nNs , more NudiX enzymes have been investigated for their decapping activity on such substrates. Recent findings indicate that NUDT12 and NUDT16 function as deNADding enzymes, removing the NAD cap from RNA^{64,161}. Additionally, seven NudiX proteins—NUDT2, NUDT7, NUDT8, NUDT12, NUDT15, NUDT16, and NUDT19—have been recognized for their deCoAping activity in vitro on RNAs that have CoA as 5' cap. Notably, mammalian NUDT2 and NUDT16 have been found to hydrolyze FAD-capped RNAs in vitro, with NUDT16 regulating the levels of FAD-capped RNAs in vivo¹⁶¹ (Table 1).

Despite the known roles of certain NudiX enzymes, many questions regarding their primary functions in cells remain unanswered. NUDT19 is known to play a crucial role in regulating CoA levels in kidneys, while overexpression of NUDT7 in mouse liver has been linked to a reduction in short-chain Acyl-CoA concentration, leading to decreased bile acids levels and diminished fatty acid oxidation in hepatocytes^{162,163}. These findings underscore the significance of these enzymes.

Future discoveries may reveal additional eukaryotic decapping enzymes, and further and more in vivo investigations on the function and specificity of these enzymes will allow us to understand their main roles inside the cells.

Table 1. Summary of the human NudIX hydrolases and their substrates on small molecules and capped RNAs; NUDT1 has similar substrate similarity with NUDT15 and NUDT18; NUDT2 has similar substrate similarity with NUDT3; NUDT5 has similar substrate similarity with NUDT9, NUDT12.

Name	Synonyms	Known Substrates
NUDT1	MTH1	8-OH-(d)GTP, 8-(OH)-(d)ATP, 2-OH-(d)ATP
NUDT2	APAH1	Np _n N (n≥4) e.g. Ap ₄ A, Gp ₄ G, Ap ₅ A, Ap ₆ A, CoA-RNA, FAD-RNA
NUDT3	DIPP1	PP-InsP ₅ , [PP] ₂ -InsP ₄ , Ap ₆ A, Ap ₅ A, capped mRNA
NUDT4	DIPP2α and DIPP2β	PP-InsP ₅ , [PP] ₂ -InsP ₄ , Ap ₆ A, Ap ₅ A
NUDT5	hYSA1, ADPRibase-II	ADP-sugars, 8-OH-(d)GDP
NUDT6	GFG, FGF-AS	ADP-Ribose
NUDT7		CoA, 3'-dephospho-CoA, CoA-RNA
NUDT8		CoA and derivatives, CoA-RNA
NUDT9	ADPRibase-I	ADP-ribose, IDP-ribose
NUDT10	DIPP3α, hAps2	PP-InsP ₅ , [PP] ₂ -InsP ₄ , Ap ₆ A, Ap ₅ A
NUDT11	DIPP3β, hAps1	PP-InsP ₅ , [PP] ₂ -InsP ₄ , Ap ₆ A, Ap ₅ A
NUDT12		NADH, NAD, Ap ₂ A, FAD, ADP-ribose, Ap ₃ A, NAD-RNA, CoA-RNA
NUDT13		NADH, Ap ₂ A
NUDT14	UGPP	UDP-glucose, ADP-ribose, ADP-glucose, GDP-glucose
NUDT15	MTH2	8-OH-(d)GTP, dGTP, CoA-RNA
NUDT16	hx29	capped mRNAs, FAD-RNA, CoA-RNA
NUDT17		?
NUDT18		?
NUDT19	hRP2, hD7rp2E	CoA and derivatives, CoA-RNA
NUDT20	hDcp2	capped mRNAs
NUDT21	CF I _m 25, CPSF5	?
NUDT22		?

1.3.2 DXO (Decapping and exoribonuclease) protein

Another RNA decapping enzyme is the mammalian enzyme DXO, formerly called Dom3Z. It has multiple functions: pyrophosphohydrolitic, decapping, and 5'-3' exoribonuclease activity^{142,164}. Regarding its decapping activity, it is known that DXO protein is involved in the recently discovered mRNA cap quality control pathway, which degrades uncapped mRNAs or mRNAs with an unmethylated cap. In vitro, DXO cleaves pyrophosphate from triphosphate 5' ends and can decap RNA with both methylated and unmethylated caps by cleaving after the first transcribed nucleotide, and it exhibits 5'-3' exoribonuclease activity¹⁴³. In vivo, the enzyme preferentially cleaves mRNAs with unmethylated caps because mature caps are protected by cap-binding proteins¹⁶⁵. DXO is also involved in the removal of the NAD cap in mammals ("deNADding"); in contrast to Nudix enzymes, DXO does not use its pyrophosphohydrolase activity but instead removes the entire NAD moiety^{62,166} (Figure 8). Theoretically, "deNADding" by DXO could also be part of the mRNA cap quality control pathway, assuming that the NAD cap results from the accidental misincorporation of NAD as the initiating nucleotide. Recent research has shown that DXO enzyme can remove the 5' end FAD and dephospho-CoA caps on RNAs, introducing a new functional role for DXO in regulating the fate of FAD-capped RNAs in mammalian cells⁷⁸.

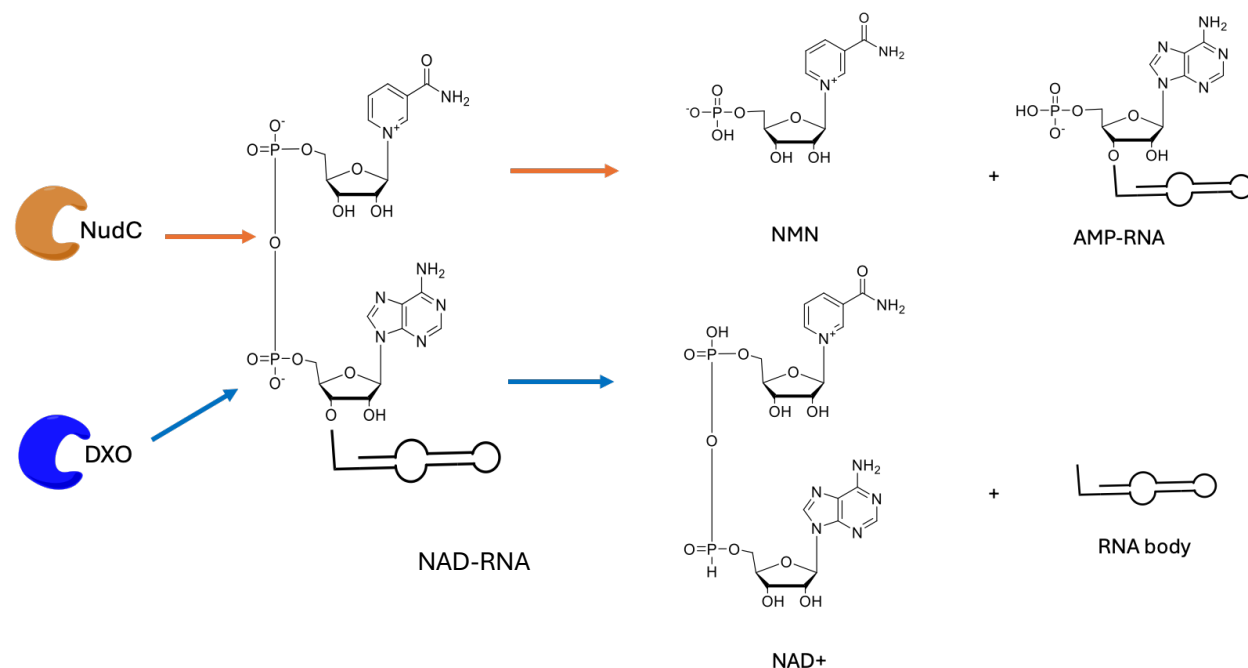


Figure 8. Schematic representation of the cleavage of NAD-RNA by NudC and DXO enzymes; NudC is cleaving between the pyrophosphate releasing Nicotinamide mononucleotide and Adenosine (NMN) and monophosphate RNA (AMP-RNA), while DXO removes the entire NAD moiety releasing the RNA body.

1.3.3 HIT (Histidine Triad Proteins)

HIT proteins, also known as histidine triad nucleotide-binding proteins, are a family of enzymes containing a histidine triad (HIT) sequence in their C-terminal domain. This sequence consists of three histidines separated by hydrophobic residues essential for nucleotide binding¹⁶⁷. HIT proteins are found across a wide range of organisms, from bacteria to humans, and are involved in various cellular processes. Among the HIT protein family, only two mammalian members have been shown to play a role in mRNA decapping: DcpS (Dcs1 in yeast) and FHIT (Hnt2/Aph1 in yeast).

The central histidine in the DcpS HIT motif is crucial for its decapping activity, characterizing the HIT motif as a novel mRNA decapping domain¹⁶⁸. This positions DcpS as the inaugural member of the HIT protein family with a clearly defined biological function. DcpS is a scavenger pyrophosphatase that hydrolyses the residual cap structure following 3' to 5' mRNA degradation¹⁶⁹. Specifically, DcpS uses cap dinucleotides or capped oligonucleotides as substrates to release m⁷GMP (N⁷-methyl GMP). The interaction between DcpS and 3' to 5' exonuclease exosome components suggests a coordinated pathway where exonucleolytic decay influences decapping activity.

Another member of the HIT protein family is FHIT (fragile histidine triad protein), a tumor suppressor protein that hydrolyses the nucleotides diadenosine triphosphate (Ap₃A) and diadenosine tetraphosphate (Ap₄A) in vitro¹⁷⁰⁻¹⁷². FHIT is believed to act as a tumor suppressor by regulating cell growth and apoptosis. Human FHIT and yeast Hnt2/Aph1 have recently been suggested to act as cap scavengers. The accumulation of undegraded cap structures such as m⁷GpppG has been proposed as a possible reason for changes in the translation of cancer-associated RNAs in FHIT-negative human cells. Additionally, other FHIT substrates, such as Ap₃A, which accumulate in FHIT-negative cells, could contribute to this complex phenotype. Although only two members of the HIT protein family are currently implicated in RNA decapping, members of all HIT subfamilies might cleave dinucleoside polyphosphates, making them potential decapping enzymes for non-canonical capped RNAs.

1.3.4 ApaH-like phosphatases

ApaH-like phosphatases are a group of enzymes that belong to the HAD (haloacid dehalogenase) superfamily. These enzymes are found across various organisms, from bacteria to humans. The name "ApaH" originates from the first enzyme of this group characterized in the bacterium *Salmonella typhimurium*, which is involved in the hydrolysis of ADP-ribose. These enzymes participate in various metabolic processes, including nucleotide, phosphate, and sugar metabolism. They typically catalyze the hydrolysis of phosphate esters but also act on sulfate esters and other types of esters. ApaH-like phosphatases play important roles in cellular metabolism by regulating the levels of

phosphorylated metabolites and are essential for maintaining cellular homeostasis. In addition to their metabolic roles, ApaH-like phosphatases have potential applications in biotechnology and medicine due to their ability to modulate cellular signaling pathways and metabolic processes. Recently, it was reported that bis(5'-nucleosyl)-tetraphosphatase (ApaH) acts as a decapping enzyme, cleaving Np_nN -RNAs⁸⁶. Moreover, ApaH enzyme also cleaves the methylated forms of Np_nN caps that bacteria use as protection against RNA degradation under starvation conditions⁸⁶.

The field of RNA decapping enzymes is currently under intensive study¹⁷³. An important future step is identifying the full set of in vivo RNA substrates for different decapping enzymes to understand their function and regulation comprehensively.

2 Aims of the study

Recent discoveries of various non-canonical RNA caps, such as dinucleoside polyphosphates (Np_nNs), nicotinamide adenine dinucleotide (NAD) and Coenzyme A (CoA) across all domains of life have prompted a reassessment of views on RNA cap function and metabolism.

The primary object of the initial part of this thesis was to investigate novel 5' RNA caps in mammalian cells using liquid chromatography–mass spectrometry (LC–MS) techniques.

Another task was to explore the connection between cellular oxidative stress and non-canonical RNA capping by exposing Expi293F cell line to menadione, a compound triggering oxidative stress.

Furthermore, we directed our attention to the biodegradation of Np_nN -RNAs by Nudix enzymes in vitro to understand their substrate specificity and cleavage activity.

Additionally, we delved into possible biosynthetic pathways of Np_nN -RNAs, focusing on putative RNA capping enzymes.

We aimed to establish a sequencing protocol employing specific enzymes to facilitate the identification of RNAs harboring dinucleoside diphosphate structures at the 5' end.

The ultimate goal of this part was to elucidate the role of non-canonical caps in mammalian cells and their connection to cellular stress reaction.

The second aim of this thesis was to characterize four distinct Nudix enzymes from plant *Arabidopsis Thaliana* (AtNUDT6,7,19 and AtNUDT27) and investigate their substrate specificity on both free Np_nNs and in vitro transcribed Np_nN -RNAs.

In this context, the substrate specificity of these Nudix enzymes could be used for the development of capturing techniques for non-canonically capped RNA, potentially aiding in the identification of the sequence and features of these RNA species in the future.

3 Materials and Methods

3.1 Cell culture, menadione treatment and RNA isolation

Cell culture

Expi293F cells were provided by Zuzana Vaitová from Kvido Stříšovský group (IOCB Prague). These types of cells derive from the 293F cell line, and they are maintained in suspension culture using Expi²⁹³ Expression Medium (A1435101) at 37°C and 8% CO₂.

Menadione treatment

0.3 x 10⁶ cells were cultured in 200 mL of Expi²⁹³ Expression Medium and 3 days post-subculture cells were directly treated with either EtOH (control samples) or 40 µM final concentration of menadione (treated samples). To quantify oxidative status of the cell's total glutathione (GSH + GSSG) and GSH/GSSG ratio were measured using the ROS-GloTM H₂O₂ assay according to the manufacturer's instruction. The menadione stock solution (Sigma Aldrich M2518-100G) was prepared in EtOH at concentration of 2 mM, and 4 mL of this solution were added into 200 mL of cell culture. After 3 h of incubation, cells were harvested by centrifugation for 5 min at 1200 rpm at 4°C and washed with ice-cold 1x PBS. The cells were stored at -80°C until the RNA isolation was performed.

Long and short RNA isolation

RNA isolation was performed using RNeasyTM from Sigma-Aldrich (R4533) according to the manufacturer's instruction. The cell lyses was performed using 1 mL of RNeasyTM for 1 x 10⁷ cells. Precipitation of DNA, proteins and polysaccharide was achieved by adding 0.4 mL of RNase-free water per mL of RNeasy. The samples were vortexed and left to rest for 15 min on ice, following centrifugation for 45 min at 6000 × g and 4°C. For the precipitation of long RNA, the supernatant was transferred to a new tube and 0.4 mL of 75% EtOH per 1 mL of the supernatant was added. The samples were vortexed and left 10 min on ice followed by centrifugation for 16 min at 6000 × g and 4°C. The resulting pellet is considered the long RNA fraction. To achieve short RNA isolation, another precipitation was done by transferring the supernatant in another tube and adding 0.8 mL of pure isopropanol per 1 mL of the supernatant. The final vortex was performed followed by incubation for 30 min at 4°C and centrifugation for 30-60 min at 6400 × g. Afterwards, the supernatant was removed, and the pellet was washed twice with 70% isopropanol. The final step was the resuspension of the pellet in 100 µL of RNase-free water. RNA was stored at -80°C for further use.

3.2 LC-MS analysis of small molecules and short RNA

3.2.1 Small molecules extraction and LC-MS analysis

Frozen Expi293F cells were combined with a solution comprising acetonitrile, methanol, and water in a ratio of 40:40:20 (V/V/V), and 100 nM ARCA cap analog (Jena Bioscience, #NU-855S) were added as an internal standard. The sample was vortexed for 30 s and centrifugated at $14000 \times g$ for 15 min, followed by filtration through a 0.2 μm membrane (Thermo Fisher Scientific, #F2517-7). The samples were analyzed on Xevo G2-XS QToF mass spectrometer (Waters). Sample preparation and analysis were conducted by my colleague Anton Škríba.

3.2.2 LC-MS analysis of short RNA

For Ap₂A identification, sRNA was digested by 50 U of Nuclease P1 from *Penicillium citrinum* (New England BioLabs (NEB), #M0660S) and 2.5 U of Shrimp Alkaline Phosphatase (rSAP, NEB, # M0371L) per 100 μg RNA in 50 mM ammonium acetate (pH 4.5) buffer for 1 h at 37°C. RNA samples incubated without Nuclease P1 were used as control. After the digestion, the samples were purified by solid phase extraction on Oasis WAX 1cc columns using Otto Positive Pressure Manifold instrument (Waters) to remove free nucleosides. Samples were evaporated on SpeedVac Concentrator and dissolved in 8 μL of 10 mM ammonium acetate in 66% acetonitrile (V/V). 5 μL was analyzed on SCIEX QTRAP 6500+ UPLC-MS instrument with ACQUITY™ Premier BEH Amide VanGuard™ FIT Column (1.7 μm , 2.1mm X 150 mm, 1/pk, Waters #186009509). Ap₂N were detected in positive multiple reaction monitoring mode (MRM), following MRM transitions of Ap₂A [M+H⁺] at m/z 677.0 into 428.0 (ADP) and 250 (adenosine), Ap₂C [M+H⁺] at m/z 653.0 into 428.0 (ADP) and 136 (adenine), Ap₂G [M+H⁺] at m/z 693.0 into 428.0 (ADP) and 136 (adenine), and Ap₂U [M+H⁺] at m/z 654.6 into 136.0 (adenine) and 119.0 (ribose). For quantification, the standard addition method was used. Sample preparation and analysis were conducted by my colleague Ondřej Nešuta. The Ap₂A quantification was done in multiple biological replicates and expressed as mean of fmol of Ap₂A per 100 μg of RNA.

3.3 Preparation of radiolabeled 35-mer RNA

3.3.1 In vitro transcription with T7 RNA polymerase (RNAP)

In vitro transcription was performed using a modification of a previously published method⁸⁶ in a 50 μL mixture containing: 80 ng μL^{-1} of template DNA (Table 2) (35A for Ap₂₋₅A, Gp₃₋₄A, NAD(H), CoA, ADP-ribose and m⁷Gp₃A, or 35G for Gp₃₋₄G, Ap₃₋₅G, and m⁷Gp₃G), 1 mM UTP, 1 mM CTP, 1 mM ATP, 0.8 mM GTP and 0.5 μL α ³²P GTP (activity: 9.25 MBq in 25 μL), 1.6 mM of Np_nNs (Ap₂₋₅A, Gp₃₋₄G, Ap₃₋₅G, m⁷Gp₃G, m⁷Gp₃A, ADP-ribose), or 8 mM of cofactors (3'-dpCoA, NAD and NADH), 5% DMSO, 0.12% triton X-100, 12 mM DTT, 4.8 mM magnesium

chloride and 10× reaction buffer for T7 RNAP (40 mM Tris-HCl, 6 mM MgCl₂, 1 mM DTT, 2 mM spermidine, pH 7.9 at 25 °C) and 62.5 units of T7 RNAP. The mixture was incubated for 2 h at 37 °C. All chemicals were either purchased from Merck or Jena Biosciences or Geni Biotech.

3.3.2 DNase I treatment

The resulting mixture was digested by DNase I (NEB) at 37°C for 1 hour, according manufacturer's instruction. The enzyme was heat deactivated at 75 °C for 10 min followed by purification on RNA miniQuick Spin Columns (Merck) or Zymo Clean RNA Clean & Concentrator Kit for further use.

3.3.3 RNA 5'-polyphosphatase treatment

The products from the previous reaction were treated with 20 units of 5'-polyphosphatase (Epicentre) for 1 h at 37 °C. All samples were purified on RNA miniQuick Spin Columns (Merck) or Zymo Clean RNA Clean & Concentrator Kit for further use.

3.3.4 Terminator™ 5'-phosphate-dependent exonuclease treatment

The RNA was treated with 1 unit of Terminator™ 5'-phosphate-dependent exonuclease (Epicentre) in reaction buffer A (supplied with the enzyme) at 30 °C for 1 h. All samples were purified on RNA miniQuick Spin Columns (Merck), or Zymo Clean RNA Clean & Concentrator Kit and concentration was measured using Nanodrop or Qubit.

3.4 Purification of recombinant enzymes

3.4.1 Purification of Plant Nudix enzymes (AtNUDT6, 7, 19 and 27)

Arabidopsis thaliana (At) NUDT6, 7, 19, and 27 genes were cloned into the pET26b vector with the C-terminal His tag (Novagen). Due to the problem with expressing the full-length AtNUDT27, it was cloned without the first 44 aa containing the chloroplast signal sequence. Proteins were overexpressed in *Escherichia coli* Rosetta DE3 cells grown for 16 hours in auto-inducing medium¹⁷⁴. Cells were lysed by sonication in buffer A (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10% glycerol and 5 mM β-mercaptoethanol) and recombinant proteins were purified on TALON Metal Affinity Resin (Clontech), eluted in buffer A supplemented with 250 mM of imidazol and fractionated by gel filtration chromatography on an ENrich™ SEC70 column (Bio-Rad). The purified proteins were concentrated to 50 μM in a buffer containing 20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 2 mM DTT, and 10% glycerol, flash frozen in liquid nitrogen, and stored at -80 °C. Protein concentration was determined at 280 nm using the molar extinction coefficient calculated based on protein composition with the ProtParam tool (<https://web.expasy.org/protparam/>). The purification of AtNUDT6, 7, 19 and 27 was

done by Joanna Kufel's group from the Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw.

3.4.2 Purification of mammalian enzymes

3.4.2.1 NUDT2 expression and purification

The human wild type full-length NUDT2 protein was generated by inserting NUDT2 gene between NdeI and HindIII cloning sites of the pET28a expression plasmid (Genscript). In this context, the NUDT2 protein is fused in-frame with an N-terminal His tag. The resulting recombinant plasmid (pET-Nudt2) was transformed into *Escherichia coli* BL21(DE3) cells. Two liters of culture were grown at 37°C in Luria-Bertani (LB) medium containing 0.03 mg/ml kanamycin until the OD₆₀₀ reached 0.5. The culture was adjusted to 0.1 mM isopropyl β-D-thiogalactoside (IPTG) and the incubation continued at 18°C for 20 h. The cells were harvested by centrifugation at 4000 × g for 15 min. The pellet was resuspended in 70 mL of lysis buffer A (20 mM HEPES, pH 7.5, 300 mM NaCl, 0.5 mM TCEP and 10% glycerol) containing one tablet of proteases inhibitor SIGMAFAST™ (Sigma-Aldrich #S8820) and cell lysis was achieved by adding lysozyme 1 mg/mL. NUDT2 was then purified on Fast Protein Liquid Chromatography (FPLC, ÄKTA Pure). The lysate was sonicated and centrifuged at 20 000 × g for 45 min at 4°C. The soluble extract was loaded on His Trap™ FF crude column, previously equilibrated with 8% buffer B (20 mM HEPES, pH 7.5, 300 mM NaCl, 0.5 mM TCEP, 10% glycerol and 250 mM imidazole). Elution of the protein was achieved using a linear gradient from 8 to 100% of buffer B. The purified protein was concentrated using Amicon Ultra filters with 10 kDa cut-off (Sigma-Aldrich, #UFC901024). Protein concentration was determined using DC™ Protein Assay Kit II (Biorad #5000112) using bovine serum albumin (BSA) as standard.

3.4.2.2 DXO expression and purification

Full-length mouse DXO recombinant protein carrying an N-terminal His-tag was expressed from plasmid pET26b in *E. coli* BL21 (DE3). Cells were grown at 37°C in LB medium containing 0.05 mg/mL kanamycin until the OD₆₀₀ reached 0.5. The culture was adjusted to 0.4 mM IPTG, and the incubation continued at 20°C for 20 h. The cells were harvested by centrifugation at 4000 × g for 15 min. The pellet was resuspended in 30 mL of lysis buffer (20 mM Tris-HCl pH 8, 300 mM NaCl and 10% glycerol) containing one tablet of proteases inhibitor SIGMAFAST™ and cell lysis was achieved by adding lysozyme 1 mg/mL. The lysate was sonicated and centrifuged at 20 000 × g for 45 min at 4°C. DXO was then purified on FPLC (ÄKTA Pure). The soluble extract was loaded on His Trap™ FF Crude column, previously equilibrated with 8% of elution buffer (20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM DTT, 10% glycerol and 250 mM imidazole). Elution of the protein was achieved using a linear gradient from 8 to 100% of buffer B. The protein was further purified by SEC in buffer C (200 mM NaCl,

20 mM Tris pH 8, 2 mM DTT, 10% glycerol) and concentrated using Amicon Ultra filters with 10 kDa cut-off. Protein concentration was determined using DC™ Protein Assay Kit II using BSA as standard.

3.4.2.3 NUDT12 expression and purification

The plasmid was generated by inserting NUDT12 gene between NdeI and HindIII cloning sites of the pET28a expression plasmid (Genscript). The plasmid was then transformed into *E. coli* BL21 (DE3) cells and one liter of culture was grown at 37°C in LB medium containing 0.05 mg/ml Kanamycin until the OD₆₀₀ reached 0.5. The culture was adjusted to 0.5 mM isopropyl β-D-thiogalactoside and the incubation continued at 20°C for 20h. The cells were harvested by centrifugation at 4000 × g for 15 min and the pellets were stored at –80 °C. All subsequent procedures were performed at 4°C. The pellet was resuspended in 50 mL of lysis buffer (150 mM NaCl, 50 mM Sodium phosphate buffer pH 8 and 10% glycerol) containing one tablet of proteases inhibitor SIGMAFAST™ and cell lysis was achieved by adding lysozyme 1 mg/ml for 1 h at 4°C. The lysate was sonicated and centrifuged at 20 000 × g for 45 min at 4°C. The soluble extract was load on His Trap™ FF Crude column, previously equilibrated with 4% elution buffer (200 mM NaCl, 50 mM Sodium Phosphate Buffer pH 8, 500 mM Imidazole and 10% glycerol). Elution of the protein was achieved using a linear gradient from 8 to 100% of buffer B. The protein was further purified by SEC in buffer (150mM NaCl, 50mM Tris-HCl pH 8, 5mM DTT). The purified protein was concentrated using Amicon Ultra filters with 3 kDa cut-off and concentration was determined using DC™ Protein Assay Kit II with BSA as a standard.

3.4.2.4 RLIG1 human ligase expression and purification

Human RNA ligase 1 (RLIG1) was expressed and purified by our postdoc Ondřej Nešuta. Briefly, coding sequence for RLIG1 was amplified from HEK293T cDNA by PCR using primers Rlig1Fwd clone and Rlig1Rev clone (Table 2). The PCR product was cleaved with Asel and XhoI and ligated into pET28a vector beforehand linearized by NdeI and XhoI restriction enzymes. Resulting pET28a_RLIG1 encoding for RLIG1 with N-terminal 6xHis tag was then transformed into production *E. coli* BL21 (DE3) strain. For RLIG1 expression, transformed *E. coli* was cultured 1L of LB medium supplemented with 0.05 mg /ml of kanamycin and incubated approx. 3h until the OD₆₀₀ reached value 0.5. At that point, the RLIG1 expression was induced by 1 mM isopropyl β-thiogalactopyranoside and incubation at 18°C for 20 h. Finally, the bacterial cells were harvested by centrifugation at 5000×g for 20 min and stored at -80°C for further protein purification.

Bacterial pellet corresponding 0.5l of bacterial culture was resuspended in the sonication buffer (50 mM sodium phosphate at pH 8, 300 mM NaHCl, 20 mM imidazole, 3 mM β-mercaptoethanol, 10 mM phenylmethylsulfonyl fluoride, 1X protease inhibitor cocktail (EDTA-free Protease Inhibitor Cocktail, Sigma Aldrich #04693116001), and 5% glycerol - 10

ml per 1 g of wet pellet and sonicated on ice for 15 min (10s ON/ 10s OFF). After the sonication, bacterial lysate was clarified by centrifugation at $7120 \times g$ and 4°C for 20 min and loaded on FPLC equipped with 1ml HisTrap crude FF column previously equilibrated with IMAC buffer (50 mM sodium phosphate at pH 8, 300 mM NaHCl, 20 mM imidazole, and 5% glycerol). After washing out unbound proteins, 6×His-RLIG1 was eluted in continuous imidazole gradient (20-500 mM). Collected fractions containing 6×His-RLIG1 were pooled, concentrated on Amicon Ultra-15 filters with 10 kDa cut-off and loaded on FPLC equipped with 1 ml HiTrap Q HP column previously equilibrated with Anex buffer (50 mM TrisHCl at pH 8 and 5% glycerol). After washing out unbound proteins, 6×His-RLIG1 was eluted in continuous NaCl gradient (0-800 mM) and collected fractions containing 6×His-RLIG1 were pooled, concentrated on Amicon Ultra-15 filters with 10 kDa cut-off. 4.8 mg of 6×His-RLIG1 were incubated with 25 U of thrombin from bovine plasma (Sigma-Aldrich, T6634-250UN) for 4.5 h at 22°C to cleave the 6×His tag. Afterwards, the RLIG1 was cleaned up again on 1 ml HiTrap Q HP column as described above. Finally, RLIG1 was buffer exchanged into 25 mM TrisHCl pH 7.6, 100 mM NaCl and concentrated on Amicon Ultra-15 filters with 10 kDa cut-off, and glycerol concentration was adjusted to 30%. The protein concentration was determined using DC™ Protein Assay Kit II with BSA as a standard.

3.5 Cleavage of capped RNA and small molecules by NudiX enzymes

3.5.1 Cleavage of capped RNA by *A. thaliana* NudiX enzymes AtNUDT6, 7, 19 and 27

The decapping assay was performed at 37°C in 10 μL reaction. To test the cleavage of the 5'-caps, the positive control contained 100 ng of the pure RNA (in vitro transcription, DNase I treatment, purified on RNA mini Quick Spin Columns), 1 μL of 10× buffer (100 mM Tris-HCl, pH 7.5, 1 M KCl, 20 mM MgCl_2 , 20 mM MnCl_2 , 20 mM DTT), and 500 nM of the recombinant AtNUDT6, 7, 19, or 50 nM AtNUDT27. Samples without enzyme were used as the negative control. The mixtures were incubated at 37°C for 1 h and purified on RNA miniQuick Spin Columns. Then 3 μL of the purified samples were mixed with 2× RNA loading dye (NEB) for analysis by 12% PAGE (600 V, 3.5 h). Typhoon FLA 9500 was used as a visualization imaging system.

3.5.2 Enzymatic assay and kinetics assay of *A. thaliana* NudiX enzymes AtNUDT6, 7, 19 and 27 on small molecules

The hydrolytic activities of AtNUDT proteins towards different Np_nNs (Ap_{2-5}A , Gp_{3-4}G , Ap_{3-5}G , $\text{m}^7\text{Gp}_3\text{G}$, $\text{m}^7\text{Gp}_3\text{A}$), 3'-dpCoA, ADP-ribose, and NAD(H) were performed using 500 nM of the AtNUDT6, 7, 19, and 27 and 400 μM of substrate in 10 mM Tris-HCl (pH 7.5), 100 mM KCl, 2

mM magnesium acetate, 2 mM DTT, at 37 °C for 1 hour. The enzyme was heat deactivated at 75 °C for 10 min followed by cooling on ice. The mixture was filtered through Microcon®Centrifugal Filter 10kDa (Sigma-Aldrich, #MRCPRT010) and was then analyzed by monitoring absorbance at 259 nm and retention time compared to standards on high-performance liquid chromatography (HPLC) using a Kinetex C18 column (4.6 × 150 mm, 5 µm, Phenomenex, #00F-4601-E0) at a flow rate of 1 ml/min. The mobile phase consisted of 100 mM triethylammonium acetate pH 7 (HPLC buffer A) and acetonitrile (HPLC buffer B). Separation was achieved using a 25-minute gradient: 100% A for 2 min; linear decrease to 95% A over 4 min; linear decrease to 70% A over 9 min; maintaining 70% A for 3 min; returning linearly to 100% A over 1 min. Chromatograms were integrated, and the percentage of substrate cleavage was calculated from corresponding AUC compared to the control without enzyme.

3.5.3 NUDT2 activity on free Ap₄A and Ap₄A-RNA

NUDT2 hydrolytic activity was performed using 50 nM of the enzyme and 200 µM of Ap₄A. The samples were analyzed as described above. The decapping assay of Ap₄A-RNA by NUDT2 was performed at 37°C for 30 minutes in 10 µL reaction, containing 100 ng of Ap₄A-RNA and different concentrations of NUDT2, respectively 50, 100, 500 nM and 1 µM. Samples were then analyzed with the same method described above for the cleavage of capped RNA.

3.5.4 NUDT12 and DXO activity on capped RNA and small molecules

The same methods described before were used to test the cleavage of both 5' RNA caps (100 ng) by NUDT12 (1 µM) and DXO (200 nM) and 200 µM of small molecules (Ap₂A and NAD).

3.5.5 Calculations of RNA decapping assay

The experiments were performed in triplicates, and the amount of capped RNA was calculated from PAGE analysis using the software ImageJ. Areas under the peaks (AUC) corresponding to the 5'-capped RNA (AUC_{ctrl}) in the control without enzyme and 5'-capped RNA (AUC_{cap}) in samples with enzyme were calculated. The percentage of 5'-capped RNA species was calculated according to: $(A_{cap}/A_{rp}) \times 100$.

3.5.6 Kinetic parameter calculations

Kinetic parameters of AtNUDT enzymes were calculated from experiments using increasing concentrations of Ap₄A (0, 5, 10, 20, 30, 50, 70, and 100 µM) and 50 nM of AtNUDT6, 7, and 27. The 500 µL mixture was incubated at 37 °C, and aliquots of 100 µL were collected at 0, 2.5, 5, 10, and 15 min and immediately heated to 75 °C for 10 min to stop the reaction. All samples were analyzed using the HPLC method described above. Calculated initial reaction velocities were plotted against Ap₄A concentration and fitted to the Michaelis–Menten model to obtain K_m and V_{max} values using Origin software (Northampton, MA). Student t-

test and analysis of variance (ANOVA) with post hoc Bonferroni analysis were performed in Microsoft Excel to assess significant differences between measured parameters where needed.

3.5.7 In vitro RLIG1 ligase assay for production of Ap₂A-RNA

The RNA ligation using RLIG1 was performed in 10 µL reaction volume for 1 h at 37°C. After the preparation of ppp-RNA using IVT, with both 35A and 35G templates, 5'-polyphosphatase treatment followed to prepare p-RNA starting either with A or G. 0.1 µM of pRNA was incubated with 1 µM of RLIG1 and 200 µM ATP in 1x RNA ligation buffer (50 mM Tris-acetate pH 7, 5 mM MgCl₂, and 1 mM DTT). Samples without RLIG1 were used as a negative control. The enzyme was heat inactivated at 95°C for 2 min. The mixture was purified on Zymo Clean RNA Clean & Concentrator and used for LC-MS analysis described previously.

3.6 Capturing method for Ap₂A-RNA for Illumina RNA Sequencing (RNASeq)

3.6.1 T4 RNA ligase 2 truncated K227Q reaction for 35-mer capped RNA

1 µM of in vitro transcribed RNA (with 5'-triphosphate, monophosphate or capped with NAD, Ap₂A, m⁷Gp₃A, or ADPRibose) was incubated with 1 µL (200 U) of T4 RNA ligase 2 truncated K227Q (NEB, #M0351), 4 µL of 50% PEG8000, and 3 µM of RNA adaptor (Table 2) in 10X T4 RNA Ligase Reaction Buffer provided with the enzyme. The ligation reaction was performed over-night at 16°C in 20 µL followed by purification on RNA miniQuick Spin Columns. Sample without T4 RNA ligase 2 were used as a negative control. Then 10 µL of the purified samples were mixed with 2x RNA loading dye and analyzed by 12% PAGE (600 V, 3.5 h). Typhoon FLA 9500 was used as a visualization imaging system.

3.6.2 mRNA Decapping enzyme (MDE), Terminator™ 5'-Phosphate-Dependent Exonuclease, *E. coli* Poly(A) Polymerase treatment for U1 capped RNA and short RNA from Expi293F cells

Either 1 µg of U1-RNA or 5 µg of the short fraction of RNA isolated from Expi293F cells were used for RNA sequencing. All samples were treated with mRNA decapping enzyme (MDE, NEB, #M0608) using the standard decapping protocol provided by the manufacture. After each reaction, samples were purified using Zymo Clean RNA Clean & Concentrator. Samples were further treated with Terminator Exonuclease (Lucigen #TER51020) for 1 h at 30°C using 10X Buffer A provided by supplier. Next, samples were treated for 1 h at 37°C with *E. coli* Poly(A) Polymerase (NEB #M0276) using a 3:7 ratio of ATP and Cordycepin (C9137-1MG). The last step of the protocol is the T4 RNA ligase 2 truncated K227Q treatment

described previously. For negative controls, NUDT12 was added prior to all treatments described above at 37°C for one hour, using 5 µM of enzyme and a cleavage buffer containing 10mM Tris-HCl pH 8.5, 100mM KCl, 2mM DTT, 2mM MgCl₂ and 2mM MnCl₂. cDNA was prepared using SuperScript™IV First-Strand Synthesis System (Thermo Fisher Scientific #18091050) using 3' RACE Adaptor (Table 2). Samples were then used for qPCR, Sanger sequencing or RNA sequencing.

3.6.3 Final PCR for Illumina RNA sequencing

For the enrichment of the final product 10 µL of the cDNA were used for PCR reaction using Phusion® High-Fidelity DNA Polymerase (NEB, #M0530). The sequences of the primers used for PCR are given in Table 2. The thermal cycling conditions were composed of initial denaturation step at 95 °C for 30 seconds, 30 cycles of 95 °C for 30 seconds, 63°C for 30 seconds, followed by final extension 72°C for 1 minute. All samples were further purified using Agencourt® RNAClean™ XP (Beckman Coulter #A66514). Samples were analyzed on Agilent TapeStation and the concentration was measured using Qubit and Nanodrop.

3.7 RT-qPCR and Western Blot analysis

3.7.1 NUDT12, DXO and RLIG1 enzyme expression in Expi293F cells

The cDNA preparation was performed using reverse transcription from 500 ng of total RNA, control or menadione-treated, using LunaScript® RT SuperMix Kit (NEB, #E3010) in 20 µL reaction volume according to manufacturer's instructions. The cDNA was then diluted by water to a volume of 50 µL and used for qPCR. The experiments were performed in technical triplicates on the LightCycler-® 480 Instrument II (Roche) using the Luna® Universal qPCR Master Mix (NEB, #M3003) with 2 µL cDNA per well. The program was as follows: Initial denaturation 1 min at 95 °C, 45 cycles of 95 °C for 15 seconds followed by annealing/extension at 55°C for 20 seconds. The Cq were called using the Abs Quant/2nd Derivative Max method in the LightCycler -® 480 software. The sequences of the primers used for qPCR are given in Table 2.

3.7.2 Ap₂A-U1, m⁷Gp₃A-U1, ppp-U1 as controls for qPCR, Sanger sequencing and RNA sequencing protocol

The cDNA was prepared using SuperScript™ IV Reverse Transcriptase (Thermo Fisher Scientific, #18090010) according to the manufacturer's instruction using U1 Reverse primer (Table 2). Then the cDNA was diluted with water to 50 µL and 2 µL were used for qPCR. The experiments were performed in technical triplicates using the same procedure described above.

3.7.3 Western Blot

For Western Blot analysis, control and menadione-treated Expi293F cells were harvested by centrifugation at 1200 rpm for 5 minutes and proteins were extracted using RIPA Lysis and Extraction Buffer (Thermo Fisher Scientific, #89900) containing 1x protease inhibitors (EDTA-free Protease Inhibitor Cocktail, Sigma Aldrich #04693116001). The protein concentration of the lysate was then measured using DC™ Protein Assay Kit II (Bio-Rad #5000112) on Tecan (Schoeller). Proteins were then loaded on 10-well protein gels (Bio-Rad, Mini-PROTEAN TGX) and run at 120V (60 µg for NUDT12 and 30 µg for DXO and GAPDH). The transfer of the proteins was done at 30V for 2 hours using PVDF membranes (Thermo Fisher Scientific). All primary antibodies were incubated with the membrane overnight at 4°C. For NUDT12 we used 1:500 dilution of Anti-NUDT12 antibody (abcam, #ab197310); for DXO we used 1:500 dilution of DOM3Z Polyclonal antibody (Proteintech Europe, #11015-2-AP); for GAPDH we used 1:5000 dilution of GAPDH Antibody (14C10) Rabbit mAb (Cell Signaling Technology #2118). Secondary antibody was added with a dilution of 1:10000 for 1 hour at room temperature (Peroxidase conjugated AffiniPure Goat Anti Rabbit IgG, Jackson Immuno research, #111-053-003). Blots were then developed using SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific, #34096) and visualized using Azure Biosystems (model c600).

3.8 Statistical Analysis

The experiments were conducted in triplicates, and statistical analysis was carried out using either a *t*-test or ANOVA (analysis of variance), with a significance value level set at $P < 0.05$.

Table 2. Names and sequences of oligonucleotides and primers.

Name	Sequence
35A	5'- CAGTAATACGACTCACTATTAGGGAAGCGGGCATGCGGCCAGCCATAG CCGATCA-3'
35G	5'- CAGTAATACGACTCACTATTGGGGAAGCGGGCATGCGGCCAGCCATA GCCGATCA-3'
DXOFwd	5'-CTTCTGCTCTGTGCTACGCA-3'
DXORev	5'-TGGCCAGGGCTGTGCATC-3'
Rlig1Fwd PCR	5'-GCCATACCTTTGGGCTCGACTA-3'
Rlig1Rev PCR	5'-GCTGGTTTGAAGTCCTCCTCAAC-3'
Rlig1Fwd clone	5'- AGGTATTAATATGAAGCGCTTGGGCTC-3'
Rlig1Rev clone	5'- AGTCTCGAGTTATACATCAAATATTATATCTTTGAGTC-3'
RNA adaptor	5'-CAAAGCACUUGCACGUCAGCCCCCCCCCCCCCCCCCCCCCCC-3'
U1 RNA	5'- AGGCTTACCTGGCAGGGGAGATACCATGATCACGAAGGTGGTTTTCCCA GGGCGAGGCTTATCCATTGCACTCCGGATGTGCTGACCCCTGCGATTTC CCCAAATGTGGGAAACTCGACTGCATAATTTGTGGTAGTGGGGGACTGCG TTCGCGCTTTCCCCTG-3'
3' RACE Adaptor	5'-GCGAGCACAGAATTAATACGACTCACTATAGGT12VN-3'
Adaptor Fwd	5'-CAAAGCACTTGCACGTCAG-3'
3' RACE outer primer	5'-GCGAGCACAGAATTAATACGACT-3'
U1Fwd	5'-CTTACCTGGCAGGGGAGA-3'
U1Rev	5'-CTACCACAAATTATGCAGTCGAG-3'
GAPDHFwd	5'-GAGTCAACGGATTTGGTCGT-3'
GAPDHRev	5'-TTGATTTTGGAGGGATCTCG-3'
NUDT12Fwd	5'-CTGGTGAAGTCCCGAGAGAG-3'
NUDT12Rev	5'-TTGAGCTACAACCCCAGCTT-3'

4 Results

4.1 PART 1: Detection of dinucleoside polyphosphates as non-canonical RNA caps in eukaryotic cells

In the past decade, the existence of dinucleoside polyphosphate capped RNAs has been a subject of intense study. It has been demonstrated that they can be incorporated as non-canonical initiating nucleotides (NCIN) by T7 RNA polymerase and *E. coli* RNA polymerase. Their presence in bacteria has been reported not only by our group but also by another independent group and confirmed by several techniques^{86,87}. Given these findings, our aim was to investigate if they exist in eukaryotic cells as well.

During my PhD, two dinucleoside polyphosphate RNA caps were the primary focus in our lab, specifically Ap₄A and Ap₂A. In the case of Ap₄A-RNA, I will provide a brief overview of the topic since it was not the main subject of my research. However, I will focus on Ap₂A-RNA in the following paragraphs, one of the central topics of my PhD thesis.

4.1.1 Diadenosine Tetraphosphate (Ap₄A) serves as a 5' RNA cap in mammalian cells

Dinucleoside tetraphosphate Ap₄A, the most extensively studied dinucleoside polyphosphate, has been reported to be highly enriched during immune activation of mast cells¹⁷⁵. This response is triggered by the binding of Immunoglobulin E (IgE) to the high-affinity receptor FcεRI8 of the mast cell, leading to the phosphorylation of serine 207 of Lysyl-tRNA synthetase (KARS). Upon phosphorylation, KARS dissociates from the multi-tRNA synthetase complex and translocates to the nucleus, where it initiates the production of Ap₄A. Ap₄A can then interact with the transcriptional repressor histidine triad nucleotide binding protein (HINT1), facilitating the transcription of previously repressed genes¹⁷⁶. This activation results in a significant increase in the concentration of Ap₄A within the nucleus, where RNA transcription occurs¹⁷⁷.

Considering all the previous reported data, our group searched for Ap₄A-capped RNA in mammalian cells. We confirmed the presence of Ap₄A in human embryonic kidney cells (HEK293T) using liquid chromatography-mass spectrometry (LC-MS) (Figure 9A, 9B). This part of the work was done by Jiří Potužník and Ondřej Nešuta. In Chapter 4.2.2, I describe my contribution to this work, where I tested both NUDT2 and DXO as decapping enzymes for Ap₄A-RNA.

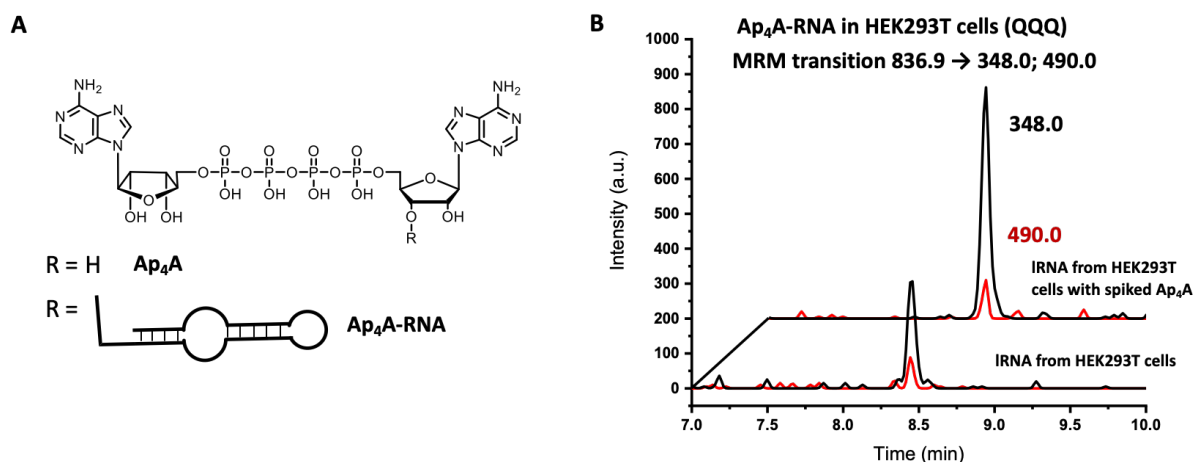


Figure 9. A) Structure of Ap₄A and Ap₄A-RNA. B) Multiple reaction monitoring (MRM) spectra of Ap₄A from IRNA after NuP1 and SAP treatment with and without spike of Ap₄A standard. Parent Ap₄A ion [M+H]⁺ at m/z 836.9 and two transitions at m/z 348.0 (black) and 490.0 (red) were monitored by MRM experiment with QQQ instrument.

4.1.2 Diadenosine diphosphate (Ap₂A) serves as 5' RNA cap in mammalian cells

4.1.2.1 In vitro incorporation of Ap₂A into RNA by T7 RNA polymerase

After the recent discovery that Np_nNs can be incorporated into RNA by T7 RNA polymerase during in vitro transcription reaction⁸⁶, we hypothesized that Ap₂A could have similar behavior. As expected, when Ap₂A was added into the reaction mixture, the in vitro transcription reaction produced a mixture of Ap₂A capped RNA (Ap₂A-RNA) and uncapped RNA (ppp-RNA) (Figure 10). After treatment with 5'-polyphosphatase and Terminator™ exonuclease the uncapped RNA was entirely digested, while Ap₂A capped RNA remained intact. The presence of the capped RNA was confirmed by PAGE analysis (Figure 10).

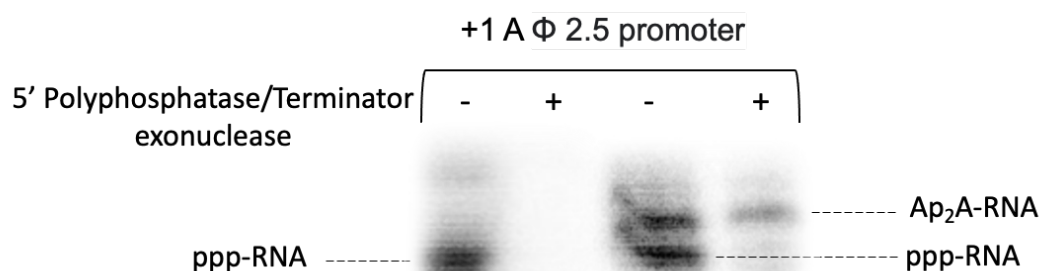


Figure 10. Polyacrylamide electrophoretic gel (PAGE—12%) analysis of the (^{32}P GTP labeled) *in vitro* transcription products (35 nt) with T7 RNAP followed by 5'-polyphosphatase and TerminatorTM exonuclease treatment.

4.1.2.2 Detection of Ap₂A capped RNAs using LC-MS in Expi293F cell line

Initially, we decided to explore the human cell line HEK293T (human embryonic kidney) as our model system. These cells are adherent and can be cultivated in Petri dishes. After several attempts to grow the cells in larger quantities to isolate enough short RNA for LC-MS, we realized that this approach was too demanding in terms of space and materials. Therefore, we decided to switch to Expi293F human cells, derived from the same 293 cell line. Expi293F cells can be cultured in suspension to larger quantities, in less time, and using fewer materials.

First, we isolated and washed the short and long fractions of RNA from Expi293F cell line to remove all non-covalently interacting molecules. After this, samples were digested by Nuclease P1 into nucleotides. Shrimp Alkaline Phosphatase (SAP) further digested the nucleotides into nucleosides, which were subsequently washed out by SPE, while leaving Ap₂A intact (Figure 11). This treatment reduces the matrix effect from co-eluting nucleotides in the Ap₂A elution time window. Next, my colleague Ondřej Nešuta established an LC-MS method for Ap₂A detection using a SCIEX QTRAP 6500+ UPLC-MS instrument with ACQUITYTM Premier BEH Amide VanGuardTM FIT Column (1.7 μm , 2.1mm X 150 mm).

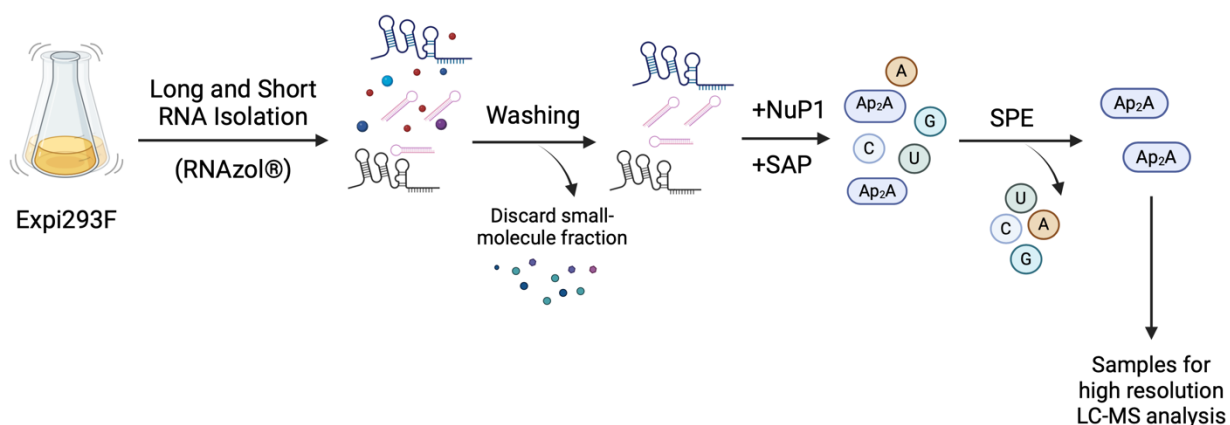


Figure 11. Scheme showing the RNA sample preparation for LC–MS analysis. Long and short RNA were isolated from Expi293F cells, washed from non-covalently bound small molecules and treated by Nuclease P1 (NuP1) and Shrimp Alkaline Phosphatase (SAP). After performing SPE (Solid Phase Extraction) samples were analyzed by LC–MS.

The possible fragments of diadenosine diphosphate (Ap₂A) that can be followed on the detector are adenosine diphosphate (ADP) and adenosine (A) with the MRM transition 677.0 → 428.0 and 677.0 → 250, respectively (Figure 12A). In the negative control samples, where the addition of Nuclease P1 was omitted, there weren't any signals of nucleotides or Np_nNs, which excluded the possibility of non-covalently bound contamination (Figure 12B). In the samples treated with NuP1, we confirmed the presence of the two possible fragments of Ap₂A in the short fraction of RNA (Figure 12C). Moreover, we observed an increase in the intensity of these two peaks at the same retention time when the standard addition method was used (Figure 12D). The used spike concentration of Ap₂A was selected to be similar to the peak area without the spike. Sample preparation and analysis were performed by my colleague Ondřej Nešuta.

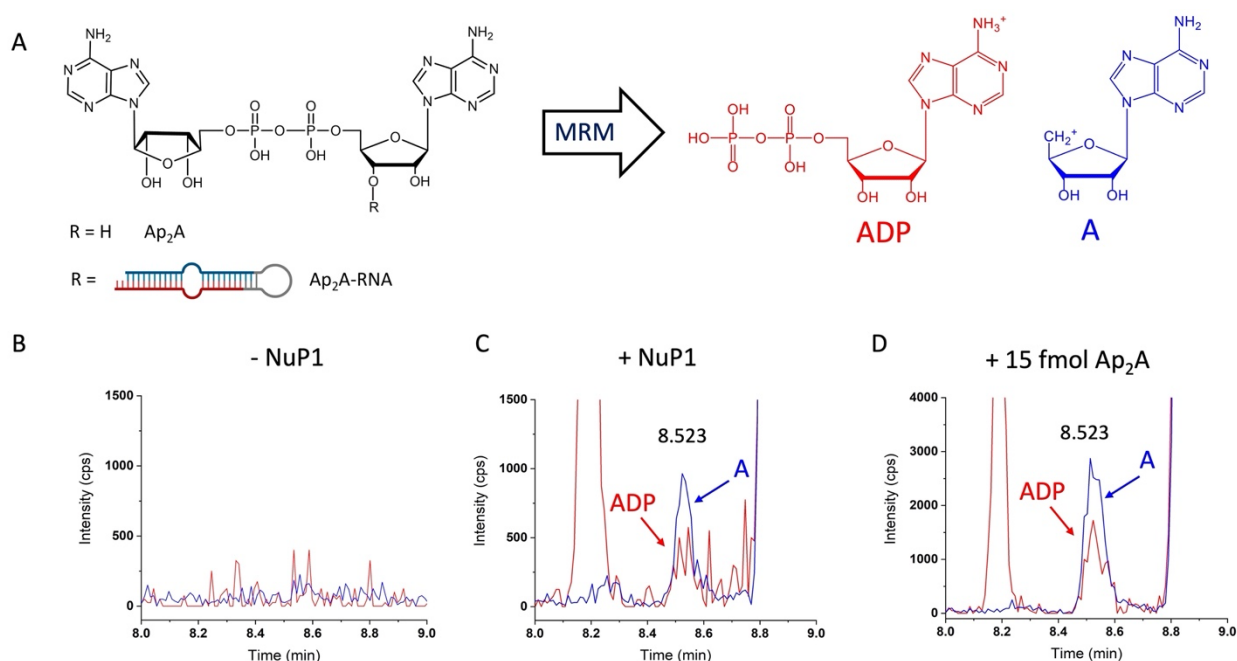


Figure 12. A) Structure of Ap₂A and Ap₂A-RNA; Multiple reaction monitoring (MRM) transition 676.1 → 428.0 was used for ADP (Adenosine diphosphate) and 676.1 → 250 for A+ (Adenosine). B) Spectra of selected fragments for Ap₂A from negative control (no NucleaseP1), C) non-spiked and D) 15 fmol spike of Ap₂A standard.

4.1.2.3 Quantification of Ap₂A-RNA after internal oxidative stress

To verify if Ap₂A-RNA levels in Expi293F cells can be influenced under different stress conditions, we performed some preliminary experiments where we exposed cells to various oxidative-stress-inducing reagents such as H₂O₂ (hydrogen peroxide), tBHQ (tert-Butylhydroquinone) and menadione. Finally, in further experiments, we used menadione, which is the most used agent for inducing oxidative stress in cells^{178–180}.

To assess the stress efficiency, we utilized the GSH/GSSG-Glo™ Assay to measure and quantify total glutathione (GSH + GSSG), as well as individual levels of GSSG and GSH-to-GSSG ratios within cultured cells. Glutathione, comprising both reduced (GSH) and oxidized (GSSG) forms, is vital for maintaining cellular health, with GSH predominantly present in healthy cells. Alterations in GSH and GSSG concentrations served as significant indicators in toxicological evaluations, highlighting potential oxidative stress consequences, such as apoptosis or cell death. In our case, the objective was to induce potential stress without inducing cell apoptosis in Expi293F cells. Therefore, we tested two mild concentrations of menadione, respectively 20 μM and 40 μM, and observed that at 40 μM menadione

concentration, the ratio between GSH and GSSG, was less than 50% compared to control cells, which indicated an increased concentration of the oxidized form of glutathione (GSSG), while the cells remained viable and alive (Figure 13).

Based on these results, in following experiments, we treated the cells with 40 μ M of menadione for 3 hours. Afterwards, we harvested the cells and isolated the RNA, which was further prepared for LC-MS measurements as described in Figure 11.

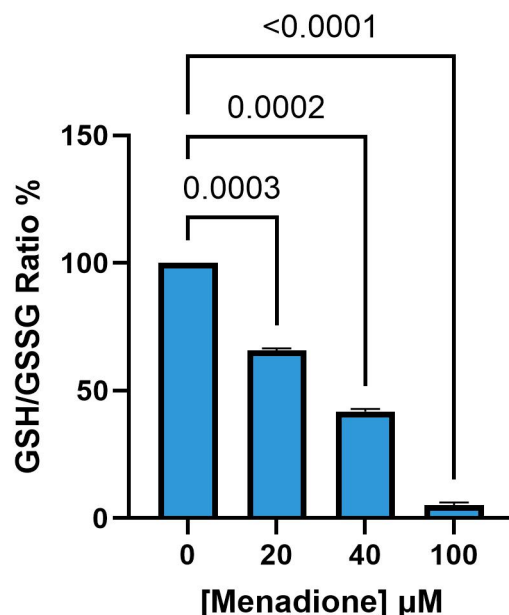


Figure 13. GSH/GSSG-Glo™ Assay on Expi293F cells with 0, 20, 40 and 100 μ M of menadione. The mean value $\pm$ SD was calculated from three replicates. Statistical differences were calculated using one-way ANOVA.

The quantification of Ap₂A-RNA in both control and menadione treated cells was performed from multiple biological replicates and expressed as the mean of fmol of Ap₂A per 100 μ g of RNA using the standard addition method (Figure 14A, 14B). Surprisingly, in both the short and long fractions of RNA, we observed approximately 60% increase of Ap₂A-RNA in cells exposed to menadione compared to the control (Figure 14C, 14D). The amount of Ap₂A was calculated to be 5.0 fmol per 100 μ g of RNA for the control cells and 8.2 fmol per 100 μ g of RNA for the menadione treated cells. In the case of long RNA, the amount of Ap₂A was calculated to be 0.5 fmol per 100 μ g of RNA for the control cells and 0.8 fmol per 100 μ g of RNA for the menadione treated cells.

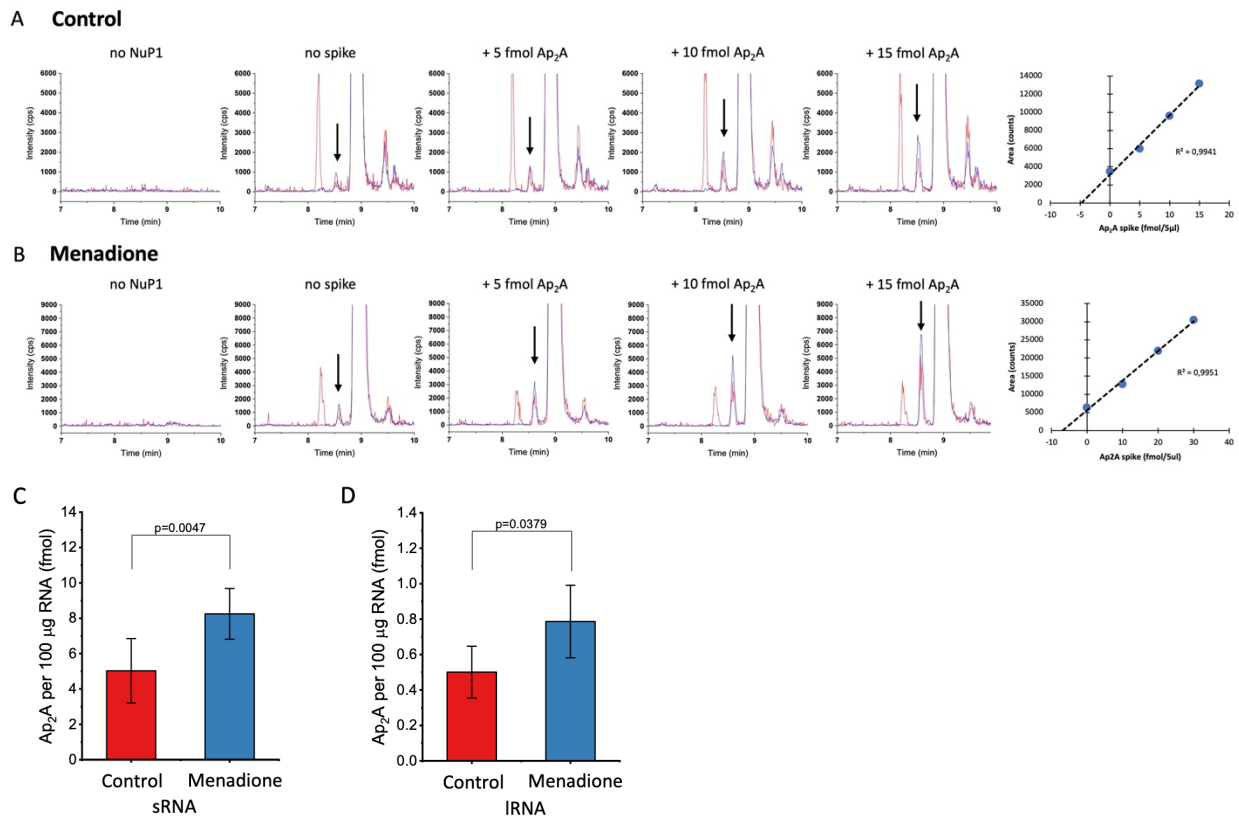
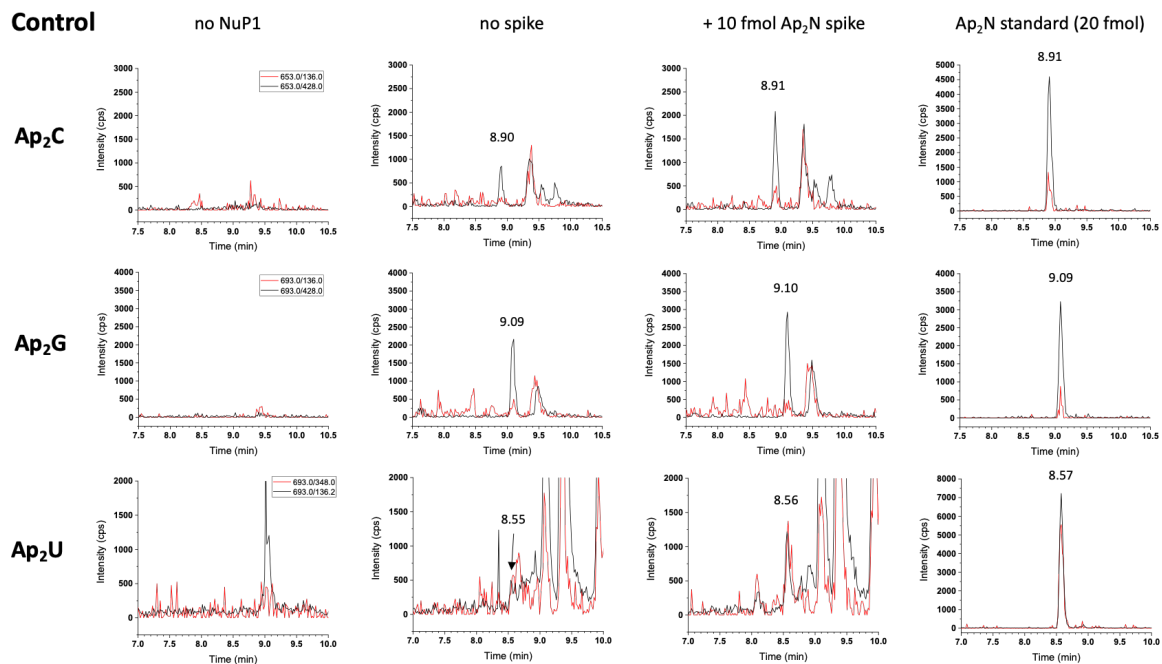


Figure 14. Calibration curve of Ap₂A-RNA quantification in digested RNA from A) control Expi293F cells and B) menadione treated cells. MRM spectra show selected fragments for Ap₂A from negative control (no NuP1), non-spiked and three spiked samples with 5, 10 and 15 fmol of Ap₂A standard. MRM transition 676.1 → 428.0 was used for ADP (adenosine diphosphate, blue) and 676.1 → 250 for A (adenosine, red). C, D) Quantification of Ap₂A-RNA in control and menadione cells for C) short RNA and D) long RNA from Expi293F. The bars represent mean values ± SD from three replicates. Statistical differences between the control and menadione were expressed as p-value calculated using t-test.

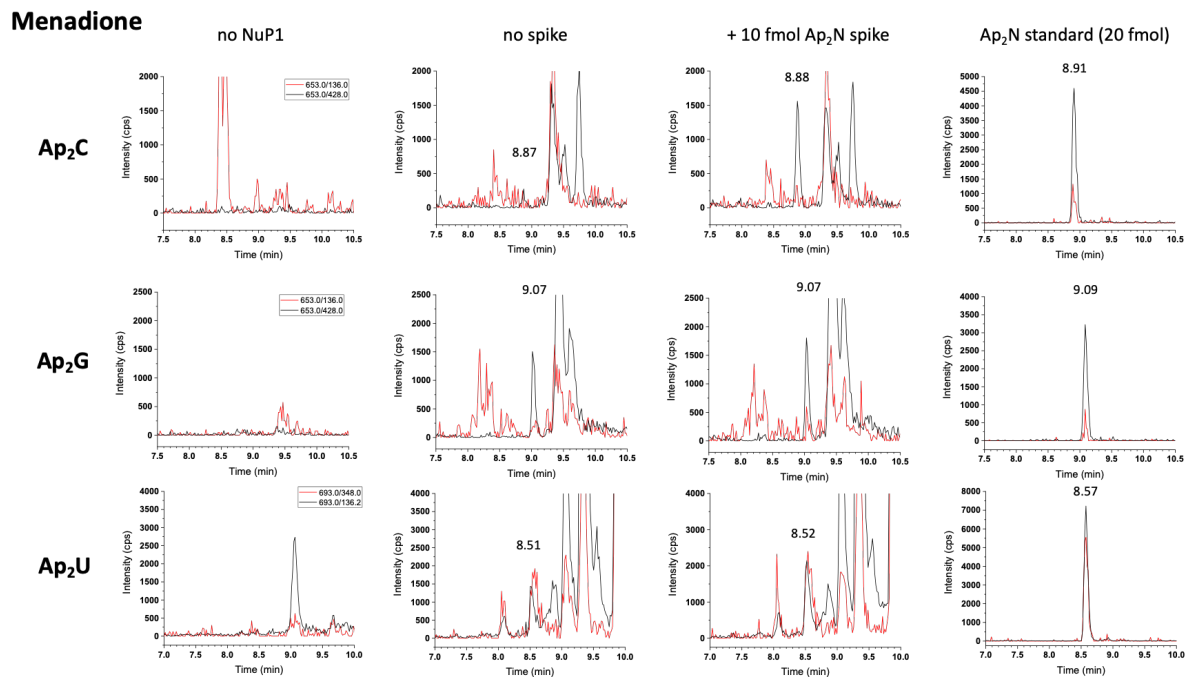
4.1.2.4 Detection and quantification of Ap₂N-RNAs in Expi293F cell line

After detecting and quantifying Ap₂A-RNA we hypothesized that other dinucleoside diphosphates such as Ap₂G, Ap₂U and Ap₂C may also be present as a part of RNA. To validate our hypothesis, we performed the same sample preparation and analysis for LC-MS. As expected, Ap₂G, Ap₂U and Ap₂C were found to be part of RNA in the isolated short fraction in both control and menadione treated cells (Figure 15A, 15B). However, after performing a relative quantification, we did not observe any statistically significant change between control and menadione samples (Figure 15C).

A



B



C

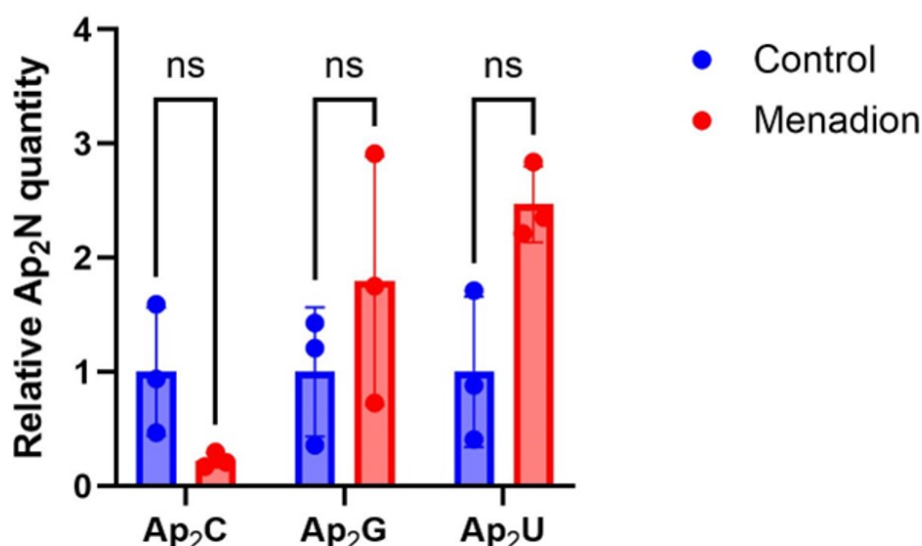
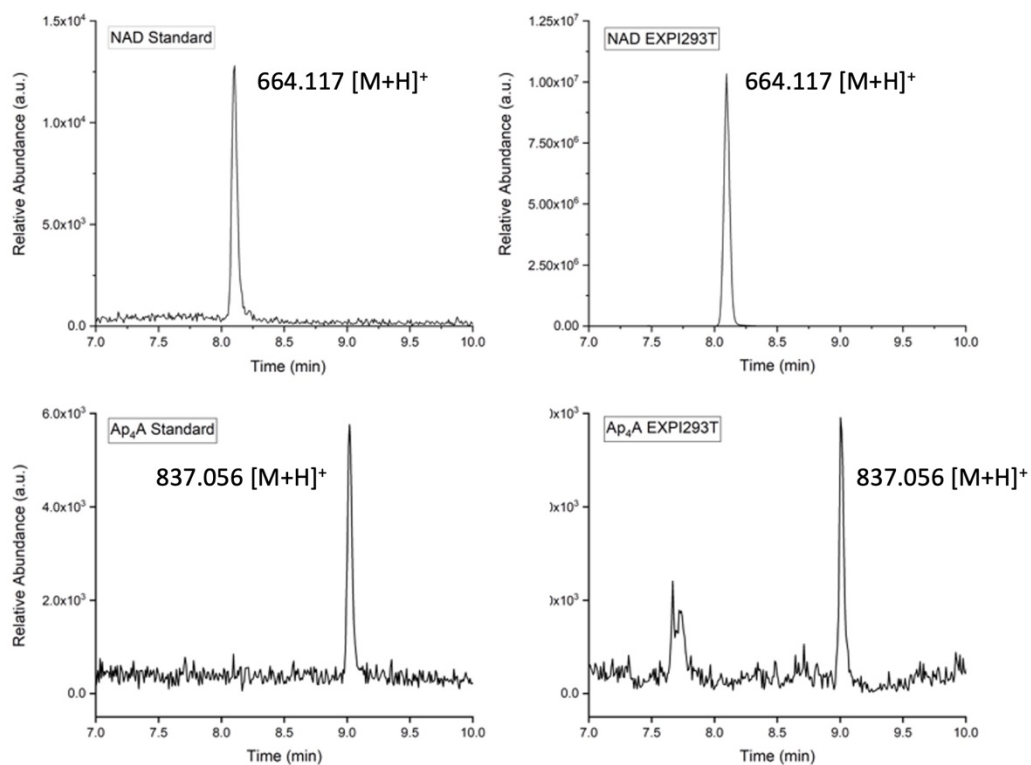


Figure 15. MRM spectra of selected fragments for Ap₂C, Ap₂G, and Ap₂U in A) control Expi293F cells and B) menadione treated Expi293F cells from negative control (no NuP1), non-spiked and 10 fmol spikes of Ap₂C, Ap₂G, and Ap₂U standards. C) Relative quantification of Ap₂N-RNAs in Expi293F. Statistical differences between the control and menadione were expressed as *p*-value calculated using *t*-test.

4.1.2.5 Detection of free NAD, Ap₄A and Ap₂N in Expi293F cell line

Knowing that RNA polymerase can incorporate Ap₂A as a NCIN in vitro, we wondered if Ap₂A can also be incorporated by RNA polymerase in vivo. For this, we first investigated the presence of Ap₂A as a free molecule in Expi293F cell line. Cells were harvested, and small molecules were extracted from the sample, then analyzed using Xevo G2-XS QToF mass spectrometer (Waters). Samples were measured and analyzed by my colleague Anton Škríba. While both NAD and Ap₄A appeared to exist as free molecules in human cells (Figure 16A), Ap₂A, Ap₂U, Ap₂C and Ap₂G were not detected by LC-MS as free molecules in Expi293F (Figure 16B). In this case, we can conclude that these molecules are probably not incorporated in vivo by RNA polymerase as NCINs, since they do not exist as free molecules in cells.

A



B

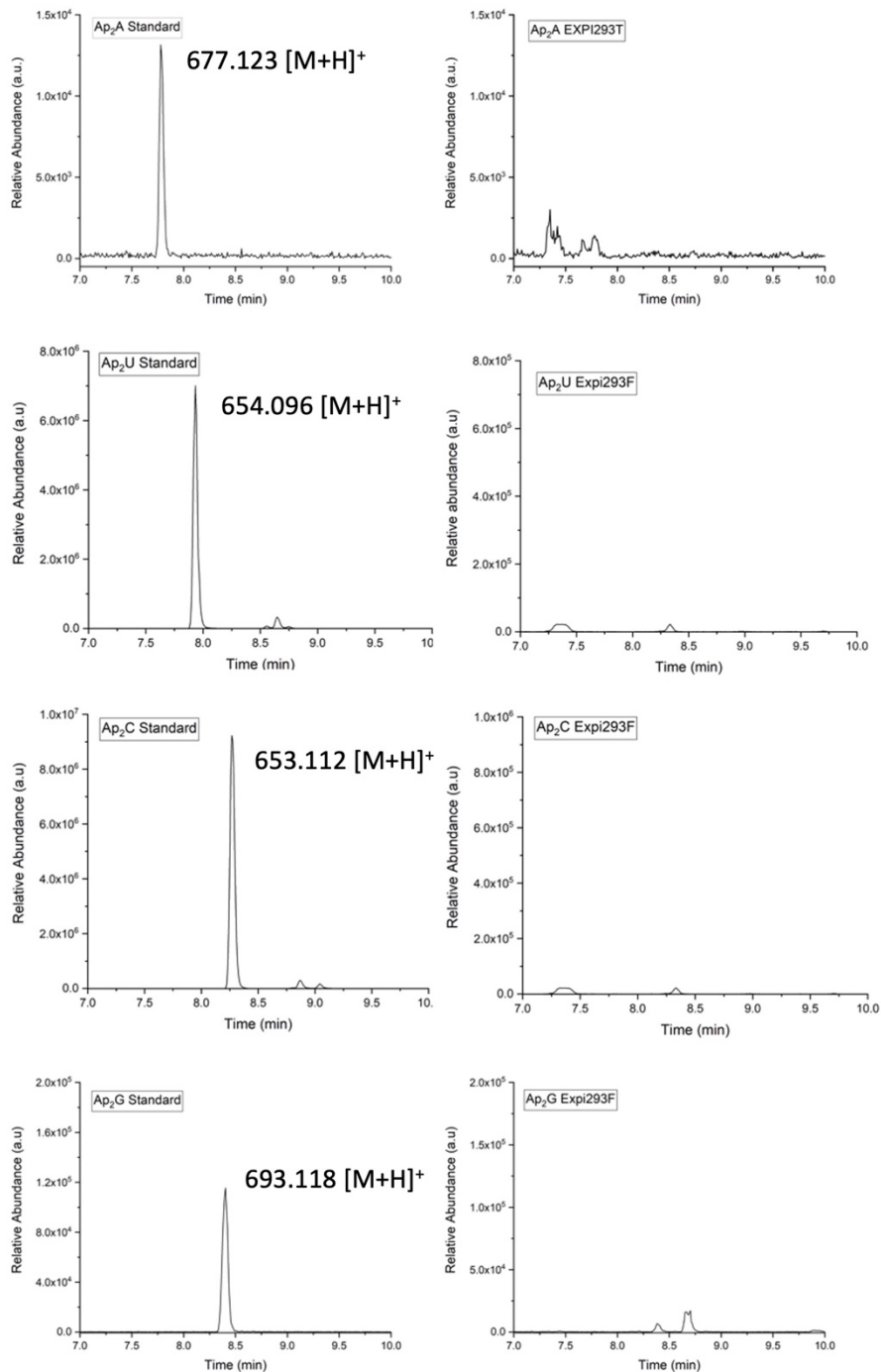


Figure 16. A) Extracted ion chromatogram of NAD and Ap₄A standards (m/z 664.117 and 837.056) and their relative presence in Expi293F cell line. B) Extracted ion chromatogram of Ap₂A, Ap₂U, Ap₂C and Ap₂G standards (m/z 677.123, 654.096, 653.112 and 693.118, respectively) and their absence in Expi293F cell line.

4.1.2.6 Human RNA ligase 1 (RLIG1) forms Ap₂N-RNA in vitro using ATP and p-RNA

Considering all the previous data, we hypothesized that there must be another enzyme responsible for the synthesis of Ap₂N-RNAs. While we were studying this non-canonical cap, a new paper reported the discovery of a new human ligase, RLIG1, involved in RNA repair. RLIG1 was shown to ligate two RNA oligonucleotides in a two-step reaction: first, the 5' monophosphate of one RNA is activated by adenylation, forming AppN-RNA, and second, AppN-RNA is ligated to 3' end of the second RNA¹⁷⁸. Since the AppN-RNA intermediate resembles Ap₂A-RNA and we observed an increase in Ap₂A-RNA under oxidative stress conditions in vivo, we speculated whether RLIG1 could be responsible for Ap₂N-RNA formation in our system.

To validate this hypothesis, my colleague Ondřej Nešuta purified RLIG1 enzyme and I was testing it in vitro using 35-mer monophosphate RNA (p-RNA) and ATP as substrates. As expected, after RNA digestion and LC-MS analysis, we confirmed the formation of Ap₂A-RNA from p-RNA starting with A by RLIG1 (Figure 17A). Moreover, considering the presence of Ap₂G-RNA in cells, we also used p-RNA starting with G and demonstrated the formation of Ap₂G-RNA (Figure 17B).

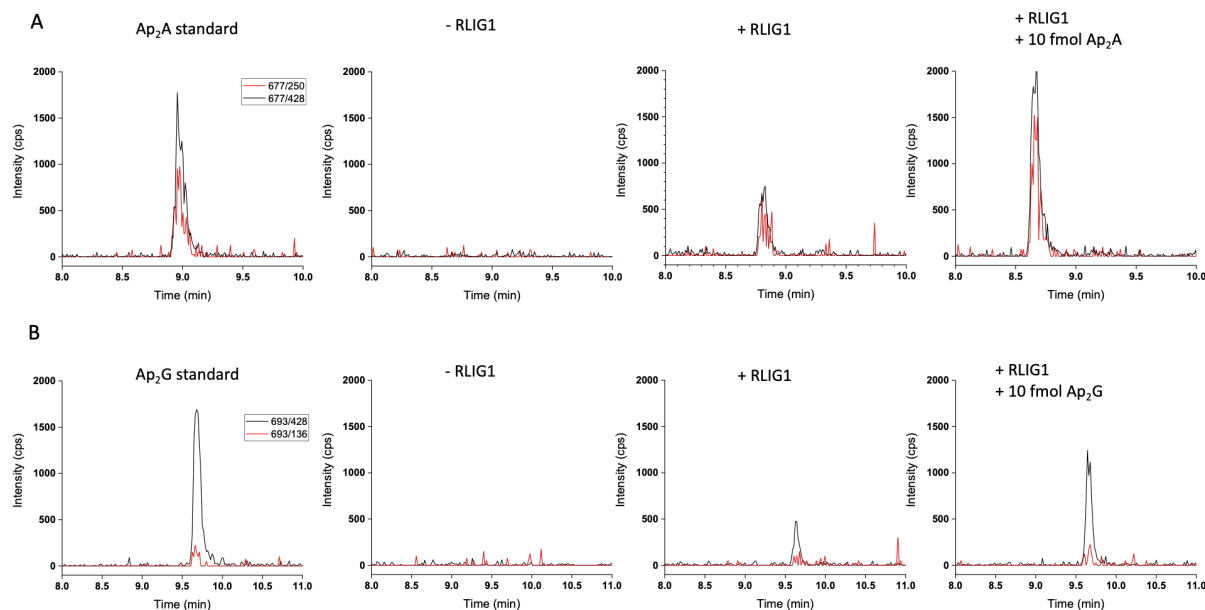


Figure 17. A) MRM spectra of selected fragments for Ap₂A in Ap₂A standard, p(A)RNA sample without RLIG1, with RLIG1, and with RLIG1 and 10 fmol Ap₂A standard addition. B) MRM spectra of selected fragments for Ap₂G in Ap₂G standard, p(G)RNA sample without RLIG1, with RLIG1, and with RLIG1 and 10 fmol Ap₂G standard addition.

4.1.2.7 Mammalian NUDT12 and mouse DXO are decapping enzymes for Ap₂A-RNA in vitro

Nudix hydrolases are enzymes responsible for degrading various molecules, including the canonical m⁷G capped RNA. The human genome encodes 24 enzymes from this family, and their conservation across species led us to speculate that mammalian Nudix enzymes might also act on non-canonical capped RNAs. Previous research has shown that mammalian Nudix proteins, such as NUDT12 and DXO, possess deNADding activity⁶⁴. NUDT12 hydrolyzes NAD-RNA into nicotinamide ribose monophosphate (NMN) and adenosine monophosphate (p-RNA), while DXO removes the entire NAD moiety from the 5' end of an NAD-capped RNA (Figure 18B).

Given the similarity between NAD-RNA and Ap₂A-RNA we hypothesized that these enzymes could also cleave Ap₂A-RNA in vitro. To test this, I isolated and purified recombinant human NUDT12 and mouse DXO (Figure 18A) and tested their activity on in vitro transcribed ³²P-labeled RNA capped with NAD and Ap₂A (Figure 18C). Both NAD-RNA and Ap₂A-RNA were completely cleaved by NUDT12, releasing monophosphate RNA (p-RNA) in both cases. In contrast, DXO removed the entire RNA cap, leaving the transcript one nucleotide shorter (Figure 18B, 18C).

We also compared the cleavage of NAD and Ap₂A capped RNAs by DXO and NUDT12. DXO cleaved both NAD and Ap₂A-RNA similarly (Figure 18D), while NUDT12 showed a significant preference for Ap₂A-RNAs (Figure 18E). Further testing of these enzymes' activity on NAD and Ap₂A as small molecules revealed slightly better activity on Ap₂A (Figure 18G, Supplementary Figure 1). DXO did not exhibit cleavage activity on free NAD or Ap₂A (Figure 18F, Supplementary Figure 1).

These results indicate that both NUDT12 and DXO cleave Ap₂A-RNA in vitro, with NUDT12 showing higher activity towards Ap₂A-RNA compared to its known substrate, NAD-RNA.

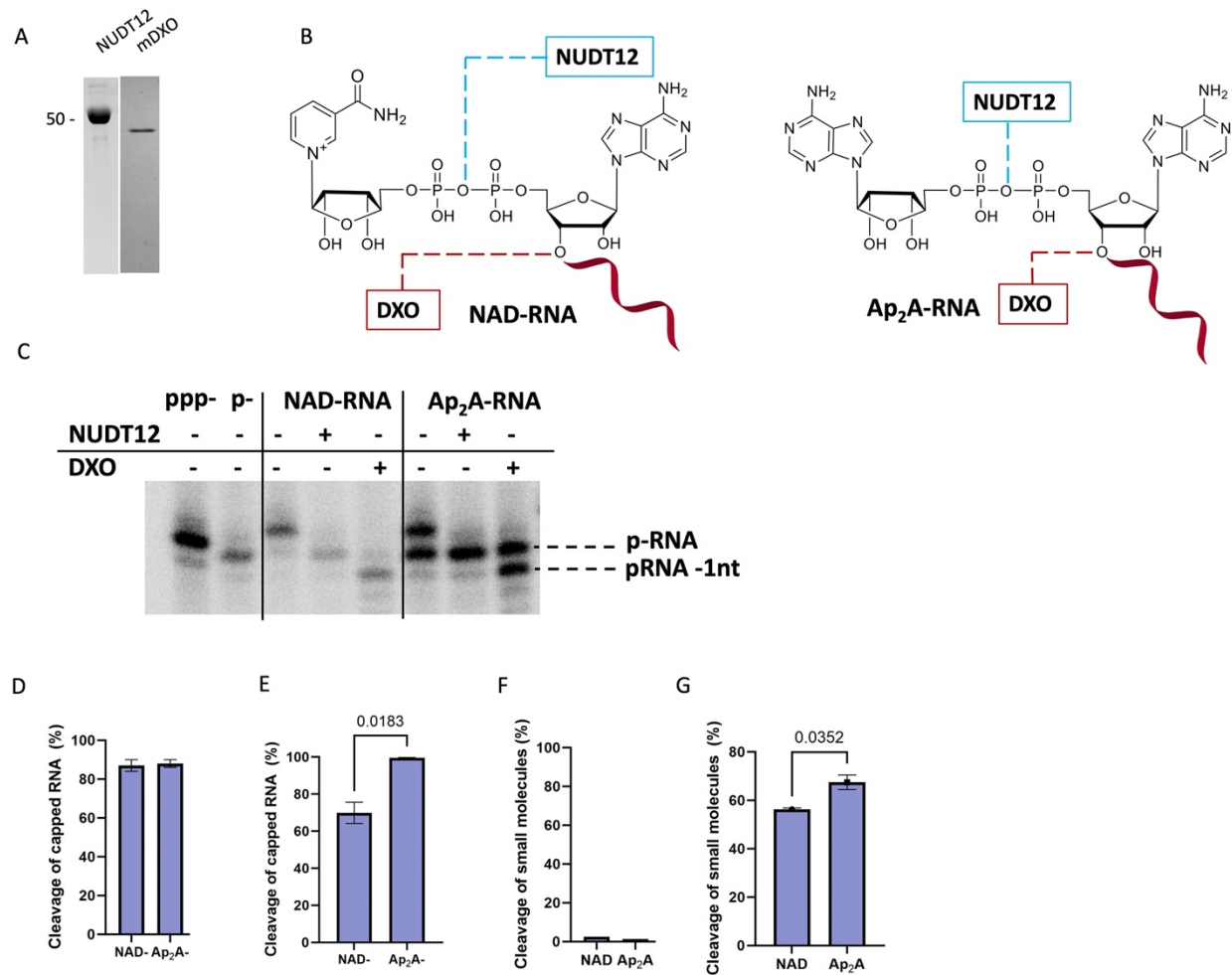


Figure 18. A) Recombinant NUDT12 and DXO proteins resolved by SDS/PAGE and stained with Coomassie blue. B) Schematic representation of the cleavage of NAD-RNA and Ap₂A-RNA by NUDT12 and DXO; C) NUDT12 and DXO, respectively 1 μ M and 200 nM, were incubated with 100 ng of ³²P-cap-labeled NAD-RNA and Ap₂A-RNA (35 nt) in decapping buffer for 1 hour at 37°C; ppp-RNA and p-RNA standards were run in parallel; D) Cleavage of DXO (200 nM) on capped RNA with NAD and Ap₂A (100 ng); E) Cleavage of NUDT12 (1 μ M) on capped RNA with NAD and Ap₂A (100 ng); F) Cleavage of DXO (200 nM) on NAD and Ap₂A (200 μ M) as small molecules; G) Cleavage of NUDT12 (1 μ M) on NAD and Ap₂A (200 μ M) as small molecules. The bars represent mean values $\pm$ SD from two replicates. Statistical differences were expressed as *p*-value calculated using *t*-test.

To investigate whether there is a correlation between the increased levels of Ap₂A-RNA in stressed cells and the expression of these decapping enzymes, we performed RT-qPCR on total RNA to measure the expression of DXO, NUDT12, and RLIG1 on a mRNA level (Figure 19A). Moreover, we performed Western Blot analysis on a protein level for DXO and NUDT12,

using GAPDH as control (Figure 19B). Nevertheless, our analysis did not reveal any change in their expression levels. This findings indicate that the observed changes in Ap₂A RNA cap levels by LC-MS are probably not caused by different amount of decapping enzymes under stress conditions.

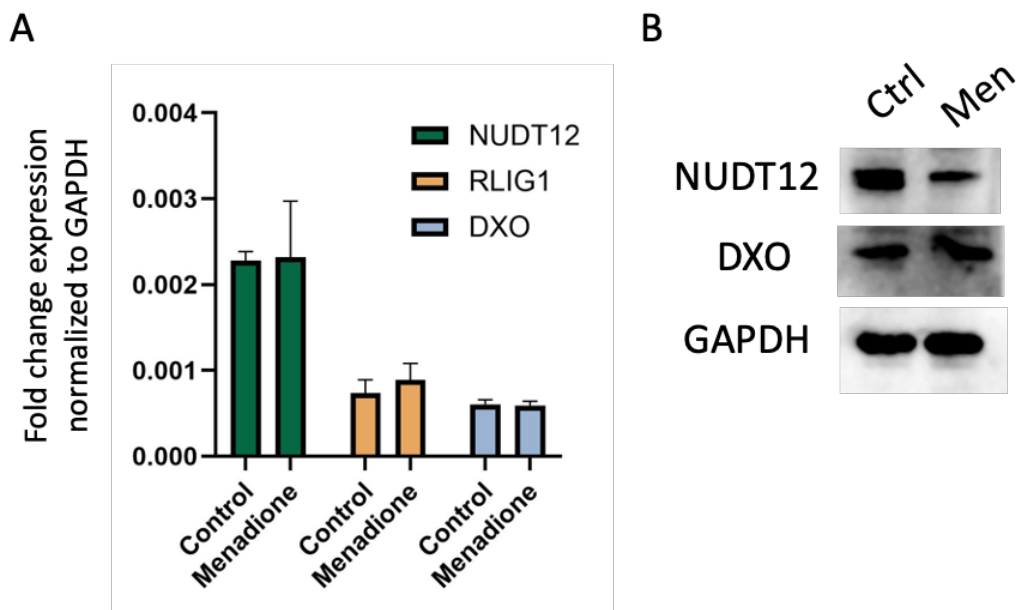


Figure 19. A) Detection of gene transcripts at the mRNA level on NUD12, RLIG1 and DXO by RT-qPCR on control and treated cells. B) Western Blot analysis for NUDT12 and DXO on control and treated cells. All transcripts were normalized by the housekeeping genes GAPDH. The bars represent mean values $\pm$ SD from three replicates.

4.1.2.8 Identification and classification of Ap₂A-RNA using T4 RNA ligase 2, truncated K227Q

After identifying and quantifying Ap₂A-RNA in Expi293F cell line, we wanted to determine the types of RNAs capped with Ap₂A molecule at the 5' end. To achieve this, we developed a method suitable for sequencing employing T4 RNA ligase 2, truncated K227Q.

We knew that it is possible to chemically prepare Ap₂A-DNA/RNA, by a reaction called pre-adenylation, and there is a ligase, respectively T4 RNA ligase 2, truncated K227Q, which specifically ligates this type of RNA or DNA to the 3' of another RNA. In our approach, the pre-adenylated RNA was specifically Ap₂A-RNA, while the other RNA molecule served as a 40 nt RNA adaptor (Table 2). Our hypothesis was that our construct of Ap₂A-RNA would be effectively ligated to the 3' end of another RNA by this ligase.

Initially, we assessed whether this ligase would selectively ligate Ap₂A-RNA to another RNA adaptor. First, we prepared in vitro transcribed 35mer RNAs and tested the activity of the ligase on ppp-RNA, p-RNA, NAD-RNA and Ap₂A-RNA (Figure 20A). We used radioactively labelled RNA, and we analyzed it on PAGE gel. As expected, only the 35 mer Ap₂A-RNA was successfully ligated to the RNA adaptor, forming an RNA product of approximately 80 nucleotides (Figure 20B). Moreover, when the adaptor was omitted from the reaction mixture, Ap₂A-RNA was the only RNA forming a circular product when treated with the ligase (Figure 20B).

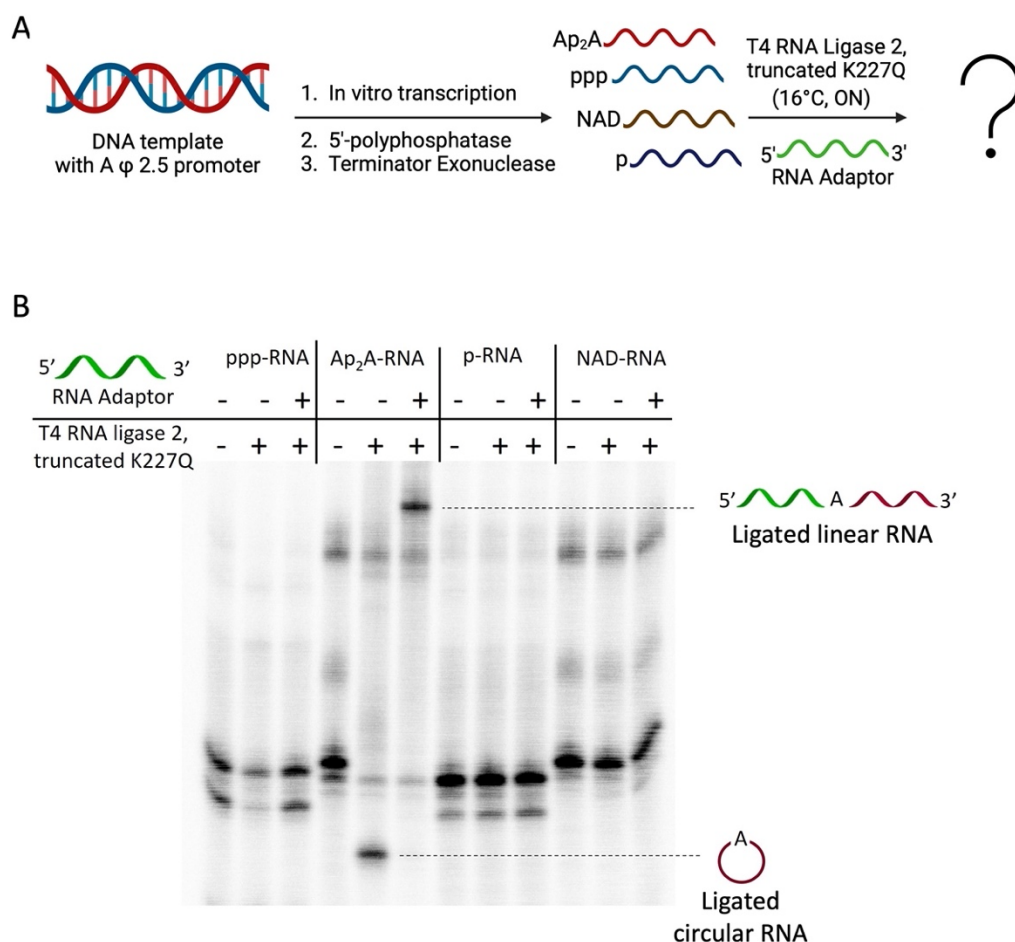


Figure 20. A) Scheme of ligation reaction in which Ap₂A-, ppp-, NAD- and p-RNAs were prepared by in vitro transcription using ³²P radiolabeled RNA; B) PAGE analysis of the ligation reaction using 100 ng of RNAs and 1 unit of T4 RNA ligase 2, truncated K227Q.

Considering the presence of other non-canonical RNA caps with structures similar to Ap₂A-RNA, such as ADPR-RNA and Ap₄A-RNA, we wanted to confirm that they would not react with the T4 RNA ligase 2, truncated K227Q. Additionally, we tested m⁷G capped RNA, the most

significant and abundant capped RNA in human cells. We showed that the only capped RNA that produced both ligated linear RNA and circular RNA was Ap₂A-RNA and the other capped RNAs were not substrates for this RNA ligase. Furthermore, when using NUDT12 as decapping enzyme for Ap₂A-RNA, the ligase failed to produce either of the two aforementioned products (Figure 21).

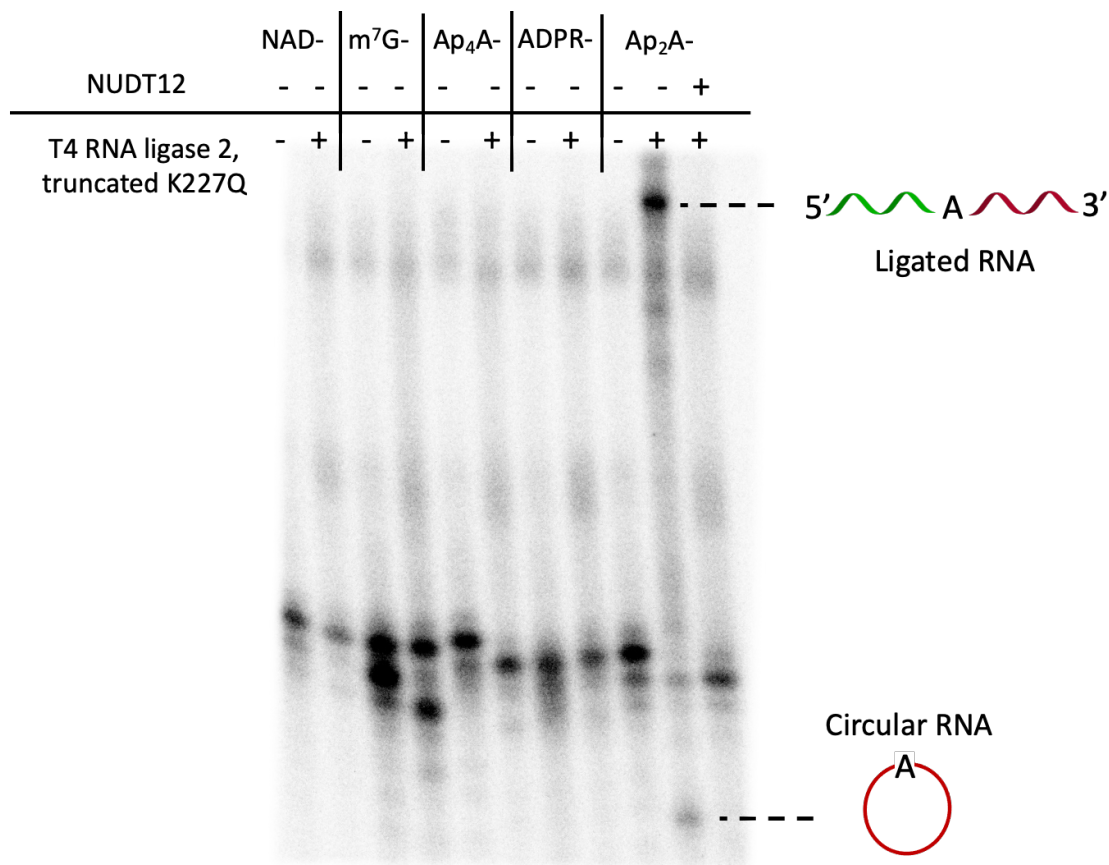


Figure 21. PAGE analysis of the ligation reaction using 100 ng of NAD-RNA, m⁷Gp₃A-RNA, Ap₄A-RNA, ADPR-RNA and Ap₂A-RNA and 1 unit of T4 RNA ligase 2, truncated K227Q; 1 μ M of decapping enzyme NUDT12 was used to confirm the cleavage of Ap₂A-RNA to p-RNA and the absence of ligation products.

After proving the selectivity of this enzyme our goal was to develop a sequencing protocol for the short, isolated RNA from cells. For this, we used model U1 RNA with the two most common caps (ppp-RNA and m⁷G-RNA) and for the negative control, Ap₂A-U1RNA was treated with NUDT12, to prove the decapping activity of the enzyme on Ap₂A-RNA. For the preparation of ppp-RNA, m⁷Gp₃A and Ap₂A-RNAs by in vitro transcription we used DNA template “U1 RNA” from Table 2. Since our primary objective was to apply this sequencing method to isolated short RNA from Expi293T cells, we included three additional reactions to avoid any nonspecific ligated products.

Firstly, all samples underwent treatment with MDE (mRNA decapping enzyme) which is selectively cleaving m⁷G-RNA and ppp-RNA to p-RNA, followed by Terminator Exonuclease treatment, which completely degraded all p-RNAs leaving Ap₂A-RNA intact. Next, we needed to install a Poly(A) strand at the 3' end of the construct because it is necessary for Illumina RNA sequencing. Moreover, using a mixture of ATP and Cordycepin we were able to blind the 3' end of the construct, preventing the formation of circular products or hetero concatemers. The final reaction involved the T4 RNA ligase 2, truncated K227Q enzyme, and the same RNA adaptor (Figure 22A). This experiment was analyzed by RT-qPCR and Illumina sequencing to confirm the proof of concept of this protocol. The RT-qPCR results demonstrated that Ap₂A-U1RNA was the only construct enriched using this protocol, compared to ppp-U1RNA and m⁷Gp₃A-U1RNA. Moreover, the negative control (Ap₂A-U1RNA cleaved by NUDT12) exhibited the same relative fold expression as ppp-U1RNA and m⁷Gp₃A-U1RNA (Figure 22B). Utilizing Illumina sequencing, we further confirmed the enrichment of Ap₂A-U1RNA reads compared to the negative control (Figure 22C). These results showed that our sequencing method, developed in vitro, is selective for Ap₂A-RNA and can be employed for in vivo short, isolated RNA.

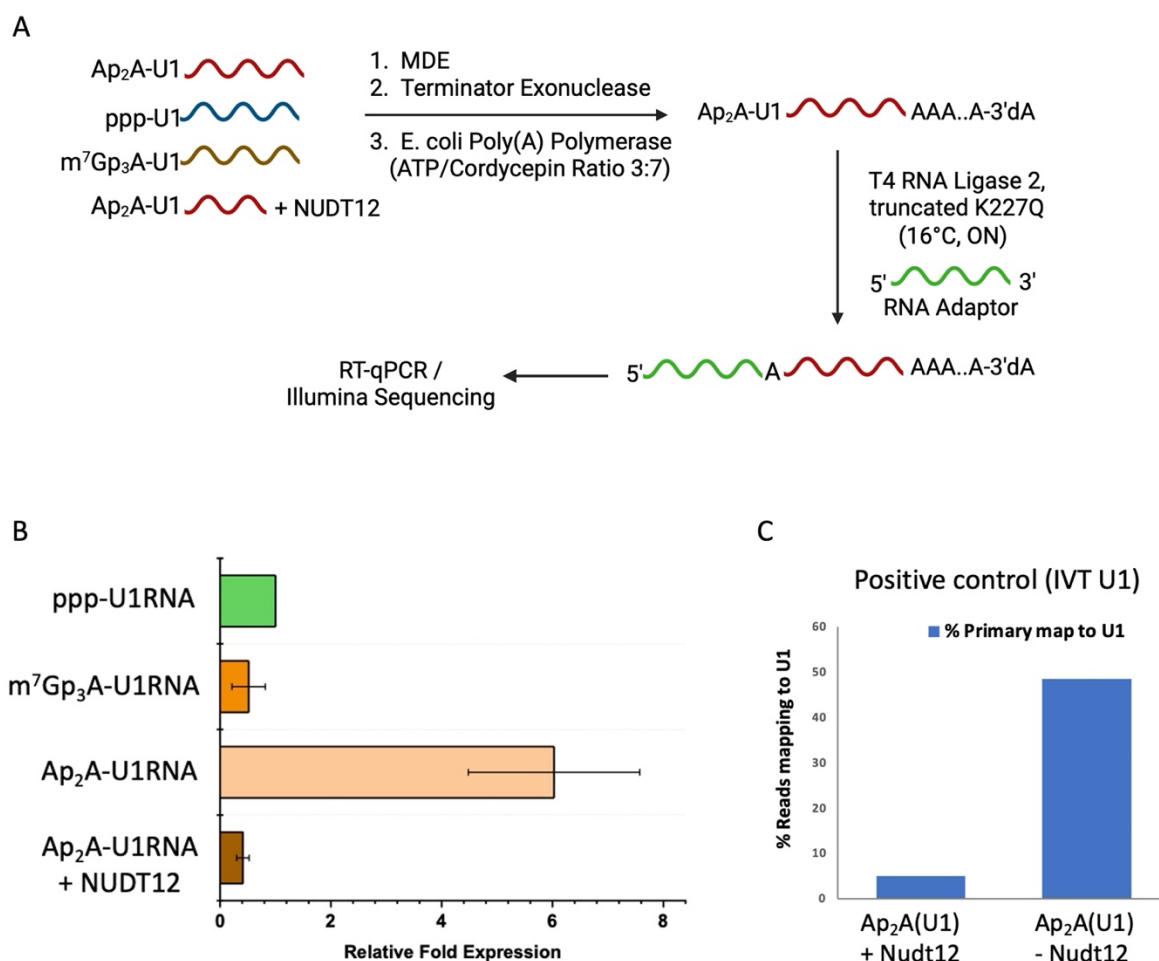


Figure 22. A) Scheme of sequencing protocol for RT-qPCR and Illumina Sequencing using ppp-U1RNA, m⁷Gp₃A-U1RNA, Ap₂A-U1RNA and Ap₂A-U1RNA + NUDT12 from IVT. B) RT-qPCR for ppp-U1, m⁷Gp₃A-U1, Ap₂A-U1, Ap₂A-U1RNA + NUDT12 normalized to pppU1-RNA using “Adaptor Fwd” and “U1 Rev” primers from Table 2. The bars represent mean values $\pm$ SD from three replicates. C) Illumina sequencing for Ap₂A-U1RNA and Ap₂A-U1RNA + NUDT12 samples mapped to U1 sequence.

Considering all the previous results, we applied the same protocol on the short fraction of RNA isolated from both control and menadione treated Expi293F cells. Having the advantage of the negative control samples treated with NUDT12, which cleaves Ap₂A-RNA to p-RNA, in the sequencing method, we were able to follow the enrichments in the samples

not treated with NUDT12, being sure that the reads were coming from Ap₂A capped RNA. The experiment was performed with three distinct biological replicates (Figure 23).

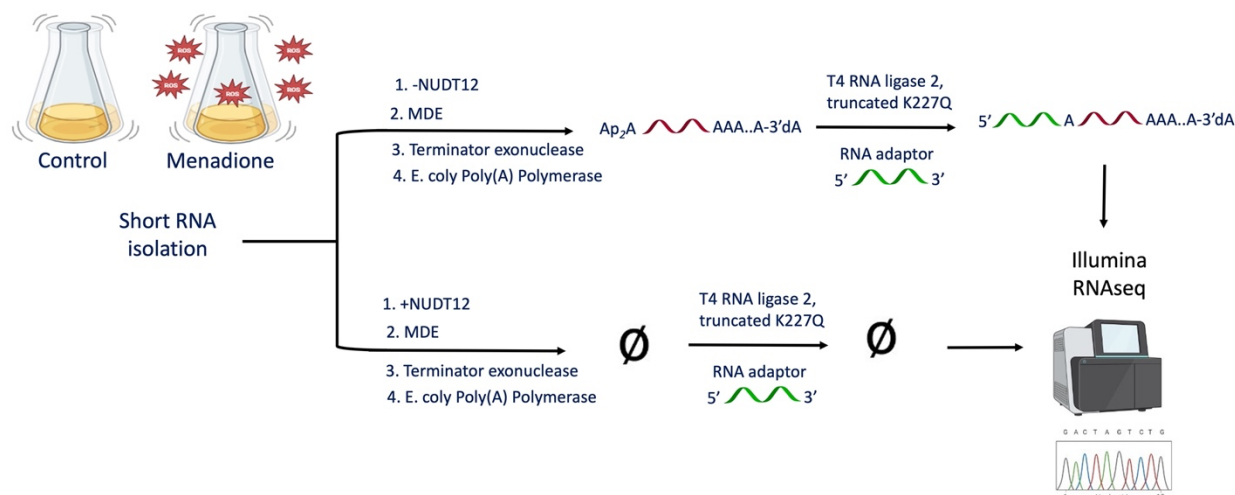


Figure 23. Scheme of sample preparation for Illumina RNA sequencing using the short fraction of RNA from both control and menadione Expi293 cells.

The sequenced libraries were mapped to the human genome, and the most significant reads we detected were 5.8S rRNA and U3 snoRNAs. 5.8S rRNA was enriched in both control and menadione samples, while SNORD3 and SNORA56 were enriched only in the control samples (Figure 24). SNORD3 and SNORA56 represent the same sequence at the same genomic location on the 17th chromosome and represent U3 snoRNA. Surprisingly, we did not observe any reads mapping to miRNAs or tRNAs.

In the next steps, we will need to confirm that the capped transcripts have indeed Ap₂A moiety using orthogonal methods, such as pull down with biotinylated probes specific to the 5.8S rRNA and to U3 snoRNAs, in combination with LC-MS analysis. Currently, all these experiments are still being optimized.

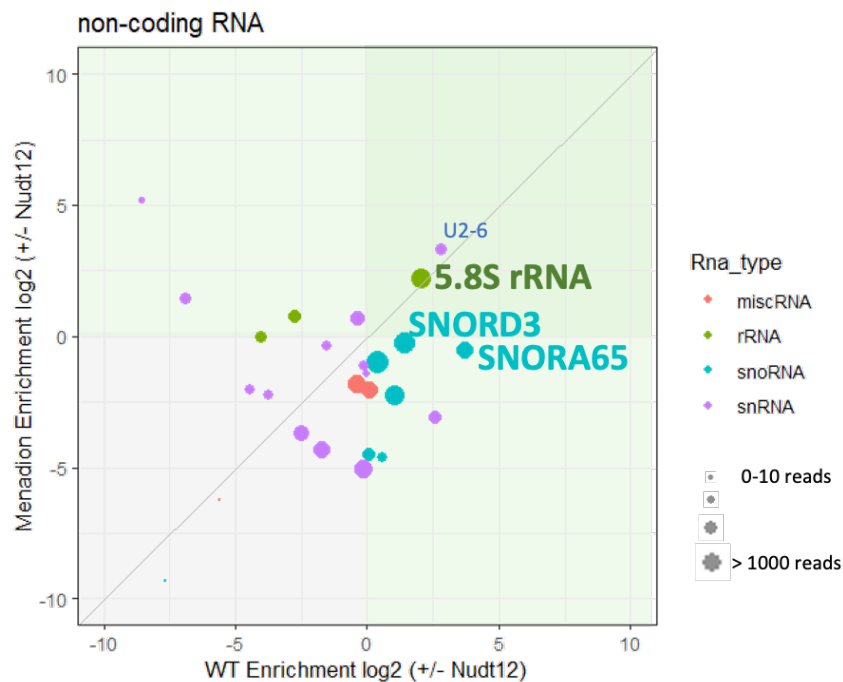


Figure 24. Scatter plot showing the enrichment of the reads (+NUDT12 vs -NUDT12) in control (x-axis) and menadione treated samples (y-axis). *Ap₂A* cap is present on 5.8S RNA, SNORA65 and SNORD3 (U3 snoRNAs).

4.2 PART 2: Decapping enzymes for non-canonical RNA caps

4.2.1 *Arabidopsis Thaliana* Plant Nudix enzymes

In this part of my research, the aim was to study the substrate specificity of some new Nudix enzymes. This study should help us to broaden the tools available for manipulation with Np_n -RNAs. Moreover, this study should indicate which other potential RNA caps can exist in eukaryotic cells. Previous studies have suggested that plants, in particular, harbor a significant number of Nudix-encoding genes, with *A. thaliana* reportedly possessing a total of 28 Nudix enzymes¹⁸¹. However, the necessity for such number of Nudix enzymes in cellular metabolism remains unclear, especially considering the apparent redundancy, where multiple enzymes may cleave the same substrate *in vitro*.

In collaboration with Joanna Kufel's group from the Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw, we decided to explore some of the *A. thaliana* Nudix proteins. Specifically, they provided us four purified enzymes: cytosolic AtNUDT6, AtNUDT7 and chloroplastic AtNUDT19 and AtNUDT27 (Figure 25). We decided to examine their substrate specificity for both small molecules and *in vitro* capped RNAs.

While it is established that some of these enzymes can cleave small molecules such as NAD(H), CoA and ADP-ribose, there is currently no research regarding their activity of Np_nNs or non-canonical capped RNAs. In addition, these proteins represent the closest plant homologs of mammalian NUDT2 and NUDT12, which have been shown to exhibit RNA non-canonical decapping activity in vitro, namely deFADding and deCoAging, respectively^{64,161}.

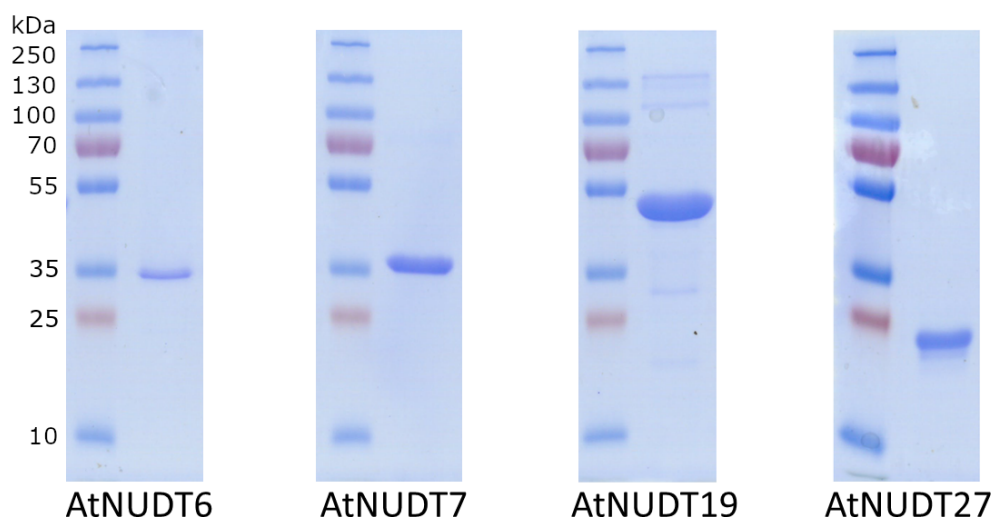


Figure 25. SDS-PAGE analysis of the protein samples after size exclusion chromatography: AtNUDT6 - 32 kDa, AtNUDT7 – 37 kDa, AtNUDT19 – 41 kDa, AtNUDT27 – 22 kDa.

4.2.1.1 AtNUDT6,7,19 and AtNUDT27 in vitro activity on capped RNA

First, we generated ³²P labelled RNA by in vitro transcription using T7 RNA polymerase in the presence of cap precursors using two DNA templates (Table 2). Template 35A contains an A at the +1 position and was used in experiments with Ap₂₋₅A, Gp₃₋₄A, NAD(H), CoA, ADP-ribose and m⁷Gp₃A. Template 35G contains a G at the + 1 position and was used in experiments with Gp₃₋₄G, Ap₃₋₅G, and m⁷Gp₃G. Asymmetrical Np_nNs such as Ap₃₋₅G can be used with both templates to yield RNA with flanking A or G, denoted as Ap₃₋₅G-RNA, or Gp₃₋₅A-RNA, respectively. Following in vitro transcription, we purified capped RNA from uncapped RNA by treating it with 5' polyphosphatase and Terminator™ exonuclease. This treatment ensured that the proportion of capped RNA in each reaction averaged 96 ± 4% (except for m⁷Gp₃A(G)-RNA, where the proportion of capped RNA was 58 ± 3%). Afterwards, we evaluated whether these enzymes could directly cleave non-canonical RNA caps from RNA (Figure 26).

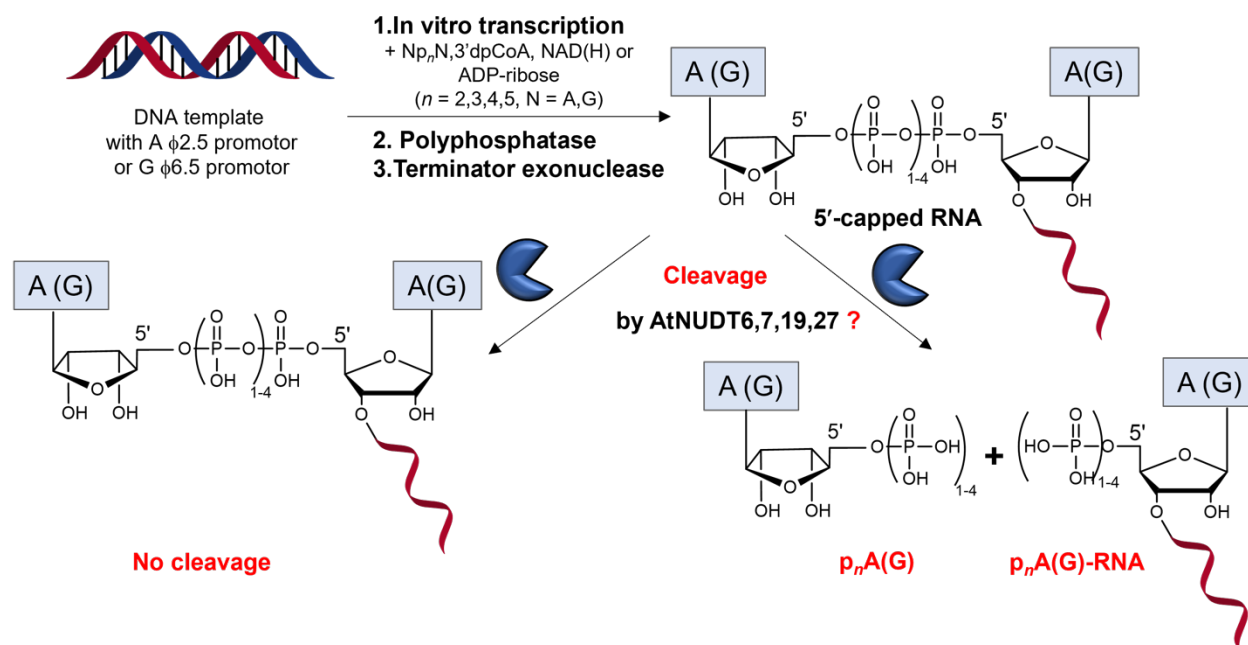


Figure 26. Screening of the RNA-decapping activity of AtNUDT6, 7, 19, and 27. Scheme of the experimental setup. ^{32}P labelled RNA was prepared by in vitro transcription with T7 RNA polymerase in the presence of the small molecules Np_nNs , NAD(H), CoA, or ADP-ribose. The side product, uncapped triphosphate RNA, was degraded by treatment with 5' polyphosphatase and Terminator™ exonuclease.

The cleavage activity of AtNUDT6, AtNUDT7, AtNUDT19, and AtNUDT27 on RNA capped with NAD(H), CoA, ADP-ribose, canonical $m^7Gp_3A(G)$, or various dinucleoside polyphosphates ($Ap_{2-5}A$, $Gp_{3-4}G$, $Ap_{3-5}G$) was monitored using PAGE analysis. The highest enzymatic activity was observed with capped RNA at a concentration of 500 nM for AtNUDT6, 7, and 19, and at a concentration of 50 nM for AtNUDT27 (Figure 27). All the screening was conducted in triplicate to ensure accurate observation of enzyme cleavage.

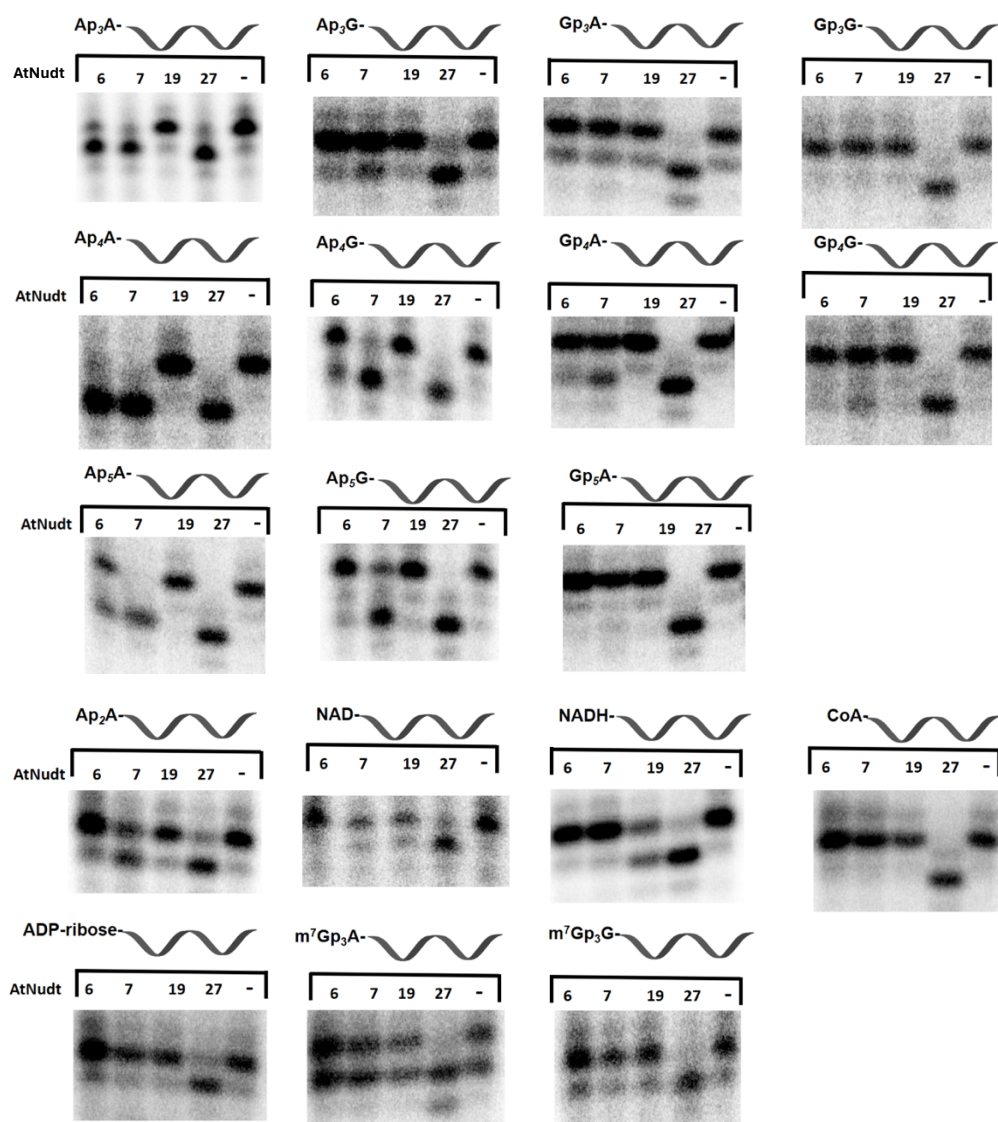


Figure 27. PAGE analysis of the RNA decapping activity of AtNUDT6,7,19 or 27 on 5'-capped ($Ap_{3-5}A$, $Ap_{3-5}G$, $Gp_{3-5}A$, $Gp_{3-4}G$, Ap_2A , NAD, NADH, CoA, ADP-ribose and $m^7Gp_3(A)G$) RNA.

To calculate the percentage of the cleavage for all enzymes, we considered the intensity of the areas under the peaks (AUC). The intensity of the 5'-capped RNA in the control without enzyme corresponds to 0% of cleavage of the substrate, and the intensity of 5'-capped RNA in samples after treatment with the enzymes is compared to the negative control.

We observed that AtNUDT6 and AtNUDT7 had a similar substrate specificity, with Ap_4A -RNA being the most favorable substrate for both. In addition, AtNUDT6 also partially cleaved (less

than 50%) Ap₃A-, Ap₅A- and Ap₄G-RNA, while AtNUDT7 cleaved Ap_{2,3,5}A-RNA, Ap_{4,5}G-RNA, and, to some extent, Gp₄A-RNA. These results suggest a preference for dinucleoside triphosphates and tetraphosphates by AtNUDT6 and AtNUDT7, considering that equal amounts of enzyme and substrate were used in the decapping assay (Table 3).

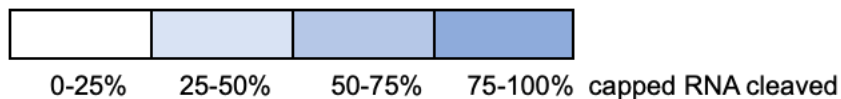
While AtNUDT6 and AtNUDT7 displayed broad substrate specificities, AtNUDT19 efficiently cleaved only the NADH-RNA, with the NAD-RNA being partially cleaved. These results align with previous publications demonstrating that AtNUDT19 cleaves free NADH and NADPH¹⁸². Notably, NAD-RNA was also partially cleaved by AtNUDT19 at 500nM enzyme concentration (Table 3).

Surprisingly, at a tenfold lower concentration (50 nM), AtNUDT27 still partially hydrolyzed the majority of capped RNA with a relatively high cleavage efficiency (60–90% after 1 h), except for Ap₃A-RNA (less than 30%) (Table 3). So far, it is known that AtNUDT27 can mediate the hydrolysis of nucleoside diphosphate Ap₅A. Our results show that the only two substrates that were cleaved more than 90% are Ap₅A-RNA and Gp₅A-RNA. This may indicate that the active site of this enzyme can accommodate and hydrolase penta-phosphates more efficiently than other Np_nNs-RNAs.

Nevertheless, our data demonstrates that AtNUDT27 can cleave not only Ap₅A-RNA but also numerous others Np_nNs that can be part of the RNA. This indicates that AtNUDT27 can be a potential general RNA-decapping enzyme in *Arabidopsis Thaliana*. Moreover, knowing that Np_nN-RNAs are a very suitable substrate for this enzyme, it is plausible that these RNA molecules may indeed be present in the plant cells.

Table 3. The cleavage of RNA capped with different small molecules *in vitro* by AtNUDT6 (500 nM), AtNUDT7 (500 nM), AtNUDT19 (500 nM), and AtNUDT27 (50 nM). Values are in % of cleaved capped RNA and represent a mean of three independent replicates $\pm$ standard deviations.

5'-cap-RNA	Nudt	6 (500 nM)	7 (500 nM)	19 (500 nM)	27 (50 nM)
Ap ₂ A-		x	53 $\pm$ 10	x	73 $\pm$ 9
Ap ₃ A-		48 $\pm$ 7	55 $\pm$ 7	x	26 $\pm$ 8
Ap ₄ A-		80 $\pm$ 11	73 $\pm$ 5	x	87 $\pm$ 4
Ap ₅ A-		15 $\pm$ 8	65 $\pm$ 17	x	90 $\pm$ 5
Ap ₃ G-		x	x	x	74 $\pm$ 11
Ap ₄ G-		35 $\pm$ 10	70 $\pm$ 20	x	85 $\pm$ 5
Ap ₅ G-		x	61 $\pm$ 12	x	85 $\pm$ 5
Gp ₃ G-		x	x	x	83 $\pm$ 5
Gp ₄ G-		x	x	x	83 $\pm$ 1
NAD-		x	x	20 $\pm$ 4	60 $\pm$ 7
NADH-		x	x	57 $\pm$ 5	65 $\pm$ 10
CoA-		x	x	x	82 $\pm$ 5
Gp ₃ A-		x	x	x	80 $\pm$ 5
Gp ₄ A-		x	30 $\pm$ 6	x	84 $\pm$ 5
Gp ₅ A-		x	x	x	90 $\pm$ 3
m ⁷ Gp ₃ A-		x	x	x	80 $\pm$ 7
m ⁷ Gp ₃ G-		x	x	x	70 $\pm$ 6
ADP-ribose		x	x	x	70 $\pm$ 2



Since Ap₄A-RNA was the best substrate for AtNUDT6, AtNUDT7 and AtNUDT27 we wanted to understand if all enzymes cleaved this molecule in the same manner. To investigate this, we prepared *in vitro* transcribed triphosphate (ppp-)RNA and monophosphate (p-)RNA as control samples for the PAGE analysis. While the main product of the Ap₄A-RNA decapping reaction by AtNUDT6 and AtNUDT7 was triphosphate (ppp-)RNA, AtNUDT27 cleaved Ap₄A-RNA into a mixture of ppp-RNA and monophosphate (p-)RNA (Figure 28). This distinctive feature of AtNUDT27 makes it different from the other two enzymes. Considering also that both AtNUDT6 and AtNUDT7 are cytosolic enzymes, while AtNUDT27 is chloroplastic, this characteristic of the enzyme might indicate a different requirement for cleavage of these products based on their cellular location.

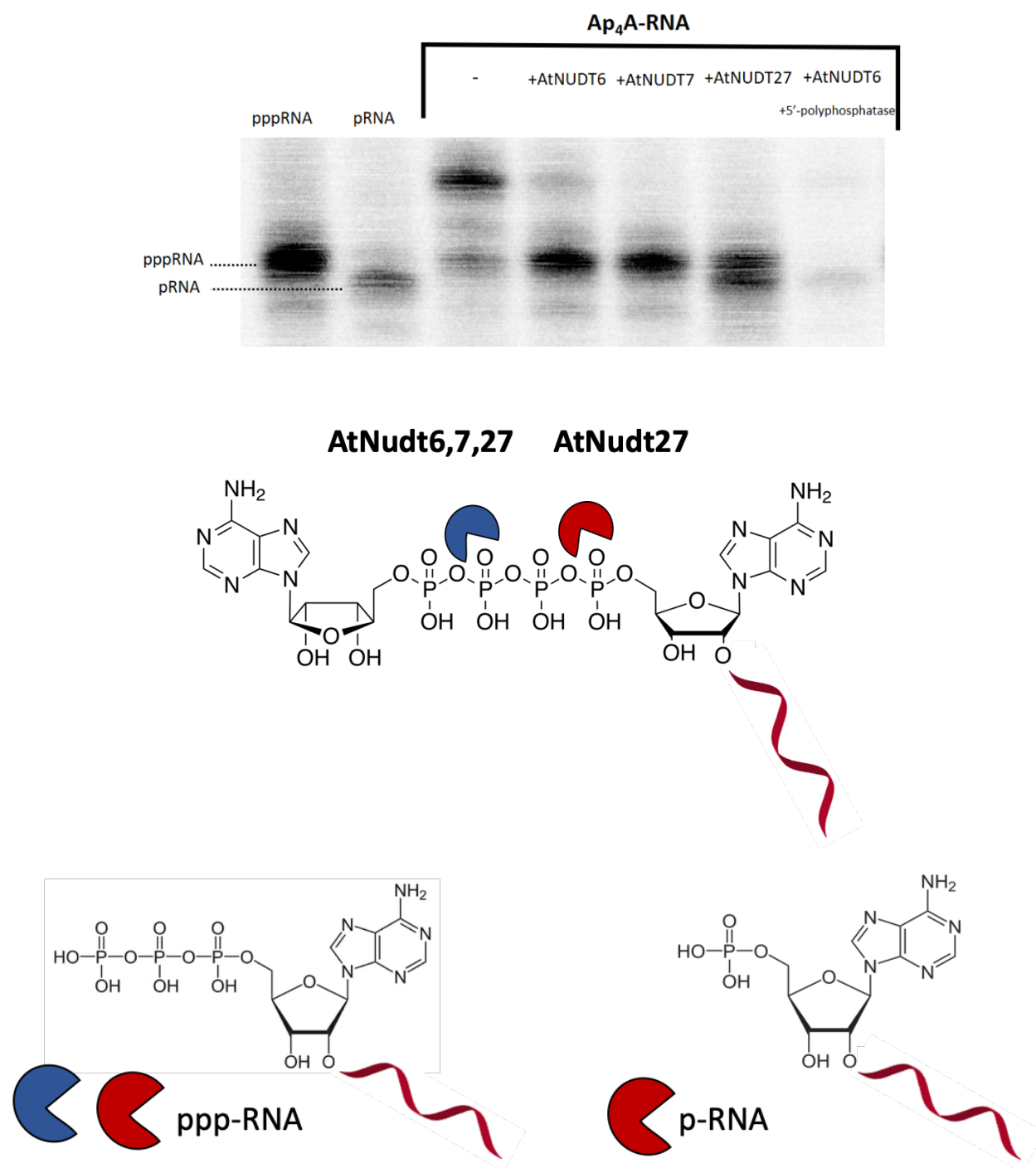


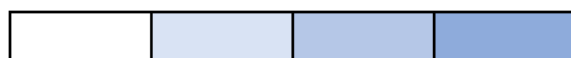
Figure 28. PAGE analysis of the RNA decapping activity of AtNUDT6, AtNUDT7 or AtNUDT27 on Ap₄A-RNA. While AtNUDT6 and AtNUDT7 lead to pppRNA, AtNUDT27 cleavage leads to mixture of pppRNA and pRNA. When 5'-polyphosphatase is added to AtNUDT6 cleaved RNA, pppRNA band disappears and new at level of pRNA is visible.

4.2.1.2 AtNUDT6,7,19 and AtNUDT27 in vitro activity on small molecules

At this point, we investigated the ability of these enzymes to cleave small molecules. While some of the small molecules are known to be in vitro substrates of these Nudix enzymes, the entire set of recently discovered non-canonical caps has not yet been tested as free compounds. Therefore, we screened all non-canonical RNA caps in their free form as substrates for the four Nudix proteins using HPLC (Supplementary Figure 2-5). Table 4 provides a summary of the cleaved substrate percentages. Interestingly, AtNUDT6 and AtNUDT7 exhibited similar substrate specificities for both small molecules and capped RNA. Surprisingly, their best substrate was not free Ap₄A, as expected based on results from assays on capped RNA (Table 3), but rather Ap₂A and NADH, both featuring a diphosphate bridge. As in the case of capped RNA, AtNUDT19 hydrolyzed free NADH with the highest efficiency. AtNUDT27 had a very high activity at 500 nM concentration and quantitatively cleaved all the free tetraphosphates (Ap₄A, Ap₄G, and Gp₄G) and Ap₅A (Table 4). This finding is intriguing, considering that the same enzymes can exhibit different behavior toward certain small molecules when compared to their behavior when those molecules attached to the RNA. This discrepancy suggests potential differences in the catalytic site of the enzymes and variations in interactions based on the presence or absence of RNA.

Table 4. Cleavage of different small molecules (400 mM) in vitro by AtNUDT6 (500 nM), AtNUDT7 (500 nM), AtNUDT19 (500 nM), and AtNUDT27 (500 nM). Values are in % of cleaved substrate and represent a mean of three independent replicates $\pm$ standard deviations.

AtNudt small molecule	6 (500 nM)	7 (500 nM)	19 (500 nM)	27 (500 nM)
Ap ₂ A-	80 $\pm$ 9	64 $\pm$ 7	16 $\pm$ 3	x
Ap ₃ A-	18 $\pm$ 5	18 $\pm$ 1	x	x
Ap ₄ A-	31 $\pm$ 1	40 $\pm$ 3	x	99 $\pm$ 0.5
Ap ₅ A-	x	31 $\pm$ 1	x	99 $\pm$ 0.5
Ap ₃ G-	x	x	x	x
Ap ₄ G-	20 $\pm$ 2	x	x	96 $\pm$ 5
Ap ₅ G-	x	x	x	89 $\pm$ 1
Gp ₃ G-	x	x	x	x
Gp ₄ G-	x	x	x	98 $\pm$ 1.5
NAD-	x	52 $\pm$ 1	x	x
NADH-	99 $\pm$ 1	83 $\pm$ 4	x	x
CoA-	53 $\pm$ 4	18 $\pm$ 6	x	x
ADP-ribose	55 $\pm$ 10	45 $\pm$ 3	x	x



0-25% 25-50% 50-75% 75-100% small molecule cleaved

4.2.1.3 Kinetic parameters of AtNUDT6,7 and AtNUDT27 for the cleavage of free Ap₄A

Considering all previous experiments, the only molecules that were consistently cleaved by the majority of enzymes were Ap₄A and Ap₄A-RNA. Moreover, concurrent with our research, the existence of Ap₄A-RNA in mammalian cells was reported in our lab. For this reason, our following experiments were focused on both Ap₄A and Ap₄A-RNA, aiming to elucidate whether these substrates are cleaved in the same manner by our tested enzymes.

In this experiment, we first measured and compared Michaelis–Menten kinetic parameters for AtNUDT6, 7, and 27 using Ap₄A as a substrate (Table 5, Figure 29B). Since AtNUDT19 did not cleave Ap₄A, it was not included in our kinetic experiments. Moreover, these parameters were already known for AtNUDT19 using NADH as a substrate¹⁸².

Although K_M values were comparable for all three enzymes, we observed that both V_{max} and K_{cat} of AtNUDT27 were significantly higher ($P < 0,05$) compared to AtNUDT6 and AtNUDT7 (Figure 29A). To measure the catalytic efficiency of all three enzymes, we calculated the K_{cat}/K_M ratio. K_{cat} describes the numbers of substrate molecules transformed into products per unit time by a single enzyme, while K_M represents the affinity of the substrate to the enzyme's active site. The K_{cat}/K_M ratio provides insight into how efficiently the enzyme catalyzes the substrate. A higher ratio indicates faster catalysis, whereas a lower ratio suggests slower catalysis. Remarkably, the K_{cat}/K_M ratio for AtNUDT27 was significantly higher compared to the other two enzymes (Figure 29A).

This finding may explain the observed behavior of AtNUDT27 being more efficient in cleaving free Ap₄A compared to AtNUDT6 and AtNUDT7.

Table 5. Kinetic parameters of AtNUDT6, AtNUDT7 and AtNUDT27 cleaving Ap₄A expressed as mean $\pm$ standard deviation.

Protein	K_M (μM)	V_{max} ($\mu M \text{ min}^{-1}$)	K_{cat} (min^{-1})	K_{cat}/K_M ($\text{min}^{-1} \mu M^{-1}$)
AtNUDT6	6.8 ± 1.5	0.5 ± 0.0	9.3 ± 0.2	1.4 ± 0.3
AtNUDT7	11.3 ± 4.9	0.4 ± 0.2	8.8 ± 3.5	0.8 ± 0.2
AtNUDT27	22.8 ± 10.4	3.6 ± 1.3	71.7 ± 25.6	3.2 ± 0.6

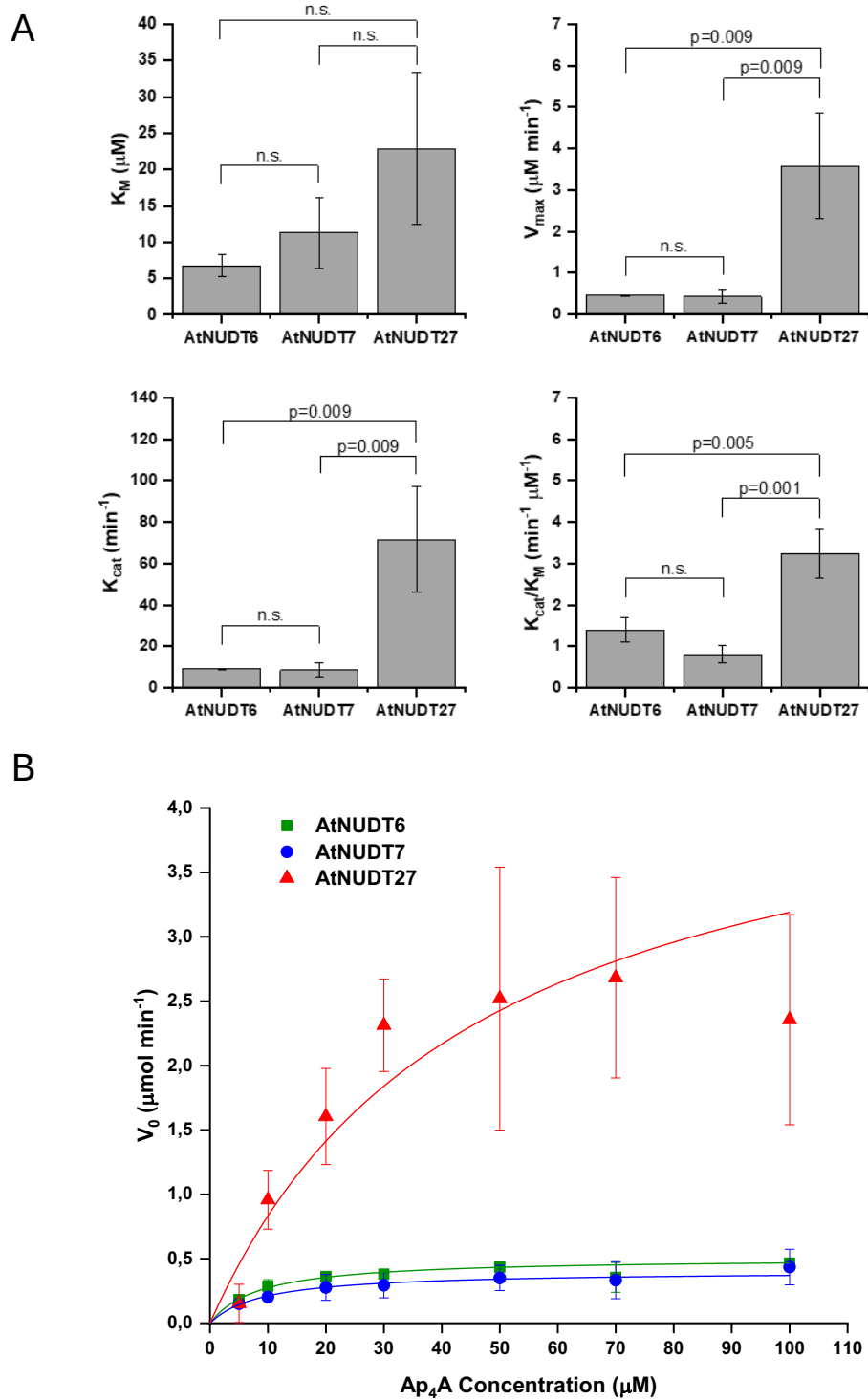


Figure 29. Kinetic studies. A) Kinetic parameters of AtNUDT6, 7, and 27 in reaction with Ap₄A. B) Michaelis-Menten plots of initial velocities for AtNUDT6, 7, and 27 in reaction with Ap₄A.

4.2.1.4 Cleavage efficiency of free Ap₄A and Ap₄A-RNA by AtNUDT6, AtNUDT7, and AtNUDT27

Our attempt to measure kinetic parameters for Ap₄A-RNA was unsuccessful due to the substantial amount of radioactively labelled RNA needed for this experiment (for each point in triplicate - large scale in vitro transcription reaction with 5' polyphosphate and Terminator™, combined with three steps of column purification). Therefore, we performed an initial screening experiment at a single concentration and time point to observe differences in substrate preferences. We monitored the cleavage of both substrates at a concentration of 2.5 μ M and at 15- and 60-minute time points. The cleavage of Ap₄A was analyzed using HPLC (Figure 30A, 30B), while decapping of Ap₄A-RNA was measured by quantifying corresponding bands of capped RNAs resolved on PAGE (Figure 30C, 30D).

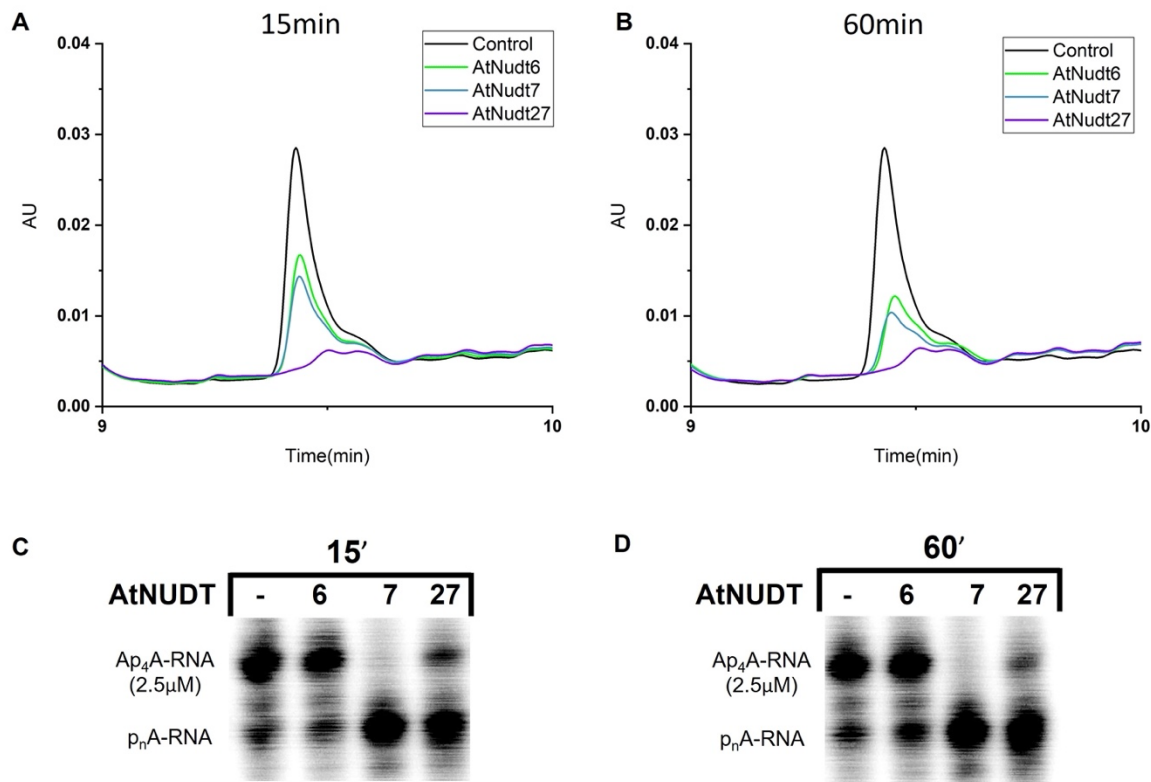


Figure 30. HPLC chromatograms of Ap₄A (2.5 μ M) before and after incubation with AtNUDT6,7 and 27 (A) 15 min, (B) 60 min. Representative example of PAGE analysis of the RNA decapping activity of AtNUDT6,7 and 27 on Ap₄A-RNA (C) 15 min, (D) 60 min.

The comparison of substrates revealed that AtNUDT27 exhibited a preference for the small molecule (Ap₄A) and cleaved it more efficiently than Ap₄A-RNA under the chosen conditions. AtNUDT6 cleaved both substrates – Ap₄A and Ap₄A-RNA – comparably well, while AtNUDT7 demonstrated higher activity on Ap₄A-RNA than on free Ap₄A (Figure 31). This finding may explain the redundancy of Nudix enzymes observed *in vivo*, where certain enzymes are used by the cell to hydrolyze free small molecules, while others cleave non-canonically capped RNA. These results could potentially reflect the *in vivo* scenario for many other organisms as well.

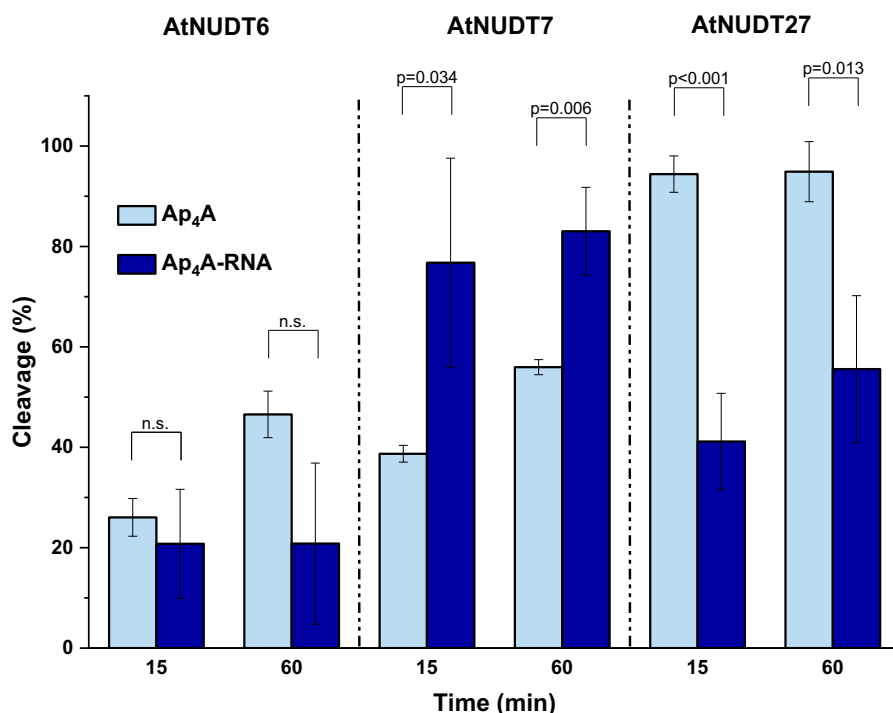


Figure 31. Column graph data representing the cleavage efficiency of free Ap₄A (2.5 μM) and Ap₄A-RNA (2.5 μM) by AtNUDT6, AtNUDT7, and AtNUDT27 (50 nM) at 15- and 60 min time points; The cleavage mean value ± SD was calculated for three replicates. Statistical differences were expressed as p-value calculated using t-test.

4.2.1.5 Inhibition of Ap₄A-RNA by AtNUDT6, 7 and 27 in presence of free Ap₄A

We also addressed whether free dinucleoside polyphosphates could inhibit the cleavage of capped RNA. To investigate this, we measured the cleavage of Ap₄A-RNA (1 μM) by AtNUDT6, 7 and 27 in the presence of free Ap₄A (1, 2 and 4 μM) (Figure 32A). Interestingly, we observed that only the activity of AtNUDT6 was affected in a dose-dependent manner. In contrast, the activity of the other two enzymes, AtNUDT7 or AtNUDT27, remained unaffected by the addition of free Ap₄A (Figure 32B). The difference between AtNUDT6 and AtNUDT7, which

exhibit comparable kinetic parameters, could potentially be explained by AtNUDT7's preference for Ap₄A-RNA over Ap₄A, as demonstrated in Figure 31. On the other hand, AtNUDT27 showed an almost 8 times higher K_{cat} for Ap₄A compared to AtNUDT6 and AtNUDT7, suggesting that the added free Ap₄A would be metabolized at a high rate, leaving sufficient enzymatic capacity to hydrolyze Ap₄A-RNA as well.

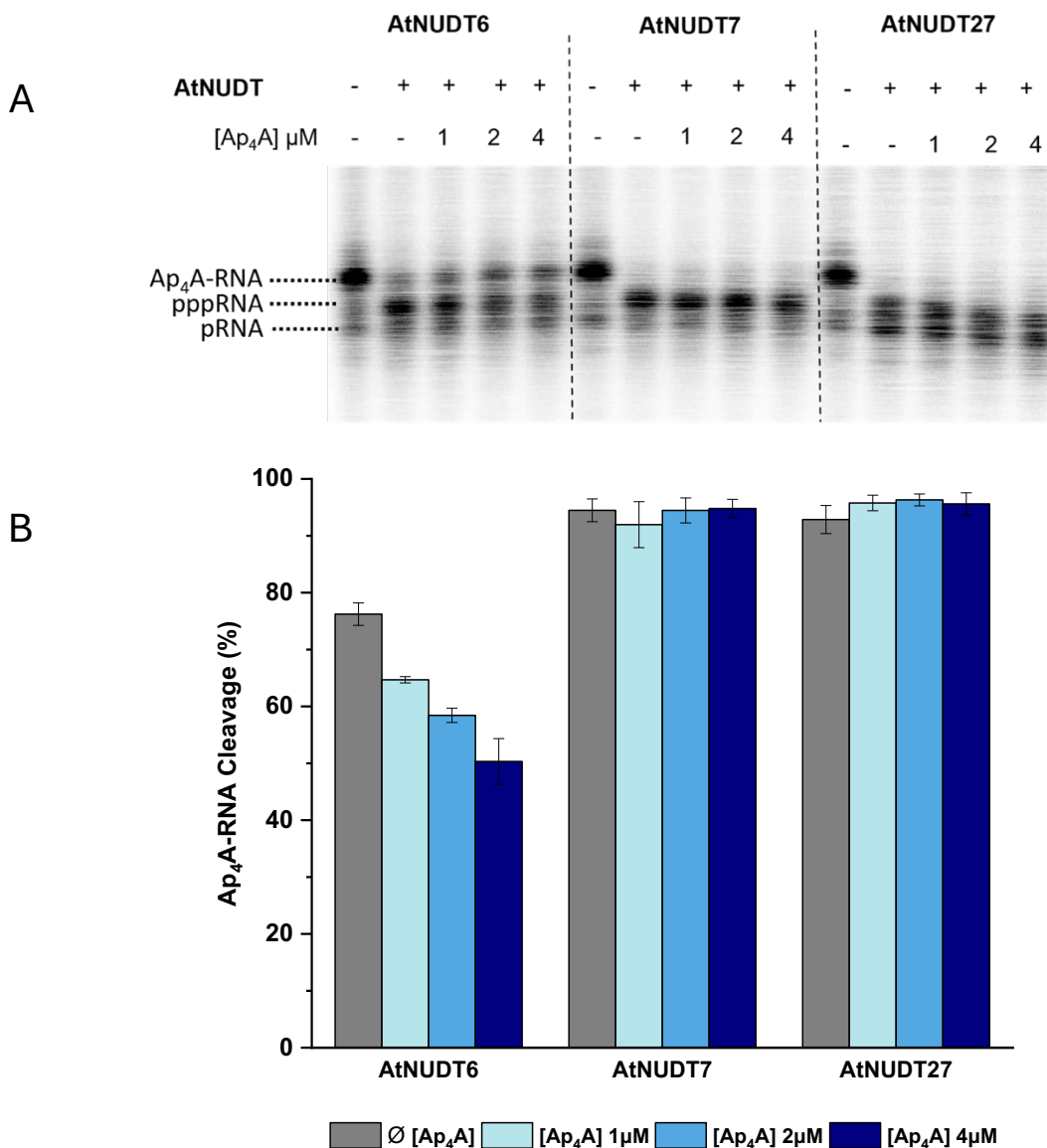


Figure 32. Inhibition study. A) Ap₄A (1,2 and 4 μ M) was added to 1 μ M Ap₄A-RNA and cleavage of ³²P labelled RNA by AtNUDT6,7 and 27 was followed by PAGE analysis; B) Column graph data represent the cleavage mean value $\pm$ SD for three replicates.

The collective data presented herein strongly suggests that plant Nudix enzymes may serve as RNA-decapping enzymes for non-canonically capped RNA. We focused on potential RNA-decapping Nudix enzymes from *A. thaliana*, where we investigated the cleavage properties of four enzymes: cytosolic AtNUDT6, AtNUDT7, and chloroplastic AtNUDT19 and AtNUDT27.

Our findings highlight AtNUDT27 as a particularly promising general RNA decapping enzyme. Furthermore, considering recent discoveries of non-canonical RNA caps across various organisms, the identification of Ap₄A-RNA as the optimal substrate for AtNUDT6,7 and AtNUDT27 suggests the potential presence of this type of capped RNA in *A. thaliana*.

These enzymes can also be used for the development of capturing techniques for the identification of Ap₄A-RNA in different cell lines. Additionally, AtNUDT19's specificity towards NADH-RNA suggests its potential utility in developing capturing techniques for the identification of NADH-RNA. Overall, our study sheds light on the diverse roles of Nudix enzymes in RNA metabolism and presents avenues for further exploration in understanding RNA modification dynamics.

4.2.2 NUDT2 and DXO decapping activity on non-canonical capped RNAs

During my PhD studies, I also worked with human NUDT2 Nudix enzyme and mouse DXO decapping exoribonuclease for different projects. NUDT2 is a member of the Nudix family, and considering that its main substrate is Ap₄A, we wanted to further explore its activity in vitro to better understand the mechanisms of its function. We isolated and purified recombinant human NUDT2 and tested its activity on Ap₄A as well as on in vitro prepared Ap₄A-RNA. The cleavage of free Ap₄A by NUDT2 produced adenosine triphosphate (ATP) and adenosine monophosphate (AMP) as expected (Figure 33A, 33B).

Next, we prepared in vitro transcribed Ap₄A-RNA and tested various conditions of Ap₄A-RNA cleavage. We observed complete decapping of Ap₄A-RNA at 500 nM concentration of NUDT2. The first resulting products of the cleavage were AMP and 5'-triphosphate RNA (ppp-RNA). At higher concentrations, the enzyme also cleaved ppp-RNA into p-RNA (Figure 33C), as was also previously reported¹⁸³.

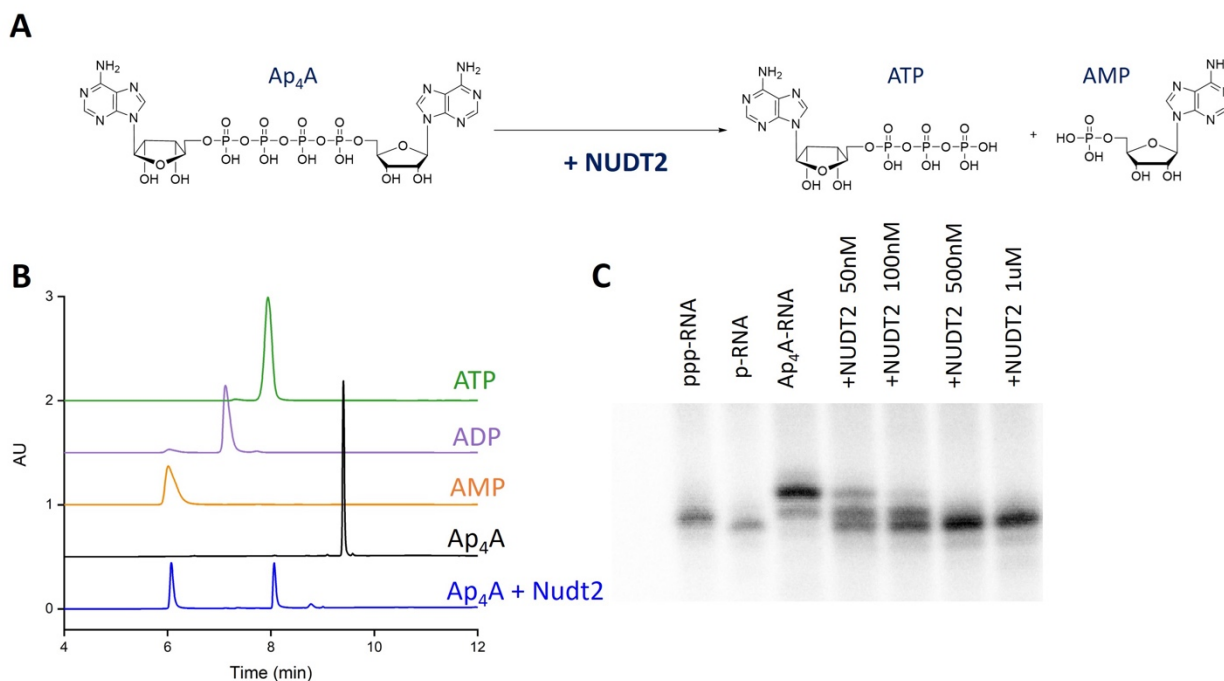


Figure 33. A) Scheme of Ap₄A cleavage by NUDT2. B) HPLC-UV chromatogram of ATP, ADP, AMP and Ap₄A incubated with or without NUDT2. C) PAGE analysis of incubated Ap₄A-RNA with various amounts of NUDT2.

In addition to Ap₄A-RNA, we tested the activity of NUDT2 on other Np_nN-RNAs, some of which were previously detected in bacteria⁸⁶. We used 100 ng of in vitro transcribed capped RNA with different Np_nNs, such as Ap₂₋₆A, Gp₂₋₃G, Ap₂₋₅G, Gp₂₋₄G and 500nM of NUDT2. Surprisingly, NUDT2 cleaved the vast majority of them, showing a lower activity towards

Np₃N-RNAs (Figure 34A, 34B). This implies that NUDT2 serves as a general decapping enzyme with broad substrate specificity. Although Ap₃A-RNA was completely cleaved, Gp₃G-RNA and Ap₃G-RNA did not appear to be good substrates for this enzyme. It will be interesting to understand if the active site of the enzyme is more prone to accept di-adenosine polyphosphates rather than di-guanosine polyphosphate.

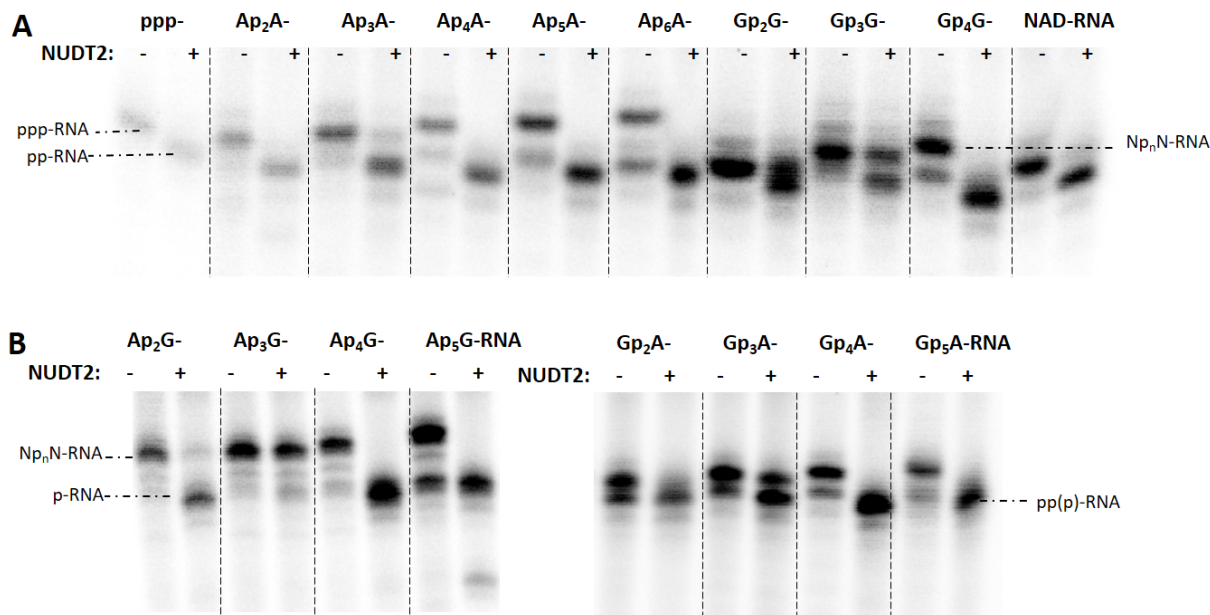


Figure 34. PAGE analysis of non-canonically capped RNAs prepared by *in vitro* transcription with NUDT2 (500 nM) and 100 ng of A) Ap₂₋₆A-RNA, Gp₂₋₃G-RNA and NAD-RNA; B) Ap₂₋₅G-RNA, Gp₂₋₄G-RNA.

Apart from the NudiX family, the human enzyme Decapping Exoribonuclease (DXO) has also been shown to decap RNA with an affinity towards NAD-RNA⁶². Unlike NUDT2, DXO does not cleave free Np_nNs. To explore this, we tested the cleavage of Ap₄A, NAD and m⁷Gp₃A using the DXO enzyme. None of these molecules were cleaved by the enzyme (Figure 35A). However, using NAD-RNA and Ap₄A-RNA as substrates, DXO interestingly cleaved Ap₄A-RNA within the polyphosphate bridge, while in case of NAD-RNA, it removed the entire RNA cap, leaving the transcript one nucleotide shorter (Figure 35B, 35C).

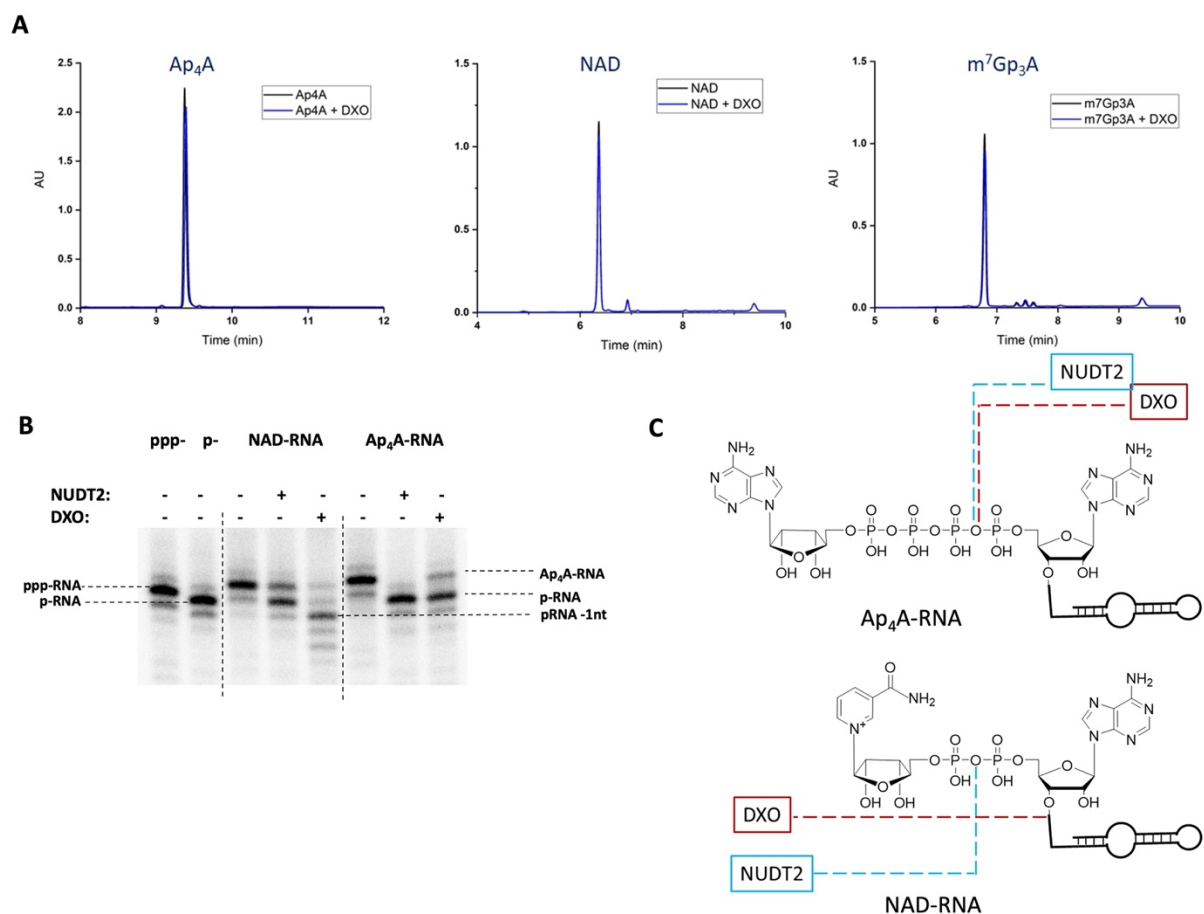


Figure 35. Cleavage of Ap₄A-RNA. A) DXO (200 nM) incubated with free Ap₄A, NAD or m⁷Gp₃A. B) Polyacrylamide gel electrophoresis (PAGE) of ³²P labeled ppp-RNA, p-RNA and NAD-RNA and Ap₄A-RNA treated with NUDT2 and DXO. Prior to the reaction with NUDT2 and DXO, NAD-RNA and Ap₄A-RNA were treated with 5'-polyphosphatase and Terminator exonuclease. C) NUDT2 and DXO cleavage on Ap₄A-RNA and NAD-RNA.

Knowing the substrate specificity of NUDT2 with capped RNA, we also wanted to explore the activity of DXO towards Np_nN-RNAs. This work was performed by high school student, Zuzana Buchová, under my supervision. We prepared ³²P labeled RNA with different Np_nNs by in vitro transcription and incubated 100 ng of the capped RNA with 200 nM of DXO. Previous research showed that NAD-capped RNA was efficiently decapped (deNADded) by DXO to RNA 5' monophosphate, which was further degraded by its 5'-3' exonuclease activity^{62,78,143}. As reported, the 5'-3' exonuclease activity was particularly well manifested at high enzyme concentration, especially above 500 nM. For this reason, we used 200 nM of DXO to observe only the decapping activity of the enzyme.

Surprisingly, DXO cleaved more than 80% of Ap₂A-, Ap₄A- (Figure 36A), Gp₂A- and Gp₄A- (Figure 36B), Gp₂G- and Gp₄G-RNA (Figure 36D), showing a preference for diphosphate and tetraphosphate capped RNAs. The cleavage product for tetraphosphates resulted in a monophosphate RNA (pRNA), while for diphosphates, the product was one nucleotide shorter, compared to pRNA. However, in the case of Ap₄G, the results were ambiguous, as not all the experiments showed consistent cleavage (Figure 35B).

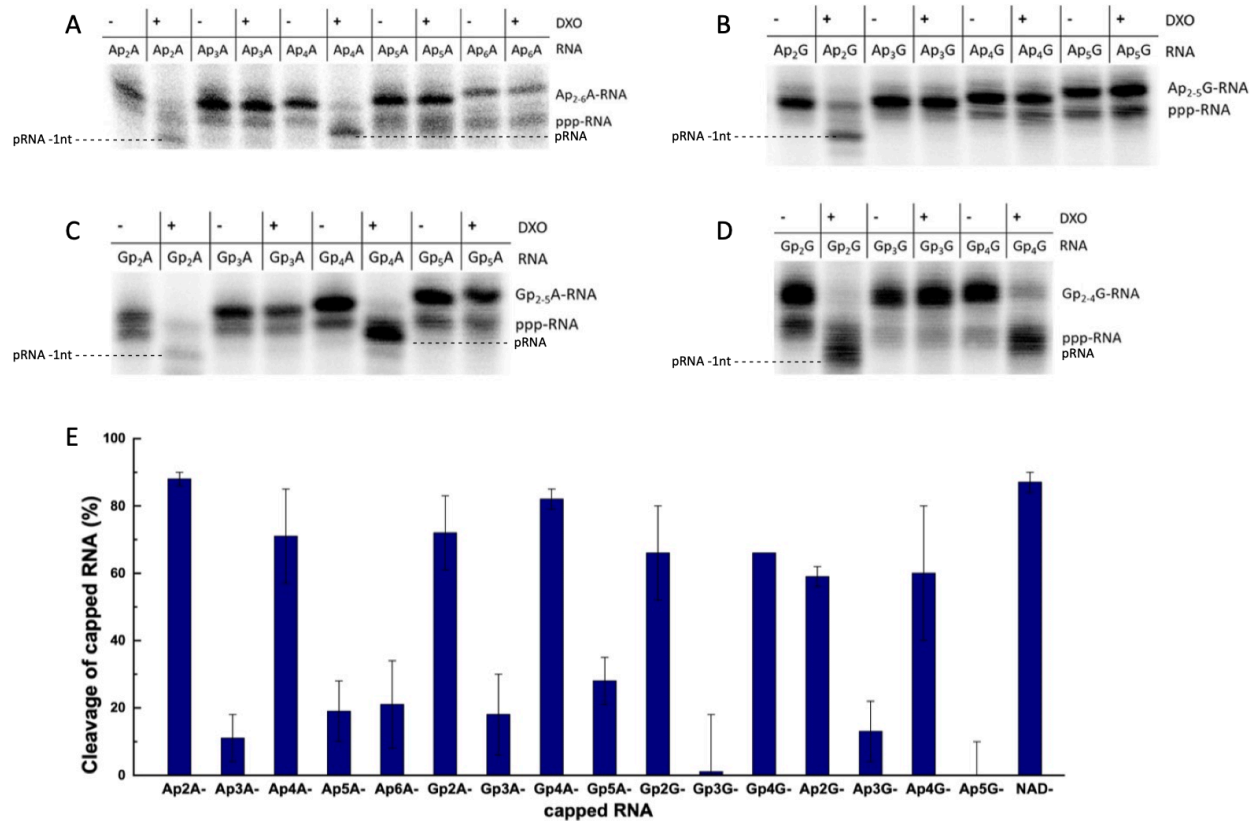


Figure 36. PAGE analysis of non-canonically capped RNAs prepared by *in vitro* transcription with DXO (200 nM) and 100 ng of A) Ap₂₋₆A-RNA, B) Ap₂₋₅G-RNA, C) Gp₂₋₅G-RNA, D) Gp₂₋₄G-RNA. E) Column graph data represent the cleavage mean value $\pm$ SD for three replicates.

Considering all the previous data, we demonstrated that both NUDT2 and DXO are decapping enzymes for newly discovered Ap₄A-RNA. Moreover, both enzymes showed cleavage activity towards non-canonical capped RNAs. Specifically, NUDT2 cleaved almost all the Np_nN-RNAs we tested, with lower activity towards capped RNAs, containing triphosphate bridge within the cap structure, such as Gp₃G, Ap₃G and Gp₃A. On the other hand, DXO showed a preference for diphosphate and tetraphosphate capped RNAs, cleaving Ap₂A, Ap₄A, Gp₂A, Gp₄A, Gp₂G and Gp₄G (Figure 36E). These results may explain why

decapping enzymes are widely distributed among all classes of organisms, possessing different substrate specificity and cellular localizations.

5 Discussion

5.1 Dinucleoside diphosphate Ap₂A as novel 5' RNA cap in mammalian cells

Non-canonical capping of RNA refers to a modification process in RNA metabolism where RNA molecules receive a chemical modification at their 5' end, comparable to canonical capping, but through mechanisms that differ from the typical capping processes observed in eukaryotic cells. It was long believed that only the eukaryotic mRNA is capped at the 5' end with a methylated guanosine that is added to the 5' end of the RNA molecule through a 5'-5' triphosphate linkage, forming what is known as a "cap structure". However, in the past decades many other modifications at the 5' end of RNAs were discovered, such as nicotinamide adenine dinucleotide (NAD), coenzyme A (CoA), flavin adenine dinucleotide (FAD), and dinucleoside polyphosphates (Np_nNs)^{57,58,69,81}.

In this study, employing LC-MS technique, we have identified a novel group of dinucleoside polyphosphates functioning as non-canonical RNA caps in the Expi293F cell line, including Ap₂A-RNA, Ap₂G-RNA, Ap₂U-RNA and Ap₂C-RNA. We mainly studied Ap₂A-RNA due to its structural resemblance to NAD-RNA, one of the extensively investigated non-canonical capped RNAs across various organisms. While NAD-RNA comprises adenine and nicotinamide nucleotides, Ap₂A consists of two adenosines linked by a diphosphate group (Figure 37).

The presence of diadenosine diphosphate Ap₂A as free molecule was reported in human platelets several years ago¹⁸⁴. Moreover, there is one more report on the existence of Ap₂A isolated from human cardiac tissue¹⁸⁵. Together with Ap₂A, the dinucleoside diphosphates Ap₂G and Gp₂G have been described as endogenous substances released from the granules of human platelets¹⁸⁶. These molecules represent a new class of growth-promoting extracellular mediators present in platelet secretory granules and in cardiac myocytes, capable of stimulating cardiac myocyte proliferation and believed to play a role in the control of cardiovascular tone.

Considering that these molecules can also be part of RNA, we investigated whether dinucleoside diphosphates might serve as substrates for RNA polymerase as free molecules, implying a co-transcriptionally incorporation mechanism. However, unlike NAD and Ap₄A, which exhibited significant signals in LC-MS using the same extraction method, we did not detect free Ap₂A, Ap₂G, Ap₂U or Ap₂C in Expi293F cells. It is possible that these

molecules are specific to platelets and cardiac myocytes. Accordingly, these findings suggest that these molecules may not be incorporated into RNA as non-canonical initiator nucleotides (NCIN) by RNA polymerase in cells. Instead, another capping mechanism might be involved in their formation.

Dinucleoside polyphosphates are often referred to as alarmones due to fact that their intracellular concentration increases in response to cellular stress. Several studies have suggested that alterations in growth conditions or exposure to environmental stressors can influence the level of certain capped RNAs^{64,86,87}. For instance, recent research involving HEK293T cells subjected to distinct environmental stress conditions, such as heat shock¹⁸⁷ and glucose deprivation¹⁸⁸, demonstrated the impact of these conditions not only on cellular NAD metabolism but also NAD-capped RNAs⁶⁴.

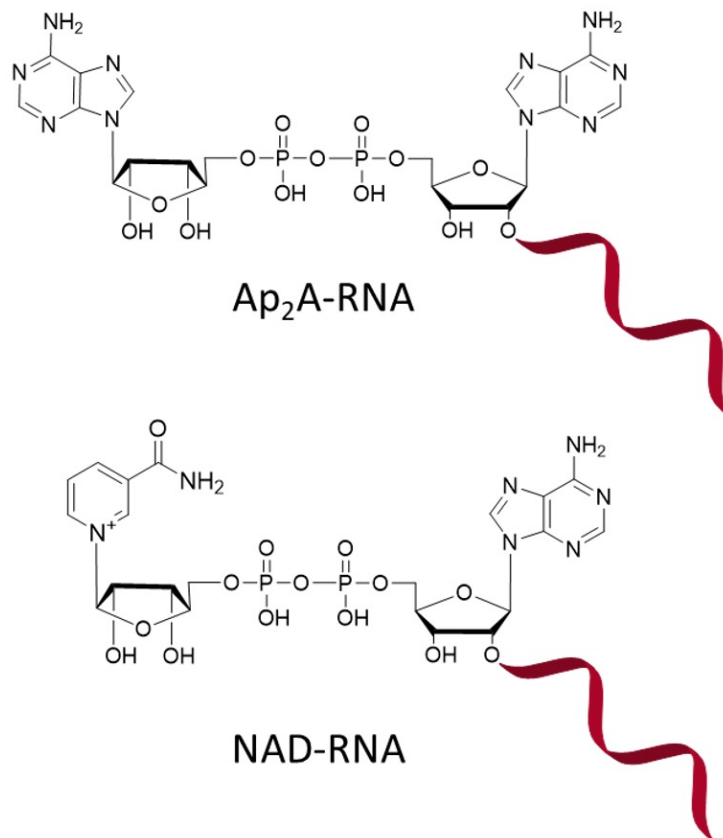


Figure 37. Structure of diadenosine diphosphate capped RNA (Ap₂A-RNA) and nicotinamide adenine dinucleotide capped RNA (NAD-RNA).

Building upon this knowledge, we exposed Expi293F cells to menadione, a compound known to induce oxidative stress. Initially, we quantified the abundance of Ap₂A-RNA in the

short fraction of RNA, and we observed a significant increase of Ap₂A RNA cap in the menadione treated samples compared to the control group. These results suggest a correlation between newly discovered Ap₂A-RNA and cellular ROS damage. Concurrently, another research group reported the discovery of a new putative human 5'-3' RNA ligase (RLIG1) that promotes ligation between 5'-PO₄ and 3'OH RNA termini via a three-step mechanism involving successive auto- and RNA-AMPylation¹⁷⁸. This discovery implies a latent RNA repair mechanism in humans. Furthermore, they showed that RLIG1 knockout (KO) cells exhibit poor resilience to ROS, with 28S rRNA degrading at higher concentrations of menadione. They attributed this RNA decay in KO cells to the absence of RLIG1, indicating that this ligase is crucial for maintaining RNA integrity under stress conditions¹⁷⁸.

We hypothesized that the increase in Ap₂A capped RNAs might be attributed to the fragmentation of longer transcripts, highlighting a complex interplay between RNA degradation and stress-induced modifications.

Therefore, we postulated that RLIG1 might be a putative capping enzyme forming Ap₂N-RNAs using fragmented RNA substrates and ATP in an initial capping mechanism. Using LC-MS technique, we observed the formation of Ap₂A-RNA and Ap₂G-RNA in vitro using monophosphate RNA (p-RNA) and ATP as substrates for the RLIG1 ligase. This finding suggests a potential novel repair mechanism wherein fragmented RNAs serve as substrate for this specific ligase, initially forming the Ap₂N moiety and subsequently ligating them to regenerate the original RNA. However, to validate this proposed model, it would be needed to demonstrate the absence or decrease of Ap₂A-RNA in RLIG1-KO cells in both control and menadione-treated conditions, indicating that Ap₂A-RNA is an intermediate of this enzyme in vivo.

It has been previously reported that two enzymes, DXO and NUDT12, are responsible for NAD-RNA degradation^{62,64}. Given the similarity between NAD-RNA and Ap₂A-RNA we hypothesized that these enzymes could also cleave Ap₂A-RNA in vitro. In our experiments, both NAD-RNA and Ap₂A-RNA were completely cleaved by NUDT12, releasing monophosphate RNA (p-RNA) in both cases. In contrast, DXO removed the entire RNA cap, leaving the transcript one nucleotide shorter. This suggests that these two enzymes could also be responsible for the degradation of Ap₂A-RNA in vivo.

Comparing the cleavage activities of DXO and NUDT12 for NAD and Ap₂A capped RNAs, we found that DXO cleaved both NAD and Ap₂A-RNA similarly. In contrast, NUDT12 showed a significant preference for Ap₂A-RNAs. Additionally, Ap₂A as a small molecule served as a more favorable substrate for NUDT12 compared to NAD, indicating that NUDT12 can cleave both Ap₂A-capped RNA and Ap₂A as a free molecule. Moreover, we confirmed that DXO is unable to cleave small molecules, such as Ap₂A and NAD, suggesting that the DXO is specifically responsible for the cleavage of capped RNAs.

We also hypothesized that the increased levels of Ap₂A-RNA in the small RNA fraction might be caused by a downregulation of one of these two enzymes (DXO and NUDT12) or an upregulation of RLIG1, the in vitro capping enzyme for Ap₂A-RNA. However, our analysis did not reveal any change in their expression levels at the mRNA or protein levels. Another approach would be to quantify Ap₂A-RNA levels of cells lacking NUDT12 or DXO. However, exploring these behaviors in vivo could be challenging due to the redundancy of decapping enzymes in human cells.

To identify the RNA transcripts capped at the 5' end with an Ap₂A moiety, we developed a sequencing protocol using T4 RNA ligase 2, truncated K227Q. This ligase specifically ligates the usually chemically pre-adenylated 5' end of DNA or RNA to the 3' OH end of RNA and does not require ATP for ligation, but only pre-adenylated linkers. The K227Q mutation in T4RNA ligase 2 reduces enzyme lysyl adenylation¹⁸⁹, which minimizes the formation of undesired ligation products such as concatemers and circles. This reduction is likely due to decreased transfer of adenylyl groups from linkers to the 5' phosphate of input RNAs¹⁹⁰. We used this enzyme's unique properties to specifically ligate an RNA adaptor to Ap₂A-RNA, optimizing the method for RNA sequencing (RNAseq). This approach allowed us to identify RNA transcripts capped with Ap₂A moiety.

Using this sequencing method on the isolated short fraction of RNA of Expi293F cells, we aimed to identify transcripts capped at the 5' end with Ap₂A. We anticipated an enrichment of these transcripts in the menadione-treated samples, compared to the control samples, given the increase in Ap₂A-RNA after ROS-induced stress. The sequenced libraries were mapped to the entire human genome, and, surprisingly, we detected the Ap₂A cap in two different RNAs: 5.8S rRNA and U3 snoRNA. Moreover, our mapping revealed that these transcripts were not fragments from longer RNAs but rather complete sequences.

Unfortunately, data analysis did not reveal any significant difference between these transcripts in control samples and menadione treated samples. On the other hand, data analysis showed that in the samples treated with NUDT12, there was a significant loss of the aforementioned transcripts, confirming the accuracy of the sequencing protocol. It is possible that the transcripts that are capped with Ap₂A, and potentially involved in the RNA repair mechanism, are only intermediate products of the ligation process, initially formed and then subsequently ligated in a second step. While LC-MS, being a very sensitive method, detected sum of these intermediates, the sequencing protocol may not have detected them due to their low abundance.

Given that the short fraction of RNA is primarily composed of tRNAs, we were surprised to see that no tRNAs were identified in our study. Since the 5' termini of tRNAs are monophosphorylated, this result could be explained by the usage of Terminator™

exonuclease, which digests RNAs with 5' monophosphate ends, thus removing all tRNAs from the samples.

To confirm the presence of the Ap₂A moiety at the 5' end of 5.8S rRNAs and U3 snoRNAs, we plan to use specific biotinylated probes to capture these two transcripts and analyze them further on LC-MS. Moreover, RT-qPCR experiments with specific primers for these sequences will allow us to measure their relative fold expression, indicating whether their expression levels are higher during menadione induced stress compared to the control cells.

Understanding non-canonical capping is essential for deciphering the full complexity of RNA biology and its regulatory mechanisms. This phenomenon can reveal how cells adapt and diversify their gene expression strategies beyond conventional pathways.

5.2 *Arabidopsis Thaliana* NudiXes have RNA decapping enzyme

The field of RNA decapping is currently gaining significant importance and has become a highly studied topic. Many 5' RNA modifications have only recently been discovered, and numerous enzymes are known to possess decapping activity in vitro. It's plausible that these enzymes may also have the potential to cleave capped RNA in vivo, perhaps under certain conditions or within specific cell types. However, the specific functions of many decapping enzymes remain unknown, as does the function of some recently discovered caps.

Among mammalian NudiX family proteins, several have been shown to exhibit activity against non-canonical caps. Notably, NUDT12 and NUDT16 showed deNADding activity (confirmed in vivo), while NUDT2 and NUDT16 displayed deFADding activity (in vitro for NUDT2 and in vivo for NUDT16). Additionally, NUDT2, NUDT7, NUDT8, NUDT12, NUDT15, NUDT16 and NUDT19 demonstrated deCoAging activity (in vitro).

From these enzymes, there are no relevant homologs in *Arabidopsis thaliana* for NUDT16 and NUDT19, while other proteins have identifiable homologs. Specifically, human NUDT2 shares homology with AtNUDT6, AtNUDT7, AtNUDT26, and AtNUDT27. Human NUDT7 and NUDT8 have homologs in AtNUDT11, AtNUDT15, and AtNUDT22. Human NUDT12 corresponds to AtNUDT19, and human NUDT15 corresponds to AtNUDT1 and AtNUDT23.

Given these reasons, our work aimed to express and purify these proteins, resulting in active enzymes for AtNUDT6, AtNUDT7, AtNUDT19 and AtNUDT27. Initially, we tested whether these enzymes could cleave canonical and non-canonical caps from RNA using in vitro transcribed RNA capped with Np_nNs, CoA, NAD(H), m⁷Gp₃A(G), and ADP-ribose. Our results revealed that AtNUDT19 exhibited high specificity and cleaved NADH-RNA almost exclusively. AtNUDT6 and AtNUDT7 were relatively specific, predominantly cleaving Ap₄A-RNA. In contrast, AtNUDT27 effectively hydrolyzed nearly all types of capped RNA at a

concentration tenfold lower than the other tested enzymes, suggesting it may function as a general RNA-decapping enzyme.

Since the in vitro substrates were not known for all of these enzymes, we also assessed their substrate specificity using free small molecules capable of forming RNA caps (Np_nNs, 3'-dpCoA, NAD(H), m⁷Gp₃A(G), and ADP-ribose). Once more, AtNUDT19 demonstrated high specificity towards NADH, consistent with its reported role of AtNUDT19 in regulating NADH and NADPH levels in chloroplasts¹⁸². Under our experimental conditions, AtNUDT19 exclusively cleaved free NADH but not NAD. Interestingly, when capped RNA was employed as a substrate, AtNUDT19 cleaved both NADH-RNA and partially NAD-RNA. This suggests a potential role for AtNUDT19 as an NAD/NADH-RNA-decapping enzyme.

Contrary to its promiscuity with capped RNA, AtNUDT27 hydrolyzed exclusively dinucleoside tetraphosphates and pentaphosphates. Our observed activity of AtNUDT27 against dinucleoside tetraphosphates contrasts with previously reported results¹⁸². In that study, the authors observed cleavage of only Ap₅A but not Ap₄A. Nevertheless, the presence of reducing agents may positively influence the activity of these enzymes¹⁹¹. Thus, this discrepancy may be explained by the presence of a higher concentration of reducing agent DTT (2 mM) in our reaction in comparison with their 1 mM DTT concentration.

AtNUDT6 and AtNUDT7 mainly cleaved free molecules with the diphosphate bridge. In fact, both AtNUDT6 and AtNUDT7 were identified as cytosolic NADH pyrophosphohydrolases participating in stress responses by modulating intracellular NADH levels in *A. thaliana*¹⁹². Ap₄A is also recognized for its involvement in stress signaling across various organisms. Since AtNUDT6, AtNUDT7, and AtNUDT27 efficiently cleaved both the free Ap₄A and the Ap₄A-RNA, we compared their preferences for cleaving these substrates. We observed that AtNUDT27 and AtNUDT6 cleaved free Ap₄A more efficiently, while AtNUDT7 favored Ap₄A-RNA.

These differences in substrate specificities suggest that in vivo, some NudiX proteins may function in degradation of free molecules, while others serve as specific decapping enzymes, likely during regulated RNA decay.

It has been shown that diadenosine polyphosphates such as Ap₃A and Ap₄A act as alarmones in *A. thaliana*, triggering a cascade of reactions leading to the production of various protective compounds¹⁹³. This finding supports the notion that plant RNA may contain other unknown RNA caps besides the canonical m⁷G structure and NAD cap. Altogether, the existence of specific enzymes that cleave NADH- and Ap₄A-RNA suggests that these RNA caps may play important roles in plant RNA metabolism, signaling, or stress response.

Our results indicate that plant Nudix enzymes may fulfil the role of RNA-decapping enzymes specific for non-canonically capped RNA. The specific cleavage properties of the individual Nudix enzymes characterized in this work can be used to develop capturing strategies for non-canonically capped RNA, thus helping us to identify the sequence and features of these RNA species.

In collaboration with Martin Lepšík from Institute of Organic Chemistry and Biochemistry (IOCB), we used molecular modeling to investigate the active site of AtNUDT7 in interaction with Ap₄A-RNA. Our results demonstrated that the active site of the enzyme can interact specifically with the Ap₄A-RNA (Figure 38).

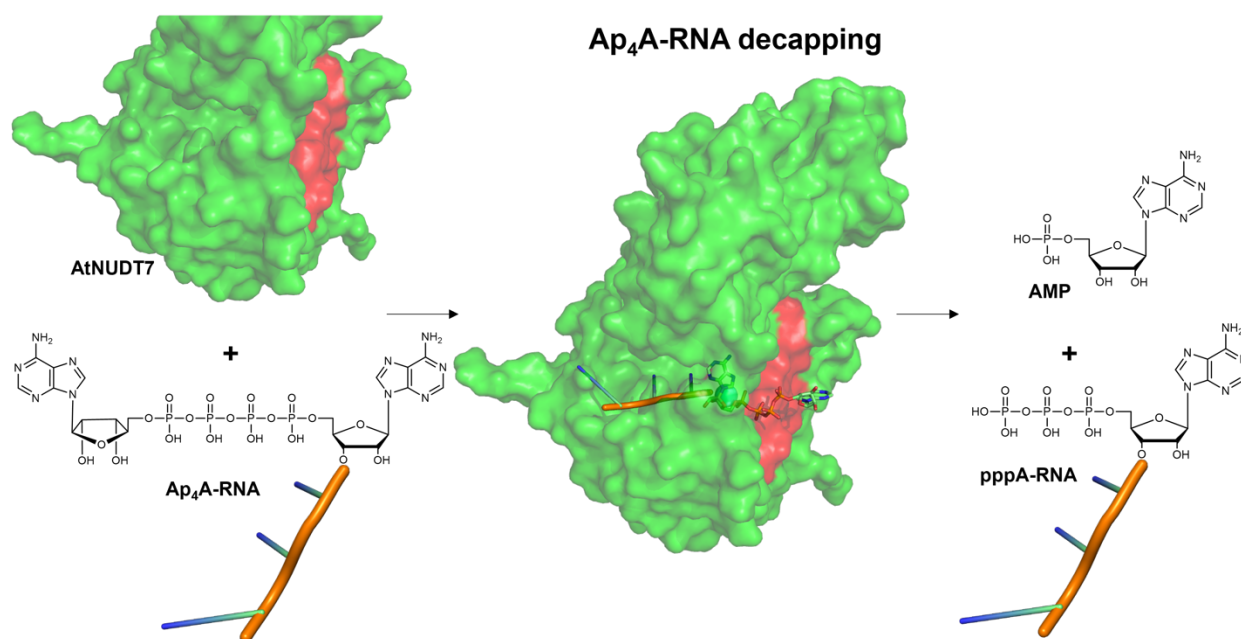


Figure 38. Molecular modelling of the interaction between AtNUDT7 and Ap₄A-RNA.

Taking all this data together, it would be interesting to isolate both short and long fractions of RNA from *Arabidopsis Thaliana* plants and search for Np_nN-RNAs, especially Ap₄A-RNA, using LC-MS technique.

5.3 Human NUDT2 and mouse DXO as decapping enzymes

Besides *A. thaliana* Nudix enzymes, I also worked with human NUDT2 Nudix enzyme and mouse DXO decapping exoribonuclease, testing their cleavage activity on non-canonical capped RNAs.

The human NUDT2 enzyme, part of the Nudix hydrolase family, plays a key role in the metabolism of various nucleotide derivatives, including diadenosine polyphosphates such as Ap₄A (diadenosine tetraphosphate). NUDT2 functions as an Ap₄A hydrolase, catalyzing the hydrolysis of Ap₄A into AMP and ATP. Ap₄A, as a free molecule, is the most abundant dinucleoside polyphosphate in eukaryotic cells and can be incorporated into RNA by RNA polymerase as non-canonical initiating nucleotide (NCIN)^{86,88}. However, Ap₄A likely has diverse roles due to its varied protein interactome¹⁹⁴.

In our group, we reported the existence of a new non-canonical RNA cap, diadenosine tetraphosphate (Ap₄A), in human and rat cell lines. It has been known that NUDT2 cleaves free Ap₄A, and NUDT2 knockout cells show a 175-fold increase in intracellular Ap₄A concentration, with significantly altered transcriptional profiles, including down-regulation of genes involved in the innate immune response and in cancer growth promotion¹⁹⁵.

Given those results, I isolated and purified recombinant human NUDT2 and tested its activity on Ap₄A and in vitro prepared Ap₄A-RNA. I demonstrated that human NUDT2 is a decapping enzyme for the newly discovered Ap₄A-RNA in vitro, observing complete decapping of Ap₄A-RNA at 500 nM of NUDT2. The products of cleavage were AMP and 5'-triphosphate RNA (ppp-RNA). At higher concentrations, NUDT2 also cleaved ppp-RNA into p-RNA, as previously reported¹⁸³. In addition, NUDT2 cleaved several other Np_nN capped RNAs, exhibiting lower activity towards Np₃N-RNAs. These results indicate that human NUDT2 is a general RNA decapping enzyme.

Apart from NudiX family, the human enzyme Decapping Exoribonuclease (DXO) has also been shown to cleave RNA with an affinity towards NAD-RNA⁶². In our research, we tested this enzyme for Ap₄A-RNA and other non-canonical RNA caps. Unlike NUDT2, DXO was unable to cleave free Np_nNs; instead, it removed the entire RNA cap, leaving the transcript one nucleotide shorter. Interestingly, Ap₄A-RNA is cleaved within the polyphosphate bridge, which contrasts its activity on NAD-RNA or Ap₂A-RNA. It remains to be seen whether DXO functions similarly in vivo. Considering all the previous data, we have shown that both NUDT2 and DXO are decapping enzymes for newly discovered Ap₄A-RNA.

Moreover, both enzymes showed a cleavage activity towards non-canonical capped RNAs. Specifically, NUDT2 cleaved almost all the tested Np_nN-RNAs, showing lower active towards triphosphates RNAs such as Gp₃G, Ap₃G and Gp₃A. On the other hand, DXO showed a preference for diphosphate- and tetraphosphate- capped RNAs, cleaving Ap₂A, Ap₄A, Gp₂A, Gp₄A, Gp₂G, and Gp₄G. Our results reported different non-canonical RNA caps as substrates for NUDT2 and DXO enzymes in vitro.

6 Conclusions

My PhD project focused on the discovery of novel 5' RNA non-canonical caps and the identification of enzymes, specifically Nudix enzyme, responsible for their degradation. The key findings of my study are summarized as follows:

1. Utilizing liquid chromatography–mass spectrometry (LC–MS), novel non-canonical RNA caps Ap₂A-RNA, Ap₂G-RNA, Ap₂C-RNA, and Ap₂U-RNA have been identified in human Expi293F cells. Notably, the concentration of Ap₂A-RNA increases in both long and short RNA fractions following treatment with menadione, a molecule known to induce reactive oxygen species (ROS)-mediated cellular stress.
2. Mouse DXO and mammalian NUDT12 enzymes have been identified as responsible for the cleavage of Ap₂A-RNA in vitro. NUDT12 exhibited higher activity towards Ap₂A-RNA compared to its known substrate, NAD-RNA.
3. The human ligase RLIG1 might act as a capping enzyme for Ap₂A-RNA and Ap₂G-RNA, as we demonstrated its ability to create Ap₂A-RNA and Ap₂G-RNA in vitro, utilizing monophosphate (p-A/G-RNA) and ATP as substrates.
4. A novel sequencing technique was developed, enabling the identification of 5.8S rRNA and U3 snoRNA as potential RNA transcripts capped at the 5' end with Ap₂A moiety.
5. AtNudt27 was identified as a promising general RNA decapping enzyme. Ap₄A-RNA was identified as the best substrate for AtNUDT6, AtNUDT7, and AtNUDT27, suggesting its potential presence in *A. thaliana*. AtNUDT6, AtNUDT7 can be utilized for the development of capturing techniques for identification of Ap₄A-RNA, while AtNUDT19 can be used to develop capturing techniques for identification of NADH-RNA.
6. Finally, mammalian NUDT2 and mouse DXO were confirmed to be decapping enzymes for recently discovered mammalian Ap₄A-RNA in vitro.

7 References

1. Higgs, P. G. & Lehman, N. The RNA World: molecular cooperation at the origins of life. *Nature Reviews Genetics* 2014 16:1 **16**, 7–17 (2014).
2. Yi, C. & Pan, T. Cellular dynamics of RNA modification. *Acc Chem Res* **44**, 1380–1388 (2011).
3. Kutter, C. & Svoboda, P. miRNA, siRNA, piRNA: Knowns of the unknown. *RNA Biol* **5**, 181–188 (2008).
4. Cappannini, A. *et al.* MODOMICS: a database of RNA modifications and related information. 2023 update. *Nucleic Acids Res* **52**, D239–D244 (2024).
5. Boccaletto, P. *et al.* MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res* **50**, D231–D235 (2022).
6. Helm, M. & Alfonzo, J. D. Posttranscriptional RNA Modifications: Playing Metabolic Games in a Cell's Chemical Legoland. *Chem Biol* **21**, 174–185 (2014).
7. Chatterjee, B., Shen, C. K. J. & Majumder, P. RNA Modifications and RNA Metabolism in Neurological Disease Pathogenesis. *Int J Mol Sci* **22**, (2021).
8. Schieweck, R., Ninkovic, J. & Kiebler, M. A. RNA-BINDING PROTEINS BALANCE BRAIN FUNCTION IN HEALTH AND DISEASE. *Physiol Rev* **101**, 1309–1370 (2021).
9. Nussbacher, J. K., Tabet, R., Yeo, G. W. & Lagier-Tourenne, C. Disruption of RNA Metabolism in Neurological Diseases and Emerging Therapeutic Interventions. *Neuron* **102**, 294–320 (2019).
10. Li, X., Xiong, X. & Yi, C. Epitranscriptome sequencing technologies: decoding RNA modifications. *Nat. Methods* **14**, 23–31 (2016).
11. Wang, Y., Zhang, X., Liu, H. & Zhou, X. Chemical methods and advanced sequencing technologies for deciphering mRNA modifications. *Chem. Soc. Rev.* **50**, 13481–13497 (2021).
12. Zhang, W., Qian, Y. & Jia, G. The detection and functions of RNA modification m6A based on m6A writers and erasers. *J Biol Chem* **297**, 100973–100974 (2021).
13. Jiang, X. *et al.* The role of m6A modification in the biological functions and diseases. *Signal Transduction and Targeted Therapy* vol. 6 Preprint at <https://doi.org/10.1038/s41392-020-00450-x> (2021).

14. RajBhandary, U. L., Stuart, A., Faulkner, R. D., Chang, S. H. & Khorana, H. G. Nucleotide sequence studies on yeast phenylalanine sRNA. *Cold Spring Harb Symp Quant Biol* **31**, 425–434 (1966).
15. Shima, H. & Igarashi, K. N 1-methyladenosine (m1A) RNA modification: the key to ribosome control. *J Biochem* **167**, 535–539 (2020).
16. Dominissini, D. *et al.* The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. *Nature* 2016 530:7591 **530**, 441–446 (2016).
17. Li, X. *et al.* Base-Resolution Mapping Reveals Distinct m1A Methylome in Nuclear- and Mitochondrial-Encoded Transcripts. *Mol Cell* **68**, 993-1005.e9 (2017).
18. Zhang, C. & Jia, G. Reversible RNA Modification N1-methyladenosine (m1A) in mRNA and tRNA. *Genomics Proteomics Bioinformatics* **16**, 155 (2018).
19. Song, H. *et al.* Biological roles of RNA m5C modification and its implications in Cancer immunotherapy. *Biomarker Research* 2022 10:1 **10**, 1–15 (2022).
20. Edelheit, S., Schwartz, S., Mumbach, M. R., Wurtzel, O. & Sorek, R. Transcriptome-wide mapping of 5-methylcytidine RNA modifications in bacteria, archaea, and yeast reveals m5C within archaeal mRNAs. *PLoS Genet* **9**, (2013).
21. Lan, J. *et al.* Functional role of Tet-mediated RNA hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature Communications* 2020 11:1 **11**, 1–15 (2020).
22. Delatte, B. *et al.* Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science* (1979) **351**, 282–285 (2016).
23. Spenkuch, F., Motorin, Y. & Helm, M. Pseudouridine: still mysterious, but never a fake (uridine)! *RNA Biol* **11**, 1540–1554 (2014).
24. Ge, J. & Yu, Y. T. RNA pseudouridylation: new insights into an old modification. *Trends Biochem Sci* **38**, 210–218 (2013).
25. Charette, M. & Gray, M. W. Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* **49**, 341–351 (2000).
26. Jack, K. *et al.* rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Mol Cell* **44**, 660–666 (2011).
27. Schwartz, S. *et al.* Transcriptome-wide Mapping Reveals Widespread Dynamic-Regulated Pseudouridylation of ncRNA and mRNA. *Cell* **159**, 148–162 (2014).
28. Li, X. *et al.* Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. *Nat Chem Biol* **11**, 592–597 (2015).

29. Dönmez, G., Hartmuth, K. & Lührmann, R. Modified nucleotides at the 5' end of human U2 snRNA are required for spliceosomal E-complex formation. *RNA* **10**, 1925–1933 (2004).
30. Zhao, X. & Yu, Y. T. Pseudouridines in and near the branch site recognition region of U2 snRNA are required for snRNP biogenesis and pre-mRNA splicing in *Xenopus* oocytes. *RNA* **10**, 681–690 (2004).
31. Zhao, Y., Rai, J. & Li, H. Regulation of translation by ribosomal RNA pseudouridylation. *Sci Adv* **9**, (2023).
32. Adachi, H. & Yu, Y.-T. Insight into the mechanisms and functions of spliceosomal snRNA pseudouridylation. *World J Biol Chem* **5**, 398 (2014).
33. BRAND, R. C., KLOOTWIJK, J., VAN STEENBERGEN, T. J. M., DE KOK, A. J. & PLANTA, R. J. Secondary methylation of yeast ribosomal precursor RNA. *Eur J Biochem* **75**, 311–318 (1977).
34. Vold, B. Modified nucleosides of *Bacillus subtilis* transfer ribonucleic acids. *J Bacteriol* **127**, 258 (1976).
35. Ouellette, A. J., Frederick, D. & Malt, R. A. Methylated messenger RNA in mouse kidney. *Biochemistry* **14**, 4361–4367 (1975).
36. Nichols, J. L. & Welder, L. THE MODIFIED NUCLEOTIDE CONSTITUENTS OF HUMAN PROSTATIC CANCER CELL (MA-160) POLY(A)-CONTAINING RNA. *Biochim Biophys Acta* **608**, 1–18 (1980).
37. Heckle, W. L., Fenton, R. G., Wood, T. G., Merkel, C. G. & Lingrel, J. B. Methylated nucleosides in globin mRNA from mouse nucleated erythroid cells. *Journal of Biological Chemistry* **252**, 1764–1770 (1977).
38. Dimock, K. & Stoltzfus, C. M. Processing and function of undermethylated chicken embryo fibroblast mRNA. *Journal of Biological Chemistry* **254**, 5591–5594 (1979).
39. Tae Suk Ro Choi, Choi, Y. C. & Henning, D. Nucleotide sequence of U-2 ribonucleic acid. The sequence of the 5'-terminal oligonucleotide. *Journal of Biological Chemistry* **250**, 3921–3928 (1975).
40. Gray, M. W. Evolutionary origin of RNA editing. *Biochemistry* **51**, 5235–5242 (2012).
41. Sun, H., Li, K., Liu, C. & Yi, C. Regulation and functions of non-m6A mRNA modifications. *Nat Rev Mol Cell Biol* **24**, 714–731 (2023).

42. Dutta, N., Deb, I., Sarzynska, J. & Lahiri, A. Inosine and its methyl derivatives: Occurrence, biogenesis, and function in RNA. *Prog Biophys Mol Biol* **169–170**, 21–52 (2022).
43. Su, A. A. H. & Randau, L. A-to-I and C-to-U editing within transfer RNAs. *Biochemistry (Mosc)* **76**, 932–937 (2011).
44. Mendoza, H. G. & Beal, P. A. Structural and functional effects of inosine modification in mRNA. *RNA* **30**, 512–520 (2024).
45. Srinivasan, S., Torres, A. G. & Ribas de Pouplana, L. Inosine in Biology and Disease. *Genes (Basel)* **12**, (2021).
46. Vasilyev, N., Gao, A. & Serganov, A. Non-canonical features and modifications on the 5'- end of bacterial sRNAs and mRNAs. *Wiley Interdiscip Rev RNA* **10**, e1509 (2019).
47. Jorgensen, S. E., Buch, L. B. & Nierlich, D. P. Nucleoside triphosphate termini from RNA synthesized in vivo by Escherichia coli. *Science (1979)* **164**, 1067–1070 (1969).
48. Furuichi, Y., LaFiandra, A. & Shatkin, A. J. 5'-Terminal structure and mRNA stability. *Nature* **266**, 235–239 (1977).
49. Muthukrishnan, S., Both, G. W., Furuichi, Y. & Shatkin, A. J. 5'-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation. *Nature* **255:5503** **255**, 33–37 (1975).
50. Züst, R. *et al.* Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nature Immunology* **2011 12:2** **12**, 137–143 (2011).
51. Despici, V. & Jaffrey, S. R. mRNA ageing shapes the Cap2 methylome in mammalian mRNA. *Nature* **614**, 358 (2023).
52. Audrito, V., Messana, V. G., Brandimarte, L. & Deaglio, S. The Extracellular NADome Modulates Immune Responses. *Front Immunol* **12**, (2021).
53. Cuenoud, B. *et al.* Effect of circadian rhythm on NAD and other metabolites in human brain. *Front Physiol* **14**, (2023).
54. Schenberger Santos, A. R. *et al.* NAD⁺ biosynthesis in bacteria is controlled by global carbon/nitrogen levels via PII signaling. *Journal of Biological Chemistry* **295**, 6165–6176 (2020).

55. Van Dao, T. *et al.* NAD-Malic Enzyme Affects Nitrogen Fixing Activity of Bradyrhizobium japonicum USDA 110 Bacteroids in Soybean Nodules. *Microbes Environ* **23**, 215–220 (2008).
56. Hottiger, M. O., Hassa, P. O., Lüscher, B., Schöler, H. & Koch-Nolte, F. Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends Biochem Sci* **35**, 208–219 (2010).
57. 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).
58. Cahová, H., Winz, M. L., Höfer, K., Nübel, G. & Jäschke, A. NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. *Nature* **519**, 374–377 (2015).
59. Vvedenskaya, I. O. & Nickels, B. E. CapZyme-Seq: A 5'-RNA-Seq Method for Differential Detection and Quantitation of NAD-Capped and Uncapped 5'-Triphosphate RNA. *STAR Protoc* **1**, (2020).
60. Shao, X. *et al.* NAD tagSeq for transcriptome-wide identification and characterization of NAD⁺-capped RNAs. *Nature Protocols* 2020 15:9 **15**, 2813–2836 (2020).
61. Walters, R. W. *et al.* Identification of NAD⁺ capped mRNAs in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **114**, 480–485 (2017).
62. Jiao, X. *et al.* 5'-end NAD⁺ cap in human cells promotes RNA decay through DXO-mediated deNADding. *Cell* **168**, 1015 (2017).
63. 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).
64. Grudzien-Nogalska, E. *et al.* Structural and mechanistic basis of mammalian Nudt12 RNA deNADding. *Nat Chem Biol* **15**, 575–582 (2019).
65. Zhang, H. *et al.* NAD tagSeq reveals that NAD⁺-capped RNAs are mostly produced from a large number of protein-coding genes in Arabidopsis. *Proc Natl Acad Sci U S A* **116**, 12072–12077 (2019).
66. Yimcharoen, M. *et al.* Mycobacterium tuberculosis Adaptation in Response to Isoniazid Treatment in a Multi-Stress System That Mimics the Host Environment. *Antibiotics* **12**, (2023).
67. Belenky, P., Bogan, K. L. & Brenner, C. NAD⁺ metabolism in health and disease. *Trends Biochem Sci* **32**, 12–19 (2007).

68. Wolfram-Schauerte, M. *et al.* A viral ADP-ribosyltransferase attaches RNA chains to host proteins. *Nature* 2023 620:7976 **620**, 1054–1062 (2023).
69. 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).
70. Löcherer, C., Bühler, N., Lafrenz, P. & Jäschke, A. Staphylococcus aureus Small RNAs Possess Dephospho-CoA 5'-Caps, but No CoAlation Marks. *Noncoding RNA* **8**, (2022).
71. Shao, X. *et al.* DpCoA tagSeq: Barcoding dpCoA-Capped RNA for Direct Nanopore Sequencing via Maleimide-Thiol Reaction. *Anal Chem* **95**, 11124–11131 (2023).
72. Huang, F. Efficient incorporation of CoA, NAD and FAD into RNA by in vitro transcription. *Nucleic Acids Res* **31**, (2003).
73. Sapkota, K. *et al.* Post-transcriptional capping generates coenzyme A-linked RNA. *RNA Biol* **21**, 1–12 (2024).
74. Huang, F. Efficient incorporation of CoA, NAD and FAD into RNA by in vitro transcription. *Nucleic Acids Res* **31**, (2003).
75. Zhou, W. *et al.* Structural insights into dpCoA-RNA decapping by NudC. *RNA Biol* **18**, 244–253 (2021).
76. Malygin, A. G. & Shemyakin, M. F. Adenosine, NAD and FAD can initiate template-dependent RNA a synthesis catalyzed by Escherichia Coli RNA polymerase. *FEBS Lett* **102**, 51–54 (1979).
77. Wang, J. *et al.* Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA. *Nucleic Acids Res* **47**, e130–e130 (2019).
78. Doamekpor, S. K. *et al.* DXO/Rai1 enzymes remove 5'-end FAD and dephospho-CoA caps on RNAs. *Nucleic Acids Res* **48**, 6136 (2020).
79. Marceau, C. D. *et al.* Genetic dissection of Flaviviridae host factors through genome-scale CRISPR screens. *Nature* 2016 535:7610 **535**, 159–163 (2016).
80. Cai, Z., Liang, T. J. & Luo, G. Effects of mutations of the initiation nucleotides on hepatitis C virus RNA replication in the cell. *J Virol* **78**, 3633–3643 (2004).
81. Sherwood, A. V. *et al.* Hepatitis C virus RNA is 5'-capped with flavin adenine dinucleotide. *Nature* **619**, 811–818 (2023).

82. Lüscher, B. *et al.* ADP-Ribosylation, a Multifaceted Posttranslational Modification Involved in the Control of Cell Physiology in Health and Disease. *Chem Rev* **118**, 1092–1136 (2018).
83. Weixler, L., Feijs, K. L. H. & Zaja, R. ADP-ribosylation of RNA in mammalian cells is mediated by TRPT1 and multiple PARPs. *Nucleic Acids Res* **50**, 9426 (2022).
84. Hopp, A. K. & Hottiger, M. O. Uncovering the Invisible: Mono-ADP-ribosylation Moved into the Spotlight. *Cells* **10**, 1–25 (2021).
85. Munnur, D. *et al.* Reversible ADP-ribosylation of RNA. *Nucleic Acids Res* **47**, 5658 (2019).
86. Hudeček, O. *et al.* Dinucleoside polyphosphates act as 5'-RNA caps in bacteria. *Nature Communications* 2020 11:1 **11**, 1–11 (2020).
87. Luciano, D. J., Levenson-Palmer, R. & Belasco, J. G. Stresses that Raise Np4A Levels Induce Protective Nucleoside Tetraphosphate Capping of Bacterial RNA. *Mol Cell* **75**, 957-966.e8 (2019).
88. Benoni, R., Culka, M., Hudeček, O., Gahurova, L. & Cahová, H. Dinucleoside Polyphosphates as RNA Building Blocks with Pairing Ability in Transcription Initiation. *ACS Chem Biol* **15**, 1765–1772 (2020).
89. 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).
90. 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).
91. Coste, H., Brevet, A., Plateau, P. & Blanquet, S. Non-adenylylated bis(5'-nucleosidyl) tetraphosphates occur in *Saccharomyces cerevisiae* and in *Escherichia coli* and accumulate upon temperature shift or exposure to cadmium. *Journal of Biological Chemistry* **262**, 12096–12103 (1987).
92. Braun, F., Recalde, A., Bähre, H., Seifert, R. & Albers, S. V. Putative Nucleotide-Based Second Messengers in the Archaeal Model Organisms *Haloferax volcanii* and *Sulfolobus acidocaldarius*. *Front Microbiol* **12**, 779012 (2021).

93. 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).
94. Krauspe, R., Parthier, B. & Wasternack, C. Synthesis of diadenosine 5',5''-P₁,P₄-tetraphosphate by organellar and cytoplasmic phenylalanyl-tRNA synthetases of *Euglena gracilis*. *FEBS Lett* **235**, 275–277 (1988).
95. 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).
96. 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).
97. Guranowski, A., Sillero, M. A. G. & Sillero, A. Firefly luciferase synthesizes P₁,P₄-bis(5'-adenosyl)tetraphosphate (Ap₄A) and other dinucleoside polyphosphates. *FEBS Lett* **271**, 215–218 (1990).
98. Guranowski, A., Just, G., Holler, E. & Jakubowski, H. Synthesis of Diadenosine 5',5''-: P₁,: P₄-tetraphosphate (appppa) from Adenosine 5'-phosphosulfate and Adenosine 5'-triphosphate Catalyzed by Yeast Appppa Phosphorylase. *Biochemistry* **27**, 2959–2964 (1988).
99. Bradford, J. F. *et al.* A., & Dinucleoside Tetraphosphate Synthetase. *Adv. Enzymol. Relat. Areas Mol. Biol* **26**, 87 (1987).
100. McLennan, A. G. Dinucleoside polyphosphates—friend or foe? *Pharmacol Ther* **87**, 73–89 (2000).
101. Kisselev, L. L., Justesen, J., Wolfson, A. D. & Frolova, L. Y. Diadenosine oligophosphates (Ap(n)A), a novel class of signalling molecules? *FEBS Lett* **427**, 157–163 (1998).
102. Flodgaard, H. & Klenow, H. Abundant amounts of diadenosine 5',5''-P₁,P₄-tetraphosphate are present and releasable, but metabolically inactive, in human platelets. *Biochemical Journal* **208**, 737 (1982).
103. Lüthje, J. & Ogilvie, A. Diadenosine triphosphate (Ap₃A) mediates human platelet aggregation by liberation of ADP. *Biochem Biophys Res Commun* **118**, 704–709 (1984).

104. Flodgaard, H. & Klenow, H. Abundant amounts of diadenosine 5',5'''-P₁,P₄-tetrphosphate are present and releasable, but metabolically inactive, in human platelets. *Biochem J* **208**, 737–742 (1982).
105. Schlüter, H. *et al.* Diadenosine phosphates and the physiological control of blood pressure. *Nature* **367**, 186–188 (1994).
106. Jankowski, J. *et al.* Identification and characterization of P(1), P(7)-Di(adenosine-5')-heptaphosphate from human platelets. *J Biol Chem* **274**, 23926–23931 (1999).
107. Schlüter, H. *et al.* Adenosine(5') oligophospho-(5') guanosines and guanosine(5') oligophospho-(5') guanosines in human platelets. *J Clin Invest* **101**, 682–688 (1998).
108. Baltzinger, M., Ebel, J. P. & Remy, P. Accumulation of dinucleoside polyphosphates in *Saccharomyces cerevisiae* under stress conditions. High levels are associated with cell death. *Biochimie* **68**, 1231–1236 (1986).
109. Guédon, G. F., Gilson, G. J. P., Ebel, J. P., Befort, N. M. & Remy, P. M. Lack of correlation between extensive accumulation of bisnucleoside polyphosphates and the heat-shock response in eukaryotic cells. *Journal of Biological Chemistry* **261**, 16459–16465 (1986).
110. Kisselev, L. L., Justesen, J., Wolfson, A. D. & Frolova, L. Y. Diadenosine oligophosphates (ApnA), a novel class of signalling molecules? *FEBS Lett* **427**, 157–163 (1998).
111. Holmuhamedov, E. L., Jovanović, S., Dzeja, P. P., Jovanović, A. & Terzic, A. Mitochondrial ATP-sensitive K⁺ channels modulate cardiac mitochondrial function. *Am J Physiol Heart Circ Physiol* **275**, (1998).
112. Klishin, A., Lozovaya, N., Pintor, J., Miras-Portugal, M. T. & Krishtal, O. Possible functional role of diadenosine polyphosphates: Negative feedback for excitation in hippocampus. *Neuroscience* **58**, 235–236 (1994).
113. Chan, P. J., Su, B. C. & Tredway, D. R. Diadenosine Tetrphosphate (Ap₄A) and Triphosphate (Ap₃A) Signaling of Human Sperm Motility. *Arch Androl* **27**, 103–108 (1991).
114. Green, A. K., Cobbold, P. H. & Dixon, C. J. Cytosolic free Ca²⁺ oscillations induced by diadenosine 5',5'''-P₁,P₃-triphosphate and diadenosine 5',5'''-P₁,P₄-tetrphosphate in single rat hepatocytes are indistinguishable from those induced by ADP and ATP respectively. *Biochemical Journal* **310**, 629–635 (1995).

115. Gasmi, L., McLennan, A. G. & Edwards, S. W. Priming of the respiratory burst of human neutrophils by the diadenosine polyphosphates, Ap4A and Ap3A: Role of intracellular Calcium. *Biochem Biophys Res Commun* **202**, 218–224 (1994).
116. Ripoll, C. *et al.* Diadenosine Polyphosphates: A Novel Class of Glucose-Induced Intracellular Messengers in the Pancreatic β -Cell. *Diabetes* **45**, 1431–1434 (1996).
117. Jovanovic, A., Zhang, S., Alekseev, A. E. & Terzic, A. Diadenosine polyphosphate-induced inhibition of cardiac KATP channels: operative state-dependent regulation by a nucleoside diphosphate. *Pflugers Arch* **431**, 800–802 (1996).
118. Jovanovic, A. & Terzic, A. Diadenosine tetraphosphate-induced inhibition of ATP-sensitive K⁺ channels in patches excised from ventricular myocytes. *Br J Pharmacol* **117**, 233–235 (1996).
119. Ripoll, C. *et al.* Diadenosine polyphosphates. A novel class of glucose-induced intracellular messengers in the pancreatic beta-cell. *Diabetes* **45**, 1431–1434 (1996).
120. Jovanovic, A., Alekseev, A. E. & Terzic, A. Intracellular diadenosine polyphosphates. A novel family of inhibitory ligands of the ATP-sensitive K⁺ channel. *Biochem Pharmacol* **54**, 219–225 (1997).
121. Green, A. K., Cobbold, P. H. & Dixon, C. J. Cytosolic free Ca²⁺ oscillations induced by diadenosine 5',5''-P₁,P₃-triphosphate and diadenosine 5',5''-P₁,P₄-tetraphosphate in single rat hepatocytes are indistinguishable from those induced by ADP and ATP respectively. *Biochem J* **310** (Pt 2), 629–635 (1995).
122. Holden, C. P., Padua, R. A. & Geiger, J. D. Regulation of ryanodine receptor calcium release channels by diadenosine polyphosphates. *J Neurochem* **67**, 574–580 (1996).
123. Pintor, J., Torres, M., Castro, E. & Miras-Portugal, M. T. Characterization of diadenosine tetraphosphate (Ap4A) binding sites in cultured chromaffin cells: evidence for a P_{2y} site. *Br J Pharmacol* **103**, 1980–1984 (1991).
124. Rodriguez del Castillo, A., Torres, M., Delicado, E. G. & Miras-Portugal, M. T. Subcellular Distribution Studies of Diadenosine Polyphosphates—Ap4A and Ap5A—in Bovine Adrenal Medulla: Presence in Chromaffin Granules. *J Neurochem* **51**, 1696–1703 (1988).
125. Klishin, A., Lozovaya, N., Pintor, J., Miras-Portugal, M. T. & Krishtal, O. Possible functional role of diadenosine polyphosphates: negative feedback for excitation in hippocampus. *Neuroscience* **58**, 235–236 (1994).

126. Jankowski, J. *et al.* Dinucleotides as Growth-promoting Extracellular Mediators: PRESENCE OF DINUCLEOSIDE DIPHOSPHATES Ap2A, Ap2G, AND Gp2G IN RELEASABLE GRANULES OF PLATELETS. *Journal of Biological Chemistry* **276**, 8904–8909 (2001).
127. Vartanian, A., Alexandrov, I., Prudowski, I., McLennan, A. & Kisselev, L. Ap4A induces apoptosis in human cultured cells. *FEBS Lett* **456**, 175–180 (1999).
128. Interleukin-2 regulation of diadenosine 5',5'''-p1 p4-tetraphosphate (Ap4A) levels and DNA synthesis in cloned murine T lymphocytes - PubMed. <https://pubmed.ncbi.nlm.nih.gov/2104244/>.
129. Baril, E. F., Coughlin, S. A. & Zamecnik, P. C. 5',5'''-P1, P4 diadenosine tetraphosphate (Ap4A): a putative initiator of DNA replication. *Cancer Invest* **3**, 465–471 (1985).
130. Rapaport, E. & Zamecnik, P. C. Presence of diadenosine 5',5''' -P1, P4-tetraphosphate (Ap4A) in mamalian cells in levels varying widely with proliferative activity of the tissue: a possible positive 'pleiotypic activator'. *Proc Natl Acad Sci U S A* **73**, 3984–3988 (1976).
131. Rapaport, E., Zamecnik, P. C. & Baril, E. F. HeLa cell DNA polymerase alpha is tightly associated with tryptophanyl-tRNA synthetase and diadenosine 5',5'''-P1,P4-tetraphosphate binding activities. *Proc Natl Acad Sci U S A* **78**, 838–842 (1981).
132. Schepers, E. *et al.* Dinucleoside polyphosphates: newly detected uraemic compounds with an impact on leucocyte oxidative burst. *Nephrology Dialysis Transplantation* **25**, 2636–2644 (2010).
133. Worbs, T., Hammerschmidt, S. I. & Förster, R. Dendritic cell migration in health and disease. *Nat Rev Immunol* **17**, 30–48 (2017).
134. Kim, B. K. *et al.* Diadenosine 5',5'''-p,p4-Tetraphosphate Deficiency in Blood Platelets of the Chédiak-Higashi Syndrome. *Blood* **66**, 735–737 (1985).
135. La Shu, S. *et al.* Ap4A Regulates Directional Mobility and Antigen Presentation in Dendritic Cells. *iScience* **16**, 524–534 (2019).
136. Despotović, D. *et al.* Diadenosine tetraphosphate (Ap4A) – an E. coli alarmone or a damage metabolite? *FEBS J* **284**, 2194–2215 (2017).
137. Zegarra, V., Mais, C.-N., Freitag, J. & Bange, G. The mysterious diadenosine tetraphosphate (AP4A). *microLife* **4**, 1–8 (2023).
138. 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).

139. Vidya, E. & Duchaine, T. F. Eukaryotic mRNA Decapping Activation. *Front Genet* **13**, 832547 (2022).
140. Charenton, C. & Graille, M. mRNA decapping: finding the right structures. *Philosophical Transactions of the Royal Society B: Biological Sciences* **373**, (2018).
141. Song, M. G., Bail, S. & Kiledjian, M. Multiple Nudix family proteins possess mRNA decapping activity. *RNA* **19**, 390–399 (2013).
142. Doamekpor, S. K., Gozdek, A., Kwasnik, A., Kufel, J. & Tong, L. A novel 5'-hydroxyl dinucleotide hydrolase activity for the DXO/Rai1 family of enzymes. *Nucleic Acids Res* **48**, 349–358 (2020).
143. Chang, J. H. *et al.* Dxo1 is a new type of eukaryotic enzyme with both decapping and 5'-3' exoribonuclease activity. *Nat Struct Mol Biol* **19**, 1011–1019 (2012).
144. Brenner, C., Biecanowski, P., Pace, H. C. & Huebner, K. The Histidine Triad Superfamily of Nucleotide-Binding Proteins. *J Cell Physiol* **181**, 179 (1999).
145. Liu, S. W. *et al.* Functional analysis of mRNA scavenger decapping enzymes. *RNA* **10**, 1412 (2004).
146. Kramer, S. The ApaH-like phosphatase TbALPH1 is the major mRNA decapping enzyme of trypanosomes. *PLoS Pathog* **13**, e1006456 (2017).
147. McLennan, A. G. The Nudix hydrolase superfamily. *Cellular and Molecular Life Sciences* **63**, 123–143 (2006).
148. Carreras-Puigvert, J. *et al.* A comprehensive structural, biochemical and biological profiling of the human NUDIX hydrolase family. *Nat Commun* **8**, (2017).
149. Ishibashi, T. *et al.* Mammalian enzymes for preventing transcriptional errors caused by oxidative damage. *Nucleic Acids Res* **33**, 3779 (2005).
150. Cai, J. P., Ishibashi, T., Takagi, Y., Hayakawa, H. & Sekiguchi, M. Mouse MTH2 protein which prevents mutations caused by 8-oxoguanine nucleotides. *Biochem Biophys Res Commun* **305**, 1073–1077 (2003).
151. Ishibashi, T., Hayakawa, H. & Sekiguchi, M. A novel mechanism for preventing mutations caused by oxidation of guanine nucleotides. *EMBO Rep* **4**, 479 (2003).
152. Rai, P. & Sobol, R. W. Mechanisms of MTH1 inhibition-induced DNA strand breaks: the slippery slope from the oxidized nucleotide pool to genotoxic damage. *DNA Repair (Amst)* **77**, 18 (2019).

153. Wang, Z., Jiao, X., Carr-Schmid, A. & Kiledjian, M. The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proc Natl Acad Sci U S A* **99**, 12663–12668 (2002).
154. Van Dijk, E. *et al.* Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *EMBO J* **21**, 6915–6924 (2002).
155. Lykke-Andersen, J. Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. *Mol Cell Biol* **22**, 8114–8121 (2002).
156. Steiger, M., Carr-Schmid, A., Schwartz, D. C., Kiledjian, M. & Parker, R. Analysis of recombinant yeast decapping enzyme. *RNA* **9**, 231 (2003).
157. Song, M. G., Li, Y. & Kiledjian, M. Multiple mRNA Decapping Enzymes in Mammalian Cells. *Mol Cell* **40**, 423–432 (2010).
158. Grudzien-Nogalska, E., Jiao, X., Song, M. G., Hart, R. P. & Kiledjian, M. Nudt3 is an mRNA decapping enzyme that modulates cell migration. *RNA* **22**, 773 (2016).
159. Lu, G. *et al.* hNUDT16: a universal decapping enzyme for small nucleolar RNA and cytoplasmic mRNA. *Protein Cell* **2**, 64 (2011).
160. SGrudzien-Nogalska, E. & Kiledjian, M. New Insights into Decapping Enzymes and Selective mRNA Decay. *Wiley Interdiscip Rev RNA* **8**, (2017).
161. Sharma, S. *et al.* Mammalian Nudix proteins cleave nucleotide metabolite caps on RNAs. *Nucleic Acids Res* **48**, 6788–6798 (2020).
162. Kerr, E. W., Shumar, S. A. & Leonardi, R. Nudt8 is a novel CoA diphosphohydrolase that resides in the mitochondria. *FEBS Lett* **593**, 1133–1143 (2019).
163. Shumar, S. A. *et al.* Nudt19 is a renal CoA diphosphohydrolase with biochemical and regulatory properties that are distinct from the hepatic Nudt7 isoform. *J Biol Chem* **293**, 4134–4148 (2018).
164. Mugridge, J. S. & Gross, J. D. Judge, jury, and executioner: DXO functions as a decapping enzyme and exoribonuclease in pre-mRNA quality control. *Mol Cell* **50**, 2–4 (2013).
165. Picard-Jean, F. *et al.* 2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO. *PLoS One* **13**, (2018).
166. Doamekpor, S. K., Sharma, S., Kiledjian, M. & Tong, L. Recent insights into noncanonical 5' capping and decapping of RNA. *J Biol Chem* **298**, 102171 (2022).
167. Han, G. W. *et al.* Crystal structure of an Apo mRNA decapping enzyme (DcpS) from Mouse at 1.83 Å resolution. *Proteins* **60**, 797–802 (2005).

168. Liu, H., Rodgers, N. D., Jiao, X. & Kiledjian, M. The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. *EMBO J* **21**, 4699 (2002).
169. Van Dijk, E., Hir, H. Le & Séraphin, B. DcpS can act in the 5'-3' mRNA decay pathway in addition to the 3'-5' pathway. *Proc Natl Acad Sci U S A* **100**, 12081–12086 (2003).
170. Pecherzewska, R. & Nawrot, B. FHIT--tumor suppressor protein involved in induction of apoptosis and cell cycle regulation. *Postepy Biochem* **55**, 66–75 (2009).
171. 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).
172. Campiglio, M. *et al.* Diadenosines as FHIT-ness instructors. *J Cell Physiol* **208**, 274–281 (2006).
173. Doamekpor, S. K., Sharma, S., Kiledjian, M. & Tong, L. Recent insights into noncanonical 5' capping and decapping of RNA. *Journal of Biological Chemistry* **298**, 102171 (2022).
174. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207–234 (2005).
175. Yu, J. *et al.* Second messenger Ap4A polymerizes target protein HINT1 to transduce signals in FcεRI-activated mast cells. *Nat Commun* **10**, (2019).
176. Ofir-Birin, Y. *et al.* Structural switch of lysyl-tRNA synthetase between translation and transcription. *Mol Cell* **49**, 30–42 (2012).
177. 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).
178. Yuan, Y. *et al.* Chemoproteomic discovery of a human RNA ligase. *Nature Communications* 2023 14:1 **14**, 1–12 (2023).
179. Kim, I. S., Sohn, H. Y. & Jin, I. Adaptive stress response to menadione-induced oxidative stress in *Saccharomyces cerevisiae* KNU5377. *Journal of Microbiology* **49**, 816–823 (2011).
180. Senejani, A. G., Magrino, J. M., Marston, A., Gregoire, M. & Dinh, K. D. Menadione Induces DNA Damage and Superoxide Radical Level In HEK293 Cells. *Cell Biology* 2019, Volume 7, Page 14 **7**, 14–22 (2019).

181. Yoshimura, K. & Shigeoka, S. Versatile physiological functions of the Nudix hydrolase family in Arabidopsis. *Biosci Biotechnol Biochem* **79**, 354–366 (2015).
182. Ogawa, T. *et al.* Molecular Characterization of Organelle-Type Nudix Hydrolases in Arabidopsis. *Plant Physiol* **148**, 1412 (2008).
183. Laudénbach, B. T. *et al.* NUDT2 initiates viral RNA degradation by removal of 5'-phosphates. *Nature Communications* 2021 12:1 **12**, 1–13 (2021).
184. Magnone, M. *et al.* Adenylic Dinucleotides Produced by CD38 Are Negative Endogenous Modulators of Platelet Aggregation *. *J Biol Chem* **283**, 24460–24468 (2008).
185. Luo, J. *et al.* Identification and characterization of diadenosine 5',5''-P₁,P₂ -diphosphate and diadenosine 5',5''-P₁,P₃-triphosphate in human myocardial tissue. *FASEB J* **13**, 695–705 (1999).
186. Jankowski, J. *et al.* Dinucleotides as Growth-promoting Extracellular Mediators: PRESENCE OF DINUCLEOSIDE DIPHOSPHATES Ap₂A, Ap₂G, AND Gp₂G IN RELEASABLE GRANULES OF PLATELETS. *Journal of Biological Chemistry* **276**, 8904–8909 (2001).
187. Raynes, R. *et al.* The SIRT1 Modulators AROS and DBC1 Regulate HSF1 Activity and the Heat Shock Response. *PLoS One* **8**, 54364 (2013).
188. Fulco, M. *et al.* Glucose restriction inhibits skeletal myoblast differentiation by activating SIRT1 through AMPK-mediated regulation of Nampt. *Dev Cell* **14**, 661–673 (2008).
189. Hafner, M. *et al.* Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* **44**, 3–12 (2008).
190. Yin, S., Ho, C. K. & Shuman, S. Structure-Function Analysis of T4 RNA Ligase 2*. *J Biol Chem* **278**, 17601–17608 (2003).
191. Olejnik, K., Murcha, M. W., Whelan, J. & Kraszewska, E. Cloning and characterization of AtNUDT13, a novel mitochondrial Arabidopsis thaliana Nudix hydrolase specific for long-chain diadenosine polyphosphates. *FEBS J* **274**, 4877–4885 (2007).
192. Ogawa, T. *et al.* Modulation of NADH Levels by Arabidopsis Nudix Hydrolases, AtNUDX6 and 7, and the Respective Proteins Themselves Play Distinct Roles in the Regulation of Various Cellular Responses Involved in Biotic/Abiotic Stresses. *Plant Cell Physiol* **57**, 1295–1308 (2016).

193. Pietrowska-Borek, M., Nuc, K., Zielezińska, M. & Guranowski, A. Diadenosine polyphosphates (Ap3A and Ap4A) behave as alarmones triggering the synthesis of enzymes of the phenylpropanoid pathway in *Arabidopsis thaliana*. *FEBS Open Bio* **1**, 1–6 (2011).
194. Krüger, L. *et al.* Chemical proteomic profiling reveals protein interactors of the alarmones diadenosine triphosphate and tetraphosphate. *Nature Communications* **2021 12:1** **12**, 1–13 (2021).
195. Marriott, A. S. *et al.* NUDT2 disruption elevates diadenosine tetraphosphate (AP4A) and down-regulates immune response and cancer promotion genes. *PLoS One* **11**, (2016).

8 Appendices

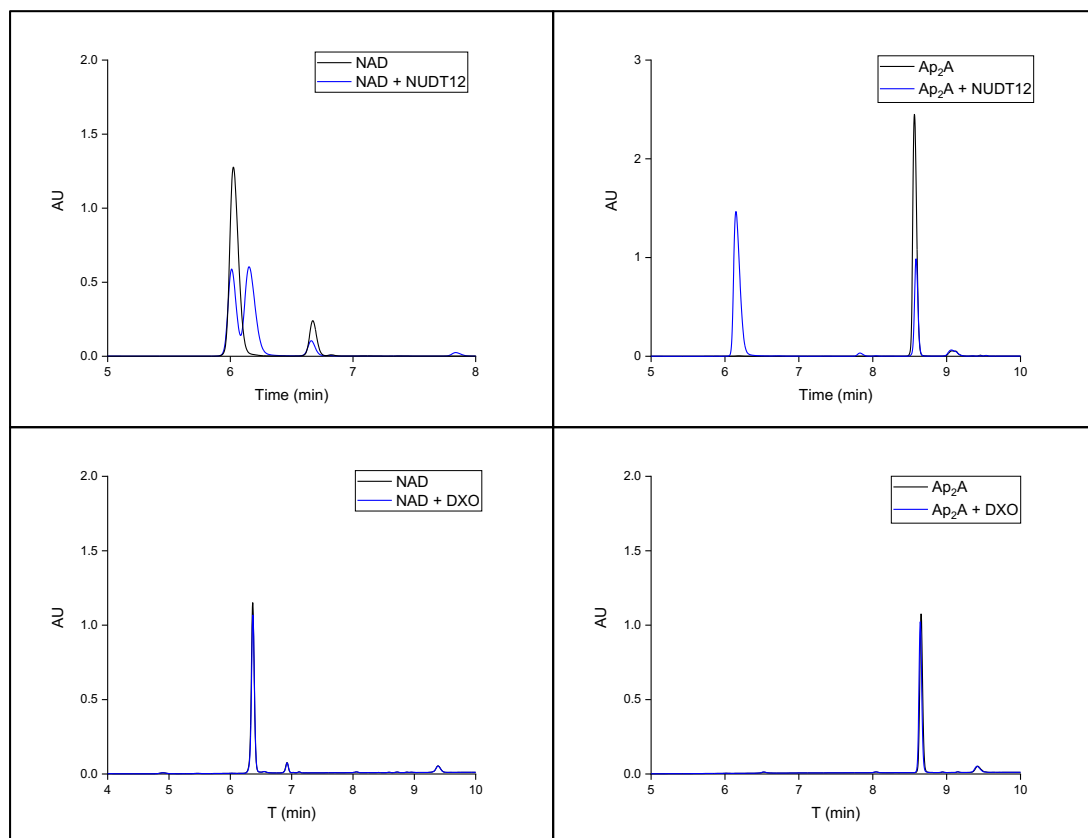
8.1 Abbreviations

A	Adenosine
ACDQ	6-amino-7-chloro-5,8-dioxoquinoline
ADAR	Adenosine diphosphate ribose
ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
Ap ₂ A	Diadenosine Diphosphate
Ap ₄ A	Adenosine tetraphosphate
ApaH	bis(5'-nucleosyl)-tetraphosphatase
ATP	Adenosine triphosphate
C	Cytosine
CHS	Béguez-Chédiak–Higashi syndrome
CoA	Coenzyme A
CTP	Cytidine triphosphate
DTT	Dithiothreitol
E. Coli	Escherichia coli
FHIT	fragile histidine triad protein
G	Guanosine
GMP	Guanosine monophosphate
H ₂ O ₂	Hydrogen Peroxide
HAD	Haloacid dehalogenase
HINT	Histidine triad nucleotide binding protein
HPLC	Hight Performance Liquid Chromatography
IVT	In vitro transcription

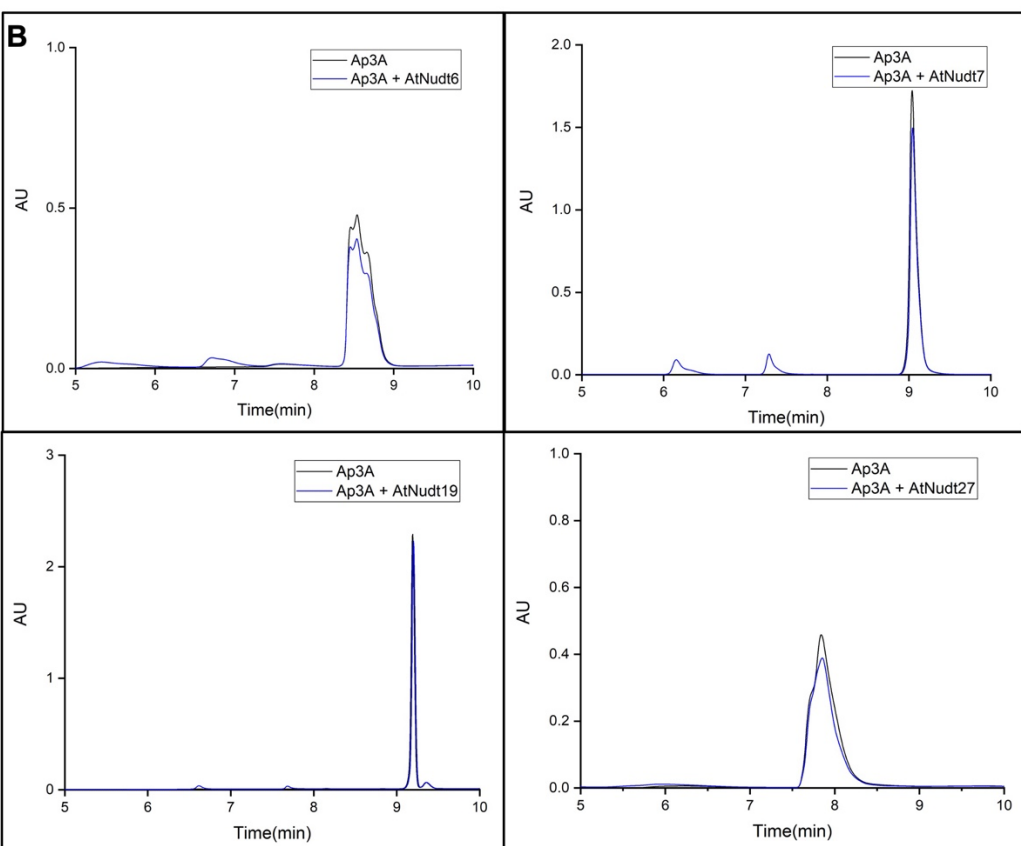
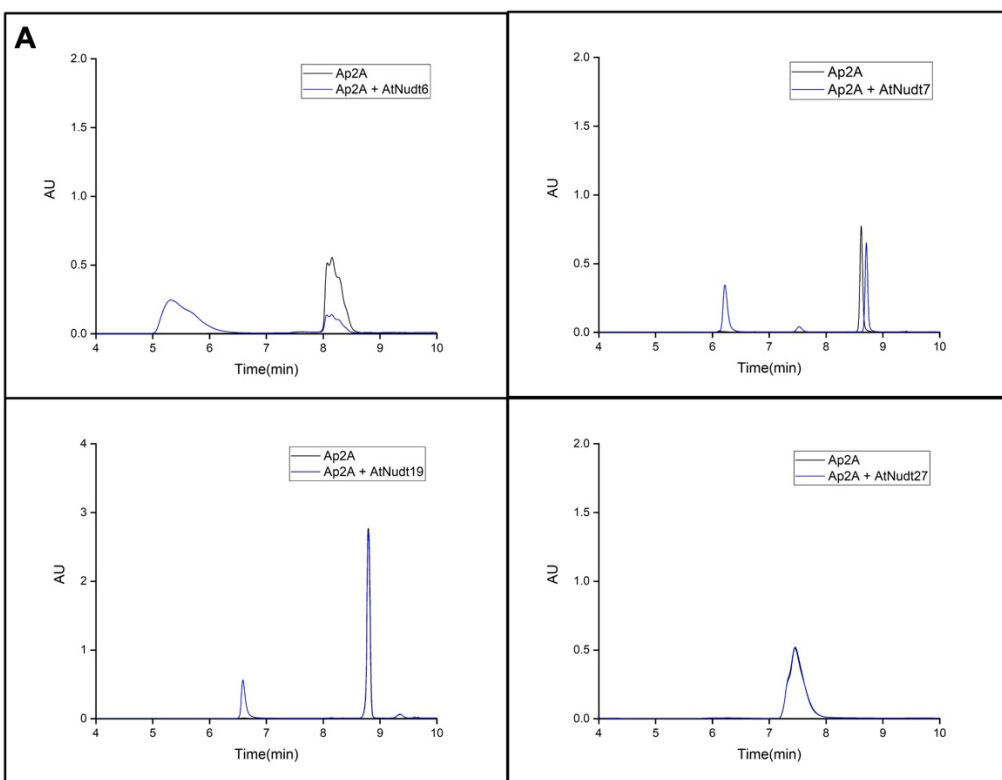
KARS	Human Lysyl-tRNA synthase
LB	Luria Bertani media
LC-MS	Liquid Chromatography-Mass Spectrometry
m ⁷ G	7-methylguanosine
MDE	mRNA decapping enzyme
MHC	Major histocompatibility complex
MITF	Microphthalmia transcription factor
mRNA	Messenger RNA
NAD	Nicotinamide adenine dinucleotide
NCIN	Non canonical initiator nucleotide
NGS	Next generation sequencing
NMN	Nicotinamide mononucleotide
Np _n Ns	Dinucleoside polyphosphates
NTP	Nucleoside triphosphate
NudiX	Nucleoside diphosphates linked to X
NuP1	Nuclease P1
PAGE	Polyacrylamide gel electrophoresis
PARPs	Poly (ADP-ribose) polymerases
PCR	Polymerase chain reaction
RIG-I	Retinoic acid inducible gene I
RNA	Ribonucleic acid
RNAP	RNA Polymerase
ROS	Reactive oxygen species
RppH	RNA 5'-pyrophosphohydrolase
RT-qPCR	Reverse transcription-quantitative polymerase chain reaction
SAP	Shrimp alkaline phosphatase

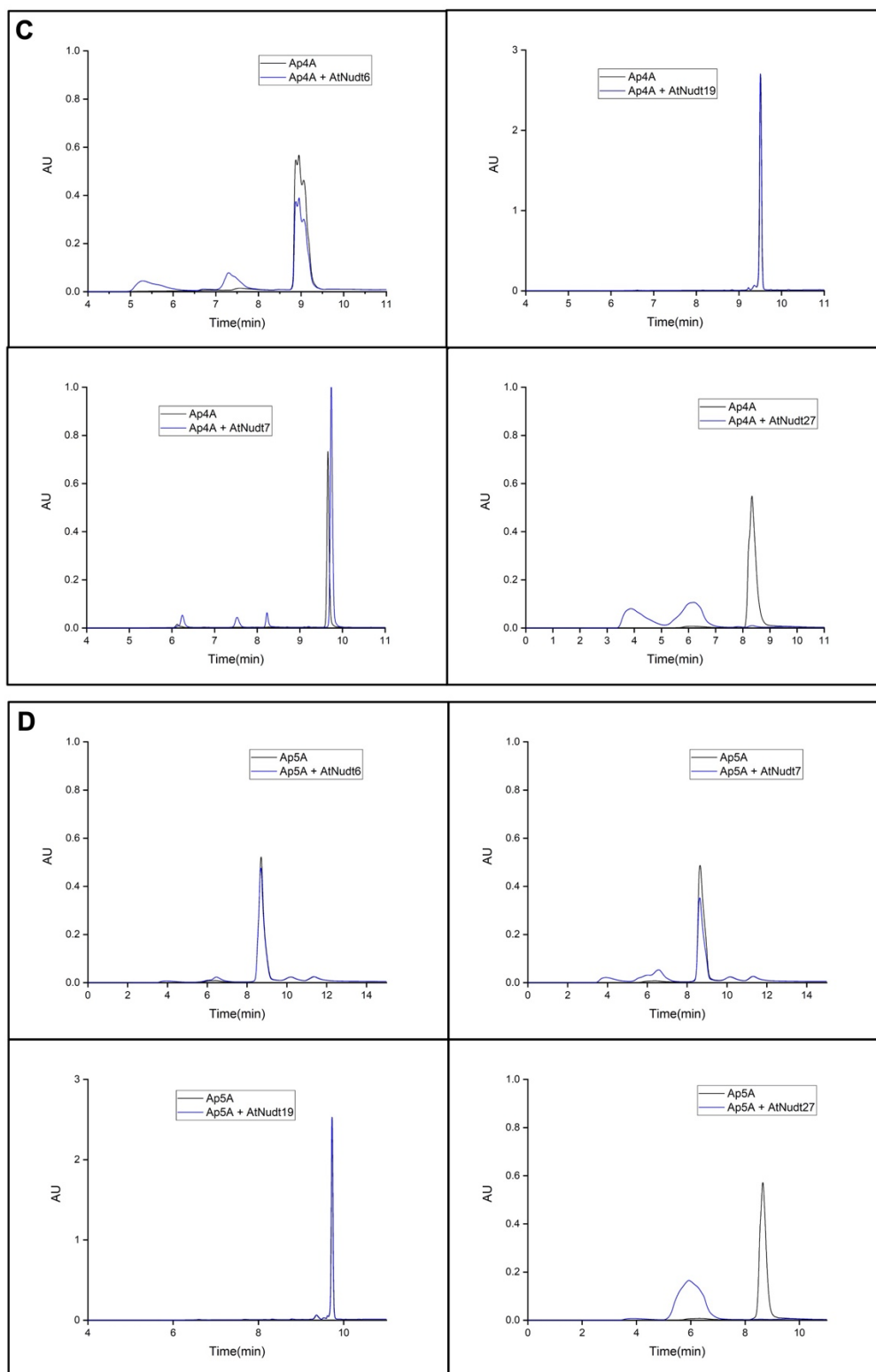
SARS	Severe acute respiratory syndrome
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
tBHQ	tert-Butylhydroquinone
TOF	Time of flight
tRNA	Transfer RNA
UTP	Uridine triphosphate
VEEV	Venezuelan equine encephalitis

8.2 Supplementary figures

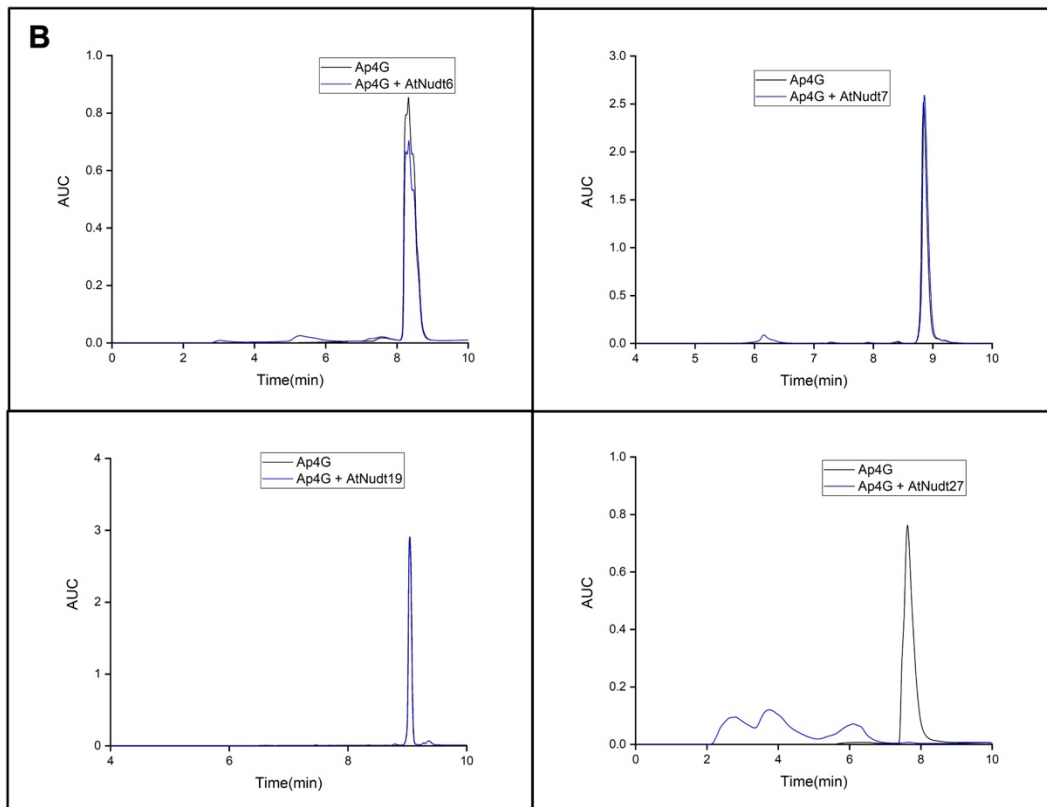
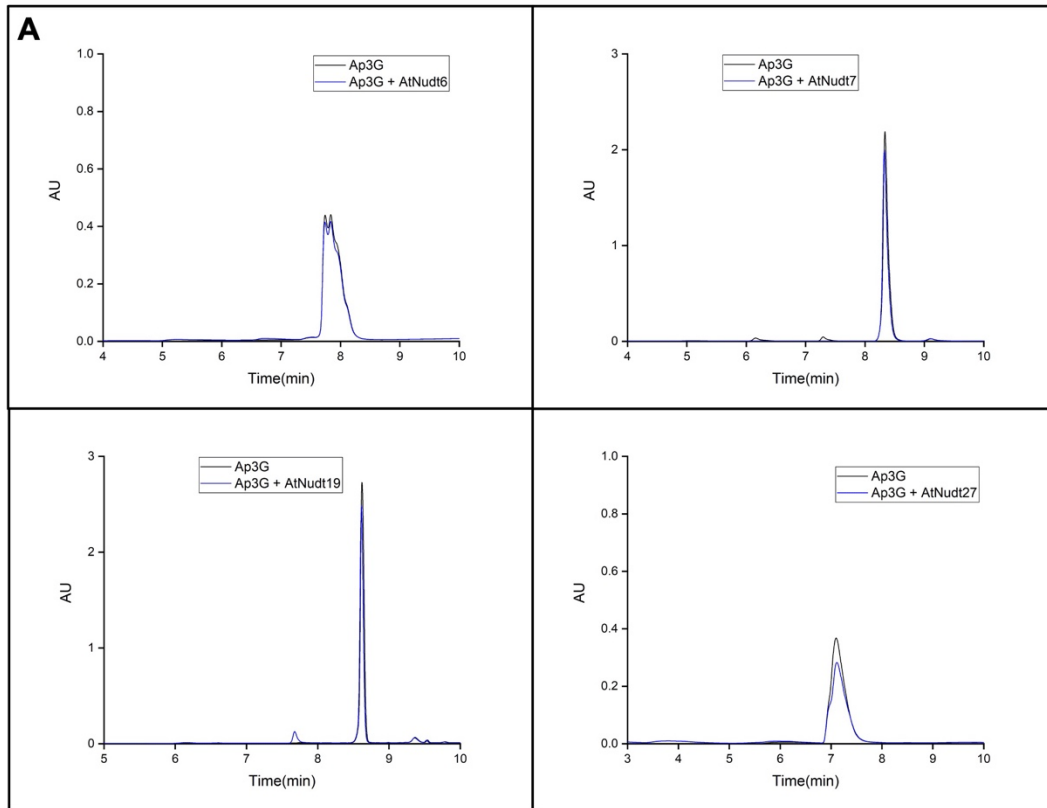


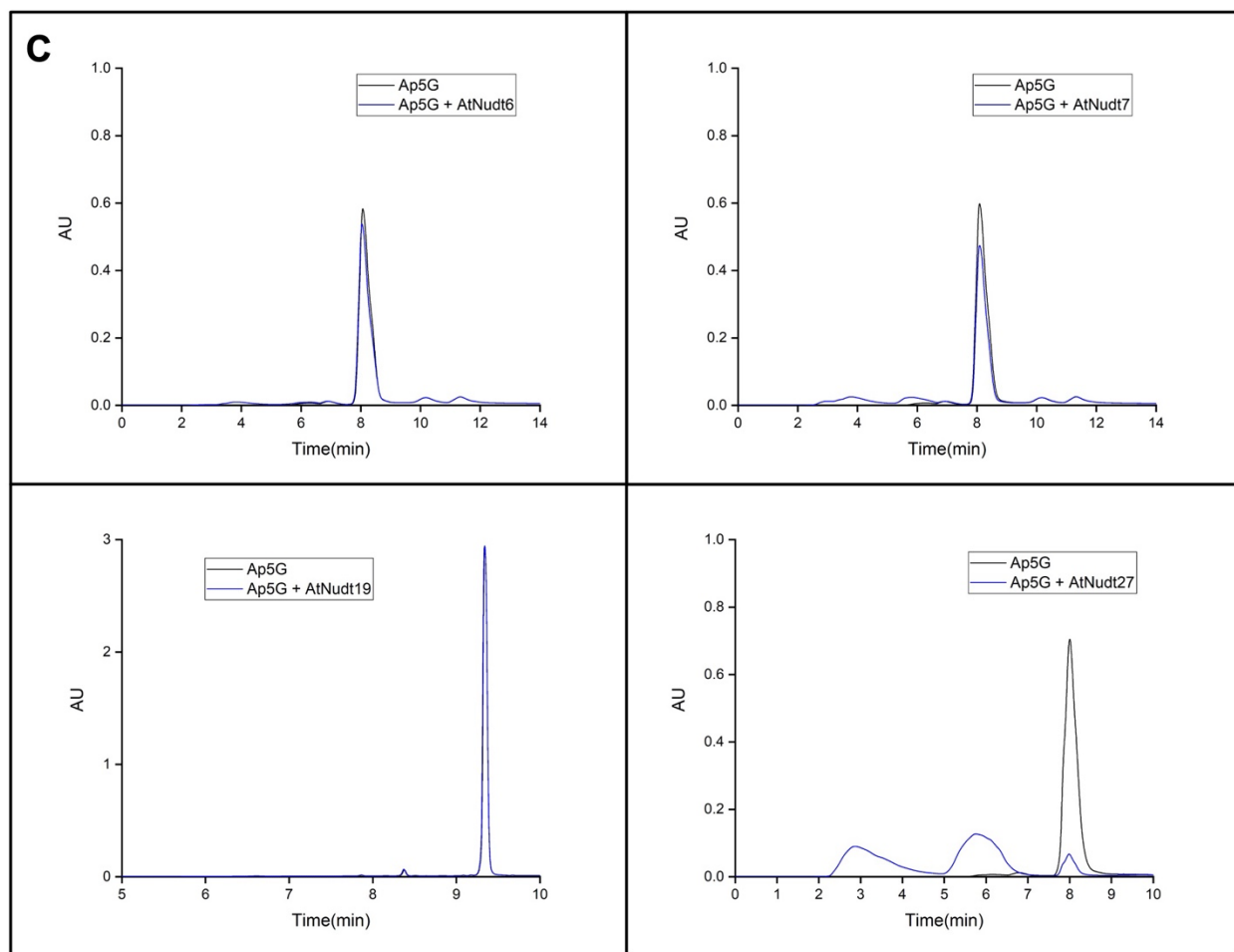
Supplementary Figure 1. HPLC chromatograms of free small molecules NAD and Ap₂A (200 μ M) before and after incubation with NUDT12 (1 μ M) and DXO (200nM)



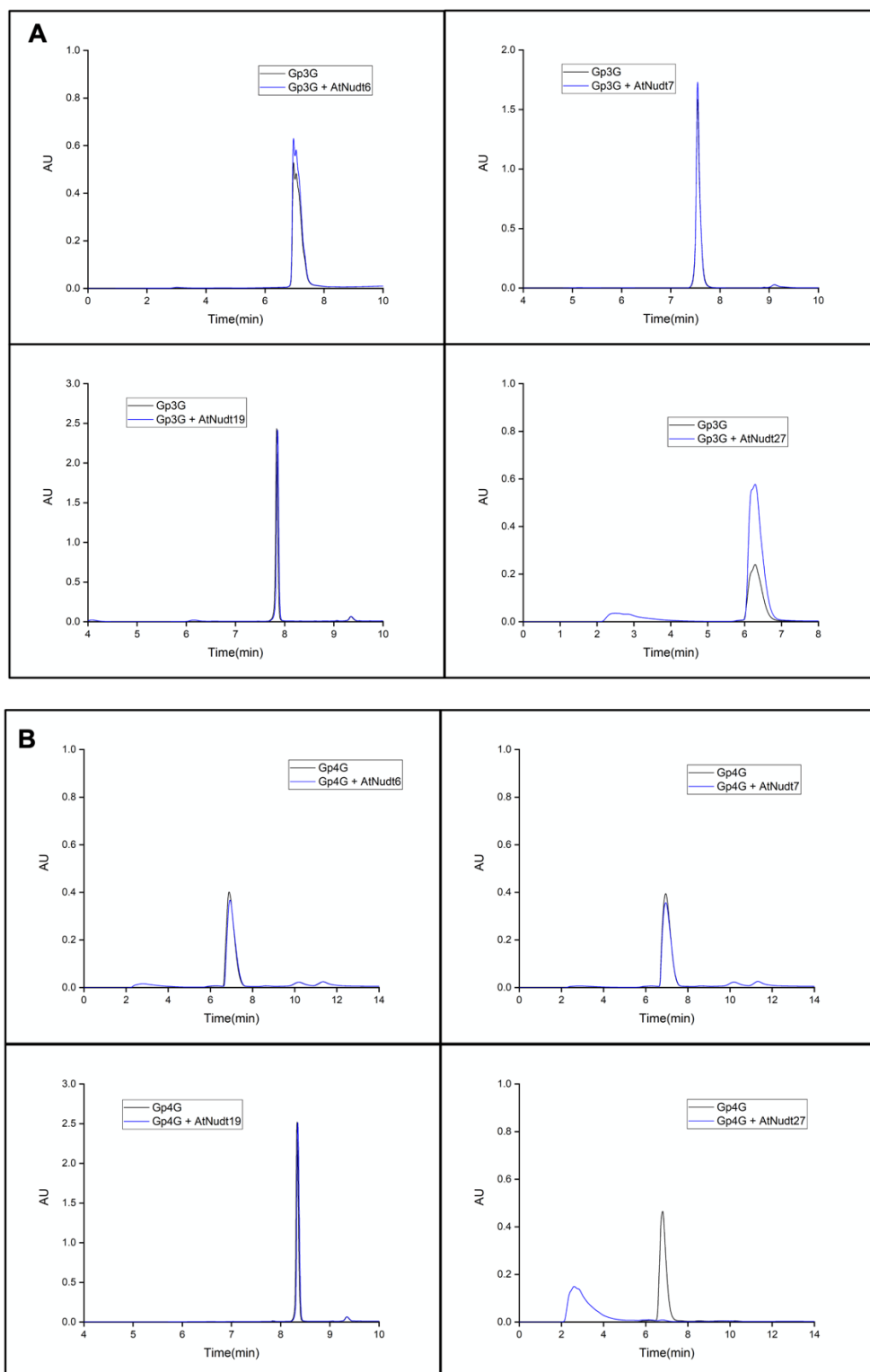


Supplementary Figure 2. HPLC chromatograms of free small molecules before and after incubation with AtNUDT6,7,19 and 27. A) Ap₂A, B) Ap₃A, C) Ap₄A, D) Ap₅A.

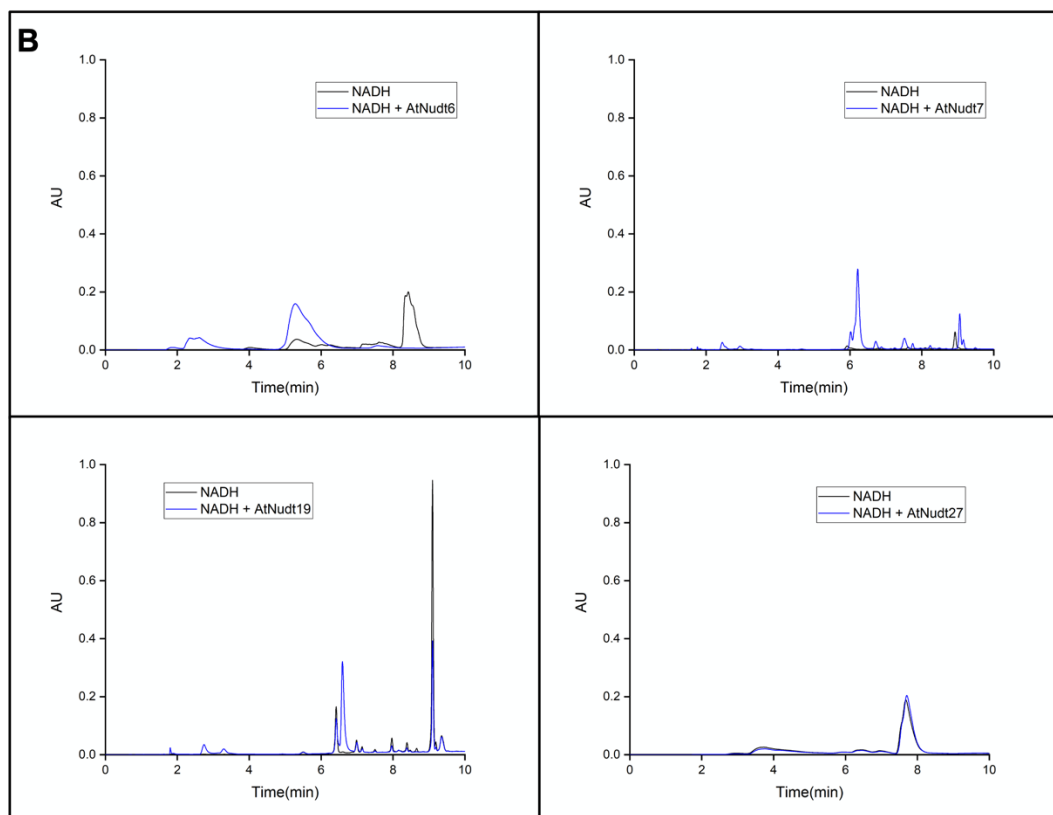
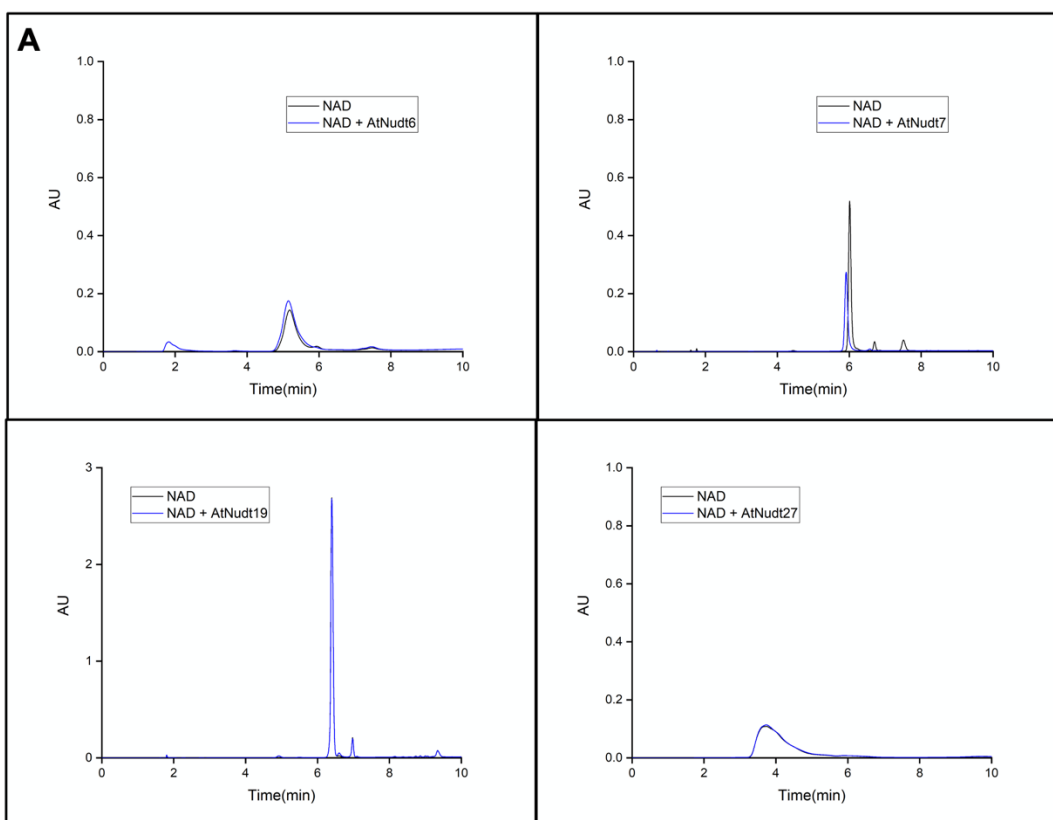


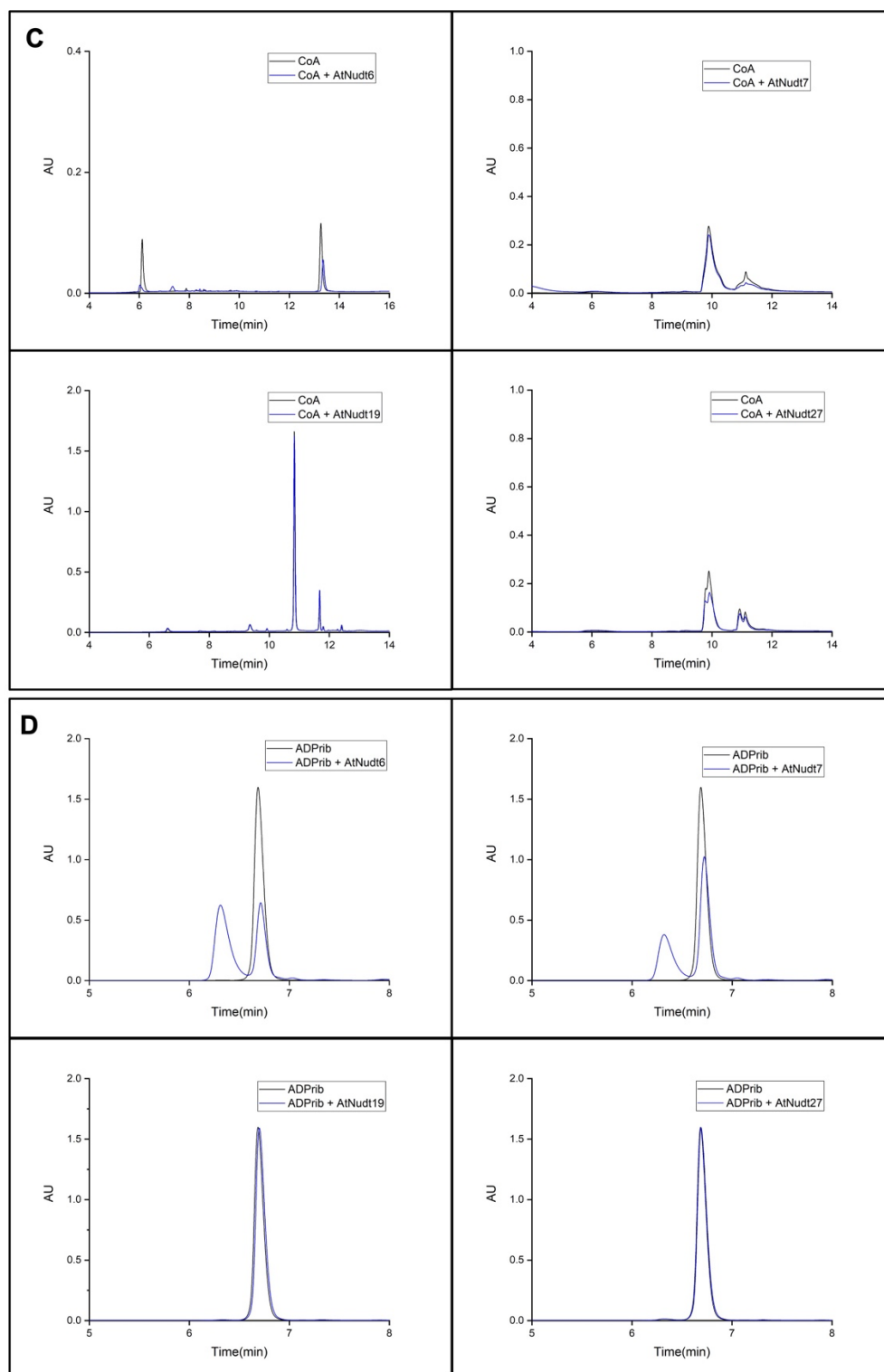


Supplementary Figure 3. HPLC chromatograms of free small molecules before and after incubation with AtNUDT6,7,19 and 27. A) Ap_3G , B) Ap_4G , C) Ap_5G .



Supplementary Figure 4. HPLC chromatograms of free small molecules before and after incubation with AtNUDT6,7,19 and 27. A) Gp₃G, B) Gp₄G.





Supplementary Figure 5. HPLC chromatograms of free small molecules before and after incubation with AtNUDT6,7,19 and 27. A) NAD, B) NADH, C) 3' dephosphoCoA, D) ADP-Ribose;

8.3 Publications

- 1) Arabidopsis thaliana Nudixes have RNA-decapping activity
Mititelu, M. B., Hudeček, O., Gozdek, A., Benoni, R., Nešuta, O., Krasnodębski, S., ... & Cahová, H. (2023). RSC Chemical Biology, 4(3), 223-228.
- 2) Diadenosine Tetraphosphate (Ap₄A) Serves as a 5' RNA Cap in Mammalian Cells
František Potužník, J., Nešuta, O., Škríba, A., Voleníková, B., **Mititelu, M. B.**, Mancini, F., ... & Cahová, H. (2023). Angewandte Chemie, e202314951.
- 3) HIV-1 Infection Reduces NAD Capping of Host Cell snRNA and snoRNA
Benoni B, Potužník JF, Škríba A, Benoni R, Trylcova J, Tulpa M, Spustová K, Grab K, **Mititelu MB**, Pačes J, Weber J, Stanek D, Kowalska J, Bednarova L, Keckesova Z, Vopalensky P, Gahurova L, Cahova H.(2024). ACS Chem Biol.
- 4) Dinucleoside diphosphate Ap₂A as novel 5' RNA cap in mammalian cells
Mititelu, M.B., Manuscript in preparation

PAPER



Cite this: *RSC Chem. Biol.*, 2023, 4, 223

***Arabidopsis thaliana* NudIXes have RNA-decapping activity†**

Maria-Bianca Mititelu,^{ab} Oldřich Hudeček,^a Agnieszka Gozdek,^c Roberto Benoni,^a Ondřej Nešuta,^{ab} Szymon Krasnodębski,^c Joanna Kufel^c and Hana Cahová^{ab*}

Recent discoveries of various noncanonical RNA caps, such as dinucleoside polyphosphates (Np_nN), coenzyme A (CoA), and nicotinamide adenine dinucleotide (NAD) in all domains of life have led to a revision of views on RNA cap function and metabolism. Enzymes from the NudIX family capable of hydrolyzing a polyphosphate backbone attached to a nucleoside are the strongest candidates for degradation of noncanonically capped RNA. The model plant organism *Arabidopsis thaliana* encodes as many as 28 NudIX enzymes. For most of them, only *in vitro* substrates in the form of small molecules are known. In our study, we focused on four *A. thaliana* NudIX enzymes (AtNUDT6, AtNUDT7, AtNUDT19 and AtNUDT27), and we studied whether these enzymes can cleave RNA capped with Np_nNs (Ap₂₋₅A, Gp₃₋₄G, Ap₃₋₅G, m⁷Gp₃G, m⁷Gp₃A), CoA, ADP-ribose, or NAD(H). While AtNUDT19 preferred NADH-RNA over other types of capped RNA, AtNUDT6 and AtNUDT7 preferentially cleaved Ap₄A-RNA. The most powerful decapping enzyme was AtNUDT27, which cleaved almost all types of capped RNA at a tenfold lower concentration than the other enzymes. We also compared cleavage efficiency of each enzyme on free small molecules with RNA capped with corresponding molecules. We found that AtNUDT6 prefers free Ap₄A, while AtNUDT7 preferentially cleaved Ap₄A-RNA. These findings show that NudIX enzymes may act as RNA-decapping enzymes in *A. thaliana* and that other noncanonical RNA caps such as Ap₄A and NADH should be searched for in plant RNA.

Received 5th October 2022,
Accepted 2nd January 2023

DOI: 10.1039/d2cb00213b

rsc.li/rsc-chembio

Introduction

Currently, more than 170 RNA modifications are known.¹ Among the least explored RNA modifications are 5' RNA caps. A typical example of a noncanonical RNA cap is nicotinamide adenine dinucleotide (NAD), which has been detected in all domains of life.²⁻⁵ Other caps, such as NADH (a reduced form of NAD),⁶ coenzyme A (CoA),⁷ and flavin adenine dinucleotide (FAD),⁸ have been detected only in some organisms. Recently, we have discovered an entirely new class of RNA caps in bacteria with the structure of dinucleoside polyphosphates (Np_nNs).⁹ The presence of free Np_nNs molecules in cells of various organisms has been known more than 50 years.¹⁰⁻¹² Their intracellular concentration increases under stress conditions,¹³ thus they are often called alarmones. For instance, Ap₃A and Ap₄A are known to act as alarmones in *Arabidopsis*

thaliana (*A. thaliana*),¹³ although their mechanism of action is unknown. The discovery of these molecules as RNA caps opened new perspectives on their function and metabolism. Free molecules such as Np_nNs, NAD, and CoA are usually cleaved by NudIX enzymes. Plants especially have a high number of NudIX-encoding genes, with a total of 28 in *A. thaliana*;¹⁴ by comparison, there are 24 in human cells and only 13 in *Escherichia coli*.¹⁵ Nevertheless, it is still unclear why so many NudIX enzymes are necessary for cellular metabolism and why many are redundant, *i.e.* two or more enzymes cleave the same substrates *in vitro*. Importantly, some NudIX enzymes have also been reported to have RNA-decapping activity.¹⁶⁻²⁰ Identification of the new decapping role of some NudIX enzymes will help us to reveal new potential RNA caps and explore their biological role.

In our search for new noncanonical RNA caps, we explored *A. thaliana* NudIX proteins and their substrate specificity. In particular, we examined four enzymes, cytosolic AtNUDT6, AtNUDT7, and chloroplastic AtNUDT19 and AtNUDT27, since small molecules such as NAD(H), CoA, ADP-ribose, and various Np_nNs are known to be their substrates *in vitro*.¹⁴ In addition, these proteins are the closest plant homologs of mammalian Nudt2 and Nudt12, which have been demonstrated to have RNA non-canonical decapping activity *in vitro*, namely deFADding and deCoAding, respectively.^{20,21} We therefore tested whether

^a Institute of Organic Chemistry and Biochemistry of the CAS, Flemingovo náměstí 2, Prague 6, Czechia. E-mail: cahova@uochb.cas.cz

^b Charles University, Faculty of Science, Department of Cell Biology, Viničná 7, Prague 2, Czechia

^c Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw, Pawlowskiego 5a, Warsaw 02-106, Poland

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d2cb00213b>



the aforesaid enzymes are capable of cleaving these noncanonical RNA caps directly from RNA. In addition, we compared their ability to cleave small molecules and noncanonically capped RNA. Our results indicate that plant Nudixes may fulfil the role of RNA-decapping enzymes specific for noncanonically capped RNA.

Results and discussion

First, we screened the *in vitro* activity of AtNUDT6, AtNUDT7, AtNUDT19, and AtNUDT27 on RNA capped with NAD(H), CoA, ADP-ribose, canonical m⁷Gp₃A(G), or various dinucleoside polyphosphates (Ap₂₋₅A, Gp₃₋₄G, Ap₃₋₅G). We prepared ³²P labelled RNA by *in vitro* transcription with T7 RNA polymerase in the presence of cap precursors using two DNA templates (Table S1, ESI†). Template 35A has at +1 position A and was used in experiments with Ap₂₋₅A, Gp₃₋₄A, NAD(H), CoA, ADP-ribose and m⁷Gp₃A. Template 35G has at position +1 G and was used in experiments with Gp₃₋₄G, Ap₃₋₅G, and m⁷Gp₃G. Asymmetrical Np_nNs such as Ap₃₋₅G can be used with both templates in order to obtain RNA with flanking A or G, Ap₃₋₅G-RNA, or Gp₃₋₅A-RNA, respectively.

Following *in vitro* transcription, we purified the capped RNA from uncapped by treatment with 5′ polyphosphatase and Terminator™ exonuclease. This treatment guaranteed that the amount of capped RNA in each reaction was in average 96 ± 4% (only in case of m⁷Gp₃A(G)-RNA the amount of capped RNA was 58 ± 3%). Afterwards, we treated capped RNA with Nudix enzymes (Fig. 1A) and followed the cleavage with PAGE analysis (Fig. 1B and Fig. S1 and S2, ESI†). We observed the highest enzymatic activity on the capped RNA at 500 nM concentration for AtNUDT6, 7, and 19 and 50 nM concentration for AtNUDT27

(Fig. S1 and S2, ESI†). AtNUDT6 and AtNUDT7 had a similar substrate specificity, with the Ap₄A-RNA being the best substrate for both of them (Fig. 1B).

In addition, AtNUDT6 also partially cleaved (less than 50%) Ap₃A-, Ap₅A- and Ap₄G-RNA. AtNUDT7 cleaved Ap_{2,3,5}A-RNA, Ap_{4,5}G-RNA, and (to some extent) Gp₄A-RNA (Table 1 and Fig. S3, ESI†). The main product of the Ap₄A-RNA-decapping reaction by AtNUDT6 and AtNUDT7 is a triphosphate (ppp-)RNA. AtNUDT27 cleaved Ap₄A-RNA into a mixture of ppp-RNA and monophosphate (p-)RNA (Fig. S4, ESI†). While AtNUDT6 and AtNUDT7 had quite wide substrate specificities, AtNUDT19 efficiently cleaved only the NADH-RNA, with the NAD-RNA being partially cleaved (Fig. 1B and Table 1 and Fig. S3, ESI†). Surprisingly, at a tenfold lower concentration (50 nM), AtNUDT27 still partially hydrolyzed the majority of capped RNA with a relatively high cleavage efficiency (60–90% after 1 h), except for Ap₃A-RNA (less than 30%) (Table 1 and Fig. S3, ESI†).

While some of the small molecules are known to be *in vitro* substrates of these Nudix enzymes, the entire set of recently discovered noncanonical caps has not yet been tested as free compounds. Therefore, we screened all the noncanonical RNA caps in their free form as substrates for the four Nudix proteins (Table 2 and Fig. S5–S8, ESI†). AtNUDT6 and AtNUDT7 had a similar substrate specificity on both the small molecules and

Table 1 Cleavage of RNA capped with different small molecules *in vitro* by AtNUDT6 (500 nM), AtNUDT7 (500 nM), AtNUDT19 (500 nM), and AtNUDT27 (50 nM). Values are in % of cleaved capped RNA and represent a mean of three independent replicates ± standard deviations

AtNUDT	6 (500 nM)	7 (500 nM)	19 (500 nM)	27 (50 nM)
5′-cap-RNA	(%)	(%)	(%)	(%)
Ap ₂ A-	x	53 ± 10	x	73 ± 9
Ap ₃ A-	48 ± 7	55 ± 7	x	26 ± 8
Ap ₄ A-	80 ± 11	73 ± 5	x	87 ± 4
Ap ₅ A-	15 ± 8	65 ± 17	x	90 ± 5
Ap ₃ G-	x	x	x	74 ± 11
Ap ₄ G-	35 ± 10	70 ± 20	x	85 ± 5
Ap ₅ G-	x	61 ± 12	x	85 ± 5
Gp ₃ G-	x	x	x	83 ± 5
Gp ₄ G-	x	x	x	83 ± 1
NAD-	x	x	20 ± 4	60 ± 7
NADH-	x	x	57 ± 5	65 ± 10
CoA-	x	x	x	82 ± 5
Gp ₃ A-	x	x	x	80 ± 5
Gp ₄ A-	x	30 ± 6	x	84 ± 5
Gp ₅ A-	x	x	x	90 ± 3
m Gp ₃ A-	x	x	x	80 ± 7
m Gp ₃ G-	x	x	x	70 ± 6
ADP-ribose	x	x	x	70 ± 2

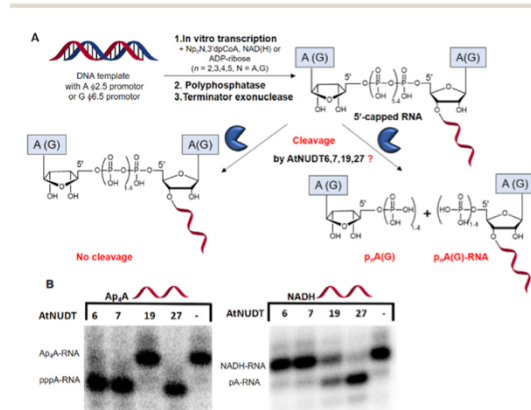
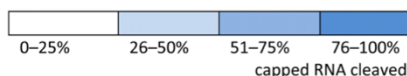
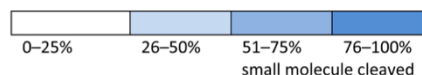


Fig. 1 Screening of the RNA-decapping activity of AtNUDT6, 7, 19, and 27. (A) Scheme of the experimental setup. ³²P labelled RNA was prepared by *in vitro* transcription with T7 RNA polymerase in the presence of the small molecules Np_nNs, NAD(H), CoA, or ADP-ribose. The side product, uncapped triphosphate RNA, was degraded by treatment with 5′ polyphosphatase and Terminator™ exonuclease. (B) A representative example of PAGE analysis of the RNA-decapping activity of AtNUDT6, 7, 19, and 27 on 5′-capped Ap₄A-RNA and NADH-RNA.

Table 2 Cleavage of different small molecules (400 μM) *in vitro* by AtNUDT6 (500 nM), AtNUDT7 (500 nM), AtNUDT19 (500 nM), and AtNUDT27 (500 nM). Values are in % of cleaved substrate and represent a mean of three independent replicates $\pm$ standard deviations

AtNUDT small molecule	6 (500 nM) (%)	7 (500 nM) (%)	19 (500 nM) (%)	27 (500 nM) (%)
Ap ₂ A	80 $\pm$ 9	64 $\pm$ 7	16 $\pm$ 3	x
Ap ₃ A	18 $\pm$ 5	18 $\pm$ 1	x	x
Ap ₄ A	31 $\pm$ 1	40 $\pm$ 3	x	99 $\pm$ 0.5
Ap ₅ A	x	31 $\pm$ 1	x	99 $\pm$ 0.5
Ap ₃ G	x	x	x	x
Ap ₄ G	20 $\pm$ 2	x	x	96 $\pm$ 5
Ap ₅ G	x	x	x	89 $\pm$ 1
Gp ₃ G	x	x	x	x
Gp ₄ G	x	x	x	98 $\pm$ 1.5
NAD	x	52 $\pm$ 1	x	x
NADH	99 $\pm$ 1	83 $\pm$ 4	56 $\pm$ 7	x
3'-dpCoA	53 $\pm$ 4	18 $\pm$ 6	x	x
ADP-ribose	55 $\pm$ 10	45 $\pm$ 3	x	x



the capped RNA. Surprisingly, their best substrate was not free Ap₄A, as expected based on the results from assays on capped RNA (Table 2 and Fig. S5–S8, ESI†) but rather Ap₂A and NADH, both of which have a diphosphate bridge. As in the case of capped RNA, AtNUDT19 hydrolyzed free NADH with the highest efficiency (Table 2 and Fig. S5–S8, ESI†). AtNUDT27 had a very high activity at 500 nM concentration and quantitatively cleaved all the free tetraphosphates (Ap₄A, Ap₄G, and Gp₄G) and Ap₅A (Table 2 and Fig. S5–S8, ESI†). To better observe the selectivity of the enzyme towards small free molecules, we decreased its concentration to 50 nM (Fig. S9, ESI†). Under these reaction conditions, AtNUDT27 preferred dinucleoside tetraphosphates (*i.e.* Ap₄A, Ap₄G, and Gp₄G) over other non-canonical RNA caps.

In addition, we compared Michaelis–Menten kinetic parameters for AtNUDT6, 7, and 27 and Ap₄A as a substrate (Fig. 2A). These parameters were already known for AtNUDT19 and NADH as a substrate.²² Although K_M values were comparable for all three enzymes, V_{max} and K_{cat} of AtNUDT27 were both significantly higher ($P < 0.05$) compared to AtNUDT6 and 7 (Fig. S11a and b, ESI†). This finding may explain observed behaviour of AtNUDT27 to be more efficient in cleaving free Ap₄A than AtNUDT6 and 7 (Table 2).

Next, we wanted to determine whether Nudix enzymes prefer free small molecules or capped RNA as substrates especially since this could reflect the situation *in vivo*. Due to the need to use a very large amount of RNA substrate for the entire series of tests, we compared only one set of substrates, free Ap₄A and Ap₄A-RNA. We observed the cleavage of both substrates at a concentration of 2.5 μM and 15- and 60 minute time points.

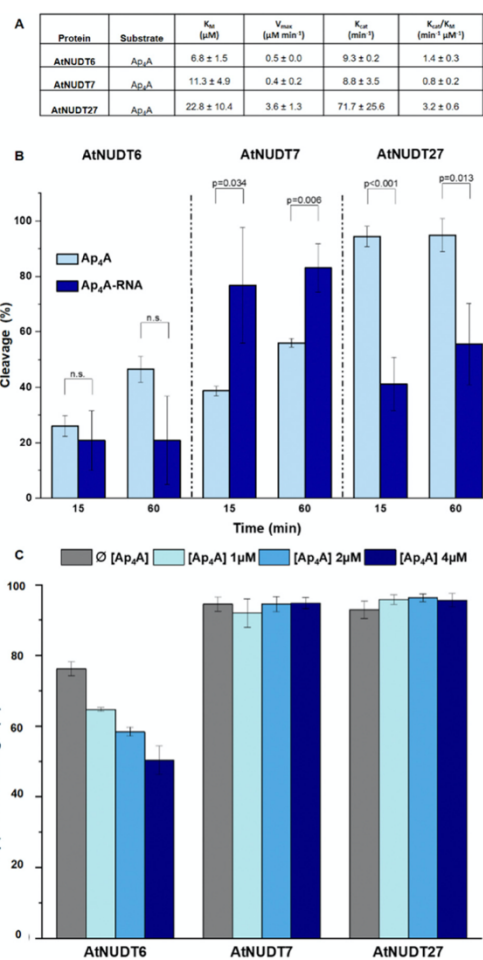


Fig. 2 (A) Kinetic parameters of AtNUDT6, AtNUDT7 and AtNUDT27 cleaving Ap₄A expressed as mean $\pm$ standard deviation. (B) Cleavage efficiency of free Ap₄A (2.5 μM) and Ap₄A-RNA (2.5 μM) by AtNUDT6, AtNUDT7, and AtNUDT27 (50 nM) at 15- and 60 min time points. (C) Inhibition study of Ap₄A-RNA (1 μM) cleavage by AtNUDT6 (500 nM), AtNUDT7 (500 nM) and AtNUDT27 (50 nM) with addition of 1, 2 or 4 μM Ap₄A.

The cleavage of Ap₄A was analysed by HPLC, while decapping of Ap₄A-RNA was measured by quantification of corresponding bands of capped RNAs resolved on PAGE (Fig. 2B and Fig. S10, ESI†). Comparison of the substrates showed that AtNUDT27 preferred the small molecule (Ap₄A) and cleaved it more efficiently than Ap₄A-RNA under the chosen conditions. AtNUDT6 cleaved comparably well both substrates – Ap₄A and Ap₄A-RNA. AtNUDT7 had higher activity on Ap₄A-RNA than on free Ap₄A. This finding may explain the redundancy of Nudix enzymes observed *in vivo*, *i.e.* some enzymes are used by the cell to hydrolyze free small molecules, while others cleave non-canonically capped RNA.

In order to understand whether free dinucleoside polyphosphates can inhibit the cleavage of capped RNA, we measured cleavage of Ap₄A-RNA (1 μ M) by AtNUDT6, 7 and 27 in presence of free Ap₄A (1, 2 and 4 μ M). We observed that only AtNUDT6 activity was affected in dose-dependent manner. The activity of other two enzymes AtNUDT7 or AtNUDT27 was not affected by addition of free Ap₄A (Fig. 2C and Fig. S12, ESI†). The difference between AtNUDT6 and AtNUDT7, which have comparable kinetic parameters, could be explained by AtNUDT7 preference for Ap₄A-RNA instead of free Ap₄A as we demonstrated in Fig. 2B. On the other hand, AtNUDT27 showed almost 8 times higher K_{cat} for Ap₄A than AtNUDT6 and 7 suggesting that added free Ap₄A would be metabolised at high rate leaving enough enzymatic capacity to hydrolyse Ap₄A-RNA as well and thus the inhibition effect was not observed.

Conclusions

In light of recent discoveries of noncanonical RNA caps in various organisms, we focused on potential RNA-decapping Nudix enzymes from *A. thaliana*. We explored the cleavage properties of four enzymes, cytosolic AtNUDT6, AtNUDT7,²³ and chloroplastic AtNUDT19 and AtNUDT27.^{14,22} First, we tested whether they could cleave canonical and noncanonical caps from RNA using *in vitro* transcribed RNA capped with Np_nNs, CoA, NAD(H), m⁷Gp₃A(G), and ADP-ribose. We found that AtNUDT19 was very specific and cleaved NADH-RNA almost exclusively. Also, AtNUDT6 and AtNUDT7 were relatively specific and cleaved mainly Ap₄A-RNA. In contrast, AtNUDT27 effectively hydrolyzed almost all types of capped RNA at a tenfold lower concentration than the other tested enzymes, making it potentially a general RNA-decapping enzyme. As the *in vitro* substrates were not known for all of these enzymes, we also tested their substrate specificity on free small molecules that can form RNA caps (Np_nNs, 3'-dpCoA, NAD(H), m⁷Gp₃A(G), and ADP-ribose). Again, AtNUDT19 was very specific towards NADH, which is in accordance with role of AtNUDT19 in regulating NADH and NADPH levels in chloroplasts, as previously reported.²² Under our experimental conditions AtNUDT19 cleaved only free NADH but not NAD. Interestingly, when capped RNA was used as substrate, AtNUDT19 cleaved NADH-RNA and also partially NAD-RNA. This may indicate its role as an NAD/NADH-RNA-decapping enzyme. Contrary to its promiscuity in the case of capped RNA, AtNUDT27 hydrolyzed dinucleoside tetraphosphates and pentaphosphates almost exclusively. Our observed activity of AtNUDT27 against dinucleoside tetraphosphates is in contrast with previously reported results.²² There, authors observed cleavage of only Ap₅A but not of Ap₄A. Nevertheless, the presence of reducing agents may positively influence the activity of these enzymes.²⁴ Thus, this discrepancy may be explained by the presence of a higher concentration of reducing agent DTT (2 mM) in our reaction in comparison with their 1 mM DTT concentration. AtNUDT6 and AtNUDT7 mainly cleaved free molecules with the diphosphate bridge. In fact, both AtNUDT6 and AtNUDT7 were demonstrated to be cytosolic NADH pyrophosphohydrolases

participating in stress responses through modulation of intracellular NADH levels in *A. thaliana*.²⁵ Ap₄A is also known to be involved in stress signalling in many organisms, and because AtNUDT6, AtNUDT7, and AtNUDT27 efficiently cleaved both the free Ap₄A and the Ap₄A-RNA, we compared their preferences on cleavage of these substrates. We observed that AtNUDT27 and AtNUDT6 cleaved the free Ap₄A more efficiently, while AtNUDT7 favoured the Ap₄A-RNA. These differences in substrate specificities suggest that *in vivo* some Nudix proteins may function in degradation of free small molecules, while others serve as specific decapping enzymes, probably during regulated RNA decay.

It has been shown that diadenosine polyphosphates such as Ap₃A and Ap₄A behave like alarmones in *A. thaliana* and trigger a cascade of reactions to yield various protective compounds.¹³ Our experiments show that plant Nudix enzymes have RNA-decapping activity and that they can cleave various noncanonical RNA caps. This finding supports the notion that plant RNA also contains other unknown RNA caps in addition to the canonical m⁷G structure and NAD cap.^{3,26} Altogether, the existence of specific enzymes that cleave NADH- and Ap₄A-RNA suggest that these RNA caps may play an important role in plant RNA metabolism, signalling, or stress reaction.

Moreover, the specific cleavage properties of the individual Nudix enzymes characterized in this work can be used to develop capturing strategies for noncanonically capped RNA and thus help us to identify the sequence and features of these RNA species.

Materials and methods

Preparation of RNA

***In vitro* transcription with T7 RNAP.** All chemicals were either purchased from Merck or Jena Biosciences and used without further purification. Oligonucleotides were purchased from Generi Biotech. *In vitro* transcription was performed using a modification of a previously published method,⁹ in a 25 μ L mixture containing: 80 ng μ L⁻¹ of template DNA (35A for Ap₂₋₅A, Gp₃₋₄A, NAD(H), CoA, ADP-ribose and m⁷Gp₃A, or 35G for Gp₃₋₄G, Ap₃₋₅G, and m⁷Gp₃G), 1 mM UTP, 1 mM CTP, 1 mM ATP, 0.8 mM GTP and 0.5 μ L α -³²P GTP (activity: 9.25 MBq in 25 μ L), 1.6 mM of Np_nNs (Ap₂₋₅A, Gp₃₋₄G, Ap₃₋₅G, m⁷Gp₃G, m⁷Gp₃A, ADP-ribose), or 8 mM of cofactors (3'-dpCoA, NAD and NADH), 5% DMSO, 0.12% triton X-100, 12 mM DTT, 4.8 mM magnesium chloride and 10 $\times$ reaction buffer for T7 RNAP (40 mM Tris-HCl, 6 mM MgCl₂, 1 mM DTT, 2 mM spermidine, pH 7.9 at 25 $^{\circ}$ C) and 62.5 units of T7 RNAP. The mixture was incubated for 2 h at 37 $^{\circ}$ C.

DNase I treatment. The DNA template was digested by DNase I to obtain pure RNA. 25 μ L of the transcription mixture, 3 μ L of 10 $\times$ reaction buffer for DNase I (100 mM Tris-HCl, 25 mM MgCl₂, 5 mM CaCl₂, pH 7.6, supplied with the enzyme), and 4 units of DNase I (NEB) were incubated at 37 $^{\circ}$ C for 60 min. The enzyme was heat deactivated at 75 $^{\circ}$ C for 10 min followed by immediate cooling on ice. All samples were purified on RNA mini Quick Spin Columns (Merck) for further use.

RNA 5'-polyphosphatase treatment. The products from the previous reaction were treated with 20 units of 5'-polyphosphatase (Epicentre) in the reaction buffer (50 mM HEPES-KOH, pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.1% β -mercaptoethanol, and 0.01% Triton X-100) for 1 h at 37 °C.

Terminator™ 5'-phosphate-dependent exonuclease treatment. Because of the incompatibility of the buffers, all samples were purified on RNA mini Quick Spin Columns before the reaction. The RNA was treated with 1 unit of Terminator™ 5'-phosphate-dependent exonuclease (Epicentre) in reaction buffer A (supplied with the enzyme) at 30 °C for 1 h.

Purification of recombinant NudiX enzymes

A. thaliana NUDT6, 7, 19, and 27 genes were cloned into the pET26b vector with the C-terminal His tag (Novagen). Due to the problem with expressing the full-length AtNUDT27, it was cloned without the first 44 aa containing the chloroplast signal sequence. Proteins were overexpressed in *Escherichia coli* Rosetta DE3 cells grown for 16 hours in auto-inducing medium.²⁷ Cells were lysed by sonication in buffer A (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10% glycerol and 5 mM β -mercaptoethanol) and recombinant proteins were purified on TALON Metal Affinity Resin (Clontech), eluted in buffer A supplemented with 250 mM of imidazol and fractionated by gel filtration chromatography on an ENrich™ SEC70 column (BioRad). The purified proteins were concentrated to 50 μ M in a buffer containing 20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 2 mM DTT, and 10% glycerol, flash frozen in liquid nitrogen, and stored at -80 °C. Protein concentration was determined at 280 nm using the molar extinction coefficient calculated on the basis of protein composition with the ProtParam tool (<https://web.expasy.org/protparam/>) (Fig. S13, ESI†).

Cleavage of capped RNA by *A. thaliana* NudiX enzymes

The decapping assay was performed at 37 °C in 10 μ L reaction. To test the cleavage of the 5'-caps, the RNA samples were divided into two parts. The positive control contained 100 ng of the pure RNA (*in vitro* transcription, DNase I treatment, purified on RNA mini Quick Spin Columns), 1 μ L of 10 $\times$ buffer (100 mM Tris-HCl, pH 7.5, 1 M KCl, 20 mM MgCl₂, 20 mM MnCl₂, 20 mM DTT), and 500 nM of the recombinant AtNUDT6, 7, 19, and 50 nM AtNUDT27. Water was used as the negative control. The mixtures were incubated at 37 °C for 1 h and purified on RNA mini Quick Spin Columns. Then 3 μ L of the purified samples were mixed with 2 $\times$ RNA loading dye (NEB) for analysis by 12% PAGE (600 V, 3.5 h). Typhoon FLA 9500 was used as a visualization imaging system.

Calculations of RNA decapping

The experiments were performed in triplicates, and the amount of capped RNA was calculated from PAGE analysis using the software ImageJ.²⁸ Areas under the peaks (AUC) corresponding to the 5'-capped RNA (A_{r_p}) in the control without enzyme and 5'-capped RNA ($A_{r_{cap}}$) in samples with enzyme were calculated. The percentage of 5'-capped RNA species was calculated according to: $(A_{r_{cap}}/A_{r_p}) \times 100$.

Enzymatic assay and kinetics assay on small free molecules

The hydrolytic activities of AtNUDT proteins towards different Np_nNs (Ap₂₋₅A, Gp₃₋₄G, Ap₃₋₅G, m⁷Gp₃G, m⁷Gp₃A), 3'-dpCoA, ADP-ribose, and NAD(H) were performed using 500 nM of the AtNUDT6, 7, 19, and 27 and 400 μ M of substrate in 10 mM Tris-HCl (pH 7.5), 100 mM KCl, 2 mM magnesium acetate, 2 mM DTT, at 37 °C. The enzyme was heat deactivated at 75 °C for 10 min followed by cooling on ice. The mixture was then analysed by monitoring absorbance at 259 nm and retention time compared to standards on high-performance liquid chromatography using a Kinetex C18 column (4.6 $\times$ 150 mm, 5 μ m, Phenomenex) at a flow rate of 1 ml min⁻¹. The mobile phase consisted of 100 mM triethylammonium acetate pH 7 (buffer A) and Acetonitrile (buffer B). Separation was achieved using a 25 minute gradient: 100% A for 2 min; linear decrease to 95% A over 4 min; linear decrease to 70% A over 9 min; maintaining 70% A for 3 min; returning linearly to 100% A over 1 min. Chromatograms were integrated, and the percentage of substrate cleavage was calculated from corresponding AUC compared to the control without enzyme.

Kinetic parameter calculation

Kinetic parameters of AtNUDT enzymes were calculated from experiments using increasing concentrations of Ap₄A (0, 5, 10, 20, 30, 50, 70, and 100 μ M) and 50 nM of AtNUDT6, 7, and 27. The 500 μ L mixture was incubated at 37 °C, and aliquots of 100 μ L were collected at 0, 2.5, 5, 10, and 15 min and immediately heated to 75 °C for 10 min to stop the reaction. All samples were analysed using the HPLC method described above. Calculated initial reaction velocities were plotted against Ap₄A concentration and fitted to the Michaelis-Menten model to obtain K_m and V_{max} values using Origin software (Northampton, MA). Student *t*-test and analysis of variance (ANOVA) with *post hoc* Bonferroni analysis were performed in Microsoft Excel to assess significant differences between measured parameters where needed.

Inhibition test with Ap₄A for AtNUDT enzymes

To test the inhibition of AtNUDT6, AtNUDT7 and AtNUDT27 enzymes in the presence of Ap₄A we measured the cleavage of Ap₄A-RNA at 37 °C in 10 μ L reaction using 1 μ M of purified Ap₄A-RNA and increased concentrations of Ap₄A (1, 2 and 4 μ M). The concentration used for AtNUDT6 and AtNUDT7 was 500 nM while for AtNUDT27 was 50 nM. The reaction was stopped using 2 $\times$ RNA Loading Dye (NEB).

Author contributions

M.-B.M., O.H., A.G., S.K. performed experiments. M.-B.M. and O.H. studied RNA and small molecule behaviour and designed experiments, A.G. and S.K. purified plant proteins. R.B. and O.N. designed kinetic assays. J.K. and H.C. designed experiments and coordinated study. M.-B.M., O.N., and H.C. wrote the manuscript.

Conflicts of interest

Authors declare no conflicts of interests.

Acknowledgements

We are grateful to all members of the Cahova Group for their help and advice. We acknowledge Martin Lepšík for figure of AtNudt7 for TOC. We acknowledge funding from the Ministry of Education, Youth and Sports (Czech Republic), programme ERC CZ (LL1603) and European Research Council Executive Agency (ERCEA) under the European Union's Horizon Europe Framework Programme for Research and Innovation (grant agreement No 101041374 - StressRNAction). J. K. acknowledges funding from National Science Centre UMO-2018/29/B/NZ3/01980.

Notes and references

- 1 P. Boccaletto, M. A. Machnicka, E. Purta, P. Piątkowski, B. Bagiński, T. K. Wirecki, V. de Crécy-Lagard, R. Ross, P. A. Limbach, A. Kotter, M. Helm and J. M. Bujnicki, *Nucleic Acids Res.*, 2017, **46**, D303–D307.
- 2 H. Cahová, M.-L. Winz, K. Höfer, G. Nübel and A. Jäschke, *Nature*, 2014, **519**, 374.
- 3 Y. Wang, S. Li, Y. Zhao, C. You, B. Le, Z. Gong, B. Mo, Y. Xia and X. Chen, *Proc. Natl. Acad. Sci.*, 2019, **116**, 12094–12102.
- 4 X. Jiao, S. K. Doamekpor, J. G. Bird, B. E. Nickels, L. Tong, R. P. Hart and M. Kiledjian, *Cell*, 2017, **168**, 1015–1027.e1010.
- 5 O. Ruiz-Larrabeiti, R. Benoni, V. Zemlianski, N. Hanišáková, M. Schwarz, B. Brezovská, B. Benoni, J. Hnilicová, V. R. Kabardin, H. Cahová, M. Vítězová, M. Převorovský and L. Krásný, *bioRxiv*, 2021, Preprint, 2021.2012.2014.472595, DOI: [10.1101/2021.12.14.472595](https://doi.org/10.1101/2021.12.14.472595).
- 6 J. G. Bird, U. Basu, D. Kuster, A. Ramachandran, E. Grudzien-Nogalska, A. Towheed, D. C. Wallace, M. Kiledjian, D. Temiakov, S. S. Patel, R. H. Ebright and B. E. Nickels, *eLife*, 2018, **7**, e42179.
- 7 W. E. Kowtoniuk, Y. Shen, J. M. Heemstra, I. Agarwal and D. R. Liu, *Proc. Natl. Acad. Sci.*, 2009, **106**, 7768–7773.
- 8 J. Wang, B. L. Alvin Chew, Y. Lai, H. Dong, L. Xu, S. Balamkundu, W. M. Cai, L. Cui, C. F. Liu, X.-Y. Fu, Z. Lin, P.-Y. Shi, T. K. Lu, D. Luo, S. R. Jaffrey and P. C. Dedon, *Nucleic Acids Res.*, 2019, **47**, e130–e130.
- 9 O. Hudeček, R. Benoni, P. E. Reyes-Gutierrez, M. Culka, H. Šanderová, M. Hubálek, L. Rulišek, J. Cvačka, L. Krásný and H. Cahová, *Nat. Commun.*, 2020, **11**, 1052.
- 10 E. Rapaport and P. C. Zamecnik, *Proc. Natl. Acad. Sci.*, 1976, **73**, 3984–3988.
- 11 P. Zamecnik, *Anal. Biochem.*, 1983, **134**, 1–10.
- 12 P. G. Zamecnik, M. L. Stephenson, C. M. Janeway and K. Randerath, *Biochem. Biophys. Res. Commun.*, 1966, **24**, 91–97.
- 13 M. Pietrowska-Borek, K. Nuc, M. Zielezińska and A. Guranowski, *FEBS Open Bio*, 2011, **1**, 1–6.
- 14 K. Yoshimura and S. Shigeoka, *Biosci., Biotechnol., Biochem.*, 2015, **79**, 354–366.
- 15 A. G. McLennan, *Cell. Mol. Life Sci. CMLS*, 2006, **63**, 123–143.
- 16 M. G. Song, S. Bail and M. Kiledjian, *RNA*, 2013, **19**, 390–399.
- 17 S. Kramer and A. G. McLennan, *Wiley Interdiscip. Rev.: RNA*, 2019, **10**, e1511.
- 18 E. Grudzien-Nogalska, X. Jiao, M. G. Song, R. P. Hart and M. Kiledjian, *RNA*, 2016, **22**, 773–781.
- 19 M.-G. Song, Y. Li and M. Kiledjian, *Mol. Cell*, 2010, **40**, 423–432.
- 20 E. Grudzien-Nogalska, Y. Wu, X. Jiao, H. Cui, M. K. Mateyak, R. P. Hart, L. Tong and M. Kiledjian, *Nat. Chem. Biol.*, 2019, **15**, 575–582.
- 21 S. Sharma, E. Grudzien-Nogalska, K. Hamilton, X. Jiao, J. Yang, L. Tong and M. Kiledjian, *Nucleic Acids Res.*, 2020, **48**, 6788–6798.
- 22 T. Ogawa, K. Yoshimura, H. Miyake, K. Ishikawa, D. Ito, N. Tanabe and S. Shigeoka, *Plant Physiol.*, 2008, **148**, 1412–1424.
- 23 T. Ogawa, Y. Ueda, K. Yoshimura and S. Shigeoka, *J. Biol. Chem.*, 2005, **280**, 25277–25283.
- 24 K. Olejnik, M. W. Murcha, J. Whelan and E. Kraszewska, *FEBS J.*, 2007, **274**, 4877–4885.
- 25 T. Ogawa, K. Muramoto, R. Takada, S. Nakagawa, S. Shigeoka and K. Yoshimura, *Plant Cell Physiol.*, 2016, **57**, 1295–1308.
- 26 H. Zhang, H. Zhong, S. Zhang, X. Shao, M. Ni, Z. Cai, X. Chen and Y. Xia, *Proc. Natl. Acad. Sci.*, 2019, **116**, 12072–12077.
- 27 F. W. Studier, *Protein Expression Purif.*, 2005, **41**, 207–234.
- 28 W. S. Rasband, *ImageJ*, U. S. National Institutes of Health, Bethesda, Maryland, USA, 1997–2018.

Nucleic Acids

Diadenosine Tetraphosphate (Ap₄A) Serves as a 5' RNA Cap in Mammalian Cells

Jiří František Potužník⁺, Ondřej Nešuta⁺, Anton Škríba, Barbora Voleníková,
Maria-Bianca Mititelu, Flaminia Mancini, Valentina Serianni, Henri Fernandez,
Kristína Spustová, Jana Trylčová, Pavel Vopalensky, and Hana Cahová*

Abstract: The recent expansion of the field of RNA chemical modifications has changed our understanding of post-transcriptional gene regulation. Apart from internal nucleobase modifications, 7-methylguanosine was long thought to be the only eukaryotic RNA cap. However, the discovery of non-canonical RNA caps in eukaryotes revealed a new niche of previously undetected RNA chemical modifications. We are the first to report the existence of a new non-canonical RNA cap, diadenosine tetraphosphate (Ap₄A), in human and rat cell lines. Ap₄A is the most abundant dinucleoside polyphosphate in eukaryotic cells and can be incorporated into RNA by RNA polymerases as a non-canonical initiating nucleotide (NCIN). Using liquid chromatography-mass spectrometry (LC-MS), we show that the amount of capped Ap₄A-RNA is independent of the cellular concentration of Ap₄A. A decapping enzyme screen identifies two enzymes cleaving Ap₄A-RNA, NUDT2 and DXO, both of which also cleave other substrate RNAs in vitro. We further assess the translatability and immunogenicity of Ap₄A-RNA and show that although it is not translated, Ap₄A-RNA is recognized as self by the cell and does not elicit an immune response, making it a natural component of the transcriptome. Our findings open a previously unexplored area of eukaryotic RNA regulation.

Dinucleoside polyphosphates (Np_nNs) are present in all types of cells and have been implicated in many biological processes.^[1] Their intracellular concentrations increase under stress conditions, which is why they are referred to as alarmones. They are produced by tRNA synthetases,^[2] and their formation has also been linked to conjugation of ubiquitin and ubiquitin-like proteins.^[3] Np_nNs are degraded by members of the Nucleoside Diphosphate-Linked Moiety X (NudiX) family of proteins.^[4] The most explored Np_nN, diadenosine tetraphosphate (Ap₄A), serves as a substrate for the AMPylation of proteins.^[5] Its role, however, is probably much more diverse with regard to its varied protein interactome.^[6]

Ap₄A was also reported to be highly upregulated during the immune activation of mast cells.^[7] Immunoglobulin E (IgE) binds to the high affinity receptor FcεR1^[8] of the mast cell, and this triggers the phosphorylation of serine 207 of the Lysyl-tRNA synthetase (KARS). Upon its phosphorylation, KARS detaches from the multi-tRNA synthetase complex and translocates to the nucleus, where it starts producing Ap₄A. Ap₄A then interacts with the transcriptional repressor histidine triad nucleotide binding protein (HINT1) and enables for the transcription of previously repressed genes.^[9]

Recently, we reported that Np_nNs can also serve as 5' RNA caps in bacteria.^[10] They are accepted by RNA polymerases and used as the first non-canonical initiating nucleotide (NCIN).^[10–11] The concentration of Np_nN-RNA in bacteria also increases under stress conditions.^[12]

Given that Ap₄A is produced in high concentrations in the nucleus and that RNA polymerases can use it as a substrate, we searched for Ap₄A capped RNA in two lines of mammalian cells, where Ap₄A was previously detected; human embryonic kidney cells (HEK293T) and rat basophilic leukemia cells (RBL-2H3) which serve as a model of mast cells (Figure 1A). We confirmed the presence of Ap₄A (Figure 1B) in both HEK293T and RBL-2H3 cells using liquid chromatography-mass spectrometry (LC-MS), specifically a quadrupole time-of-flight mass spectrometer (Q-TOF) (Figure 1C, D). We detected similar amounts of Ap₄A in both of the cell lines used (Figure 1F, Figure S1). We assumed that an increase in Ap₄A could potentially increase the amount of Ap₄A capped RNA in the same way that bacterial Np_nN-RNA capping increased in response to

[*] J. František Potužník, ⁺ Dr. O. Nešuta, ⁺ Dr. A. Škríba, B. Voleníková, M.-B. Mititelu, F. Mancini, V. Serianni, H. Fernandez, K. Spustová, Dr. J. Trylčová, Dr. P. Vopalensky, Dr. H. Cahová
Chemical Biology of Nucleic Acids, Institute of Organic Chemistry and Biochemistry of the CAS
Flemingovo náměstí 2, Prague 6 (Czechia)
E-mail: cahova@uochb.cas.cz

J. František Potužník, ⁺ M.-B. Mititelu, F. Mancini, V. Serianni
Department of Cell Biology, Faculty of Science, Charles University
Viničná 7, Prague 2 (Czechia)

[†] These authors contributed equally to this work.

© 2023 The Authors. Angewandte Chemie International Edition published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution Non-Commercial NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

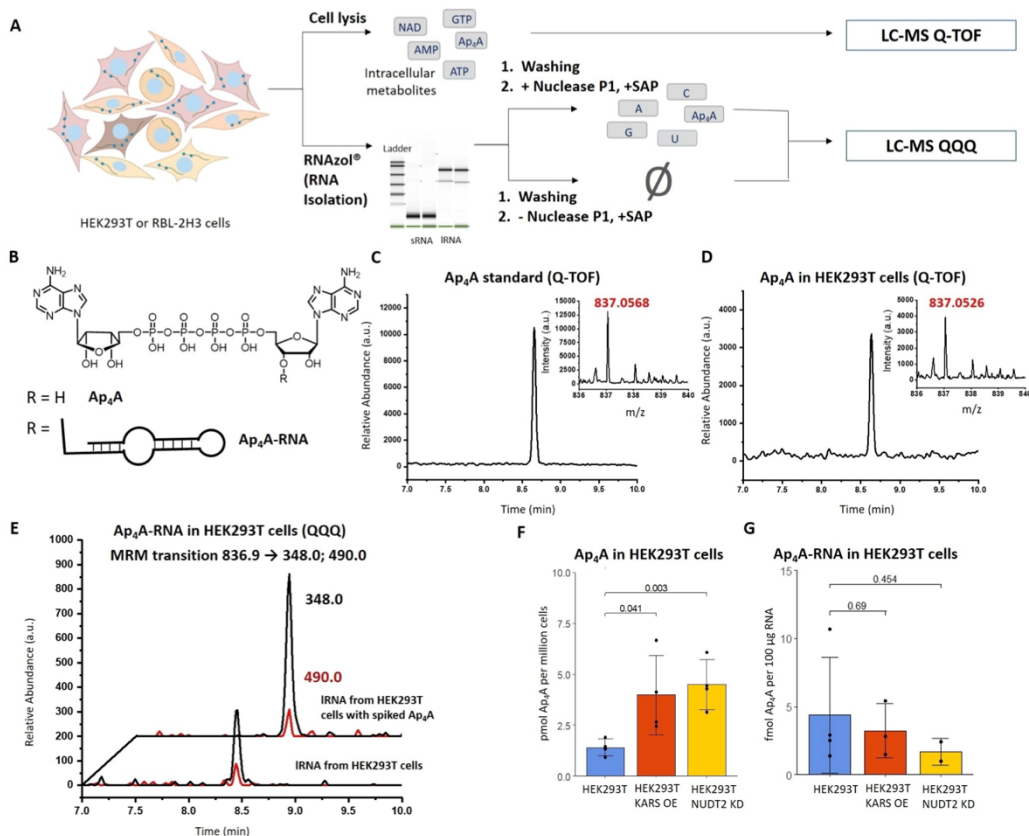


Figure 1. Detection and quantification of Ap₄A and Ap₄A-RNA in mammalian cells. **A**) Scheme of sample preparation for LC-MS analysis. **B**) Structure of Ap₄A and Ap₄A-RNA. **C**) Extracted Ion Chromatogram (EIC) and mass spectra of Ap₄A standard. **D**) EIC and mass spectra of Ap₄A from HEK293T cell lysate measured on the Q-TOF instrument. **E**) EIC of Ap₄A from IRNA after NuP1 and SAP treatment with and without a spike of Ap₄A standard. Parent Ap₄A ion [M + H]⁺ m/z 836.9 and two transitions m/z 348.0 (black) m/z 490.0 (red) were monitored by MRM experiment with QQQ instrument. **F**) Quantification of Ap₄A in HEK293T, HEK293T KARS OE and HEK293T NUDT2 KD cells. All of the experiments were performed in tetraplicates. **G**) Quantification of Ap₄A-RNA in HEK293T, HEK293T KARS OE and HEK293T NUDT2 OE cells. The error bars in **F** and **G** represent the standard deviation and the p-values from pairwise t-test are shown above the compared groups. All of the experiments were performed on a minimum of two biological replicates (HEK293T measured in tetraplicate, HEK293T KARS OE measured in triplicate).

Np₄N increase.^[10] For that reason, we stimulated RBL-2H3 cells with IgE and an antigen.

Surprisingly, this activation of the cells caused an insignificant decrease in the concentration of Ap₄A (Figure S1), which was in contrast to previously reported data.^[8] In order to have a system with overproduced Ap₄A, we generated two mutant HEK293T cell lines. One with the inducible expression of phosphomimetic KARS-mut (S207D),^[9] which has the enzyme KARS constantly located in the nucleus where it only produces Ap₄A (HEK293T KARS OE). The other was a line with a knockdown of the enzyme Nudix hydrolase 2 (HEK293T NUDT2 KD). It was reported that NUDT2 cleaves Ap₄A and that NUDT2

knock-out cells have 175-fold higher intracellular concentration of Ap₄A and have a significantly changed transcriptional profile with down-regulated genes involved in the innate immune response and in cancer growth promotion.^[13] We confirmed the increase of intracellular Ap₄A in both of the mutant HEK293T cell lines (Figure 1F). We then focused on the detection of the capped Ap₄A-RNA. We isolated the fraction of RNA longer than 200 nucleotides (nt) (IRNA) from both cell lines (Figure S2) and washed it using 3 kDa filters to remove non-covalently bound small molecules. The RNA was subsequently digested into nucleotides and RNA caps by Nuclease P1 (NuP1). Simultaneously, we added Shrimp Alkaline Phosphatase (SAP) to

remove phosphate groups from the digested nucleotides while leaving Ap₄A intact (Figure 1A). This treatment reduces the matrix effect coming from the co-eluting nucleotides in the Ap₄A elution time window. While our experiments show that NuP1 partially cleaves Ap₄A in the RNA matrix (Figure S3), we are currently not aware of a more suitable enzyme that can completely remove intact Ap₄A from the rest of the RNA.

The samples were then subjected to analysis by LC-MS using a triple quadrupole mass spectrometer (QQQ) for the identification and quantification of Ap₄A in RNA. As a control, we used an identical sample without NuP1 treatment (–NuP1; Figure S4–7). We detected and quantified Ap₄A-RNA in HEK293T cells (4.4 fmol/100 µg, Figure 1E, Figure S4–6). We detected Ap₄A-RNA in the RBL-2H3 cell line but its very low abundance did not allow for proper quantification (Figure S7). We also wanted to explore whether there is a correlation between the intracellular concentration of Ap₄A and Ap₄A-RNA. To our surprise, the amount of measured Ap₄A capped RNA remained similar in both of our mutant cell lines with increased Ap₄A (3.2 fmol/100 µg in HEK293T KARS OE and 1.7 fmol/100 µg in HEK293T NUDT2 KD, Figure 1G). Given this constant level of Ap₄A-RNA, there is likely a more intricate mechanism of Ap₄A-RNA capping and decapping.

Since an increase in the intracellular concentration of Ap₄A did not lead to its increased incorporation as an NCIN into RNA, we wanted to explore alternative mechanisms of Ap₄A-RNA formation. The triphosphate at the 5' end of uncapped RNA may potentially react in the same way that ATP does in the reaction with lysyl-AMP catalyzed by KARS to create Ap₄A. To confirm if KARS catalyzes the reaction of ppp-RNA with lysyl-AMP to create Ap₄A-RNA, we tested it in vitro. While human KARS formed Ap₄A under our conditions (Figure S8), we did not observe any RNA capping (Figure S9). This either means, that other capping enzymes are responsible for Ap₄A-RNA formation, or it is co-transcriptionally regulated by an unknown mechanism.

Another way to regulate Ap₄A-RNA could involve posttranscriptional degradation by decapping enzymes. We have previously shown that bacterial^[10] and plant^[14] Nudix enzymes cleave Np_nN-RNA. The human genome encodes 24 enzymes from this family, and given their conservation across species,^[15] we postulated that mammalian Nudixes may have a similar activity. NUDT2 is a member of the Nudix family, and considering the change in Ap₄A in the NUDT2 KD cell line, which is not reflected in the Ap₄A-RNA concentration, we wanted to further explore its activity in vitro to better understand the mechanisms of its function. We isolated and purified recombinant human NUDT2 and tested its activity on Ap₄A as well as on the in vitro prepared Ap₄A-RNA (Figure S10). The cleavage of Ap₄A by NUDT2 produced adenosine triphosphate (ATP) and adenosine monophosphate (AMP) (Figure S10). We then tested various conditions of Ap₄A-RNA cleavage and observed complete decapping of Ap₄A-RNA at 500 nM concentration of NUDT2. The resulting products of the cleavage were ATP and 5'-monophosphate RNA (p-RNA).

At higher concentrations, the enzyme also cleaved ppp-RNA into p-RNA (Figure S10) as was also previously reported.^[16] In addition to Ap₄A-RNA, we tested the activity of NUDT2 on other Np_nN-RNAs, some of which were previously detected in bacteria^[10] (Figure S11). NUDT2 cleaved the vast majority of them, with a lower activity towards Np₃N-RNAs. This implies that NUDT2 serves as a general decapping enzyme with broad substrate specificity. However, this observed in vitro activity is in contrast with the measured amounts of Ap₄A-RNA in the NUDT2 KD line (Figure 1G). This phenomenon could be explained by a substrate redundancy between different members of the Nudix family^[15] or a potential upregulation of other RNA decapping enzymes upon NUDT2 KD.

Apart from the Nudix family, the human enzyme Decapping Exoribonuclease (DXO) has also been shown to decap RNA with an affinity towards NAD-RNA.^[17] Unlike NUDT2, DXO does not cleave Np_nNs (Figure S12), instead, it removes the entire RNA cap and leaves the transcript one nucleotide shorter. Interestingly, the Ap₄A-RNA is cleaved within the polyphosphate bridge, which is in stark contrast to its activity on NAD-RNA (Figure 2, Figure S12).

After confirming the presence of the capped Ap₄A-RNA in mammalian cells and identifying its decapping enzymes, we searched for its interacting proteins. Based on the fact that HINT1 binds Ap₄A with minimal hydrolytic activity,^[7] we tested whether Ap₄A-RNA could be a better substrate for hydrolysis or just an interaction partner. We prepared recombinant HINT1 along with in vitro tran-

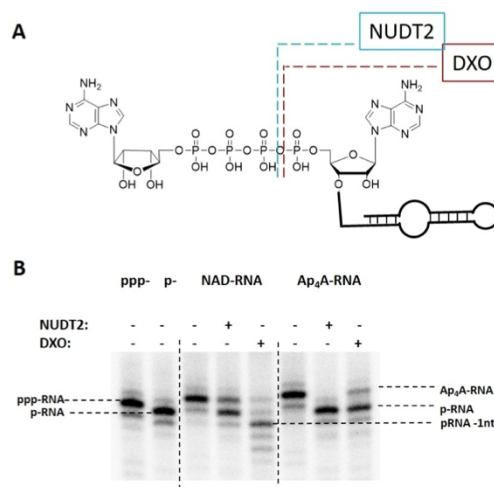


Figure 2. Cleavage of Ap₄A-RNA. A) NUDT2 and DXO cleave Ap₄A-RNA. B) Polyacrylamide gel electrophoresis (PAGE) of ³²P labeled ppp-RNA, p-RNA and NAD-RNA and Ap₄A-RNA treated with NUDT2 and DXO. Prior to the reaction with NUDT2 and DXO, NAD-RNA and Ap₄A-RNA were treated with 5'-polyphosphatase and Terminator exonuclease. Experiments were performed in triplicate. Uncropped image of the gel Figure S16.

scribed model Ap₄A-RNA, m⁷Gp₃Am-RNA (cap1-RNA), and ppp-RNA as controls. Using PAGE and an electrophoretic mobility shift assay (EMSA), we observed an interaction between Ap₄A-RNA and HINT1 ($K_D = 2.22 \pm 0.16 \mu\text{M}$) but no hydrolysis (Figure S13). However, the same interaction pattern was also observed with control cap1-RNA ($K_D = 1.88 \pm 0.65 \mu\text{M}$) and ppp-RNA ($K_D = 2.08 \pm 0.58 \mu\text{M}$) (Figure S14A). This data points to a non-specific interaction of HINT1 with any type of 5' RNA cap or with the RNA backbone.

Another RNA binding protein interacting with the 5' end of RNA is the eukaryotic translation initiation factor 4E (eIF4E). eIF4E binds to the cap1 of mRNA, facilitates the recruitment of ribosomes, and initiates translation.^[18] Since we observed Ap₄A-RNA in the long fraction of isolated RNA, we hypothesized that Ap₄A may function as another mRNA cap and that it can also be recognized by eIF4E. We performed an EMSA with recombinant eIF4E. While we observed the formation of a complex with Ap₄A-RNA, it was significantly weaker ($K_D = 0.92 \pm 0.09 \mu\text{M}$) than with the canonical cap1-RNA ($K_D = 0.20 \pm 0.04 \mu\text{M}$) and similar to ppp-RNA ($K_D = 1.09 \pm 0.17 \mu\text{M}$) (Figure S14B). Even though the interaction of eIF4E with Ap₄A-RNA was weaker than with its canonical partner, it does not mean that Ap₄A-RNA is not translatable. To ascertain the translatability of Ap₄A-mRNA we established an *in vivo* assay using HEK293T cells. We prepared *Renilla* luciferase (RLuc) mRNA with the three caps: Ap₄A, cap1, ppp-, and firefly luciferase (FFLuc) with cap1 (as a normalizer) by *in vitro* transcription. After the degradation of uncapped RNA by 5' polyposphatase and Terminator exonuclease, the mRNA was purified by HPLC to remove dsRNA, which was reported to trigger the innate immune system^[19] (Figure 3A). We transfected HEK293T cells with the purified RLuc and FFLuc mRNA and measured the activity of both luciferases after 24 hours (Figure 3B). Consistently with our EMSA results, we did not observe any translation from Ap₄A-RNA (Figure 3C) and ppp-RNA, while cap1-RNA was translated efficiently. To exclude the possibility that Ap₄A-RNA is decapped by NUDT2 and degraded prior to its translation, we used our HEK293T NUDT2 KD cells for the same *in vivo* translational assay. The results showed no translation of Ap₄A-RNA, while cap1-RNA was translated efficiently (Figure 3D). To eliminate the possibility that other decapping enzymes are involved in Ap₄A-RNA degradation, we also performed an *in vitro* translational assay employing retic lysate. This experiment allowed us to assess the inherent ability of Ap₄A-RNA to be translated without interference from other intracellular processes. Consistently with the previous assays, cap1-RNA was efficiently translated, and control ppp-RNA and Ap₄A-RNA were not (Figure 3E).

Key to any RNA function within the cell is its stability. For that reason, we compared the stability of cap1-, ppp-, and Ap₄A-mRNA in HEK293T cells at various time points (4, 6, 8 and 24 h) post transfection by RT-qPCR. The results show that all of the differently capped RNAs are degraded at the same rate (Figure 3F).

Ap₄A and NUDT2 have also been associated with innate immunity. For that reason, we tested if the presence of Ap₄A-RNA induces an innate immune response. We measured the expression of several immune genes involved in nucleic acid sensing (IFIT-1, 2, 3, 5, MDA-5, and RIG-I) 6 hours after the transfection with Ap₄A-mRNA, cap1-mRNA and crude (non HPLC purified) ppp-RNA as positive control. We transfected HeLa cells instead of HEK293T cells as those have a dysfunctional innate immune response.^[20] We did not observe any difference in the expression of these genes between the two types of capped RNAs at 100 ng of transfected RNA (Figure 3G). This shows that Ap₄A-RNA is recognized as self by the cell and does not elicit an intracellular immune response.

We are the first to present the discovery of the new Ap₄A RNA cap in eukaryotes. It is also the first example of a dinucleoside polyphosphate functioning as a cap in higher organisms. Using LC-MS, we have shown that Ap₄A is present in the two different cell lines used—HEK293T and RBL-2H3. The overexpression of the phosphomimetic KARS-mut and NUDT2 KD leads to an increase in Ap₄A, but this change is not reflected in an increase of Ap₄A-RNA. This discrepancy suggests that the incorporation of Ap₄A into RNA is regulated by another mechanism and is not simply stoichiometric. We tested if human KARS caps ppp-RNA with ATP, leading to Ap₄A-RNA in a similar manner in which it creates Ap₄A. *In vitro* experiments showed Ap₄A formation but no RNA capping activity. We therefore presume, that non-canonical transcription initiation is the most probable way of forming Ap₄A-RNA.

We have also shown that in addition to cleaving Ap₄A, NUDT2 cleaves Ap₄A-RNA along with a number of other Np_nN capped RNAs, making it a general RNA decapping enzyme. Interestingly, our *in vitro* results do not correspond to the amount of Ap₄A-RNA in the NUDT2 KD line. This disparity might be explained by a general substrate redundancy between various Nudix enzymes. In addition to NUDT2, we identified DXO as another Ap₄A-RNA decapping enzyme *in vitro*. It is important to note that the DXO cleavage site of Ap₄A-RNA is different to the one of DXOs preferred substrate NAD-RNA.^[17a] It remains to be seen whether DXO functions the same way *in vivo*.

HINT1 was reported to interact with Ap₄A (with a negligible hydrolytic activity). While we observed an interaction between Ap₄A-RNA and HINT1, the interaction was non-specific, and the protein also interacted with cap1-RNA and ppp-RNA. Because we detected Ap₄A-RNA in the long RNA fraction, we also explored its interaction with eIF4E and its translatability. The interaction with eIF4E was minimal compared to the canonical cap1-RNA and comparable with non-capped ppp-RNA. To exclude the possibility of non-canonical translation initiation^[21] without eIF4E, we performed *in vivo* translational assays. Ap₄A-RNA was not translated in the HEK293T cells, nor was it translated in NUDT2 KD cells, which means that decapping of the mRNA by NUDT2 prior to translation does not play a role. The inability of Ap₄A RNA to be translated is further supported by the *in vitro* translational assay, as there is no visible Ap₄A-RNA translation in contrast to cap1-RNA.

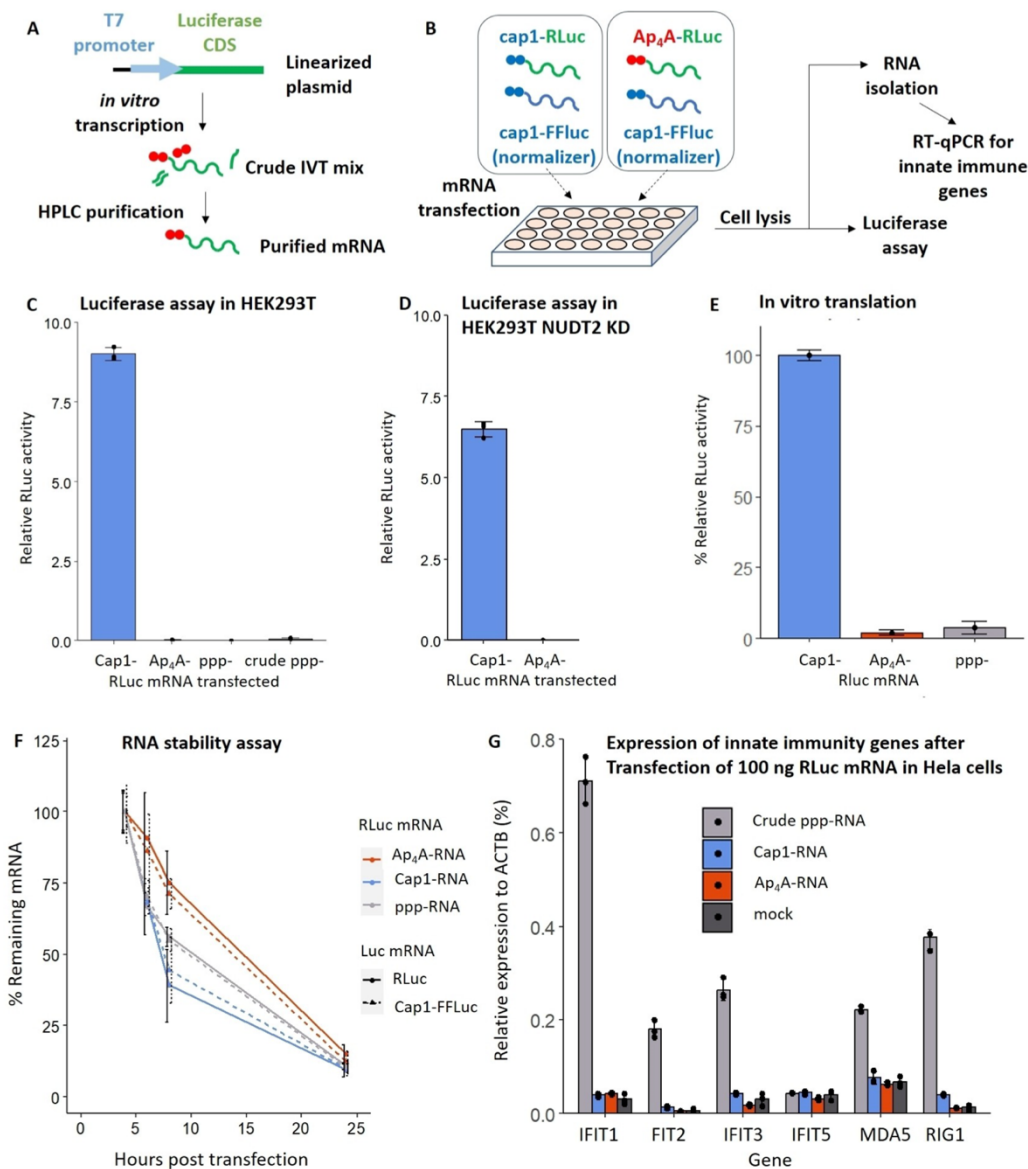


Figure 3. Testing Ap₄A-RNA translatability, cellular stability and its recognition by the innate immune system. A) Preparation of capped mRNA for transfection. B) Schematic representation of mRNA transfection and processing of the cells for luciferase assay and RT-qPCR. C) Relative RLuc activity normalized to FFLuc at 24 hours post transfection in HEK293T cells. D) Relative RLuc activity normalized to FFLuc at 24 hours post transfection in HEK293T NUDT2 KD cells. E) In vitro translation of FFLuc using rabbit reticulocytes. F) The amounts of transfected capped RLuc and FFLuc mRNA after 4, 6, 8, and 24 h post transfection. The expression values normalized to the housekeeping Actin Beta (ACTB) gene expression were recalculated to the mean amount detected at 4 h represented as 100%. G) Expression of antiviral response genes 6 hours post transfection with 100 ng of RLuc mRNA per one well in a 24-well plate in HeLa cells. The expression was assessed by RT-qPCR and normalized to the expression of the ACTB housekeeping gene. Mock transfection contained water instead of mRNA. The error bars in C–F represent the standard deviation. The experiments in C and D were measured in technical triplicate. The experiment in E was performed in a duplicate, measured in a technical duplicate. The experiments in F–G were performed in a biological triplicate, measured in a technical duplicate.

Interestingly, we observed minimal differences in the stability of cap1- and Ap₄A-RNA which suggests that the Ap₄A-RNA is not degraded and has the potential to interact with unknown intracellular partners.

The 5' termini of RNA are important for discerning between mature self RNA and immature or foreign (viral) RNA. Pattern recognition receptors like RIG-I, MDA5, or IFITs were reported to recognize the 5' ends of RNA.^[22] Ap₄A and NUDT2 are also implicated in innate immunity. We, therefore, measured the expression of nucleic acid sensing genes after Ap₄A-RNA transfection of HeLa cells. Interestingly, we did not observe any changes in these genes showing that Ap₄A-RNA is recognized as self by the cellular immune system.

In summary, we discovered a new type of RNA cap in mammalian cells and identified possible mechanisms of its regulation. The exact role of Ap₄A-RNA, however, remains unclear. Even though Ap₄A-RNA is not translated under any of the conditions tested, we cannot exclude alternative translational pathways triggered under stress, especially in the context of Np_nNs acting as alarmones. It can also mean that Ap₄A serves as a cap of long non-coding RNAs and has an entirely different function. We are currently developing a sequencing method for Ap₄A-RNA as well as a method to identify its protein interacting partners, both of which are necessary to shed light on the function of this newly discovered RNA modification.

Supporting Information

The authors have cited additional references within the Supporting Information, Refs. [22–27].

Acknowledgements

We would like to thank the laboratory of Dr. Petr Dráber for RBL-2H3 cells, Dr. Petr Špaček for the plasmid bearing the eIF4E coding sequence, prof. Joanna Kufel for the plasmid bearing the DXO coding sequence, Dr. Zuzana Keckesova for her help with plasmid design, Dr. Pawel J. Sikorski for his help with the immunological experiments. We also thank Zuzana Buchova for her help with the DXO experiments and Dr. Roberto Benoni for laying the foundations of the LC-MS method. Figure 1 and the TOC were created with BioRender.com We acknowledge funding from the European Research Council Executive Agency (ERCEA) under the European Union's Horizon Europe Framework Programme for Research and Innovation (grant agreement No 101041374—StressRNAction). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: Ap₄A · Ap₄A-RNA · Diadenosine Tetraphosphate · RNA Cap · RNA Modification

- [1] A. G. McLennan, *Pharmacol. Ther.* **2000**, *87*, 73–89.
- [2] a) M. Wright, N. Boonyalai, J. A. Tanner, A. D. Hindley, A. D. Miller, *FEBS J.* **2006**, *273*, 3534–3544; b) C. V. Dang, J. A. Traugh, *J. Biol. Chem.* **1989**, *264*, 5861–5865; c) R. T. Guo, Y. E. Chong, M. Guo, X. L. Yang, *J. Biol. Chem.* **2009**, *284*, 28968–28976; d) O. Goerlich, R. Foeckler, E. Holler, *Eur. J. Biochem.* **1982**, *126*, 135–142; e) H. Belrhali, A. Yaremchuk, M. Tuko, C. Berthet-Colominas, B. Rasmussen, P. Bösecke, O. Diat, S. Cusack, *Structure* **1995**, *3*, 341–352.
- [3] K. H. Götz, M. Mex, K. Stuber, F. Offensperger, M. Scheffner, A. Marx, *Cell Chem. Biol.* **2019**, *26*, 1535–1543.e1535.
- [4] A. G. McLennan, *Cell. Mol. Life Sci.* **2006**, *63*, 123–143.
- [5] M. Frese, P. Saumer, Y. Yuan, D. Herzog, D. Höpfner, A. Itzen, A. Marx, *Angew. Chem. Int. Ed.* **2023**, *62*, e202213279.
- [6] L. Krüger, C. J. Albrecht, H. K. Schammann, F. M. Stumpf, M. L. Niedermeier, Y. Yuan, K. Stuber, J. Wimmer, F. Stengel, M. Scheffner, A. Marx, *Nat. Commun.* **2021**, *12*, 5808.
- [7] J. Yu, Z. Liu, Y. Liang, F. Luo, J. Zhang, C. Tian, A. Motzik, M. Zheng, J. Kang, G. Zhong, C. Liu, P. Fang, M. Guo, E. Razin, J. Wang, *Nat. Commun.* **2019**, *10*, 4664.
- [8] N. Yannay-Cohen, I. Carmi-Levy, G. Kay, C. M. Yang, J. M. Han, D. M. Kemeny, S. Kim, H. Nechushtan, E. Razin, *Mol. Cell* **2009**, *34*, 603–611.
- [9] Y. Ofir-Birin, P. Fang, S. P. Bennett, H.-M. Zhang, J. Wang, I. Rachmin, R. Shapiro, J. Song, A. Dagan, J. Pozo, S. Kim, A. G. Marshall, P. Schimmel, X.-L. Yang, H. Nechushtan, E. Razin, M. Guo, *Mol. Cell* **2013**, *49*, 30–42.
- [10] O. Hudeček, R. Benoni, P. E. Reyes-Gutierrez, M. Culka, H. Šanderová, M. Hubálek, L. Rulíšek, J. Cvačka, L. Krásný, H. Cahová, *Nat. Commun.* **2020**, *11*, 1052.
- [11] R. Benoni, M. Culka, O. Hudeček, L. Gahurova, H. Cahová, *ACS Chem. Biol.* **2020**, *15*, 1765–1772.
- [12] D. J. Luciano, R. Levenson-Palmer, J. G. Belasco, *Mol. Cell* **2019**, *75*, 957–966.e958.
- [13] A. S. Marriott, O. Vasieva, Y. Fang, N. A. Copeland, A. G. McLennan, N. J. Jones, *PLoS One* **2016**, *11*, e0154674.
- [14] M.-B. Mititelu, O. Hudeček, A. Gozdek, R. Benoni, O. Nešuta, S. Krasnodebski, J. Kufel, H. Cahová, *RSC Chem. Biol.* **2023**, *4*, 223–228.
- [15] J. Carreras-Puigvert, M. Zitnik, A.-S. Jemth, M. Carter, J. E. Unterlass, B. Hallström, O. Loseva, Z. Karem, J. M. Calderón-Montaño, C. Lindskog, P.-H. Edqvist, D. J. Matuszewski, H. Ait Blal, R. P. A. Berntsson, M. Häggblad, U. Martens, M. Studham, B. Lundgren, C. Wahlby, E. L. L. Sonhammer, E. Lundberg, P. Stenmark, B. Zupan, T. Helleday, *Nat. Commun.* **2017**, *8*, 1541.
- [16] B. T. Laudenbach, K. Krey, Q. Emslander, L. L. Andersen, A. Reim, P. Scaturro, S. Mundigl, C. Dächert, K. Manske, M. Moser, J. Ludwig, D. Wohlleber, A. Kröger, M. Binder, A. Pichlmair, *Nat. Commun.* **2021**, *12*, 6918.

- [17] a) X. Jiao, S. K. Doamekpor, J. G. Bird, B. E. Nickels, L. Tong, R. P. Hart, M. Kiledjian, *Cell* **2017**, *168*, 1015–1027.e1010; b) B. Benoni, R. Benoni, J. Trylcova, K. Grab, J. Pačes, J. Weber, D. Stanek, J. Kowalska, L. Bednarova, Z. Keckesova, L. Gahurova, H. Cahova, *bioRxiv preprint* **2022**, 2022.2011.2010.515957.
- [18] A. Batool, S. Aashaq, K. I. Andrabi, *J. Cell. Biochem.* **2019**, *120*, 14201–14212.
- [19] P. J. Sikorski, M. Warminski, D. Kubacka, T. Ratajczak, D. Nowis, J. Kowalska, J. Jemielity, *Nucleic Acids Res.* **2020**, *48*, 1607–1626.
- [20] C. B. Ferreira, R. P. Sumner, M. T. Rodriguez-Plata, J. R. Saiyaah, R. S. Milne, A. J. Thrasher, W. Qasim, G. J. Towers, *Mol. Ther.-Methods Clin. Dev.* **2020**, *17*, 209–219.
- [21] J. J. D. Ho, S. Lee, *Trends Biochem. Sci.* **2016**, *41*, 821–823.
- [22] V. Hornung, J. Ellegast, S. Kim, K. Brzozka, A. Jung, H. Kato, H. Poeck, S. Akira, K.-K. Conzelmann, M. Schlee, S. Endres, G. Hartmann, *Science* **2006**, *314*, 994–997.

Manuscript received: October 5, 2023

Accepted manuscript online: November 7, 2023

Version of record online: November 22, 2023

