Univerzita Karlova

1. lékařská fakulta

Studijní program: Biomedicína

Studijní obor: Biochemie a patobiochemie

UNIVERZITA KARLOVA
1. lékařská fakulta

Ing. Barbora Benoni

Studium funkce NAD-RNA v HIV-1 infekci

Studies of NAD-RNA and its function in HIV-1 infection

Disertační práce

Školitelka: Ing. Hana Cahová, Ph.D.

Praha, 2024

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně, a že jsem řádně uvedla a citovala všechny použité prameny a literaturu. Současně prohlašuji, že práce nebyla využita k získání jiného nebo stejného titulu.

Souhlasím s trvalým uložením elektronické verze mé práce v databázi systému meziuniverzitního projektu Theses.cz za účelem soustavné kontroly podobnosti kvalifikačních prací.

V Praze, 25.06.2024

Barbora Benoni

Identifikační záznam

BENONI, Barbora. *Studium funkce NAD-RNA v HIV-1 infekci. [Studies of NAD-RNA and its function in HIV-1 infection.]* Praha, 2024. 145 stran, 3 přílohy. Disertační práce. Univerzita Karlova, 1. lékařská fakulta, Ústav organické chemie a biochemie AV ČR, v.v.i.. Vedoucí závěrečné práce Ing. Hana Cahová, Ph.D.

Acknowledgments

I would like to thank all the people who supported me through this journey. My heartfelt thanks go to my family, my friends, and my colleagues. And of course, to my supervisor Ing. Hana Macíčková Cahová, Ph.D. for the possibility of working on this interesting project and her invaluable advices.

I would like to thank all the labmembers that contributed to the project realisation. Especially, I would like to thank M.Sc. Jiří František Potužník, M.Sc. Anton Škríba, Ph.D., and Ing. Pavel Vopálenský, Ph.D. for helping me with my experiments in the time of my absence. My sincere thanks also go to RNDr. Jana Trylčová, Ph.D. and Ing. Kristina Spustová, awesome laboratory assistants. RNDr. Jana Trylčová, Ph.D. was a great help and teacher in the BSL3 laboratory. I am grateful to Maria-Bianca Mititelu, M.Sc. and Zuzana Buchová for providing me with the DXO enzyme. I would like to thank Roberto Benoni, Ph.D. and M.Sc. Martin Hubálek, Ph.D. for teaching me a great deal about LC-MS methods. I greatly appreciate the advice of the members of Jan Weber Group and Josef Cvačka Group, their help has been invaluable to me. I would like to thank M. Sc. Katarzyna Grab for the TMG cap synthesis and Valentina Cutano, Ph.D. for her guidance in the western blot analysis.

I would also like to thank the experts who were involved in the validation survey for this dissertation thesis M.Sc. Lenka Gahurová, Ph.D., M.Sc. Lucie Bednárová, Ph.D., prof. M.Sc. David Staněk, Ph.D., and M.Sc. Zuzana Kečkéšová, Ph.D. This work would not be possible without the excellent bioinformatics data analyses of M.Sc. Lenka Gahurová, Ph.D. and meticulous CD spectra measurement of Lucie Bednárová, Ph.D.

My biggest thanks goes to my mum *Bohumila Svojanovská* for her continuous encouragement throughout the years of study and writing this thesis.

A special thanks goes to 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) for providing the funding for the work.

Computational resources were supplied by the project "e-Infrastruktura CZ" (eINFRA CZ LM2018140) supported by the Ministry of Education, Youth and Sports (Czech Republic). Data are accessible in the GEO database under accession GSE191019.

Abstrakt

Viry se vyznačují jednoduchou strukturou a jsou dobře popsány ve vědecké literatuře. Na základě studií provedených na virech bylo učiněno mnoho objevů týkajících se RNA modifikací.

HIV-1 infekce snižuje množství buněčného nikotinamidadenindinukleotidu (NAD). Nedávné studie ukázaly, že NAD může sloužit jako 5' nekanonická čepička některých RNA u bakterií, kvasinek, rostlin a savců. NAD ovlivňuje RNA stabilitu a účinnost RNA translace. O funkci NAD čepičky se však ví stále málo. Tato práce se zabývá NAD čepičkou v souvislosti s HIV-1 infekcí, přičemž v rámci jiné práce jsme studovali m^1A modifikaci v HIV-1.

Zjistili jsme, že HIV-1 infekce ovlivňuje nejen buněčné hladiny NAD, ale také množství NAD čepiček na sRNA. Pomocí NAD captureSeq jsme identifikovali čtyři snRNA (U1, U4ATAC, U5E a U7) a čtyři snoRNA (SNORD3G, SNORD102, SNORA50A a SNORD3B), které po HIV-1 infekci ztrácejí NAD čepičky. Zvláště zajímavá je U1 snRNA, která má sekvenci komplementární s HIV-1 pre-mRNA a váže se na ni při sestřihu. Zjistili jsme, že NAD čepička destabilizuje komplex HIV-1 pre-mRNA a U1 snRNA.

DXO je enzym odstraňující NAD čepičku. Připravili jsme buňky se zvýšenou produkcí DXO a sledovali jsme množství NAD-RNA ve spojitosti s infektivitou HIV-1. Zvýšená produkce DXO způsobuje snížené množství NAD-RNA a zvýšenou infektivitu HIV-1. Naopak zvýšení množství buněčného NAD vede ke snížené infektivitě HIV-1.

Tato práce identifikuje nové sRNA s NAD čepičkou v lidských buňkách a navrhuje, že NAD-RNA snižuje infektivitu HIV-1 a může hrát roli v antivirové obraně.

Klíčová slova: RNA modifikace, NAD, U1 snRNA, RNA-seq, LC-MS, DXO, HIV-1

Abstract

Viruses have a simple structure and are well described in scientific literature. Based on studies conducted on viruses, many discoveries regarding RNA modifications have been made.

HIV-1 infection reduces the amount of cellular nicotinamide adenine dinucleotide (NAD). Recent studies have shown that NAD can serve as a 5' non-canonical cap for some RNAs in bacteria, yeast, plants, and mammals. NAD capping affects RNA stability and the efficiency of RNA translation. However, surprisingly little is known about the function of NAD cap. This work focuses on NAD capping in the context of HIV-1 infection, while in another study, we also investigated the m^1A modification in HIV-1.

We found that HIV-1 infection affects not only cellular levels of NAD but also the amount of NAD caps on sRNA. Using NAD captureSeq, we identified four snRNAs (U1, U4ATAC, U5E, and U7) and four snoRNAs (SNORD3G, SNORD102, SNORA50A, and SNORD3B) that lose their NAD caps after HIV-1 infection. Particularly interesting is U1 snRNA, which has a sequence complementary to HIV-1 pre-mRNA and binds to it during splicing. We discovered that the NAD cap destabilizes the complex between HIV-1 pre-mRNA and U1 snRNA.

DXO is an NAD decapping enzyme. We prepared cells with DXO overexpression and examined the amount of NAD-RNA in connection with HIV-1 infectivity. DXO overexpression causes reduced amounts of NAD-RNA and increased HIV-1 infectivity. In contrast, repletion of cellular NAD leads to decreased HIV-1 infectivity.

This work identifies new sRNAs with NAD caps in human cells and proposes that NAD-RNA decreases HIV-1 infectivity and may play a role in antiviral defense.

Key words: RNA modifications, NAD, U1 snRNA, RNA-seq, LC-MS, DXO, HIV-1

Acronyms

5' SS	5' splice-site
7SL RNA	Signal recognition particle RNA
ADAR	Adenosine deaminase acting on RNA
ADPRC	ADP-ribosyl cyclase
Ago2	Argonaute 2
AIDS	Acquired Immune Deficiency Syndrome
ALKBH5	AlkB homolog 5
APB	Acryloylaminophenylboronic acid
ARE	AU-rich instability element
AtNUDX23	Nudix pyrophosphohydrolase 23
BSL3	Biosafety Level 3
CBC	Cap-binding complex
CD38	Cyclic ADP ribose hydrolase
circRNA	Circular RNA
CiRNA	Intronic circRNA
ciRS-7	Circular RNA sponge for miR-7
CMTr1, 2	Cap methyltransferase 1, 2
CTD	C-terminal domain
CuAAC	Copper-catalyzed azide-alkyne cycloaddition
D	Dihydrouridine
DMSO	Dimethyl sulfoxide
dpCoA	Dephospho-coenzyme A
DXO	Decapping exoribonuclease
DTT	Dithiothreitol
EC	Elongation complex
EciRNA	Exonic circRNA
EIciRNA	Exon-intron circRNA
eIF3	Eukaryotic initiation factor 3
ESI	Electrospray ionisation
FAD	Flavin adenine dinucleotide
FHIT	Fragile histidine triad protein
FTO	Fat mass and obesity-associated protein
GTase	RNA guanylyltransferase
HCV	Hepatitis C virus
HIT	Histidine triad
HIV-1	Human immunodeficiency virus 1
HRM	High-resolution melting
I	Inosine
KD	Knock-down
lncRNA	Long non-coding RNA
LTR	Long terminal repeat
m^1A	1-methyladenosine
m^5C	5-methylcytosine
m^6A	N^6-methyladenosine
m^6Am	N^6,2'-O-dimethyladenosine
m^7G	7-methylguanosine
METTL3	Methyltransferase-like 3

mRNA	Messenger RNA
miRNA	Micro RNA
mtRNA	Mitochondrial RNA
N7MTase	Guanine-N7 methyltransferase
NAD	Nicotinamide adenine dinucleotide
NA	Niacin
NCIN	Non-canonical initiating nucleotide
ncRNA	Non-coding RNA
Nm	2'-O-methylation
NMD	Nonsense-mediated decay
NMN	Nicotinamide mononucleotide
Np_nN	Dinucleoside polyphosphate
NR	Nicotinamide riboside
OE	Overexpression
PABP	Poly(A) RNA-binding protein
PAP	Poly(A) polymerase
PARP	Poly (ADP-ribose) polymerase
PBS	Primer binding site
Pi	Inorganic phosphate
PIC	Preinitiation complex
piRISC	piRNA-induced silencing complex
piRNA	PIWI-interacting RNA
PPi	Pyrophosphate
pRNA, ppRNA, pppRNA	Mono-, Di-, Triphosphate RNA
PTGS	Post-transcriptional gene silencing
r.t.	Room temperature
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNA pol	RNA polymerase
RNAi	RNA interference
ROS	Reactive oxygen species
RppH	RNA 5'-pyrophosphohydrolase
rRNA	Ribosomal RNA
RRE	Rev response element
RT	Reverse transcriptase
RTPase	RNA triphosphatase
SAH	S-adenosyl-L-homocysteine
SAM	S-adenosyl-L-methionine
scaRNA	Small Cajal body-specific RNA
SD	Standard deviation
SINE	Short Interspersed Nuclear Element
SIRLOIN	SINE-derived nuclear RNA localization
siRNA	Small interfering RNA
Sirt	Sirtuin
SNORA	H/ACA box snoRNA
SNORD	C/D box snoRNA
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
snRNP	Small nuclear ribonucleoprotein

SRP	Signal recognition particle
TAR	Transactivation response
Tat	Trans-activator of transcription
TBP	TATA-binding protein
$TCID_{50}$	50% tissue culture infectious dose
TdT	Terminal deoxynucleotidyl transferase
TE	Transposable element
TI	Threshold of infectivity
Tm	Melting temperature
TMG	^{2,2,7}trimethylguanosine
TRBP	Transactivation response element RNA-binding protein
tRNA	Transfer RNA
TSS	Transcription start site
UBF	Upstream binding factor
UDP-Glc	Uridine diphosphate glucose
UDP-GlcNAc	Uridine diphosphate N-acetylglucosamine
WT	Wild-type
XTT	Yellow tetrazolium salt
Ψ	Pseudouridine

Contents

1 Introduction

To date, more than 170 different types of RNA modifications with a wide variety of chemical properties have been identified. These modifications are present in all types of RNA molecules, including transfer RNA (tRNA), ribosomal RNA (rRNA), messenger RNA (mRNA), and various non-coding RNAs (ncRNAs). Our understanding of RNA modifications and their functions is broadening due to the increasing variety of capturing techniques and the number of studied organisms (Boccaletto et al., 2022). However, little is known about the function of RNA modifications in the context of host-pathogen interactions. Several studies have shown that Human immunodeficiency virus 1 (HIV-1) uses RNA modifications to regulate its gene expression during infection and replication (Fukuda et al., 2021; Ringeard et al., 2019). For example, N^1-methyladenosine (m^1A) is known to affect RNA stability and translation, contributing to the regulation of viral gene expression and replication (Burnett and McHenry, 1997).

HIV-1 infection causes reduced amounts of free nicotinamide adenine dinucleotide (NAD) in human cells (Murray et al., 1995). NAD plays a vital role in redox reactions as a co-substrate for various enzymes. The relatively recent discovery that NAD can serve as a 5' non-canonical RNA cap has opened a new and exciting area of RNA biology research (Chen et al., 2009). Several methods have been developed for capturing and sequencing NAD-capped RNAs (NAD-RNAs) providing information on the types and possible roles of NAD-RNAs. While in bacteria, the NAD cap protects mRNA from degradation, in mammals and plants, it may promote mRNA decay. In plants, NAD-mRNAs were detected in polysomal fractions, suggesting that they are translated, while NAD-mRNAs in mammals do not appear to have the translation capacity (Cahová et al., 2015; Hu et al., 2021; Zhang et al., 2019). Despite extensive research on NAD-RNAs in various organisms, relatively little is known about the function of NAD on specific RNAs and in particular processes. However, NAD cap was also detected in ncRNAs, suggesting specialized roles that remain elusive.

Since HIV-1 is a well-studied and characterized retrovirus, we focused our research on the physiological responses of human cells to HIV-1 infection. More specifically, our objective was to identify NAD-RNAs affected by the infection. As part of another project, we studied m^1A modification in HIV-1 particles. The literature review summarizes types of RNAs, their processing, and recent discoveries in the field of RNA modifications. The main focus is on NAD-RNAs and the biological processes connected to this modification.

1.1 RNA types

The central dogma of molecular biology describes how the information contained in the DNA is transcribed into mRNA and transported from the nucleus to the cytoplasm. There, tRNA and rRNA helps to translate it into a sequence of amino acids (Diercks et al., 2021).

Until recently, RNA was thought to serve only as a messenger between DNA and proteins. However, RNA has various functions, including driving chemical reactions, carrying genetic information, regulating gene expression, RNA splicing, and modifying other RNA types (Diercks et al., 2021). These functions are carried out by non-coding RNAs (ncRNAs) such as microRNAs (miRNAs) and small interfering RNAs (siRNAs), which inhibit mRNA translation by binding to specific sites, small nuclear RNAs (snRNAs) involved in mRNA splicing, small nucleolar RNAs (snoRNAs) involved in rRNA modification (Mattick and Makunin, 2006), PIWI-interacting RNAs (piRNAs) that silence transposable elements (Siomi et al., 2011), circular RNAs (circRNAs) that act as molecular sponges (Hansen et al., 2013), signal recognition particle RNA (7SL RNA) involved in protein localization (Gussakovsky and McKenna, 2021), and long non-coding RNAs (lncRNAs), which function as nuclear organization factors (Rinn and Guttman, 2014) (Fig. 1).

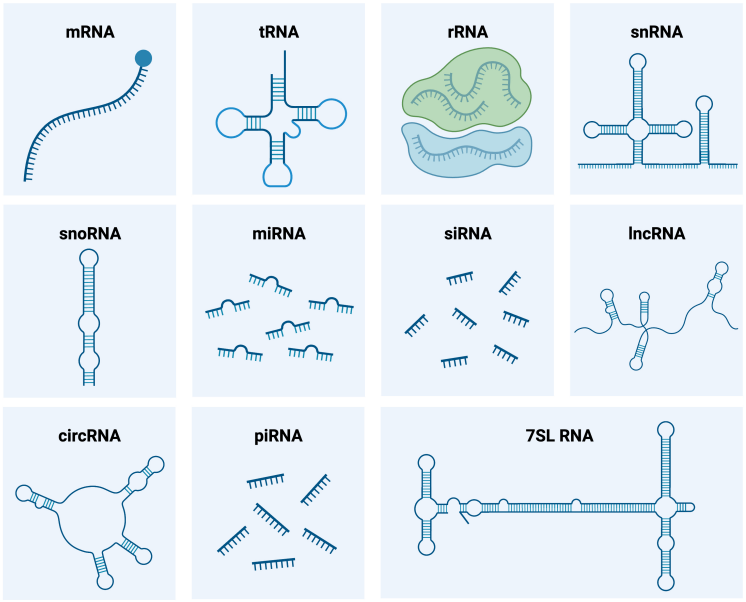


Fig. 1: Overview of the types of RNA discussed in the text and their 2D structures.

1.1.1 mRNA

mRNAs are the most heterogeneous type of RNA in terms of sequence and size, ranging from hundreds to several thousands of nucleotides. The mRNA content of a cell depends on its physiological stage and cell type (Valleriani et al., 2011; Wang et al., 2013). The NCBI Reference Sequence project and the NIH Mammalian Gene Collection have identified around 30,000 non-redundant human mRNA sequences, with an average of 10-15 molecules of each species per cell (Furey et al., 2004).

mRNAs undergo post-transcriptional modifications, primarily through methyl group additions: (A) on the nucleobase: N^6-methyladenosine (m^6A) and 5-methylcytosine (m^5C), (B) to the ribose: 2'-O-methylation (Nm), and (C) both: N^6,2'-O-dimethyladenosine (m^6Am). Other modifications include pseudouridine (Ψ) and inosine (I) (Boccaletto et al., 2022) (Fig. 2).

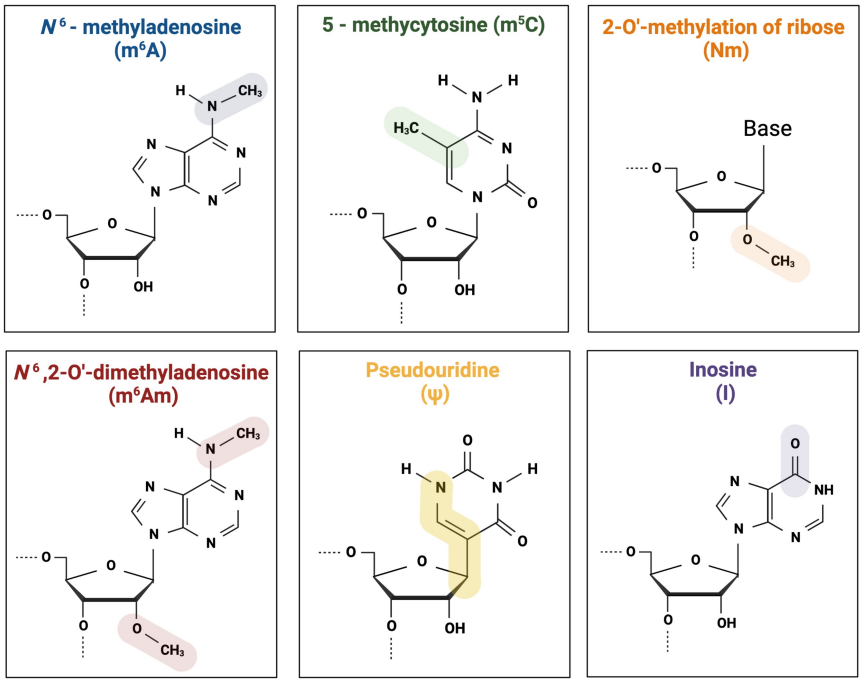


Fig. 2: Structures of some mRNA modifications. The sites of modifications are highlighted in colour.

1.1.2 rRNA

Most RNA molecules in cells are rRNAs (around 80%), which, along with proteins, form ribosomes. In eukaryotes, one ribosome consists of a small subunit (40S) and a large subunit (60S). During transcription by RNA pol I, pre-rRNA is folded, nucleolytically processed, and modified. rRNAs contain a large number of Ψ modifications clustered in

functionally important regions (Baßler and Hurt, 2019; Ni et al., 1997)

The small subunit contains 18S rRNA bound to 33 ribosomal proteins, while the large subunit is composed of 5S, 5.8S, and 25S/28S rRNAs associated with 46 ribosomal proteins. Each type of rRNA is represented in one copy per ribosome (Baßler and Hurt, 2019).

1.1.3 tRNA

The primary role of tRNAs is to transfer amino acids during protein synthesis. However, specific tRNAs also participate in other cellular functions, such as priming DNA synthesis in retroviruses, plant pararetroviruses, and retrotransposons. Canonical tRNA molecules are around 76 nucleotides long and fold into a cloverleaf-like secondary structure (Marquet et al., 1995). This structure consists of four arms: the acceptor stem, dihydrouridine (D) loop, anticodon loop, and TΨC loop. The anticodon triplet in the anticodon stem-loop recognizes the codon in the mRNA template. The tertiary structure of tRNA is L-shaped, stabilized by interactions between loops and additional modifications. Mitochondrial tRNAs are shorter, ranging from 59 to 75 nucleotides (Schimmel, 2018; Suzuki et al., 2011).

In eukaryotes, tRNA maturation involves several steps (Fig. 3). Pre-tRNAs, transcribed by RNA pol III, have extra sequences at their 5' and 3' ends. RNase P endonuclease removes the 5' leader sequence, while endonucleases and exonucleases remove the 3' trailer sequence. tRNA nucleotidyl transferase then adds the CCA sequence to the 3' end. Intron sequences within the pre-tRNAs are spliced out by endonucleases, and the ends are joined by ligase (Hopper and Phizicky, 2003). tRNAs undergo numerous posttranscriptional modifications (Fig. 4), helping to form mature tRNAs and fold into their unique tertiary structures (Pan, 2018).

If tRNAs lack some modifications in various stages of their maturation, they are degraded through a system called rapid tRNA decay (Alexandrov et al., 2006). In yeast, the decay system interacts with the translation machinery and acts on tRNAs lacking modifications (e.g. m^5C and $m^{2,2}G$ at position 26) (Dewe et al., 2012). In *Saccharomyces cerevisiae*, 5' 7-methylguanosine (m^7G) cap structures stabilize pre-tRNAs. Inhibition of RNase P leads to accumulation of capped pre-tRNAs, while depletion of the capping enzyme Ceg1p reduces their levels. Inhibition of 5' exonucleases also results in more capped pre-tRNAs (Ohira and Suzuki, 2016).

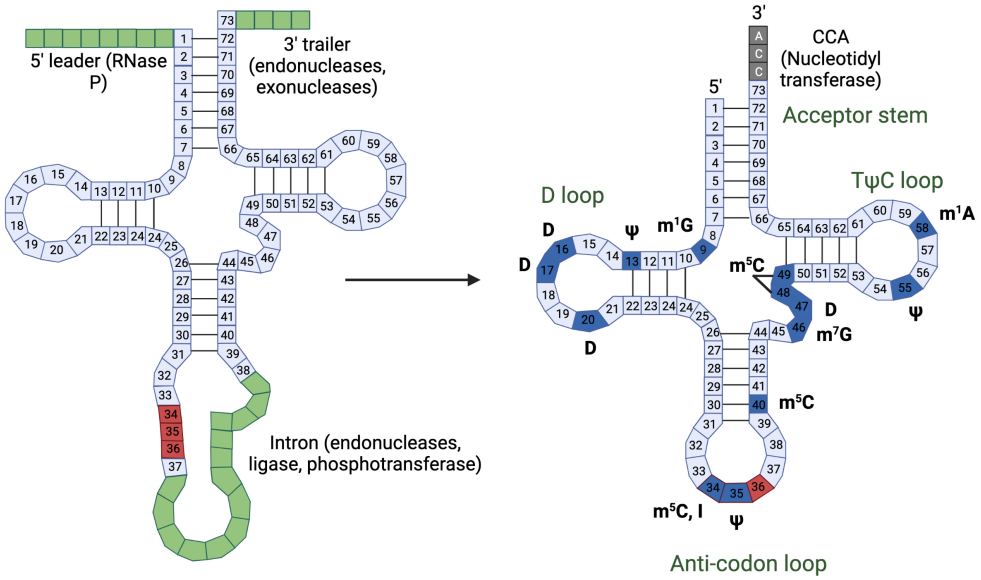


Fig. 3: A schematic secondary structure of a pre-tRNA and a mature tRNA with the positions of several modifications, using the usual tRNA numbering system. Nucleotides part of the mature tRNA are light blue, parts removed during the maturation are green, and the anticodon is red. The names of the four arms are highlighted in green, and the positions of the tRNA modifications discussed in the text are in dark blue.

tRNA modifications assist in retroviral RNA annealing or encapsidation, and are important for interaction with aminoacyl-tRNA synthetase (Giegé, 2008). Misregulated m^7G_{46} methylation causes growth deficiency in yeast and is linked to microcephalic primordial dwarfism in humans (Shaheen et al., 2015). Mutations in TRMT10A, affecting m^1G_9, are associated with microcephaly, intellectual disability, and young adult-onset diabetes (Narayanan et al., 2015; Yew et al., 2016). D at positions 16 and 17 promotes less stable base-pairing than uridine (Hiley et al., 2005). m^5C affects codon-anticodon interactions and enhances tRNA stability (Agris, 1996). I_{34} in the anticodon can pair with U, C, or A, expanding tRNA's reading capacity (Grosjean et al., 2010) and Ψ was shown to affect RNA structure and stability (Newby and Greenbaum, 2002).

m^1A_{58} has a critical role in maturation and stability of tRNA methionine (Anderson et al., 2000). Furthermore, the m^1A_{58} modification serves as a stop signal in retroviral DNA production. Retroviruses use the first 18 nucleotides of tRNA, counting from the 3' end, as primers for replication. Because the bulky m^1A_{58} modification is located at the position 19 from the 3' end, it halts synthesis after 18 nucleotides. This termination triggers strand transfer and subsequent synthesis of the entire retroviral DNA (Burnett

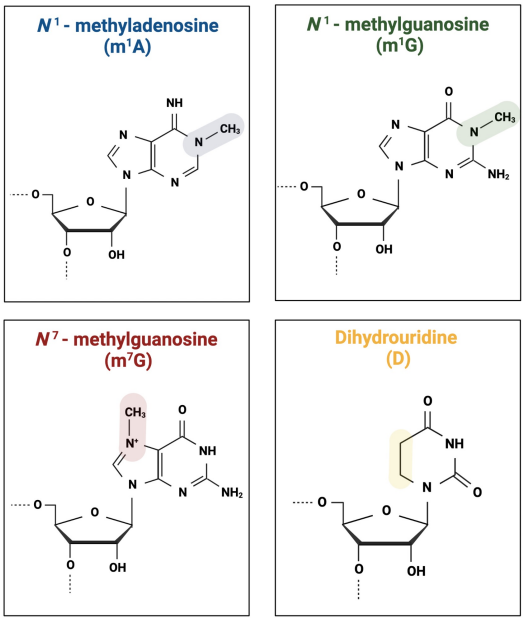


Fig. 4: Structures of some tRNA modifications. The sites of modifications are highlighted in colour.

and McHenry, 1997).

The retroviral reverse transcriptase requires a primer to initiate DNA synthesis, and the most widely used tRNAs are $tRNA^{Trp}$, $tRNA^{Pro}$, $tRNA_{1,2}{}^{Lys}$, $tRNA_3{}^{Lys}$, $tRNA_1{}^{Met}$ (Dahlberg, 1980; Marquet et al., 1995). Different retroviruses use various tRNAs as primers, with no direct link between the primer tRNA and the host or the disease type. In all retroviruses, the primer tRNA's 3' end binds to the primer binding site (PBS) on the viral RNA, located downstream of the U5 region. For example, HIV-1 selectively co-packs $tRNA_3{}^{Lys}$, whose 3' terminal 18 nucleotides anneal to the PBS (Isel et al., 1993; Ratner et al., 1985). Similarly, Rous sarcoma virus uses $tRNA^{Trp}$ as a primer (Harada et al., 1975), and Moloney murine leukemia virus uses $tRNA^{Pro}$ (Harada et al., 1979).

1.1.4 snRNA

snRNAs are 150 nucleotides long, uridine-rich RNAs. They are highly abundant in the nucleus and primarily involved in intron splicing. The main snRNAs (U1, U2, U4, U5, and U6), along with several Sm proteins and other snRNA-specific proteins, form small nuclear ribonucleoprotein (snRNP) complexes that facilitate pre-mRNA splicing. These complexes recognize specific intron sequences through base-pairing interactions, excise the introns, and ligate the exons together (Karijolich and Yu, 2010; Maniatis and Reed, 1987).

In addition to the major snRNAs, there are minor spliceosomal snRNA species (U11, U12, U4atac, and U6atac) that splice a minor class of introns (Will and Lührmann, 2005). Besides splicing, snRNPs assist on regulation of gene expression, miRNAs synthesis, and contribute to catalytic reactions. During maturation, snRNAs acquire various modifications that affect their structure, stability, and interactions with other RNAs and proteins. The most abundant internal modifications are Ψ and Nm (Karijolich and Yu, 2010).

U1 snRNA

U1 snRNA recognizes 5' SS of exons through complementary base-pairing. The 5'-terminal sequence of U1 snRNA is highly conserved across species (Zhuang and Weiner, 1986). During transcription, U1 snRNA acquires an m^7G cap and is exported to the cytoplasm, where it is hypermethylated to ^{2,2,7}trimethylguanosine (TMG cap). In the cytoplasm, it associates with Sm proteins, and is imported back to the nucleus (Hamm et al., 1990; Urlaub et al., 2001). The U1 snRNP complex includes U1 snRNA, seven Sm proteins, and three U1-specific proteins (U1 70K, U1 A, and U1 C, Fig. 5).

The human U1 snRNA 5' region is densely modified with two Nm (Am_1 and Um_2) and two conserved Ψ (Ψ_5 and Ψ_6) (Karijolich and Yu, 2010). Thermodynamic studies of Ψ in RNA duplexes show it neither destabilizes nor stabilizes double-strand formation (Hall and McLaughlin, 1991). Human 5' SS has three classes based on U1 snRNA base-pairing strength: strong, intermediate, and weak. Ψ minimally affects strong 5' SS but may influence weaker ones (Roca et al., 2005).

In HIV-1, U1 snRNP prevents polyadenylation at the 5' LTR (long terminal repeats), crucial for viral gene expression, by interacting with the major splice donor site (Ashe et al., 1997). This interaction likely involves U1 snRNP and cap-binding proteins, blocking access to the poly(A) site (Lewis et al., 1996). U1 70K protein alone is insufficient for this inhibition, but mutations in U1 stem-loop 1 (the binding site for U1 70K protein) suppress HIV-1 polyadenylation (Vagner et al., 2000). U1 snRNA also binds to the 5' splice site in env mRNA, protecting it from degradation and over-splicing. It is crucial for HIV-1 that most of its transcripts remain unspliced or partially spliced (Kammler et al., 2001; Lu et al., 1990).

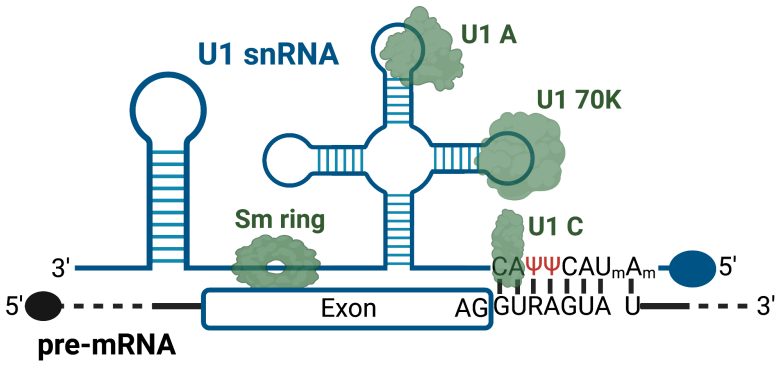


Fig. 5: Secondary structure of human U1 snRNA with snRNA-pre-mRNA interactions shown. Three proteins specific to U1 (U1 70K, U1 A and U1 C) and Sm ring are shown in green, U1-Ψ5 and U1-Ψ6 are highlighted in red. Filled circles depict the cap structures. Remarkably, many modified nucleotides of U1 snRNA lie within the sequence that base-pairs with pre-mRNA.

U2 snRNA

U2 snRNA is the most modified snRNA, containing ten Nm residues, thirteen Ψ, and a TMG cap. In HeLa cells, modifications within the first 20 nucleotides are necessary for interaction of U2 with U1 snRNP. Nm modifications at positions 1, 2, 12, and 19 are all required for efficient splicing, whereas the TMG cap is not essential for pre-mRNA splicing (Dönmez et al., 2004).

U4, U5, U6 snRNA

Both U4 snRNA and U5 snRNA have TMG cap, while U6 snRNA contains a cap structure with monomethylated gamma phosphate (Singh and Reddy, 1989). Despite lacking TMG cap, U6 snRNA is found in immune precipitates with other snRNAs when anti-TMG IgG is used (Bringmann et al., 1984). This is due to extensive intermolecular sequence complementarity between U4 and U6 snRNAs leading to their coexistence in a snRNP particle (Hashimoto and Steitz, 1984).

1.1.5 snoRNA

snoRNAs are noncoding RNAs, 60-300 nucleotides long, typically transcribed from intronic regions by RNA pol II. There are two main classes: C/D box snoRNAs (SNORDs), involved in ribose methylation of rRNA, and H/ACA snoRNAs (SNORAs), responsible for pseudouridylation (Esteller, 2011). A third class, small Cajal body-specific RNAs (scaRNAs), guides site-specific methylation and pseudouridylation in U1, U2, U4, and U5 snRNAs (Darzacq et al., 2002).

SNORDs have two conserved motifs, box C (RUGAUGA) near the 5' end and box D (CUGA) near the 3' end, along with less conserved C' and D' boxes (Fig. 6). They recognize target RNA via a 9-21 nt guide region adjacent to box D or D', which is complementary to the substrate region modified by the methyltransferase fibrillarin (Samarsky et al., 1998). SNORAs contain an ACA box at the 3' end and two hairpin structures connected by a hinge region with ANANNA motif (box H). Each hairpin has a pseudouridylation pocket complementary to the target RNA, positioning the substrate uridine 14-16 nt upstream from the H or ACA box for pseudouridylation by dyskerin (Ganot et al., 1997; Kiss, 2001).

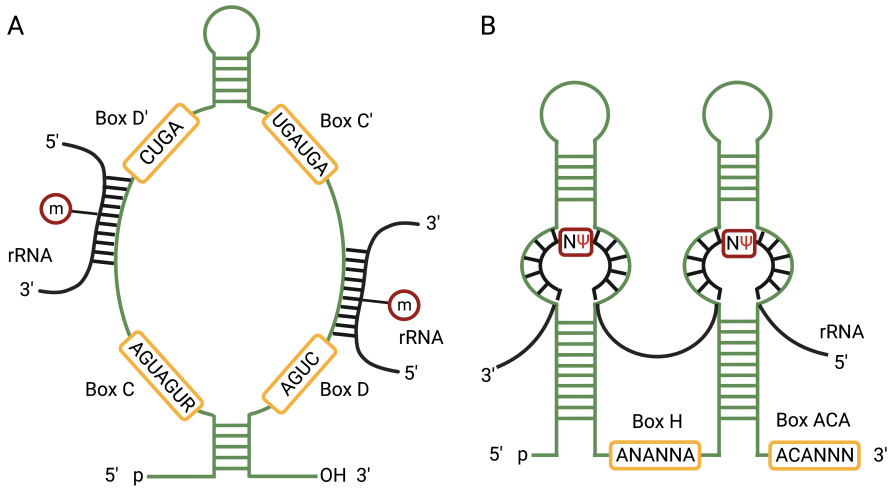


Fig. 6: Structures of the box C/D and box H/ACA snoRNAs. A) The C/D box snoRNA structural features: the C, C', D and D' sequences are highlighted by yellow boxes, upstream of the box D/D' are the guide sequences complementary to target rRNAs, the methylation (5 nt upstream of box D/D') is depicted by red circle. B) The H/ACA snoRNA structural features: the H and ACA sequences are highlighted by yellow boxes, two pseudouridylation pockets interact with the site of pseudouridylation (NΨ highlighted in red) of the target rRNA.

1.1.6 miRNA

miRNAs are endogenous 21-22 nt long RNAs that act in eukaryotes through RNA interference (RNAi). They can block the mRNA from being translated, or accelerate its degradation. miRNAs bind to 3' UTR of mRNA with partially mismatched sequence complementarity (Lee et al., 1993; Reinhart et al., 2000).

The biogenesis of miRNA (Fig. 7) starts from a pri-miRNA precursor, which is processed in the cell's nucleus into an imperfect shorter stem-loop structure (pre-miRNA) by the nuclear RNase III Drosha (Lee et al., 2003). The pre-miRNA is then exported into the cytoplasm by Exportin 5 (Yi et al., 2003). In the cytoplasm, the

loop region of the pre-miRNA is cleaved by RNase Dicer and RNA-binding cofactor TRBP (transactivation response element RNA-binding protein) into 18-25 nt miRNA duplexes (Bernstein et al., 2001). Only one of these two strands is then incorporated into the RNA-induced silencing complex (RISC) (Martinez et al., 2002), which contains a protein called Argonaute 2 (Meister et al., 2004). Such an active complex guided by the miRNA can then silence gene expression, and since the complementarity of miRNA is not perfect, one complex can downregulate several different transcripts (Lim et al., 2005).

miRNAs regulate viral infections and antiviral responses. Five miRNAs, miR-28, miR-125b, miR-150, miR-223, and miR-382, are enriched in resting memory CD4+ T-cells, which are less permissive to infection than activated CD4+ T-cells. These miRNAs target several HIV-1 mRNAs, inhibiting HIV-1 production (Huang et al., 2007). Conversely, some miRNAs enhance HIV-1 replication; miR-132 is upregulated following CD4+ T cell activation and increases viral replication in the Jurkat CD4+ T cell line (Chiang et al., 2013). miRNAs can be derived from both, coding and non-coding regions of the viral genome (Kaul et al., 2009).

1.1.7 siRNA

siRNAs are 20-25 nt long RNAs mostly produced by the breakdown of viral RNA as a part of antiviral defense. miRNAs and siRNAs were discovered independently but their biogenesis, assembly into RNA-protein complexes, and ability to regulate gene transcripts negatively in diverse eukaryotes are closely related. Unlike miRNAs, siRNAs degrade target mRNA completely, and this gene silencing requires perfect sequence complementarity (Fig. 7). siRNAs were first discovered in plants and named as the "post-transcriptional gene silencing" (PTGS) (Matzke and Matzke, 1995).

RNAi, first described in Caenorhabditis elegans, has become a tool for inhibiting cellular and viral gene expression (Fire et al., 1998). siRNA can inhibit HIV-1 production by targeting mRNAs for the viral Gag protein, showing potential for therapeutic intervention (Novina et al., 2002). Studies targeting different steps of HIV-1 replication with siRNA have reported positive results (Capodici et al., 2002; Coburn and Cullen, 2002). Since siRNA effects are transient, another system was developed for the therapeutic use of RNAi, which is short hairpin RNA delivered by lentiviral vector (Lee et al., 2005). Despite this, HIV-1 can develop resistance, as seen with siRNA targeting

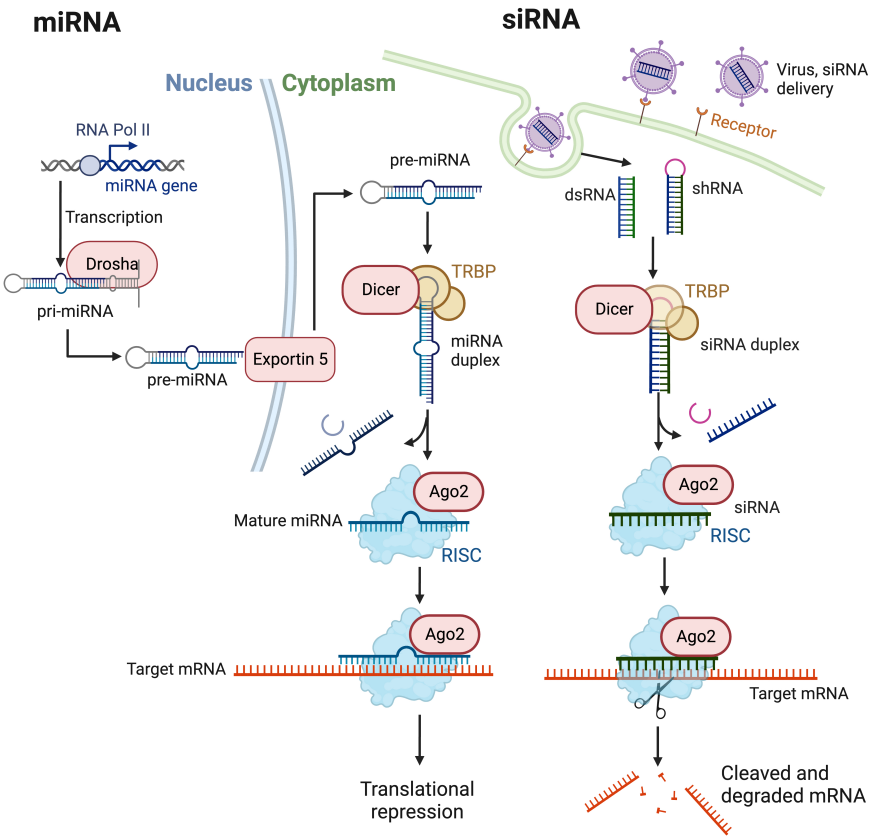


Fig. 7: Comparison of the miRNA and siRNA pathways. miRNAs are derived from large precursor pri-miRNAs transcribed by RNA pol II. In the cytoplasm, both pre-miRNAs and viral-derived dsRNAs share a common processing pathway, which involves Dicer/TRBP and RISC. The loop regions of pre-miRNA and the dsRNAs/short hairpins are trimmed by Dicer to generate miRNA and siRNA duplexes. One strand of the duplexes is incorporated into the RISC, while the other is targeted for degradation. Complex guided by the miRNA represses the translation of target mRNA, while transcripts complementary to siRNAs are targeted to degradation. In both cases, the gene expression is silenced.

the Nef gene, which led to sequence modifications. Additionally, HIV-1 encodes viral siRNA precursors that restrict antiviral activity in human cells, with the Tat protein partially suppressing Dicer's function (Bennasser et al., 2005; Das et al., 2004).

1.1.8 circRNA

circRNAs are non-coding RNAs that form covalently closed loop structures. They were first discovered as viroids, which are uncoated infectious RNA molecules pathogenic to some higher plants (Sanger et al., 1976). Since then, only a handful of other circRNAs were discovered and they were thought to be by-products of spliceosome-mediated splicing errors (Cocquerelle et al., 1993). Only recently, thanks to the improved deep sequencing technologies and bioinformatics, they were discovered also in mammalian

cells and established as a special novel type of endogenous non-coding RNA (Salzman et al., 2012).

CircRNAs are divided into three categories based on their source: exonic circRNAs (EciRNAs), exon-intron circRNAs (EIciRNAs), and intronic circRNAs (CiRNAs). Most circRNAs are formed by back-splicing, a special type of alternative splicing. There are two proposed pathways of circRNA formation from exons, where the main difference is the order of splicing steps (Fig. 8) (Barrett et al., 2015; Ma et al., 2020; Zhang et al., 2014). CiRNAs are generated from intronic lariat precursors that are processed by exonucleases (Zhang et al., 2013).

Since circRNAs possess no free ends, they resist miRNA-mediated exonucleolytic degradation and can act as sponges, inhibitors of miRNA function. ciRS-7 (circular RNA sponge for miR-7) contains more than 70 selectively conserved miRNA target sites. It strongly suppresses miR-7 activity, resulting in increased levels of miR-7 targets (Hansen et al., 2013).

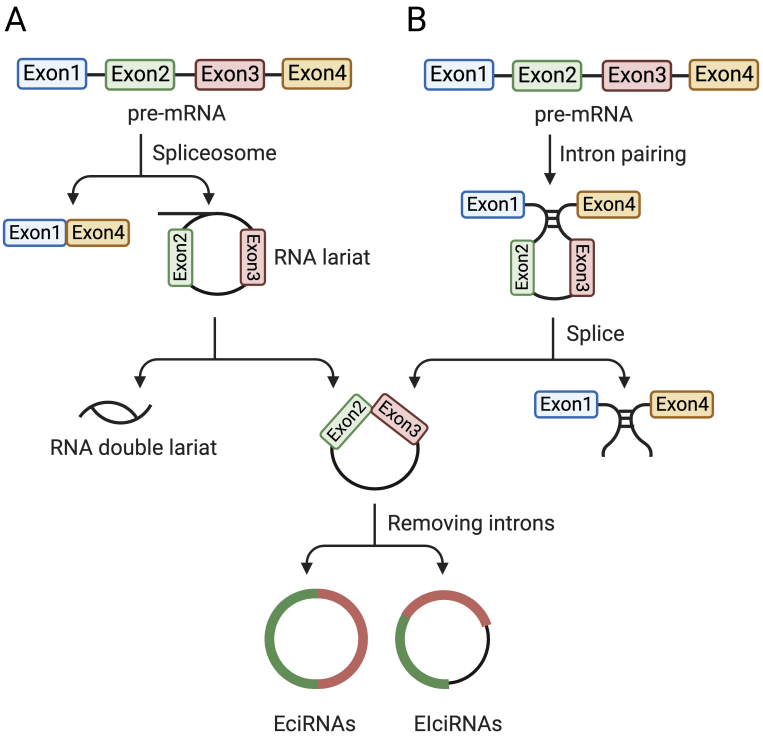


Fig. 8: Biosynthesis of EciRNA and EIciRNA. A) Exon-skipping or lariat intermediate circularization: pre-mRNA is spliced, causing exon1 and exon4 to covalently bond, while the sequence between the exons becomes an RNA lariat containing several exons and introns. Some introns of the circRNA are removed resulting in EciRNA or EIciRNA. B) Direct back-splicing or intron-pairing-driven circularization: the upstream intron pairs with the downstream intron and then the 2'-hydroxyl of one intron reacts with the 5'-phosphate of the second intron. Some introns of the circRNA are removed resulting in EciRNA or EIciRNA.

1.1.9 lncRNA

lncRNAs are RNAs longer than 200 nucleotides that do not code proteins. They can bind to DNA, RNA, and proteins, modulating gene expression in various biological and pathological contexts. lncRNAs regulate gene expression at transcriptional and post-transcriptional levels through mechanisms like RNA editing, pre-mRNA splicing, mRNA stability, chromatin modification, and miRNA-induced repression. They are uniquely expressed in different tissues and specific cancer types (Iyer et al., 2015; Statello et al., 2021). Most lncRNAs are transcribed by RNA pol II, spliced similarly to mRNAs, and often bear 5' end m^7G caps and 3'-end poly(A) tails, although many remain inefficiently processed and are retained in the nucleus (Grossi et al., 2020).

1.1.10 piRNA

piRNAs are 24-32 nt ncRNAs expressed mainly in mammals, specifically in germ cells. In most animals, their 3' termini are modified by 2'-O-methylation resulting in increased stability (Kirino and Mourelatos, 2007; Ohara et al., 2007). piRNAs are processed through complex mechanisms from ssRNA precursors that are mainly transcribed from two genomic sources, unannotated large blocks known as piRNA clusters and active transposable elements (TEs). piRNAs associate with the PIWI subfamily of Argonaute proteins to form piRNA-induced silencing complexes (piRISCs) (Siomi et al., 2011). piRNAs were first discovered in *Drosophila* where PIWI mutations led to the abolishment of germline stem cell division thus the name origin: P-element-induced wimpy testis (Lin and Spradling, 1997).

The most known role of the piRISCs in the germ line is TE silencing. TE are selfish mobile genetic elements that can cause genome damage and failure in gonadal development and fertility. Therefore, organisms needed to develop a mechanism that represses TE to maintain genome integrity. It was first indicated that piRNA in *Drosophila* functions in a TE surveillance pathway and it can adapt its response upon contact with active TE (Siomi et al., 2011).

1.1.11 7SL RNA

7SL RNA is the component of the signal recognition particle (SRP) which functions in protein secretion. 7SL RNAs are around 300 nt long and consist of an Alu domain and S domain, which is a 155 nt long sequence that is unique to 7SL RNA, separated by the linker

region. Alu sequences are the most abundant human repetitive DNA sequences, which were derived from 7SL RNA (or DNA) by a deletion of the central 7SL-specific sequence (Gussakovsky and McKenna, 2021; Ullu and Tschudi, 1984). SRP complex is comprized of one molecule of 7SL RNA and six different SRP proteins (Halic et al., 2004). Interestingly, it was found that in addition to genomic RNA, HIV-1 particles package some host RNAs such as 7SL RNAs, tRNAs and U6 snRNAs. They are selectively encapsidated through independent mechanisms and 7SL RNAs are packaged in HIV-1 particles as efficiently as viral RNAs (Houzet et al., 2007; Onafuwa-Nuga et al., 2006).

1.2 RNA processing

RNA processing involves a complex of highly regulated steps that begins with the RNA transcription and continues until the production of mature RNA. This process includes several key stages such as capping, splicing, and polyadenylation of RNA.

1.2.1 Transcription

Transcription is the process where RNA pol transcribes segments of DNA into RNA. The template DNA strand is the noncoding strand, while the coding strand has the same sequence as the new RNA molecule. In bacteria, a single type of DNA-dependent RNA pol performs transcription, while eukaryotic cells use three main types: RNA pol I, II, and III (Darst, 2001; Roeder and Rutter, 1969).

Transcription consists of three stages: initiation, elongation, and termination. Although the basic principles are similar across RNA pol I, II, and III, each has specific characteristics. RNA pol I, responsible for up to 60% of cellular transcriptional activity, synthesizes one large rRNA precursor and recognizes only one promoter. RNA pol III transcribes many copies of a few genes, while RNA pol II transcribes a wide variety of genes with diverse promoters and interacting factors (Laferté et al., 2006; Liu et al., 2011b).

RNA pol II

RNA pol II is responsible for the transcription of mRNA, snRNA, miRNA and circRNA (Lee et al., 2004; Mattick and Makunin, 2006). Their genes are controlled by two distinct families of cis-acting regulatory DNA sequences (Fig. 9). Regulatory elements of the first family consist of promoter and nearby elements located close to the transcription start

site (TSS) of the genes. The second family are distal regulatory elements, which can be located far from the TSS but due to the DNA structure are spatially close to the promoter. These elements can interact with DNA-binding proteins and include enhancers, silencers, insulators, or locus control regions (Blackwood and Kadonaga, 1998).

RNA pol II transcription initiation begins with the assembly of the preinitiation complex (PIC) at the promoter sequence near the TSS. The PIC comprises six general transcription factors (TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH), hypo-phosphorylated RNA pol II, and the Mediator complex (He et al., 2016; Roeder, 1996). TFIIB recruits RNA pol II, with its unphosphorylated C-terminal domain (CTD), along with TFIIE, TFIIF, and TFIIH to the promoter. Phosphorylation of the CTD promotes the release of the Mediator and the transition to elongation, where the CTD also recruits capping enzymes (Gonatopoulos-Pournatzis and Cowling, 2014; Hsin and Manley, 2012).

DNA strands are separated to form a single-stranded bubble in an ATP-dependent process and RNA pol II tightens its grip on the nascent RNA (Luse and Jacob, 1987). During elongation, negative elongation factors can pause transcription, preventing reinitiation (Liu et al., 2011b; Saunders et al., 2006; Shao and Zeitlinger, 2017). Once these factors are released, elongation resumes, adding nucleotides in the 5' to 3' direction. At termination, RNA pol II slows down at the terminator sequence, recruiting 3' end cleavage and polyadenylation complexes, eventually releasing the pre-mRNA (Eaton and West, 2020; Kuehner et al., 2011).

RNA pol I

The sole function of RNA pol I is the synthesis of the large rRNA precursor (pre-rRNA). RNA pol I initiation requires the formation of PIC at the DNA promoter. This includes the upstream binding factor (UBF), which is known to bend the DNA, recruit RNA pol I, and stabilize the binding of other factors (Jantzen et al., 1990; Learned et al., 1985). Genes for pre-rRNA are simultaneously transcribed by dozens of RNA pol I separated by stretches of DNA generating 'Christmas tree' like structures (Miller Jr and Beatty, 1969).

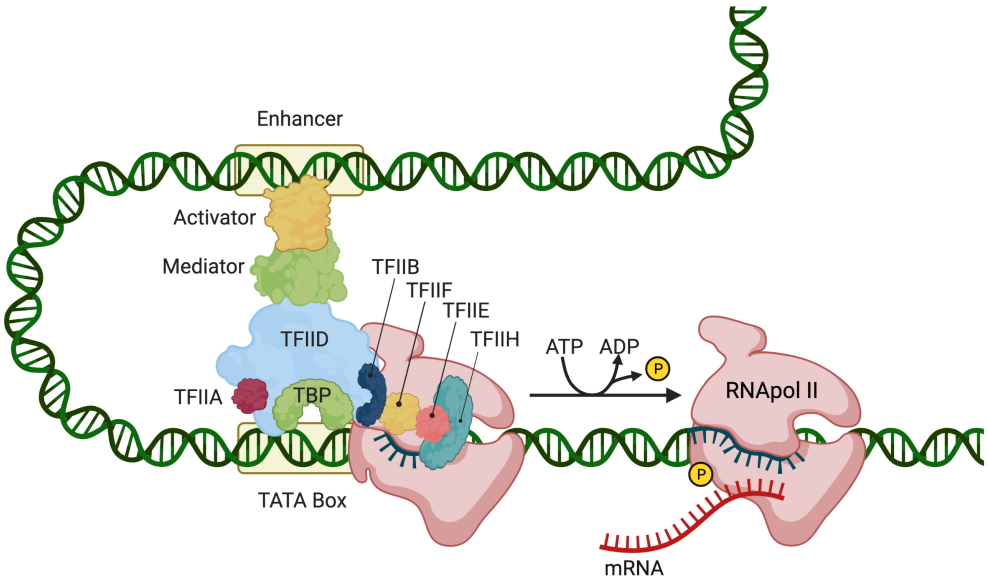


Fig. 9: Assembly of the pre-initiation complex. The binding of an activator to its enhancer sequence is crucial for subsequent interaction with a mediator and recruitment of the general transcription factors. TFIIB facilitates the recruitment of hypo-phosphorylated RNA pol II with TFIIE, TFIIF and TFIIH to the promoter to assemble PIC. Unphosphorylated CTD of RNA pol II binds to the mediator, while its phosphorylation causes the release of the mediator and RNA pol II from the promoter area.

RNA pol III

RNA pol III synthesizes a broad spectrum of ncRNAs like tRNA, 5S rRNA, 7SL RNA, U6 snRNA, and some pre-miRNAs (Borchert et al., 2006). The introns were found only in tRNA genes between 37 and 38 nt of mature tRNAs (Marck and Grosjean, 2002). Transcription termination sites are marked by at least four T residues on the coding DNA strand (Braglia et al., 2005), which results in a so-called 3' trailer of 2-5 U residues (Maraia and Lamichhane, 2011). RNAs transcribed by RNA pol III lack the 5' end cap. However, 7SK RNA contains methyl groups at the 5' end (Shumyatsky et al., 1990) and a similar structure was also detected on U6 snRNA (Epstein et al., 1980; Singh and Reddy, 1989).

1.2.2 RNA capping

The initial product of RNA pol II carries a 5' triphosphate end (pppRNA). Capping begins once the first 25-30 nucleotides are incorporated into the nascent transcript (Fig. 10). RNA triphosphatase (RTPase) hydrolyze the 5' γ-phosphate of pppRNA to produce 5' diphosphate pre-mRNA (ppRNA) and inorganic phosphate (Pi). RNA guanylyltransferase (GTase) then reacts with guanine triphosphate (GTP) to form a

covalent GMP intermediate, which links to the 5'-ppRNA, creating Gppp-RNA. The guanine at the N7 position is then methylated by RNA guanine-N7 methyltransferase (N7MTase) using S-adenosyl-L-methionine (SAM), forming Cap 0 and releasing S-adenosyl-L-homocysteine (SAH). The Cap 0 structure, common in eukaryotic cells and many viruses but absent in archaea and bacteria, protects nascent mRNA from degradation by nucleases (Ramanathan et al., 2016; Shuman, 2000).

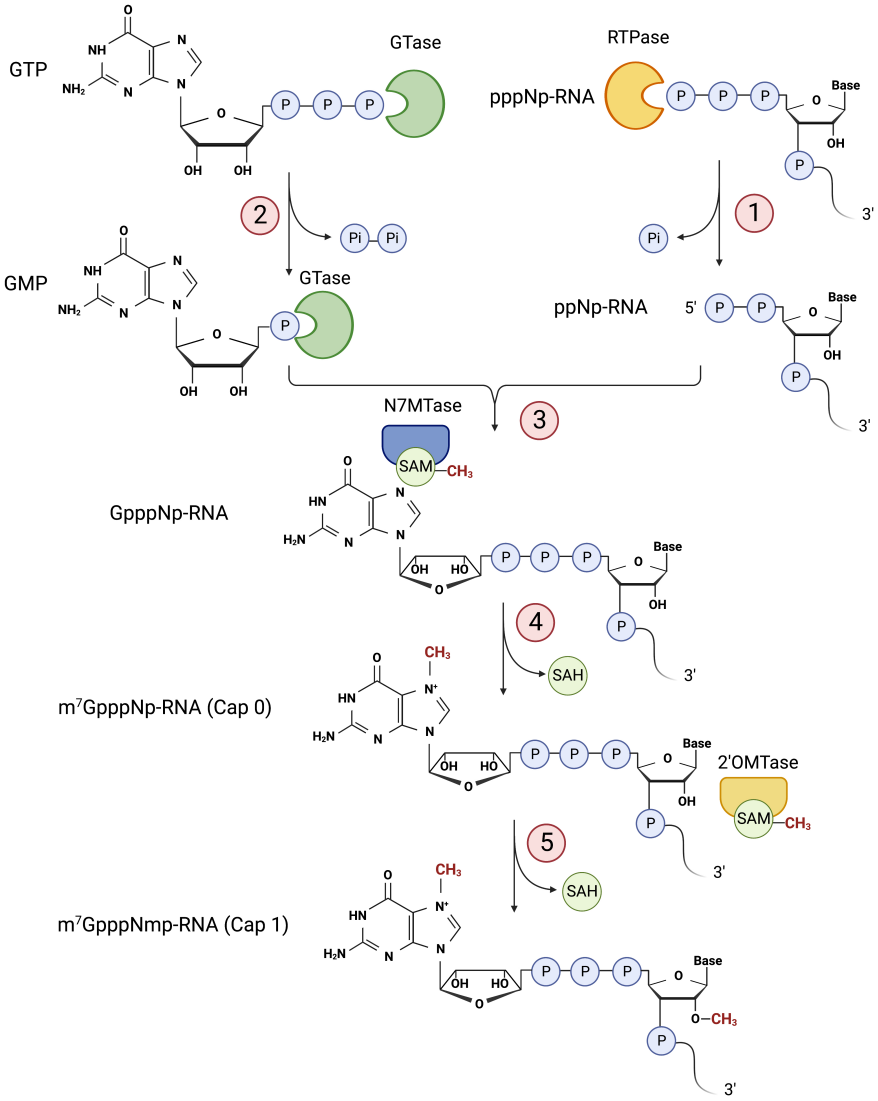


Fig. 10: Enzymatic steps of RNA capping. γ-phosphate is removed from pppNp-RNA (N: the first transcribed nucleotide, p: phosphate group) by RTPase. This generates ppRNA and Pi. GTase consumes GTP molecule and transfers GMP to the ppRNA, forming a reverse 5'-to-5' phosphate linkage (GpppNp-RNA). Next, the N7MTase adds a methyl group from the methyl donor SAM to the N7 amine of the guanosine cap to form the Cap 0 structure (m^7GpppNp-RNA). The 2'-hydroxyl of the first nucleoside can be further methylated by 2'-O methyltransferase to yield Cap 1 (m^7GpppNmp-RNA, where m denotes a methyl group).

Subsequent modifications occur at the 2'-O position of the first and second nucleotides. The m^7G-specific 2'-O methyltransferase CMTr1 methylates the first nucleotide to generate Cap 1 (Fig. 11). Cap 1 on dsRNA prevents interactions with RIG-I and MDA5, cytoplasmic sensors that trigger cellular type I interferon response to viral infections (Schuberth-Wagner et al., 2015; Züst et al., 2011). Further methylation in the cytoplasm at the 2'-O position of the second nucleotide by CMTr2 results in Cap 2 (Ramanathan et al., 2016; Smietanski et al., 2014). In mouse embryos, both CMTr1 and CMTr2 are essential for mouse embryonic development and Cap 1 and Cap 2 have essential roles in gene regulation (Dohnalkova et al., 2023). Cap 2 accumulates on host long-lived mRNAs through a slow conversion from Cap 1. While increased amounts of mRNAs with Cap 1 lead to activation of RIG-I, the mRNAs with Cap 2 has reduced ability to bind and activate RIG-I, thereby preventing activation of the interferon signaling pathway (Despic and Jaffrey, 2023).

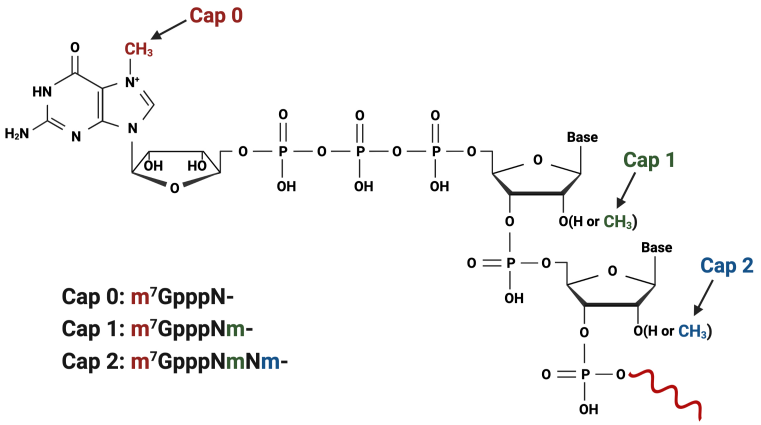


Fig. 11: Canonical mRNA caps in eukaryotes. The Cap 0 structure (m^7GpppN-) is composed of N7-methylguanosine linked to the first nucleoside of mRNA through a triphosphate bridge. The methyl group of Cap 0 is highlighted in red. The Cap 0 structure is further methylated at the 2'-O positions of the ribose of the first and second nucleoside of mRNA to generate the Cap 1 ($m^7GpppNm$-, where N denotes the first transcribed nucleotide and m denotes a methyl group) and Cap 2 ($m^7GpppNmNm$-) structures, respectively. The additional methyl group of Cap 1 is highlighted in green and Cap 2 in blue.

m^6Am, a prevalent modification found on Cap 1 and Cap 2, is a dynamic and reversible epitranscriptomic modification (Wei et al., 1975). It enhances mRNA stability by protecting it from the mRNA-decapping enzyme Dcp2. On the other hand, the fat mass and obesity-associated protein (FTO) preferentially demethylates m^6Am rather than m^6A, which leads to reduced stability of m^6A containing mRNAs (Mauer et al., 2017).

In humans, some snRNAs, snoRNAs, and telomerase RNAs have Cap 0 which is further hypermethylated to TMG cap (Fig. 12). snRNA bound to Sm core proteins is exported from the nucleus to the cytoplasm, where its cap structure is hypermethylated by trimethylguanine synthase 1 (Mattaj, 1986). The N7 methylation precedes N2 methylation, requiring the transfer of two methyl groups from SAM to form the TMG cap (Hausmann and Shuman, 2005). TMG cap serves as an import signal recognized by snurportin 1, allowing the mature snRNA to re-enter the nucleus (Huber et al., 1998).

Proper TMG cap formation, especially in U1 snRNA, is crucial for efficient pre-mRNA splicing (Mouaikel et al., 2002). In nematodes, TMG-capped coding RNAs can bind eIF4E and undergo translation (Liu et al., 2011a). Some mammalian selenoprotein mRNAs also undergo TMG capping and can be translated (Wurth et al., 2014).

Human U6 and 7SK sRNAs bear γ-methyl triphosphate cap. γ-methyltransferases catalyze the transfer of methyl group from SAM to a γ-phosphate oxygen at the unprocessed 5' end of RNAs. Gamma-methyltransferases are unable to methylate ppRNA, which ensures that transcripts that have been processed by RTPase are protected (Gupta et al., 1990; Singh and Reddy, 1989).

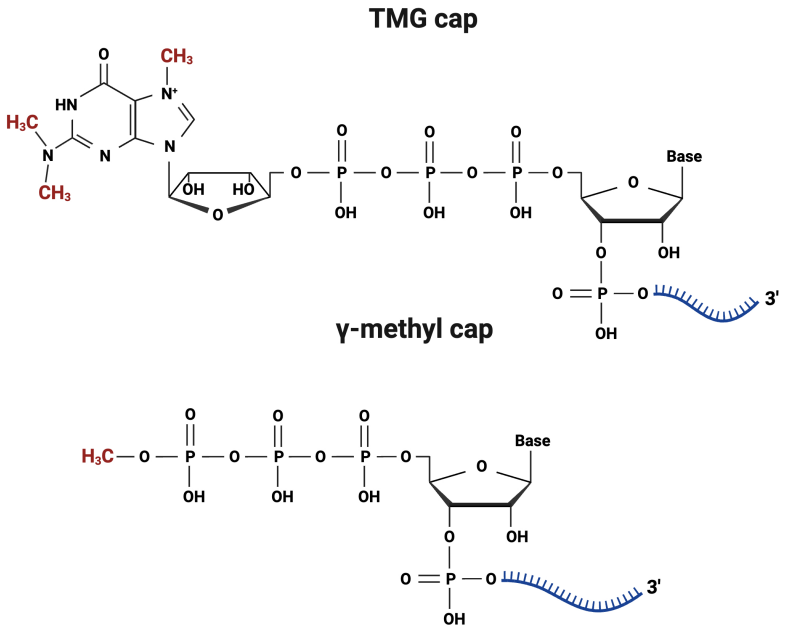


Fig. 12: Canonical caps of small RNAs in eukaryotes. TMG cap bears three methylations on guanine, two at position N2 and one at position N7. γ-methyl triphosphate cap is methylated only at the position of γ-phosphate oxygen. The methyl groups are highlighted in red.

RNA capping occurs also in the cytoplasm of eukaryotic cells. During mRNA decay, the cap is hydrolyzed by decapping enzymes Dcp2 or NudT16. The uncapped mRNA accumulates within P-bodies and can be either degraded by the 5'-3' exonuclease Xrn1 or re-capped and translated (Brengues et al., 2005; Peculis et al., 2007; Schoenberg and Maquat, 2012; Van Dijk et al., 2002). Cytoplasmic re-capping affects only a portion of mRNA transcriptome, encoding proteins involved in nucleotide binding, RNA and protein localization, and the mitotic cell cycle. This cyclical process of decapping and re-capping, termed cap homeostasis, allows uncapped RNAs to be stored in P-bodies under stress and re-enter translation upon recovery, conserving cellular energy (Mukherjee et al., 2012).

1.2.3 RNA splicing

Eukaryotic pre-mRNA transcripts contain exons and introns. During splicing, introns are removed, and exons are ligated together by spliceosomes, which are complexes of snRNAs and proteins (Fig. 13). Introns are defined by conserved 5' and 3' splice sites and a branch point, typically an adenosine. The U1 snRNP recognizes the 5' splice site, and U2 snRNP binds to the branch point near the 3' splice site, forming complex A. This marking occurs co-transcriptionally (Black et al., 1985; Listerman et al., 2006).

Next, U4/U6-U5 tri-snRNP is incorporated, forming complex B. The spliceosome is activated when ATP-dependent Prp28p destabilizes U1 snRNP, releasing it from the 5' splice site and allowing U4/U6 duplex disassembly. Helicase Brr2 unwinds U4 from U6, leading to U4 snRNP release and U6 snRNP interaction with U2 snRNP, forming complex C. This interaction is crucial for the spliceosome's catalytic core. A transesterification reaction occurs at the branch point, cleaving the 5' splice site and forming a lariat-3' exon intermediate. A second transesterification at the 3' splice site ligates the exons. The mature mRNA is produced, and the lariat intron is excised (Madhani and Guthrie, 1992; Small et al., 2006; Staley and Guthrie, 1999; Wilkinson et al., 2020).

m^7G cap and the CBC are essential for efficient splicing, as they interact with splicing machinery components. Depletion of CBC inhibits splicing and reduces U1 snRNP recruitment to the 5' splice site (Izaurralde et al., 1994; Lewis et al., 1996). Mutations in U1 snRNA, often found in cancers, create novel splice junctions and alter the splicing of multiple genes, including cancer drivers like Musashi RNA binding protein 2, leading to mis-splicing or increased excision of the 5' splice site (Shuai et al., 2019).

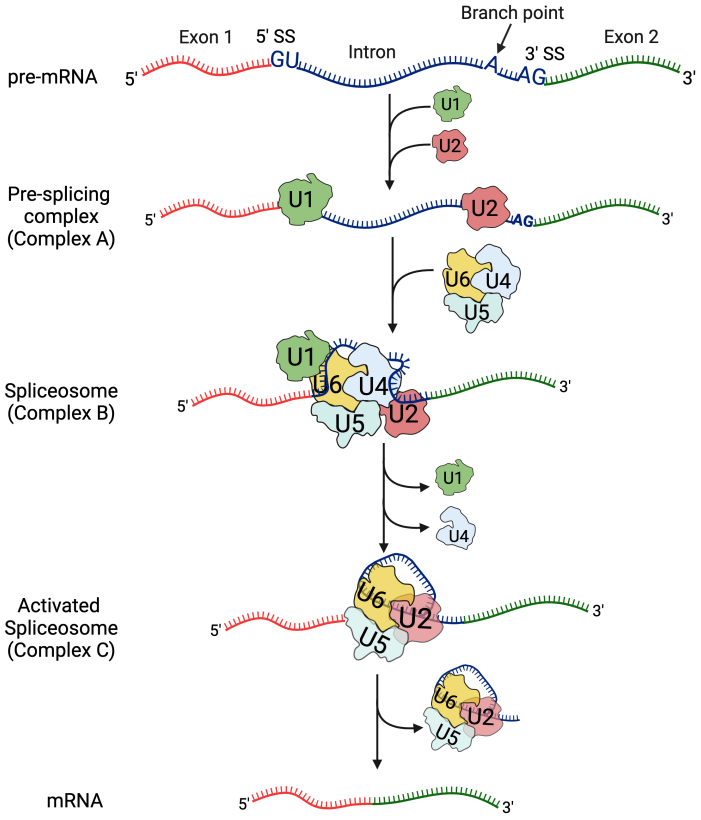


Fig. 13: Major spliceosome assembly and splicing of pre-mRNA. 5' SS and 3' SS with the conserved residues as well as the branch point adenosine are indicated in the pre-mRNA.

1.2.4 RNA polyadenylation

RNA polyadenylation is the addition of approximately 200 adenosine nucleotides at the 3' end, known as the poly(A) tail (Fig. 14). Poly(A) tail protects mRNA from degradation and allow its transport to the cytoplasm and translation. Polyadenylation begins after transcription termination and involves two steps: endonucleolytic cleavage at the 3' end and poly(A) tail addition (Slomovic et al., 2006).

The polyadenylation signal, a conserved AAUAAA sequence, is recognized by the cleavage and polyadenylation specificity factor. Additional upstream or downstream sequences enhance cleavage efficiency. For instance, the U- or GU-rich sequence downstream of the cleavage site is recognized by the cleavage stimulation factor (Hu et al., 2005; Wahle and Keller, 1992).

Once the polyadenylation factors bind, cleavage occurs 10-30 residues downstream of the polyadenylation signal. Poly(A) polymerase (PAP) adds adenosine residues to the 3' end. Poly(A) RNA-binding proteins (PABPs) bind the new poly(A) tail, protecting

mRNA from degradation (Murthy and Manley, 1995; Wahle and Keller, 1992; Wigington et al., 2014). The poly(A) tail regulates mRNA decay. Newly transcribed RNAs receive a tail of about 230 nucleotides, gradually shortened by deadenylases (Eisen et al., 2020; Munroe and Jacobson, 1990).

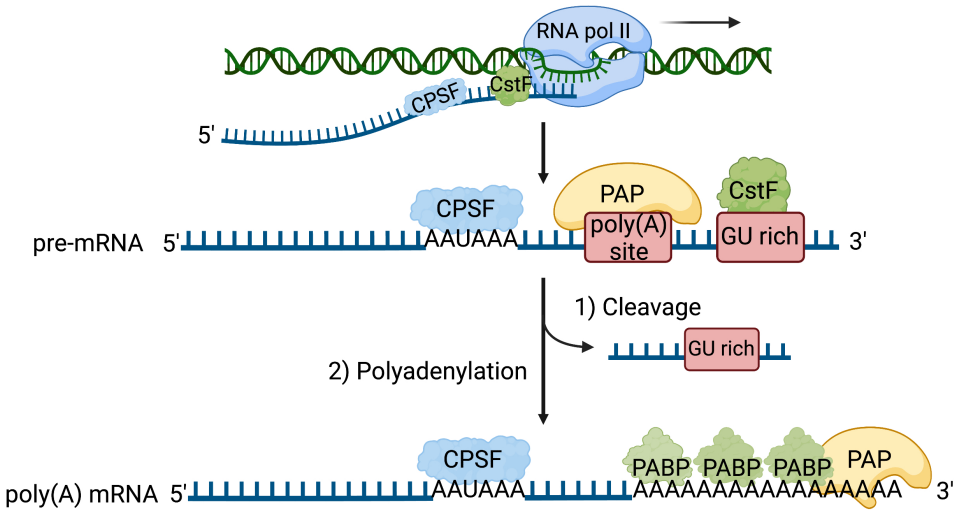


Fig. 14: The steps of poly(A) tail synthesis. The CPSF and CstF factors recognize the polyadenylation signal and GU-rich sequence located upstream and downstream of the cleavage site, respectively, while the pre-mRNA is still being synthesized. They interact with the PAP and cleave pre-mRNA between the AAUAAA sequence and the downstream GU-rich element. The upstream fragment is then polyadenylated, while the downstream fragment is degraded. PABPs interact with the poly(A) tail to function as its shield.

1.2.5 RNA degradation

RNA degradation is crucial for gene expression control, elimination of aberrant RNAs, and antiviral defenses. Eukaryotic RNAs are degraded by three main classes of enzymes: endonucleases (internal), 5' exonucleases (hydrolyze RNA from the 5' end), and 3' exonucleases (degradation from the 3' end). There are three pathways: (A) deadenylation-dependent decay, initiating with the removal of the 3' poly(A) tail; (B) deadenylation-independent decay, often starting with 5' cap removal; and (C) endonuclease-mediated decay, beginning with internal RNA cleavage (Fig. 15) (Schaeffer et al., 2009).

A) Deadenylation-dependent decay

The deadenylation process follows a two-phase model: Pan2/Pan3 removes the poly(A) tail until PABP can no longer bind, then Ccr4-Not removes the remaining adenines, leading to rapid mRNA degradation (Brown and Sachs, 1998; Yamashita et al., 2005).

RNA decapping

Decapping is regulated both biochemically and spatially, with most mRNA decay factors associated with nontranslated mRNAs in P-bodies (Brengues et al., 2005). Bulk mRNA degradation is believed to occur in P-bodies, though aggregation in P-bodies is not required for mRNA decay (Eulalio et al., 2007). mRNAs can be also degraded co-translationally (Pelechano et al., 2015).

Dcp2, a Nudix superfamily hydrolase, is a key enzyme in the 5'-3' decay pathway. Activated by coactivator Dcp1 and factor EDC4, Dcp2 cleaves the 5' cap of mRNA, releasing m^7GDP and pRNA. EDC4 binds Dcp1, Dcp2, and Xrn1, coupling decapping with Xrn1-mediated degradation. Nudt16 and Nudt3 have also decapping activity and regulate mRNA stability (Chang et al., 2014; Dunckley and Parker, 1999; Grudzien-Nogalska et al., 2016).

DXO (decapping exoribonuclease) functions preferentially on incompletely capped RNAs and possesses both decapping and exoribonuclease activities. DXO decapping activity generates pRNA and its exonucleolytic activity enables further pRNA degradation (Jiao et al., 2013). 2'-O-methylation on the cap structure drastically reduces the affinity of DXO for RNA (Picard-Jean et al., 2018). DXO removes non-canonical caps like NAD, FAD (flavin adenine dinucleotide), and dpCoA (dephospho-coenzyme A), processes known as deNADing, deFADing, and deCoAping, respectively (Doamekpor et al., 2020). DXO has a higher affinity for NAD-RNA compared to m^7G-RNA, leading to efficient degradation (Jiao et al., 2017).

Some Nudix proteins, like Nudt16, can also hydrolyze non-canonical caps such as NAD, FAD, and dpCoA (Sharma et al., 2020). Nudt2 can cleave Ap_4A (František Potužník et al., 2024), whereas Dcp2 cannot hydrolyze non-canonical caps (Jiao et al., 2017).

Exoribonuclease digest

The XRN family includes two exoribonucleases: Xrn1, located in the cytoplasm and often found in P-bodies, and Xrn2, located in the nucleus. Xrn1 digests pRNA in the 5'-3'

direction but cannot degrade pppRNA or m^7G-RNA. It recognizes the 5' monophosphate of RNA through a pocket which excludes larger 5' groups (Jones et al., 2012; Pellegrini et al., 2008). Interestingly, Xrn1 functions as a deNADing enzyme, degrading NAD-RNA using the same catalytic site as for pRNA hydrolysis. Its activity on pRNA is 20 times more efficient than on NAD-RNA. Like DXO proteins, Xrn1 releases intact NAD from NAD-RNA, allowing NAD regeneration (Sharma et al., 2022).

Exosomal degradation

The RNA exosome structure and composition are highly conserved across species (Zinder and Lima, 2017). The riboexonucleases EXOSC10 and DIS3 degrade a variety of RNAs in the 3'-5' direction following deadenylation (Butler, 2002). In some cases, RNA is deadenylated and degraded by 3'-5' exoribonucleolytic activity, but the cap structure remains. A scavenger decapping enzyme, DcpS, hydrolyzes the residual cap and releases m^7GMP (Wang and Kiledjian, 2001). It is a conserved member of the histidine triad (HIT) superfamily and cannot function on RNAs that are less than ten nucleotides long (Liu et al., 2002). Another member of the HIT family, FHIT (fragile histidine triad protein), is involved in the degradation of residual 5' cap structures in humans. FHIT cleaves m^7GpppG and produces a mixture of m^7GMP and m^7GDP (Taverniti and Séraphin, 2015).

B) Deadenylation-independent degradation

Several pathways of RNA decay are deadenylation-independent, including nonsense-mediated decay (NMD) (Lykke-Andersen, 2002), AU-rich element (ARE)-mediated decay (Bakheet et al., 2001), and miRNA-coupled decapping (described previously) pathways (Rehwinkel et al., 2005). NMD is an mRNA surveillance mechanism that degrades aberrant mRNAs with premature translation termination codons (Lykke-Andersen, 2002). Likewise the general degradation pathway, NMD involves decapping, deadenylation, and 5'-3' exonucleolytic degradation by Xrn1 (Lejeune et al., 2003).

AREs act as regulators of mRNA stability. They are located in the 3' UTRs of many mRNAs encoding transcription factors, cytokines and proto-oncogenes. They can recruit ARE-binding proteins and subject the mRNAs to ARE-mediated decay (Barreau et al., 2005). Tristetraprolin is a key ARE binding factor that interacts with the Dcp2 protein

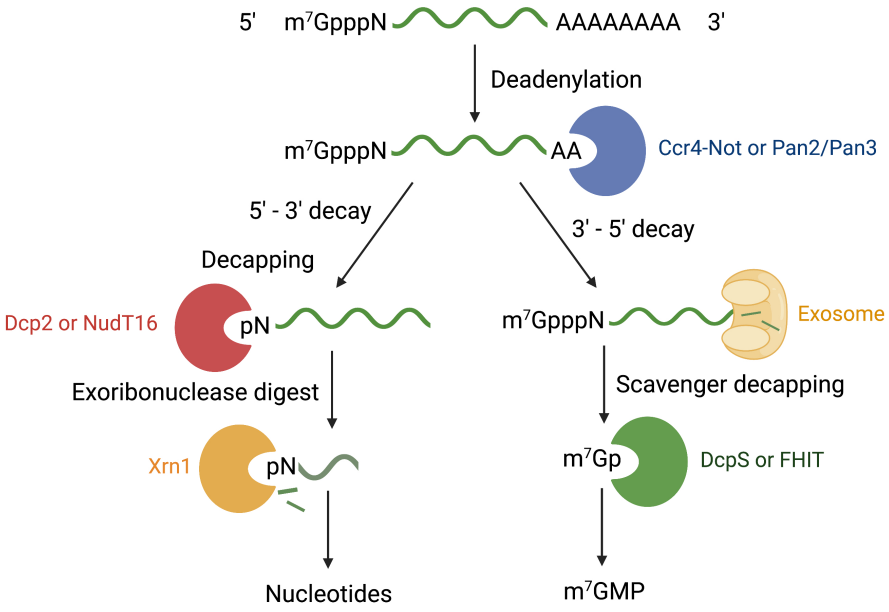


Fig. 15: Summary of the general deadenylation-dependent RNA degradation model. The decay of mRNAs starts with deadenylation by the Ccr4-Not complex and Pan2/Pan3 complex, followed by decapping or degradation in the exosome. After the exosomal degradation, the residual cap is hydrolyzed by scavenger enzymes DcpS or FHIT. In the 5'-3' decay pathway, the deadenylated RNA is decapped by one of the canonical decapping enzymes Dcp2 or Nudt16. The last step is the digest by 5'-monophosphate-dependent exoribonuclease Xrn1.

and promotes ARE-dependent Dcp2 decapping (Franks and Lykke-Andersen, 2007). Nudt16 is also involved in ARE-mediated decay, although it is utilised differentially than Dcp2 (Li et al., 2011).

C) Endonuclease-mediated decay

Endonuclease-mediated decay starts with the mRNA's endonucleolytic cleavage, producing two fragments, each with one unprotected end. The RNA 5' and 3' products are subsequently degraded by the exosome and Xrn1, respectively (Tomecki and Dziembowski, 2010). Several endonucleases have been characterized, and their activation often depends on specific signals such as stress stimuli. For instance, Polysomal ribonuclease 1 is mainly responsible for the degradation of albumin mRNA (Dompenciel et al., 1995) while ribonuclease MRP for mitochondrial RNA and pre-rRNA (Lygerou et al., 1996).

1.3 RNA modifications

Eukaryotic RNA modifications have been known and studied since the 1970s. However, they were primarily detected on tRNAs and rRNAs due to their abundance. Lately, they

have been identified also on mRNAs and various ncRNAs. To date, over 170 different RNA modifications, internal and non-canonical caps, have been identified (Boccaletto et al., 2022).

1.3.1 Internal modifications

Internal RNA modifications are chemical alterations on nucleotides within the RNA molecules. These modifications can affect RNA functions, stability, localization, translation efficiency, and interaction with other molecules. Following part mentions only a few representatives of RNA modifications.

m^6A

m^6A is the most abundant internal modification in mRNA, with around one m^6A per 2000 nucleotides (Desrosiers et al., 1974; Dominissini et al., 2012). It does not affect coding capacity but can destabilize complexes with uracil (Dai et al., 2007; Roost et al., 2015). The enzymes able to install the methylation are called writers, erasers are enzymes removing the methylation, and readers are enzymes able to interacting with it. The writer complex consists of the catalytic subunit methyltransferase-like 3 (METTL3), the substrate-recognizing subunit METTL14 and many accessory subunits (Bokar et al., 1997) (Fig. 16). During the process, a methyl group is transferred from methyl donor S-adenosylmethionine (SAM) into nascent pre-mRNA at the N6 position of the adenosine base (Schwartz et al., 2014b). FTO is the first eraser that was discovered to have demethylation activity towards m^6A (Jia et al., 2011). AlkB homolog 5 (ALKBH5) demethylates m^6Am present in 5' caps (Mauer et al., 2017; Zheng et al., 2013).

Most readers of the m^6A belongs to the YTHD family. YTHDF2 was shown to bind to m^6A, which results in the re-localization of the mRNA from the translatable pool to mRNA decay sites (Wang et al., 2014). On the contrary, YTHDF1 and YTHDF3 promote translation of m^6A-containing transcripts by interacting with translation machinery (Li et al., 2017a; Meyer et al., 2015; Wang et al., 2015). Another m^6A reader, the eukaryotic initiation factor 3 (eIF3), also promotes the translation. It was shown that a single 5' UTR m^6A binds eIF3, and such mRNA can be translated even in a cap-independent manner. eIF3 can recruit the preinitiation complex to initiate translation in the absence of the cap-binding factor eIF4E (Meyer et al., 2015).

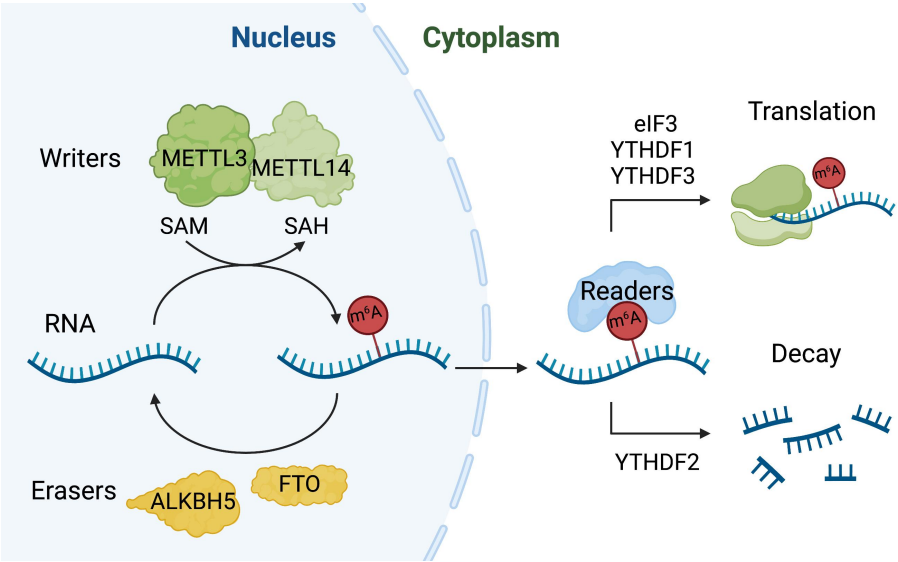


Fig. 16: Summary of m^6A modification machinery. The methyltransferase complex consists of METTL3 and METTL14. SAM is used as a methyl donor for the methylation. This modification can be removed by erasers FTO and ALKBH5. The methyltransferases and demethylases act in the nucleus. In the cytoplasm, the m^6A modification is recognized by readers, which binds to the methylation site on RNA and can induce either translation or decay.

m^1A

m^1A was thought to be a highly prevalent modification in mRNA. Recently, it was discoered that its prevalence was overestimated due to cross-reactivity of detection antibodies (Grozhik et al., 2019). m^1A modification is spread on tRNA, where it occurs at several different positions. In human cells, positions 9, 14 and 58 are conserved m^1A sites in tRNAs (Li et al., 2017b). m^1A prevents Watson-Crick base pairing causing the nucleosides at positions 9, 14, and 58 to form non-Watson-Crick base pairs. This is crucial for establishing the tRNA three-dimensional structure (Smoczynski et al., 2024).

Nm and Ψ

Nm and Ψ are the most abundant rRNA modifications. They occur close to functionally important sites and are critical for ribosome assembly and accurate protein synthesis (Birkedal et al., 2015). These modifications are introduced mainly by two classes of snoRNPs, box C/D and box H/ACA snoRNPs, which mostly guide the modification of only a single rRNA nucleotide (Fig. 6). The box C/D snoRNPs catalyze 2'-O-methylation, which is initiated by base-pairing of the snoRNA adjacent to the D/D box (conserved sequence motif) with the pre-rRNA (Van Nues et al., 2011).

In snRNAs, these modifications are mainly present in regions important for pre-mRNA splicing, including the sites of RNA-RNA interactions (Karijolich and Yu, 2010). For example, U1 snRNA contains Ψ at positions 5 and 6 within the sequence base pairing with pre-mRNA. It seems to be crucial in the recognition process of the 5' splice site of pre-mRNA (Roca et al., 2005). The mechanism of snRNA modification installing is the same as for rRNA, by box C/D and box H/ACA snoRNPs (Karijolich and Yu, 2010).

In mRNAs, Ψ is regulated by environmental signals such as serum starvation and plays a role in enhancing transcript stability (Carlile et al., 2014; Schwartz et al., 2014a). Nm at the first and second transcribed nucleotides prevents mRNA degradation by DXO (Picard-Jean et al., 2018).

Inosine

Adenosine to I conversion is carried out by adenosine deaminase acting on RNA (ADAR) proteins, requiring a duplex structure at modification sites (Zinshteyn and Nishikura, 2009). I, structurally similar to guanosine, preferentially base-pairs with cytidine, affecting dsRNA stability (Wright et al., 2018).

I within the tRNA anticodon region can pair with U, C, or A expanding tRNA's reading capacity during translation (Grosjean et al., 2010; Pan, 2018).

m^6Am

m^6Am is a dynamic and reversible epitranscriptomic modification (Wei et al., 1975). It is found in 30-40% of mRNAs adjacent to the m^7G cap, where it has protective effect against the mRNA decapping enzyme Dcp2 (Mauer et al., 2017; Wei et al., 1975). However, FTO preferentially demethylates m^6Am rather than m^6A (Mauer et al., 2017).

m^5C

m^5C is enriched in untranslated regions in mRNA and near Argonaute binding regions, serving as a DNA damage code that promotes homologous recombination (Chen et al., 2020; Squires et al., 2012). In tRNA, m^5C affects the codon-anticodon interactions (Agris, 1996).

1.3.2 Non-canonical RNA caps

Traditionally, RNAs in eukaryotic cells possess m^7G cap or TMG cap. However, recent research has revealed the existence of alternative cap structures known as non-canonical caps. These caps are added to RNAs by alternative mechanisms and can influence the stability, translation efficiency, and gene expression regulation of RNAs. The non-canonical caps originate from abundant coenzymes such as NAD, FAD, or dpCoA, and additionally from metabolites like dinucleoside polyphosphates (Np_nN).

FAD, UDP-Glc and UDP-GlcNAc caps

It was shown that bacterial RNA pol can initiate transcription with the cellular cofactor FAD, uridine diphosphate glucose (UDP-Glc), and uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) *in vitro* (Fig. 17) (Julius and Yuzenkova, 2017). This expanded the repertoire of possible RNA caps, suggesting that FAD might also serve as an RNA cap. Thanks to a new technique for caps detection, called CapQuant (systems-level mass spectrometry-based technique), it was confirmed that FAD, UDP-Glc and UDP-GlcNAc function as caps in human cells. Apart from human cells, these caps were also found in *S. cerevisiae*, *Escherichia coli*, DENV-2 virus (Wang et al., 2019a), and hepatitis C virus (HCV) (Sherwood et al., 2023).

FAD-capQ was developed and used for FAD cap detection in human cells. FAD caps were found on sRNAs with less than 200 nucleotides. DXO decapping enzyme releases FAD caps from RNAs and produces 5'-monophosphate RNA. Upon disruption of the DXO gene, the total levels of FAD caps were 2-fold higher, which shows a functional role for DXO in regulating the fate of FAD-RNAs (Doamekpor et al., 2020). Both mammalian Nudt2 and Nudt16 hydrolyze FAD-RNAs *in vitro*. Like DXO, disruption of the gene encoding Nudt16 led to increased levels of short FAD-RNAs in cells. In contrast, the levels of FAD-RNAs in cells lacking Nudt2 remained unchanged. It was suggested that Nudt2 either does not function on FAD-RNAs in cells or modulates only a small subset of RNAs (Sharma et al., 2020).

Recently, a technique for selective capturing of FAD-RNAs was described (Sherwood et al., 2023). It combines the enzymatic activity of Nudix pyrophosphohydrolase 23 (AtNUDX23), which is specific for FAD, with the CapZyme-seq method (Vvedenskaya et al., 2018). It was found that 75% of the 5' termini of HCV(+) and HCV(–) are FAD capped. It was discovered that the FAD cap serves as protection against RIG-I

mediated immune responses, especially during early steps of the viral life cycle (Sherwood et al., 2023).

The enzymes that can hydrolyze UDP-Glc-RNAs and UDP-GlcNAc-RNAs neither the function of these metabolite caps were not discovered yet (Sharma et al., 2020).

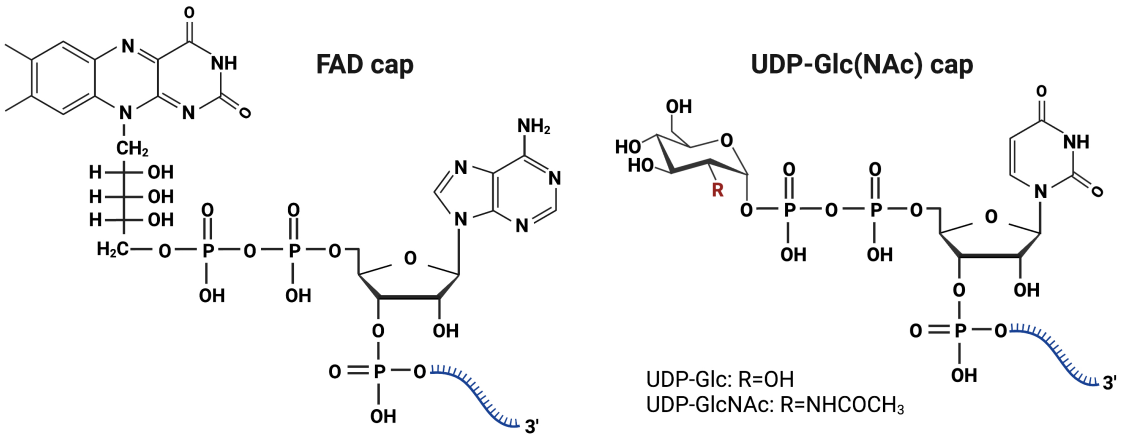


Fig. 17: Chemical structures of non-canonical caps FAD, UDP-Glc and UDP-GlcNAc.

dpCoA cap

dpCoA has been identified as a cap (Fig. 18) in *E. coli*, *Streptomyces venezuelae* (Kowtoniuk et al., 2009), and recently in RNA from mouse's liver (Shao et al., 2023). In *E. coli*, dpCoA caps were found only in smaller RNAs. Additionally, other small molecule-RNA conjugates such as acetyl-, succinyl-, and methylmalonyl-dpCoA were reported (Kowtoniuk et al., 2009). The CapQuant detection technique did not confirm dpCoA caps in RNA from dengue virus, *E. coli*, yeast, mouse tissues, nor human cells (Wang et al., 2019a). Conversely, the dpCoA tagseq method identified 44 types of dpCoA-RNAs in mouse liver RNA, but subsequent revisions revealed 58% of these were likely artifacts from nanopore sequencing chimeric reads (Shao et al., 2023).

dpCoA can be incorporated into RNA during transcription initiation by *E. coli* RNA pol and processed by the bacterial decapping enzyme NudC to 5'-monophosphate RNA (Bird et al., 2016). The DXO decapping enzyme also hydrolyzes dpCoA-RNAs similarly to deFADing. Crystal structures of mouse DXO with FAD and CoA showed that the flavin of FAD and the pantetheine group of CoA bind to the same region (Doamekpor et al., 2020). Seven Nudix proteins (Nudt2, Nudt7, Nudt8, Nudt12, Nudt15, Nudt16, and Nudt19) were identified to possess deCoAping activity *in vitro* and potentially *in vivo* (Sharma et al., 2020).

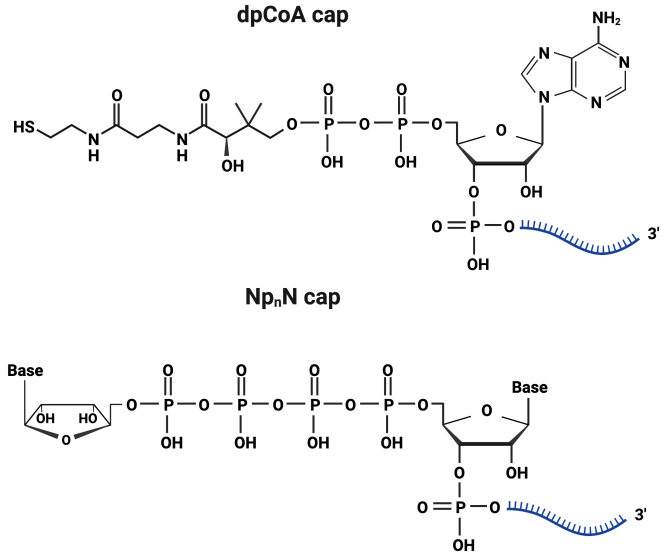


Fig. 18: Chemical structures of non-canonical caps dpCoA and Np_nN (variant with four phosphate groups).

Np_nN caps

Np_nN caps, originating from metabolites, were discovered in bacteria (Hudeček et al., 2020; Luciano et al., 2019) (Fig. 18). Np_nNs such as Ap_3A, Ap_4A, Ap_3G, and Ap_4G were initially identified as stress molecules or alarmones, accumulating rapidly in bacteria under oxidative stress (Bochner et al., 1984). RNA pol has been shown to incorporate Np_nNs as non-canonical initiating nucleotides *in vitro*.

The presence of Np_nN caps was confirmed by LC-MS technique in fractions of sRNA from *E. coli*. Six previously unknown Np_nN-caps (Ap_3A, m^6Ap_3A, Ap_3G, m^7Gp_4Gm, Ap_4A, mAp_5G) were detected in exponential phase and three additional caps (mAp_4G, mAp_5A, $2mAp_5G$) in late stationary phase. RppH (RNA 5'-pyrophosphohydrolase from the Nudix family) and ApaH (bis(5'-nucleosyl)-tetraphosphatase) are well-known bacterial phosphohydrolases, which can also cleave the novel RNA caps. Methylation of Np_nN caps protects them from RppH decapping, while ApaH hydrolyzes even methylated caps, suggesting that methylation stabilizes Np_nN-RNAs against RppH decapping during starvation (Hudeček et al., 2020).

Ap_4A, Up_4A, Cp_4A and Gp_4A caps were reported on mRNAs and sRNAs in *E. coli* under disulfide stress (Luciano et al., 2019).

Lysyl-tRNA synthetase synthesizes Ap_3A and Ap_4A *in vitro* (Zamecnik et al., 1966), and aminoacyl-tRNA synthetases are involved in Ap_4 biosynthesis during exponential

growth and heat shock. Overproduction of these synthetases in *E. coli* significantly increases intracellular Ap_4N concentrations (Brevet et al., 1989). It was demonstrated that *E. coli* RNA pol and lysyl-tRNA synthetase can add Np_4A caps *in vitro*. *E. coli* RNA pol prefers to initiate transcription with Ap_4A over ATP or NAD, and Gp_4A is preferred over ATP(Luciano and Belasco, 2020).

Ap_4A caps have been found in human and rat cell lines. The amount of Ap_4A caps does not correlate with free cellular Ap_4A levels. Despite associations with innate immunity, Ap_4A-RNAs do not trigger an immune response in HeLa cells and are recognized as self by the cell. In HEK293T cells, Ap_4A-RNAs are not translated (František Potužník et al., 2024).

1.3.3 NAD cap

NAD is a well-known coenzyme involved in cellular metabolism. It exists in oxidised (NAD+) or reduced (NADH) form, and it has been demonstrated that both can be introduced on RNAs as a cap *in vitro* (Huang, 2003). While NAD cap (Fig. 19) was detected *in vivo* across all organisms and many types of RNAs, NADH was only detected in mitochondrial RNAs.

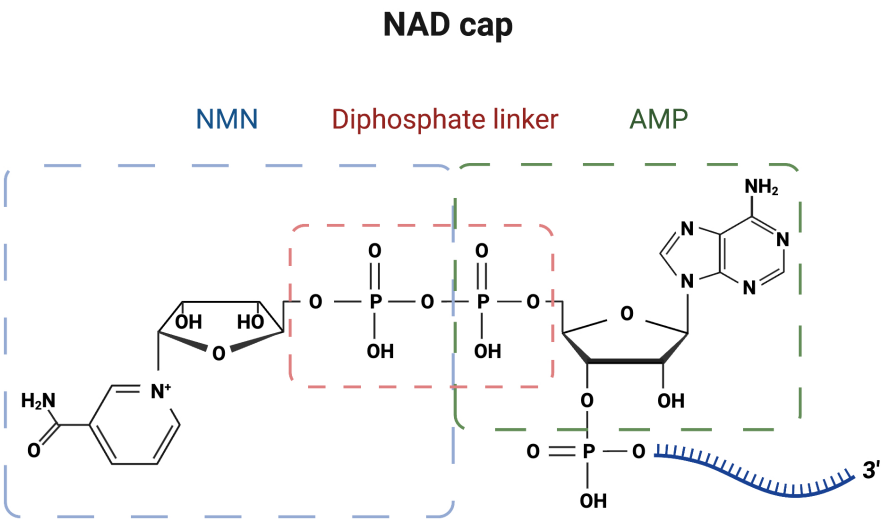


Fig. 19: Chemical structure of NAD non-canonical cap. NAD comprise a nicotinamide mononucleotide (NMN) and AMP moiety linked via two phosphate groups.

NAD caps in prokaryotes

NAD cap was first detected in RNA from *E. coli* and *S. venezuelae* using LC-MS. NAD cap was predominantly detected in sRNAs at a high abundance of 3,000 copies per cell in *E. coli* (Chen et al., 2009).

NAD captureSeq method was developed to identify NAD-RNAs (Cahová et al., 2015; Winz et al., 2017). It was first applied on total RNA from *E. coli* and revealed NAD caps in subsets of sRNAs, including RNAI, GadY, GcvB, ChiX, McaS, and CopA, and sRNA-like 5'-terminal fragments of some mRNAs. NAD was one of the first cap structures found in prokaryotes. NAD cap stabilizes RNA against RppH decapping and RNase E cleavage *in vitro*. NudC, Nudix family member in *E. coli*, is the main NAD decapping enzyme, initiating RNase E-mediated RNA decay (Cahová et al., 2015).

NAD caps were detected in *Bacillus subtilis*, predominantly in mRNAs. One of the identified mRNAs, veg, was used for the *in vitro* assays. RNA pol was able to incorporate NAD into veg-mRNA, and RppH was able to decap NAD-veg-mRNA in the presence of Mn_{2+} but not Mg_{2+} ions. Knockout of RppH did not result in increased NAD capping *in vivo* but led to upregulation of expression of approximately 600 genes. Since there is no NudC ortholog in *B. subtilis*, it is speculated that there might be additional NAD cap-metabolizing enzymes (Frindert et al., 2018).

NAD caps in plants

NAD-RNAs are widespread in *Arabidopsis thaliana*. Using NAD captureSeq were identified NAD caps on mRNAs encoded by the nuclear and mitochondrial genomes but not the chloroplast genome. NAD-RNAs from nuclear genes were spliced, polyadenylated, and associated with actively translating polysomes, indicating they can be translated (Wang et al., 2019b). NAD caps were identified in unopened flower buds and 12-day-old seedlings in mRNAs, pri-miRNAs, snoRNAs, and lncRNAs. Minimal overlap between NAD-mRNAs in flower buds and seedlings suggests developmental-specific regulation of NAD capping (Yu et al., 2021).

NAD tagSeq, using Oxford Nanopore sequencing, was applied on RNA from *A. thaliana*. Most of the detected NAD-RNAs originated from protein-coding genes. These transcripts had shorter 5' UTRs than canonical m^7G-RNAs. However, they were polyadenylated and showed translation start and end signatures, suggesting that they can be translated (Zhang et al., 2019).

The SPAAC-NAD-Seq method improved the sensitivity of NAD captureSeq, detecting a greater number of NAD-RNAs, predominantly of low abundance, in *A. thaliana*. This optimized technique retained full-length sequence information, unlike other methods that captured truncated RNAs (Hu et al., 2021).

NAD caps in yeast

NAD-RNAs were found in *S. cerevisiae* in subsets of snRNAs, mtRNAs and some pre-mRNAs (Walters et al., 2017). mtRNAs enriched in the NAD captureSeq dataset were those not undergoing 5' end processing, suggesting they originate from transcription by mitochondrial RNA pol with the NAD moiety added during transcription initiation (Walters et al., 2017). Up to 60% of mtRNAs were capped with NAD or NADH, with levels correlating to intracellular NAD and NADH levels. This indicates that NAD capping might be a mechanism to sense and respond to alterations in NAD metabolism (Bird et al., 2018).

NAD cap was detected in budding *S. cerevisiae* in short mRNA fragments and at low numbers. The few full-length NAD-mRNAs were not translated by cytosolic ribosomes *in vitro*. Therefore, in contrast with previous works, it was suggested that NAD capping in yeast is undesirable and largely accidental (Zhang et al., 2020).

NAD caps in mammalian cells

In human HEK293T cells, mRNAs, snoRNAs, and scaRNAs were found to possess NAD caps. DXO knockdown led to increased NAD-RNA levels. Unlike m^7G cap, NAD cap was not found to promote translation. NAD-mRNAs and m^7G-mRNAs have similar half-lives, indicating translation deficiency in NAD-RNA is not due to instability (Jiao et al., 2017).

NAD-capQ, a novel detection method, was applied to RNA from *E. coli*, *S. cerevisiae*, and human cells. Deletion of NAD decapping enzymes led to increased NAD-RNA levels. Cellular NAD levels correspondingly influenced NAD-RNA levels (Grudzien-Nogalska et al., 2018). Up to 15% of human mtRNAs have NAD caps, higher than other RNAs, likely due to +1A promoters in mitochondrial genes. NADH caps were found only on mitochondrial RNAs, suggesting distinct roles for NAD and NADH caps in mitochondria (Bird et al., 2018).

The ONE-seq method revealed that as mice age, cellular NAD and NAD-RNA levels

decrease. However, some NAD-RNAs were more enriched in aged mouse livers, for instance RNAs involved in metabolic pathways and oxidation-related processes. On the other hand, the NAD-RNAs enriched in the young mice are involved in pathways related to nucleotide excision repair and RNA splicing (Niu et al., 2023).

Biosynthesis of NAD-RNAs

It has been proposed that NAD capping occurs after transcription initiation by one or more enzymes that participate in NAD biosynthesis in a manner analogous to the addition of m^7G cap (Luciano and Belasco, 2015). Initially, NAD incorporation into RNA was observed *in vitro* with T7 RNA pol, but not with *E. coli* RNA pol (Chen et al., 2009). However, it was later shown that NAD can serve as a non-canonical initiating nucleotide (NCIN) during transcription initiation for both bacterial RNA pol and eukaryotic RNA pol II *in vitro* and *in vivo*. Previous discrepancies were attributed to high ATP concentrations in earlier assays, which out-competed NAD (Bird et al., 2016). This was confirmed by competition assays showing that NAD capping efficiency depends on the NAD/ATP ratio (Bird et al., 2016; Vvedenskaya et al., 2018).

The efficiency of NAD capping is influenced by promoter DNA sequences at and upstream of TSS. NAD-mediated initiation occurs at promoters with A at TSS +1 (+1A promoters), while it does not occur at +1G promoters. The nicotinamide moiety can form a "pseudo-base pair" with template-strand T at position -1 where two H-bonds of nicotinamide pair with Watson-Crick H-bonding atoms of the template T (Bird et al., 2016; Frindert et al., 2018). *S. cerevisiae* and human mitochondrial RNA pol can cap with NAD and NADH *in vitro* (Bird et al., 2018).

CapZyme-seq is a high-throughput sequencing method using NCIN-decapping enzymes NudC and Rai1 developed to identify the promoter consensus sequences. The promoter sequence from -3 to +4 for NAD capping was determined as follows: HRR<u>A</u>SWW (+1 A is underlined). The H at position -3 means that capping efficiency is higher for A, T, and C than for G; R at positions -2 and -1 represents G or A; at position +2, the capping efficiency is higher for G and C than for A and T, yielding the consensus S; and W stands for A or T (Vvedenskaya et al., 2018).

snoRNAs and scaRNAs are intron-derived and exonucleolytically processed RNAs. Therefore, NAD-capping cannot be explained only by NCIN-mediated transcription. Mammalian cells likely possess another, yet unknown, mechanism of NAD-capping that

occurs after transcription initiation (Jiao et al., 2017).

NADcapPro and circNC were applied to RNA from *S. cerevisiae*. Full-length mRNAs and snoRNAs were identified to bear NAD caps. NAD-RNAs were associated with mitochondrial ribosomes, indicating potential translation in mitochondria. TSS for NAD-mRNAs and m^7G-mRNAs were found to differ, suggesting post-transcriptional NAD capping of snoRNAs and mRNAs (Sharma et al., 2023).

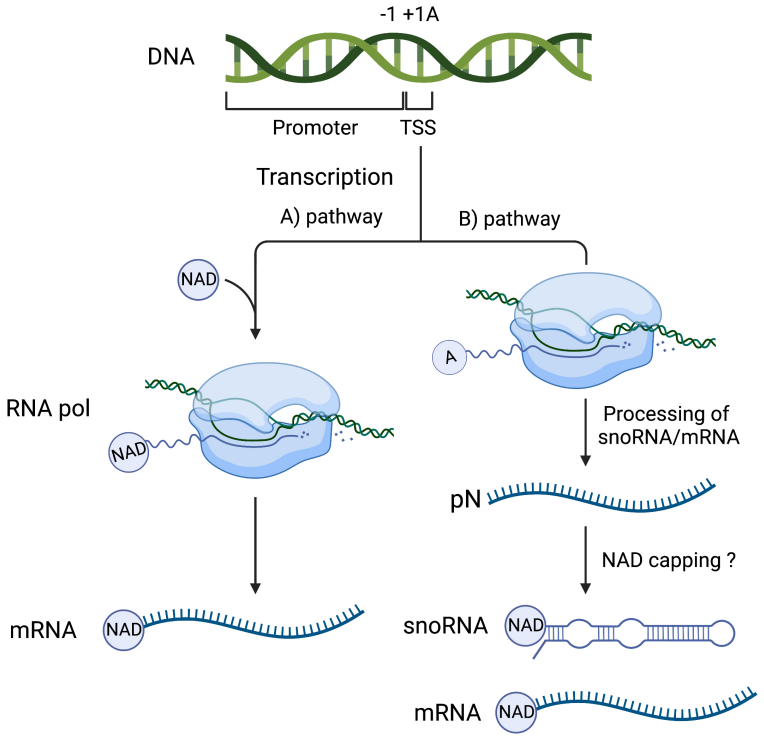


Fig. 20: Summary of NAD capping mechanism. A) RNA can be capped by NAD co-transcriptionally by RNA pol. B) Another possible pathway is the post-transcriptional addition of the NAD cap. This mechanism is less studied, and enzymes involved in this pathway are not known.

NAD decapping

The enzymes responsible for NAD decapping in bacteria, RppH and NudC, belong to the same family as the m^7G decapping enzyme Dcp2. The Nudix hydrolases are phosphohydrolases that cleave nucleotide diphosphates linked to a moiety X. NudC was first described in *E. coli* as hydrolyze, which processes NAD or NADH into nicotinamide mononucleotide (NMN or NMNH) and AMP (Frick and Bessman, 1995).

RppH removes pyrophosphate from the 5' end of pppRNA and ppRNA, leading to RNase E-dependent degradation in *E. coli* or RNase J1-dependent degradation in *B. subtilis* (Bischler et al., 2017; Cahová et al., 2015; Frindert et al., 2018). The NAD cap

in bacteria, like the m^7G cap in eukaryotes, protects RNAs. NAD-RNAI is significantly more stable against RNase E and RppH compared to pRNAI and pppRNAI (Cahová et al., 2015). Conversely, NudC decaps NAD-RNAI and NADH-RNAI *in vitro* but is inactive against pppRNAI. After NudC decapping, pRNAI is degraded by RNase E. *In vivo*, NudC deletion in *E. coli* doubles the amount of NAD-RNAI (Cahová et al., 2015). NudC prefers NAD-RNAs over NAD or NADH, suggesting NAD-RNAs are its primary substrates (Höfer et al., 2016). In *B. subtilis*, RppH removes NMN from veg NAD-mRNA in the presence of Mn^{2+}, but this reaction is slower than dephosphorylation of veg pppRNA. RppH depletion does not affect NAD-RNA levels, indicating the presence of other NAD-decapping enzymes (Frindert et al., 2018).

A. thaliana contains AtDXO1, which functions as the decapping enzyme of NAD-RNA. Knockdown of AtDXO1 led to growth retardation and pale-green pigmentation. DXO lacking plants accumulated several RNAs, including quality control siRNAs from improperly degraded mRNAs (Kwasnik et al., 2019). AtDXO1 has both NAD-RNA decapping and 5'-3' exonuclease activities, but does not hydrolyze the m^7G cap in *A. thaliana* (Pan et al., 2020).

In *S. cerevisiae*, three NAD decapping enzymes exist: Rai1, Dxo1, and Npy1. Rai1 and Dxo1 are orthologs of DXO, while Npy1 is a homolog of bacterial NudC. Deletion mutants for all combinations revealed that these enzymes target different NAD-RNA populations in their specific cellular locations. A hierarchical order of NAD-RNA processing was established, with Rai1 acting first. Rai1 deletion had the most significant negative impact on growth compared to wild-type and other mutants (Zhang et al., 2020).

Recently, additional NAD decapping enzymes were identified in yeast: Xrn1 and Rat1 (Xrn2 in mammals). *In vitro*, Xrn1 and Rat1 hydrolyze NAD-RNAs, releasing intact NAD and degrading the RNA. The Xrn1 knockout strain showed a 1.5-fold increase in total NAD caps compared to the wild-type. While nuclear-encoded NAD-RNAs levels were unchanged, mitochondrially encoded NAD-RNAs were significantly increased. Super-resolution microscopy confirmed that Xrn1 is associated with mitochondria, suggesting it directly modulates mitochondrial NAD-RNAs (Sharma et al., 2022).

The primary deNADing enzyme in human cells is DXO, which decaps NAD-RNAs by hydrolyzing the phosphodiester bond between the first and second encoded nucleotides, producing pRNA (Fig. 21). NAD-RNAs are less stable than m^7G-RNAs and pppRNAs in control HEK293T cells indicating that NAD cap promotes RNA decay. DXO preferentially

targets NAD-mRNAs, NAD-snoRNAs, and NAD-scaRNAs (Jiao et al., 2013, 2017).

Nudt12 was initially identified as an NAD/H diphosphatase with a substrate preference for the reduced form (Abdelraheim et al., 2003). It is a homolog of bacterial NudC and hydrolyzes NAD-RNA to produce NMN and 5' pA-RNA. Comparison of NAD-RNA from Nudt12 KO cells and DXO KO cells showed that Nudt12 and DXO have a pronounced preference for different sets of RNAs (Grudzien-Nogalska et al., 2019).

Nudt16 is an NAD decapping enzyme, hydrolyzing both free NAD and NAD-RNA. Nudt16 KO cells showed a 1.3-fold increase in total NAD-RNAs, indicating its role in regulating NAD-RNA levels (Sharma et al., 2020). Recently, human Xrn1 was also identified as a deNADing enzyme, releasing intact NAD and degrading the RNA *in vitro* (Sharma et al., 2022).

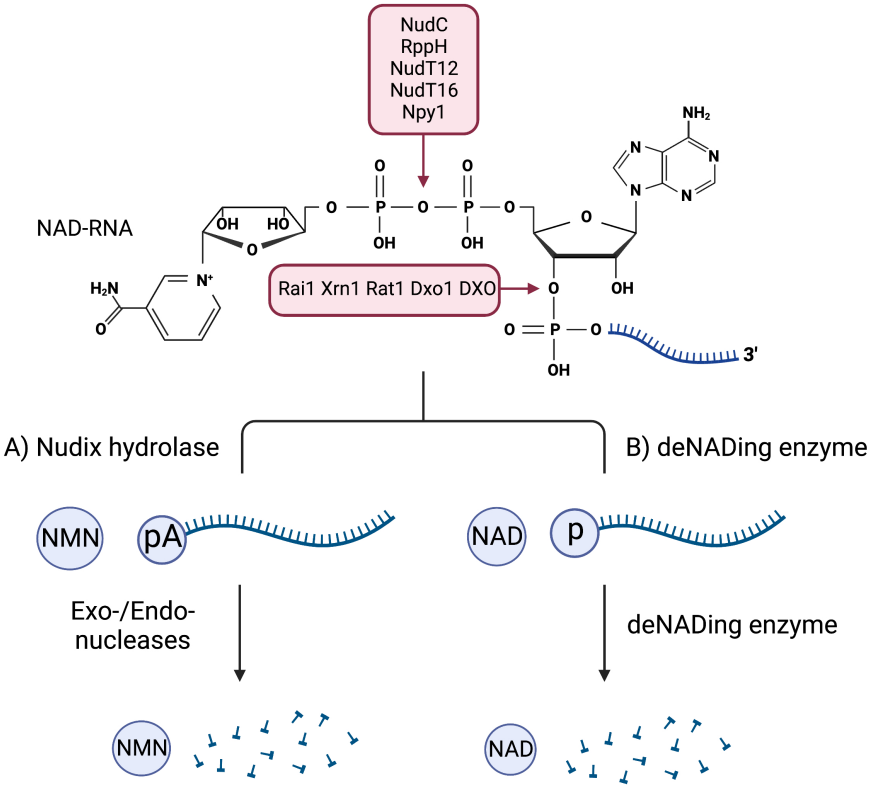


Fig. 21: Schematic overview of the known pathways of NAD cap removal. The cleavage sites of the Nudix hydrolases and the DXO/Rai1 enzymes are indicated with the red arrowheads. A) Nudix hydrolases remove the NMN moiety, and the resulting RNA is degraded by exonucleases and endonucleases. B) DXO/Rai1 enzymes possess deNADing and exoribonuclease activity. The cleavage site for the deNADing activities of Xrn1 and Rat1 is the same as that of DXO/Rai1 enzymes.

1.4 Viruses

Viruses are small obligate intracellular parasites that can multiply only in living cells of animals, plants, or bacteria (bacteriophages). Viruses consist of genetic material, either DNA or RNA, surrounded by a protein capsid, and some possess an additional lipid envelope derived from the host cells. A complete virus particle is called a virion. Its main function is to deliver its genome into the host cell. The viruses use the host cell's machinery to express the viral genome and to make new virions (Johnson, 2019).

1.4.1 Retroviruses

Retroviruses are RNA viruses that replicate via reverse transcription of RNA into linear double-stranded DNA. Retroviruses permanently alter the genetic landscape of the infected cell by DNA integration into the host genome (Johnson, 2019). Retroviruses are fundamentally simple systems to study and many discoveries had been made on them. Therefore, we decided to use retroviruses as a model system to investigate the RNA modifications. HIV-1 infection is responsible for decrease in cellular NAD level (Murray et al., 1995) and for this reason we focused our research on HIV-1.

1.4.2 HIV-1

HIV-1 belongs to the lentivirus subfamily of retroviruses that primarily infect cluster of differentiation 4+ T-lymphocytes (CD4+ T cells) and macrophages. The loss of CD4+ T cells, which are white blood cells essential for immune response, causes acquired ammune deficiency syndrome (AIDS) (Swanstrom and Coffin, 2012). HIV-1 is an enveloped virus that contains two copies of single-stranded RNA, 7SL RNAs, tRNAs, and integrase, protease and reverse transcriptase enzymes (Fig. 22). The capsid shell comprises around 200 capsid protein hexamers and 12 capsid protein pentamers (Zhao et al., 2013). HIV-1 replicates by modulating the host transcription and translation machinery. HIV-1 acquired a large array of regulatory elements that target key transcription and translation factors to ensure its expression (Roebuck and Saifuddin, 1999).

1.4.3 HIV-1 life cycle

The key stages of the HIV-1 life cycle are viral entry, reverse transcription, integration, transcription, translation, assembly, and budding (Fig. 23). HIV-1 contains an envelope

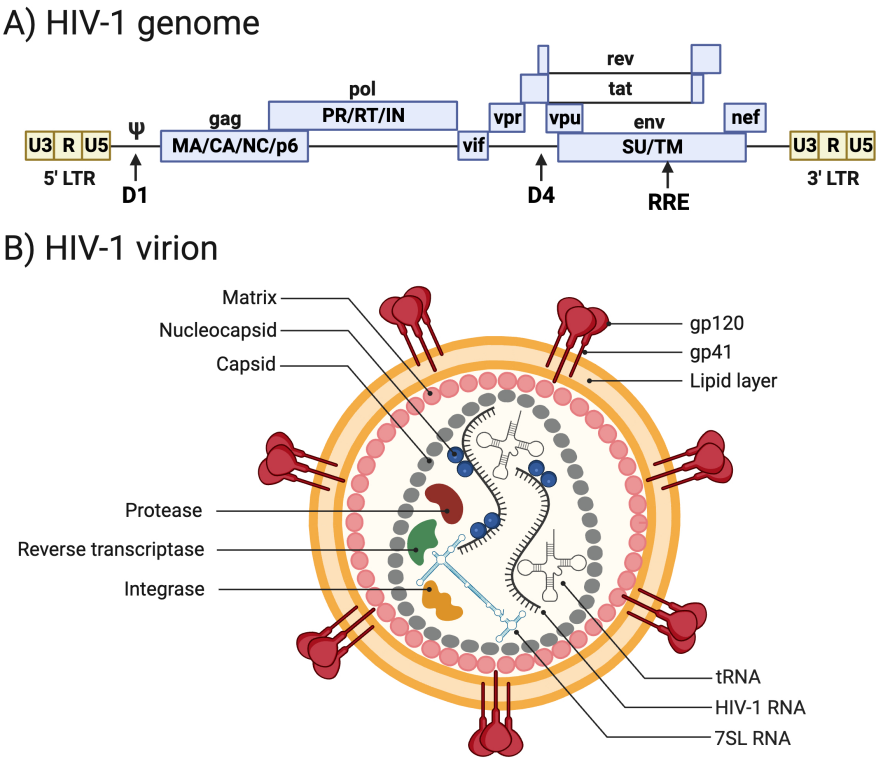


Fig. 22: Schematic representation of the HIV-1 genome and a mature HIV-1 particle. A) Diagram of the composition of the HIV-1 genome shows relative locations of the gag, pol, env, vif, vpr, vpu, nef, tat, and rev open reading frames. Gag contains the matrix (MA), capsid (CA), nucleocapsid (NC) and p6 domains. Env is composed of the surface (SU, gp120) and the transmembrane (TM, gp41) domains. Under the pol gene are indicated the protease (PR), reverse transcriptase (RT), and integrase (IN) coding regions. The γ indicates the position of the RNA packaging signal, and 5' and 3' LTR are shown with their U3, R, and U5 regions. B) Schematic representation of a mature HIV-1 particle. It consists of two genomic RNAs covered in nucleocapsid.

glycoprotein, Env, produced as a precursor gp160, which is trimerized and then cleaved into gp120 (receptor-binding) and gp41 (fusion). The virion attaches to cells via gp120 binding to CD4 and coreceptors (CCR5 and CXCR4), leading to structural changes in gp41 and fusion with the host cell membrane, allowing viral capsid entry (Chen, 2019).

Until very recently, it has been believed that the release of the enzymes from viral capsid occurs in the cytosol. However, new evidence indicates that the capsid enters the cell nucleus and releases the enzymes there. The precise molecular mechanism of capsid nuclear import and uncoating are not known yet (Müller et al., 2022).

The HIV-1 gRNA is converted into double-stranded complementary DNA (cDNA) by RT, which has DNA polymerase and RNase H activities. HIV-1 uses host $tRNA_3^{Lys}$ as a primer for DNA synthesis (Kleiman, 2002). HIV-1 co-packs around 770 copies of host tRNAs, among which Lys tRNAs are the most abundant (Šimonová et al., 2019). RNA in

the newly generated RNA-DNA duplex is degraded apart from the purine-rich sequence, which serves as primers for the second DNA strand (Hu and Hughes, 2012). The final DNA is longer than the RNA genome because it contains two LTRs (U3 and U5) derived from the RNA's 3' and 5' ends (Hu and Hughes, 2012).

The viral cDNA is integrated into the host genome via integrase, involving 3' processing and strand transfer. Integrase removes a pGT dinucleotide at the 3' end of each LTR and mediates a nucleophilic attack by the 3' hydroxyl residues of the viral DNA on phosphodiester bridges in the target DNA. 3' end of viral DNA is ligated to the host DNA and the flanking end is cleaved and the gap is then repaired by host DNA pol and DNA ligase (Lesbats et al., 2016).

The integrated HIV-1 genome is about 9,700 bp and its transcription by host RNA pol II generates full-length pre-mRNAs (Roebuck and Saifuddin, 1999). Transcription, initiated from a single TSS at the 5' LTR, is enhansed by the Tat (Trans-activator of transcription) protein binding to TAR (transactivation response) element. HIV-1 RNAs are m^7G capped and polyadenylated. Caps in some unspliced or partially spliced RNAs are hypermethylated to TMG-RNAs (Freed, 2015; Yedavalli and Jeang, 2010).

Most viral pre-mRNAs remain unspliced to encode Gag and Gag-Pol polyproteins and serve as gRNA. Fully spliced RNAs encode Tat, Rev, and Nef, while partially spliced RNAs encode Env, Vif, Vpr, and Vpu (Kuzembayeva et al., 2014; Purcell and Martin, 1993).

U1 snRNA binding is essential for HIV-1 env mRNA stability, independent of splicing (Kammler et al., 2001). U1 snRNA influences the interactions at the 3' end of the introns, contributing to the alternative selection of the 3' splice sites (Hoffman and Grabowski, 1992; Robberson et al., 1990). Modifications in the sequence of U1 snRNA inhibit viral protein expression and lead to reduced HIV-1 replication. The use of U1 snRNA mutants was even suggested as a possible strategy for HIV therapeutic (Knoepfel et al., 2012; Mandal et al., 2010; Sajic et al., 2007). HIV-1 can tolerate a wide range of aberrant splicing but excesivelly high amount of splicing results in the loss of the gRNA and decreased HIV-1 production (Bilodeau et al., 2001).

Completely spliced transcripts are exported from the nucleus by a regular cellular mRNA export pathway. The partially spliced and unspliced HIV-1 mRNAs are exported via Rev protein binding to RRE, circumventing the usual cellular export restrictions. The transcripts retained in the nucleus are mostly subjected to nuclear degradation or spliced

to completion (Meyer and Malim, 1994).

Unspliced HIV-1 mRNA can be translated by classical cap-dependent or non-canonical mechanisms (leaky scanning, frameshifting, shunt, and cap-independent mechanisms). The Gag precursor polyprotein contains the matrix, capsid, nucleocapsid and p6 domains. The GagPol precursor polyprotein contains matrix, capsid, nucleocapsid, protease, reverse transcriptase and integrase domains. The newly synthesized viral proteins and RNAs assemble at the plasma membrane to form a new viral particle. Gag recruits two gRNAs by binding to the packaging signal. The 5' UTR of viral gRNA contains primer-binding site element, which binds Lys tRNA to be packaged into virions. Gag then multimerises into a multiprotein until it reaches the plasma membrane (Freed, 2015).

Env recruits the endosomal sorting complex required for transport (ESCRT) I. After assembly, a membrane scission reaction releases the nascent particle from the cell surface. This step is mediated by the cellular ESCRT III machinery, which is hijacked by Gag. After the release, the viral particle undergoes a maturation step, during which the Gag polyprotein is cleaved by a viral protease, forming the conical capsid core (Freed, 2015).

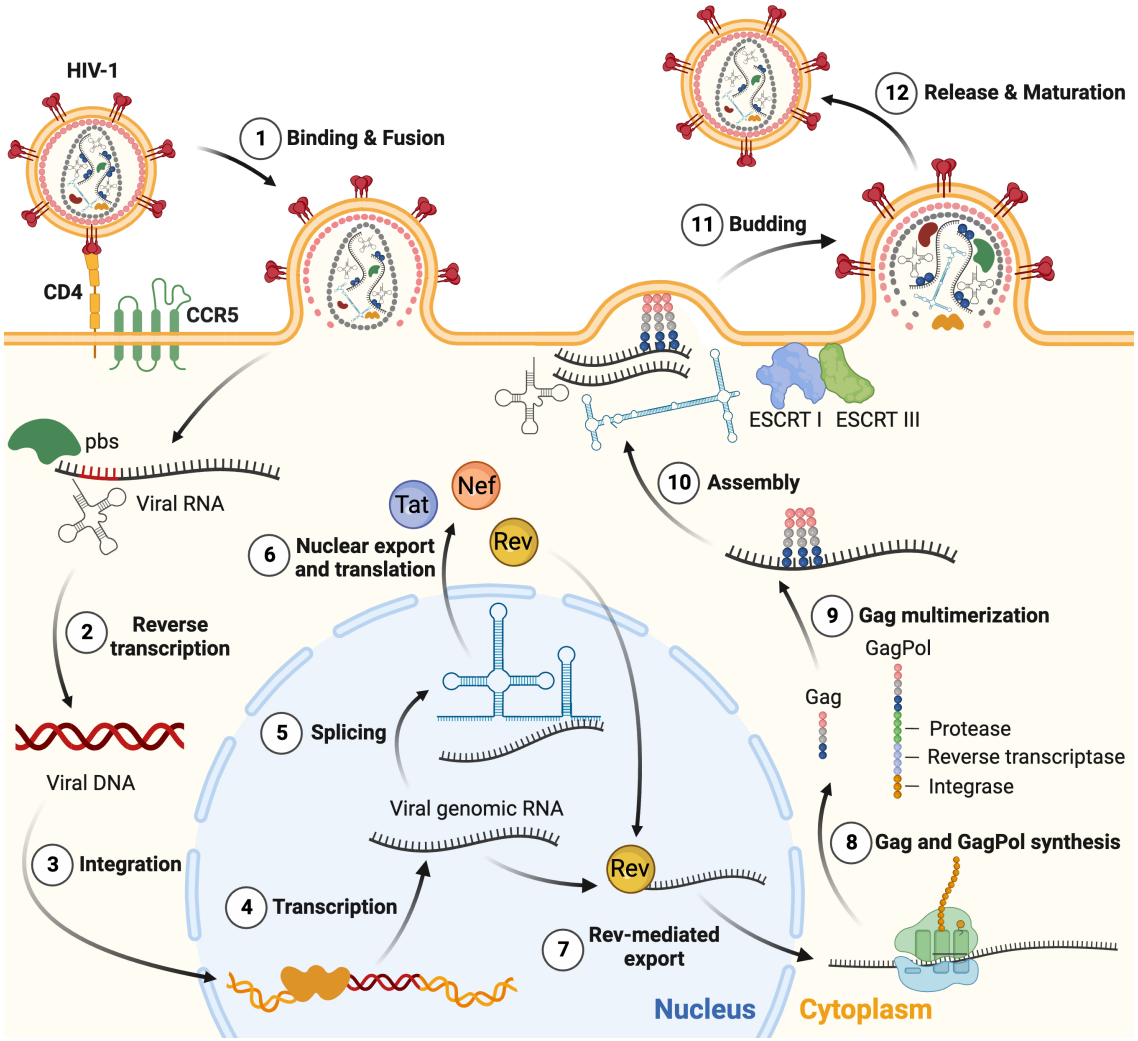


Fig. 23: Schematic overview of the HIV-1 replication cycle. 1) The infection begins when HIV-1 binds to the CD4 receptor via gp120/gp41 complex. The viral and cellular membranes fuse, and the viral core enters the host cytoplasm. 2) Viral RNA is converted into DNA by the viral RT using tRNA as a primer. 3) Integrase catalyzes the insertion of DNA into the host cell chromosome. 4) The transcription of viral gRNA is mediated by host RNA pol II, yielding full-length pre-mRNAs. 5) Some RNAs undergo splicing by the host spliceosome, resulting in partially and fully spliced RNAs. 6) Fully spliced RNAs are exported from the nucleus and translated by host machinery into Rev, Tat, and Nef. 7) Rev is imported into the nucleus and mediates the nuclear export of intron-containing transcripts. 8) The unspliced viral RNA is translated into the Gag or GagPol precursor polyprotein. 9) Gag recruits the viral genomic RNAs, begins to multimerise and reaches the plasma membrane. 10) The ribonucleoprotein complex binds to the plasma membrane and interacts with other host RNAs for packaging into the viral particle. 11) Assembly and budding continue until all components are incorporated into the budding particles. Env and Gag recruit ESCRT I and ESCRT III for the membrane scission reaction. 12) The viral protease cleaves the Gag and GagPol polyprotein precursors shortly after virus release, forming the conical capsid core.

1.4.4 NAD role in HIV-1 infection

NAD is one of the most common metabolites in the human body and plays a vital role in redox reactions and signaling. Most NAD molecules are recycled from precursors such

as nicotinamide (NAm), NMN, nicotinamide riboside (NR) and niacin (NA) (Xie et al., 2020). It has been reported that HIV-1 infection causes a decrease in cellular NAD and NADH levels, called pellagra. It varies with viral load and HIV strain. NAm supplement increased intracellular NAD levels in HIV-1 infected cells, so the depletion can be reversed (Murray et al., 1995). In addition, NAm inhibits HIV-1 infection in cell culture in a dose-dependent manner (Murray and Srinivasan, 1995).

NAD acts as a cofactor for several essential enzymes, including cyclic ADP ribose hydrolases (CD38), sirtuins (Sirts), and Poly (ADP-ribose) polymerases (PARPs). CD38 is a membrane glycoprotein and one of the major consumer of NAD in mammals. It is involved in cell signaling and immune functions, and catalyzes the synthesis of NAm and ADP-ribose using NAD. Its expression is highly upregulated in response to viral infections, and therefore, it was studied for potential anti-HIV effects. It was found that CD38 down-modulates the binding of HIV-1 gp120 to CD4 receptor by directly binding to gp120 (Savarino et al., 2003).

Sirts are a class of deacetylases with antiviral properties requiring NAD as a cofactor. Nuclear Sirts 1 and 6, and mitochondrial Sirts 3 and 4 inhibit glycolysis and promote mitochondrial function. Additionally, Sirt 3 inhibits generation of reactive oxygen species (ROS). NAD removal by CD38 inactivates Sirts, which leads to increased mitochondrial production of ROS, thus reduced cell viability (Guarente, 2014).

PARPs catalyze ADP-ribosylation of target macromolecules by transfering the ADP-ribose group from NAD. PARPs play an essential role in the regulation of cell death by contributing to the apoptosis-inducing factor-dependent cell death. In response to infection and the oxidative stress-induced DNA damage, the amount of PARPs increases and their activity consumes large amounts of NAD. Sustained PARP activation can cause NAD depletion and subsequent cell death (Cantó et al., 2015).

NAD plays a critical role in many physiological processes, and its levels impact energy metabolism, DNA repair, epigenetic modification, inflammation, circadian rhythm and stress resistance. NAD has been proven to be a nucleotide analogue in RNA capping. NAD amount fluctuates in response to cellular stresses and multiple viruses were found to decrease NAD concentration (Xie et al., 2020). Therefore, we decided to study the effect of HIV-1 infection on NAD capping.

2 Aim of the work

To date, more than 170 different types of RNA modifications have been identified (Boccaletto et al., 2022). NAD as a cap structure was detected by LC-MS analysis in *E. coli* and *S. venezuelae* in 2009 (Chen et al., 2009). Since then, it has been found in bacteria, human cells, plants, and archaea (Cahová et al., 2015; Gomes-Filho et al., 2023; Jiao et al., 2017; Wang et al., 2019b). However, little is known about the function of NAD cap in the context of host-pathogen interactions. HIV-1 uses RNA modifications to regulate its gene expression during infection and replication (Fukuda et al., 2021; Ringeard et al., 2019). Furthermore, it has been proven that HIV-1 infection reduces the total amount of cellular NAD (Murray et al., 1995). Therefore, we aimed to study the relationship between viral infection and NAD capping.

The objectives of this study:

1. Prepare sufficient quantities of HIV-1 virions for subsequent RNA isolation.
2. Generate cell lines with DXO overexpression (OE) and DXO knockdown (KD).
3. Prepare sufficient quantities of control MT-4 cells and MT-4 cells under all tested conditions for RNA isolation.
4. Employ NAD CaptureSeq method for the detection of NAD-RNAs in control and HIV-1 infected cells.
5. Assess the levels of free NAD in both control and HIV-1 infected MT-4 cells.
6. Verify the presence of NAD-capped RNA in HIV-1 virions via LC-MS analysis.
7. Measure the amount of NAD caps in control, HIV-1 infected, DXO OE, and DXO KD MT-4 cells using LC-MS.
8. Measure the NAD cap content on U1 snRNA using LC-MS.
9. Compare the stability of the complex formed by HIV-1 RNA with TMG-capped and NAD-capped U1 snRNA.
10. Evaluate the infectivity of HIV-1 in control, DXO OE, and DXO KD MT-4 cell.
11. Examine the splicing pattern in control and infected MT-4 cells under various conditions.

3 Experimental section

3.1 Instruments and materials

3.1.1 Instruments

- NanoDrop OneC UV-Vis Spectrophotometer, Thermo Fisher
- LightCycler 480 II, Roche
- SYNAPT G2, Waters
- Xevo G2–XS QTof mass spectrometer, Waters
- Typhoon biomolecular imager, GE HealthCare
- CD Spectropolarimeter J-815, Jasco
- EnVision plate reader, Perkin Elmer
- Tecan microplate reader, Schoeller
- Azure c600 imager, Azure Biosystems
- 4200 TapeStation System, Agilent
- Qubit Fluorometer, Thermo Fisher

3.1.2 Chemicals

Chemicals were obtained from Roth, Jena Bioscience, Merck or Roche at molecular biology quality or above if not stated otherwise. Enzymes were purchased from ThermoFisher Scientific, NEB or Epicenter. Lentiviral vectors for transient transfection (shRNA knockdown DXO vector VB900039-9959kyv or DXO gene expression vector VB900001-2728fsj) were produced by VectorBuilder. Antibodies were purchased from Proteintech Europe and Jackson ImmunoResearch. Cell culture plasticware and cellulose-acetate filters were obtained from VWR. Oligonucleotides were synthesised by Generi Biotech. RT-qPCR Kit (Luna Universal One-Step kit) was acquired from New England BioLabs.

3.1.3 Oligonucleotides

Tab. 1: Oligonucleotides used in this study. Rev: reverse, fwd: forward, prm: primer. Promoter sequences and indexes are underlined.

Name	Sequence (starting at 5' end)
RNA 5' adaptor	rApp-CNNNNNAGATCGGAAGAGCACACGTCTG-(C3)
Rev transcription prm	CAGACGTGTGCTCTTCCGAT
cDNA anchor sense	p-CAGATCGGAAGAGCGTCGTGT-(C3)
cDNA anchor antisense	ACACGACGCTCTTCCGATCTGGG
NAD Seq Fwd PCR prm	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC GCTCTTCCGATCT
NAD Seq Rev PCR prm 1	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCA GACGTGTGCTCTTCCGATCT
NAD Seq Rev PCR prm 2	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCA GACGTGTGCTCTTCCGATCT
NAD Seq Rev PCR prm 3	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCA GACGTGTGCTCTTCCGATCT
NAD Seq Rev PCR prm 4	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCA GACGTGTGCTCTTCCGATCT
NAD Seq Rev PCR prm 5	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCA GACGTGTGCTCTTCCGATCT
NAD Seq Rev PCR prm 6	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCA GACGTGTGCTCTTCCGATCT
NAD Seq Rev PCR prm 7	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCA GACGTGTGCTCTTCCGATCT
NAD Seq Rev PCR prm 8	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCA GACGTGTGCTCTTCCGATCT
SNORD3G Fwd	GCGTTCTCTCCTGAGCATGA
SNORD3G Rev	ACCACTCAGACTGTGTCCTCT
SNORD3B Fwd	TCTGAACGTGTAGAGCACCG
SNORD3B Rev	CTCTCCCTCTCACTCCCCAA
SNORA50A Fwd	GCACTGCCTTTGAACCTGATG
SNORA50A Rev	GAGCAGTTCAGTTGAGAGCTG
SNORD102 Fwd	GACTGTTTTTTTTGATTGCTTG
SNORD102 Rev	AGCCGGTGAAATGTGT
U1 Fwd	TGGCAGGGGAGATACCATGA
U1 Rev	TACGCAGTCGAGTTTCCCAC
U4 ATAC Fwd	CCATCCTTTTCTTGGGGTTGC
U4 ATAC Rev	GCAAAAGCTCTAGTTGATGCGG
U5E Fwd	AGTTTCTCTTCATATCGCA
U5E Rev	AAAATGCAAGACCTCT
U7 Fwd	GTGTTACAGCTCTTTTAG
U7 Rev	TCCGGTAAAAAGCCAG
Fwd prm for U1 template	TAATACGACTCACTATTAGGCTTACCTGGCAGGGGAG
Rev prm for U1 template	CAGGGGAAAGCGCGAACGCAG
U1 snRNA 20 nt	TAATACGACTCACTATTAGGCTTACCTGGCAGGGGAG
HIV-1 mRNA 20 nt	TAATACGACTCACTATAGGGTAAGCCGTACATGTAAT
Biotinylated U1 probe	GCAATGGATAAGCCTCGCCC

3.2 Methods

3.2.1 Cell culture

The human CD4+ T-cell line MT-4 (NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH from Dr. Douglas Richman) and the human embryonic kidney HEK293T cell line (American Type Culture Collection, LGC Standards, UK) were cultured under standard conditions at 37 °C under a humidified (>90%) atmosphere of 5% CO_2/95% air. The MT-4 cell line was cultured in RPMI 1640 medium with L-Glutamine and 25 mM Hepes supplemented with 10% FBS, penicillin (100 U/mL) and streptomycin (100 µg/mL). The HEK293T cell line was cultured in Dulbecco's Modified Eagle Medium (DMEM) high glucose supplemented with 10% FBS, penicillin (100 U/mL) and streptomycin (100 µg/mL). The cells were passaged every 3-4 days.

3.2.2 HIV-1 culture

HIV-1 is an infectious agent and working with it requires a safe environment. All experiments with MT-4 cells (both infected and uninfected by HIV-1) were conducted in the Biosafety Level 3 (BSL3) laboratory. Once the samples were lysed with RNAzol, they were safely transported to another laboratory with lower safety requirements.

MT-4 cells were infected with a cell-free HIV-1 strain NL4-3 at an MOI of 1, which was generated by transient transfection of HEK293T cells with a pNL4-3 plasmid (obtained through NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH from Dr. Malcolm Martin). The infected cultures were subsequently expanded by co-cultivation: 48 h post-infection, cell culture supernatants containing viral particles and infected cells were added to naive MT-4 cells ($5*10^5$ cells per mL) at a ratio of 1:9. The co-culture was synchronised by three successive additions of infected culture supernatant to uninfected MT-4 cells ($5*10^5$ cells per mL, the ratio of 1:9, 27 h interval). For the NAD repletion experiments, the RPMI medium was supplemented with 10 mM of nicotinamide (NAm), the NAD precursor, as described in Murray and Srinivasan (1995). The cell cultures were incubated for an additional 40 hours and then harvested. MT-4 cells infected and uninfected were collected by centrifugation (225 × g, 5 min, at 20 °C). Cells pellets were washed with phosphate-buffered saline and cells were lysed with RNAzol reagent. Virus-containing supernatants were filtered with 0.45 µm pore size cellulose-acetate filter and stored at −80 °C for further experiments. HIV-1 titer was measured by 10-fold serial infection of

TZM-bl cells (NIH AIDS Reagent Program, Division of AIDS, NIAID, NIH from Dr. John C. Kappes and Dr. Xiaoyun Wu) in triplicate, calculated by Reed-Muench method and expressed as 50% tissue culture infectious dose.

3.2.3 Sucrose cushion viral RNA isolation

HIV-1 particles were concentrated from the cleared culture medium by centrifugation through a cushion of 20% (wt./wt.) sucrose in phosphate-buffered saline (90,000 × g, 90 min, 4 °C). The pellet was resuspended in RNase/DNase buffer (Tris HCl 100 mM, $MgCl_2$ 25 mM, $CaCl_2$ 25 mM) with DNase I (10 U/mL), RNase I (200 U/mL) and RNase A (20 mg/mL) and incubated 2 h at 37 °C. RNase/DNase treatment was stopped by adding RNAzol. These samples were utilized for my research project and for the project analysing m^1A modification in HIV-1 (Šimonová et al., 2019). A mock medium was prepared in the same way from uninfected MT-4 cells. The presence of viral RNA was confirmed on LightCycler 480 II (Roche) by Luna Universal One-Step RT-qPCR Kit according to the manufacturer's protocol.

3.2.4 HIV-1 RNA digestion and LC-MS analysis

HIV-1 RNA was digested by Nuclease P1 (1 U/µg of RNA) in 50 mM ammonium acetate buffer (pH 4.5) or mock treated with water (negative control) for 1 h at 37 °C. Digested samples were filtered with Microcon-10 kDa columns, and the filter was washed once with 100 µL of water. The wash and the filtrate were combined, dried using SpeedVac, and resuspended in 10 L of LC-MS water. For each sample, 8 L of the dissolved material was injected. LC-MS was performed using a Waters Acquity UPLC SYNAPT G2 instrument with an Acquity UPLC BEH Amide column (1.7 µm, 2.1 mm × 150 mm, Waters). The mobile phase A consisted of 10 mM ammonium acetate (pH 9), and the mobile phase B of 100% acetonitrile. The flow rate was kept at 0.25 mL/min, and the mobile phase composition gradient was the following: 80% B for 2 min; linear decrease to 68.7% B over 13 min; linear decrease to 5% B over 3 min; maintaining 5% B for 2 min; returning linearly to 80% B over 2 min. For the analysis, electrospray ionisation (ESI) was used with a capillary voltage of 1.80 kV, a sampling cone voltage of 20.0 V, and an extraction cone voltage of 4.0 V. The source temperature was 120 °C) and the desolvation temperature 550 °C), the cone gas flow rate was 50 L/h and the desolvation gas flow rate 250 L/h. The detector was operated in negative ion mode.

3.2.5 Preparation of DXO KD and DXO OE cells from lentiviral vectors

The lentiviral vectors were generated using HEK293T cells by a transient three-plasmid transfection as previously described (Soneoka et al., 1995). The cells were transfected with 1 µg of the pMDG envelope plasmid, 1 µg of the pR8.2 packaging plasmid, and 1.5 µg of the lentiviral vector (shRNA knockdown DXO vector VB900039-9959kyv or DXO gene expression vector VB900001-2728fsj, VectorBuilder) in the mixture of Opti-MEM (200 µL) and X-tremeGENE™ HP DNA Transfection Reagent (18 µL). Vector particles were collected from medium 48 h and 72 h post-transfection. MT4 cells ($5*10^6$) were infected with 1 mL of collected media for DXO OE and 2 mL for DXO KD, incubated for 1 hour and transferred into flasks with 5 mL of RPMI media. After 24 hours, the cells were centrifuged (225 × g, 5 min, at 20 °C), washed with phosphate-buffered saline, and resuspended in 10 mL of RPMI with 10 µL of puromycin (2mg/mL) as selection marker.

3.2.6 NAD/NADH measurement

NAD total levels in cells were assessed using the NAD/NADH Quantification Kit according to the manufacturer's instructions (Sigma-Aldrich). Control, HIV-1 infected, DXO OE, and HIV-1 infected DXO OE MT-4 cells, all with or without NAm supply, were washed with cold phosphate-buffered saline and $2*10^5$ of each was pelleted (centrifuging at 2,000 r.p.m. for 5 min). Cells were lysed with NADH/NAD Extraction Buffer by 2 cycles of freeze/thawing (20 min on dry ice followed by 10 minutes at room temperature (r.t.)). Insoluble material was removed by centrifuging (13,000 × g, 10 min) and lysates were deproteinised by filtering through a 10 kDa cut-off spin filters. 50 µL of the extract together with NADH standards (0 (blank), 20, 40, 60, 80, and 100 pmol/well) were transferred to a 96-well flat-bottom plate, all in duplicates. NAD was converted to NADH using NAD Cycling Enzyme Mix and the mix was incubated for 5 min at r.t. 10 µL of NADH Developer was added into each well, and the reaction was incubated for 4 hours at r.t. The absorbance was measured at 450 nm (A_{450}), and this experiment was conducted in three biological replicates. As background value was taken the value obtained for the 0 NADH standard and was subtracted from all the sample values. The values obtained from the NADH standards were plotted to generate a standard curve. The NAD total concentration of each sample was determined from the plotted standard curve. The amount of NAD total was divided by the cell number and

expressed as pmol of NAD per 1000 cells.

3.2.7 Isolation of sRNA and lRNA fractions from the cell culture

sRNA and lRNA fractions were purified according to the RNAzol manufacturer's protocol. The cells were first lyse in RNAzol by repeated pipetting. DNA, protein and polysaccharide were precipitated by adding RNase-free water, incubating for 10 minutes at r.t., and centrifugating (6,000 × g, 30 min, 4 °C). The supernatant was transferred into a new tube, and 75% ethanol was added to precipitate lRNA. After 10 min of incubation at r.t. and centrifugation (6,000 × g, 16 min, 4 °C), the resulting white pellet was washed twice with 75% ethanol. The supernatant containing sRNA (shorter than 200 bases) was transferred to a new tube and mixed with 0.8 volume of 100% isopropanol. The sample was incubated (30 min, r.t.) and centrifuged (6,000 × g, 30 min, 4 °C). The resulting pellet was washed twice with 70% isopropanol. The RNA pellets were solubilised in RNase-free water. The RNA concentration was determined on NanoDrop ONE, and the RNA sample quality control was performed on a TapeStation.

3.2.8 U1 snRNA capture and quantification

Annealing of 5 nmol of biotinylated U1 probe (Tab. 1) to 50 µg of sRNA was performed in a buffer supplemented with 10 mM Tris, 50 mM NaCl and 1 mM EDTA, pH 7.8, by a temperature slowly decreasing from 65 °C to 25 °C. Dynabeads™ MyOne™ Streptavidin C1 (ThermoFisher Scientific) were prepared based on manufacturer protocol. Briefly, the beads were resuspended and 100 µL was transferred to a new tube and washed 3 times with 1x binding and washing buffer (5 mM Tris-HCl, 0.5 mM EDTA, 1M NaCl, pH 7.5). Since the beads are not supplied in RNase-free solutions, they had to be washed twice in solution A (DEPC-treated 0.1 M NaOH and 0.05 M NaCl) and twice in solution B (DEPC-treated 0.1 M NaCl). After washing, the beads were resuspended in twice the original volume of 2x binding and washing buffer, and an equal volume of biotinylated U1 probe annealed to RNA was added. Samples were incubated for 15 mins at r.t. using gentle rotation. The beads were washed twice with 1x binding and washing buffer and resuspended in RNase-free water. To dissociate the captured U1 snRNA, samples were incubated for 10 minutes at 75 °C. The concentration of released U1 snRNA was measured on NanoDrop, and the purity of samples was validated via 12.5% PAGE (Fig. 57).

3.2.9 Sample preparation and LC-MS measurement

RNA was washed 3 times with 8.3 M urea, 2 times with RNAse-free water, twice with 4.15 M urea, and at the end 4 times with RNAse-free water to remove non-covalently bound NAD (Frindert et al., 2018; Zhang et al., 2020). 50 µg of sRNA or 5 µg of U1 snRNA was incubated in buffer containing 50 µM ammonium formate (pH 7.0), 10 mM DTT and 10 mM $MgCl_2$, with 20 pmol of NudC pyrophosphatase and 1 U of shrimp alkaline phosphatase. The samples were brought to 40 µL with RNAse-free water and incubated for 30 min at 37 °C. 10 µL of 1.5 nM nicotine amide riboside-d4 internal standard (Torronto Research Chemicals) was spiked to the sample. The mixture was filtered through 10 kDa Vivacon 500 filter (Sartorius, 14 000 × g, 15 min, at 10 °C) and mixed with 2 equivalents of acetonitrile (LC-MS grade) in HPLC vial.

Samples were separated with HPLC (Acquity H-class, Waters) using Xbridge Premier BEH amide column (2.5 µm, 4.6 mm × 150 mm, Waters) heated to 35 °C. Autosampler was kept at 10 °C, injection volume was 50 µL, and the flow rate was 1 mL/min. Mobile phase A contained 10 mM ammonium formate (pH 9.0) in 90% (v/v) acetonitrile and mobile phase B 10 mM ammonium formate (pH 9.0) in ultrapure water (18.2 MΩ.cm, Purelab Chorus system, Elga). The gradient of separation was following: 0 min 10% (v/v) B, 2 min 10% B, 8 min 33% B, 10 min 33% B, 10.1 min 10% B, 15 min 10% B. Analytes were detected using Xevo G2–XS QTof mass spectrometer (Waters) equipped with an electrospray ionization source. The spray voltage was 3 kV, sampling cone 20 and source offset 40 V. Source temperature was kept at 150 °C and desolvation gas at 500 °C. Gas flow rates were 50 L/h for cone gas and 1000 L h^{-1} for desolvation gas. The mass spectrometer operated in positive ion mode, covering a 50-600 m/z range in sensitive mode with the Tof-MRM function and a 255 to 123 transition. The scan time was set to 1 second, and the collision gas energy was 3. Nicotine amide riboside was identified by matching the retention time with a spiked internal standard and quantified using an external calibration curve. This curve consisted of spiked sample blanks (which underwent the same procedure as the analysed RNA) with four calibration points (4.6, 14, 42, and 125 pM concentrations of nicotine amide riboside), fitted using linear regression. Each point was normalized to the isotopically labeled internal standard and represented an average of two separate injections. These expreriments were done by M.Sc. Jiří František Potužník, M.Sc. Anton Škríba, Ph.D.

3.2.10 NAD captureSeq library preparation

NAD captureSeq libraries from four biological replicates for each sample were performed similarly as described (Cahová et al., 2015) (Fig. 24). Only a short fraction of RNA from MT-4 cells infected by HIV-1 and MT-4 cells uninfected were used for the library preparation. The most critical step is the enzyme reaction catalysed by ADP-ribosyl cyclase (ADPRC), an enzyme specific mainly for NAD-capped RNAs. Samples not treated by ADPRC were used as a negative control for non-specifically bound RNAs. 100 µg of sRNA for each sample was incubated with 4-pentyn-1-ol (10%) and ADPRC (2.5 µg) in 50 mM Na-HEPES (pH 7.0) and 5 mM $MgCl_2$ for 30 min at 37 °C. The reaction was stopped by phenol/ether extraction, and RNA was ethanol precipitated. NAD-capped RNA was then biotinylated (biotin-PEG3-azide, 250 mM) via a copper-catalyzed azide-alkyne cycloaddition (CuAAC) in a freshly prepared mixture of CuSO4 (1 mM), THPTA (0.5 mM) and sodium ascorbate (2 mM) for 30 min at 25 °C, with shaking at 350 r.p.m. The biotinylated RNA was purified by phenol/ether extraction and ethanol precipitated. Mobicol Classic columns (MoBiTec, GmbH) were assembled and 50 µL of streptavidin Sepharose (GE Healthcare) was transferred to each column. After washing the beads by adding three times 200 µL of immobilisation buffer (10 mM Na-HEPES, 1 M NaCl and 5 mM EDTA, pH 7.2) to each column and centrifuge at 16,100 g for 1 min at r.t., the beads were blocked with acetylated BSA (100 µg/mL) in 100 µL of immobilisation buffer for 20 min at 20 °C, with shaking at 1,000 r.p.m. After three more washes, the RNA was immobilised on the beads for 1 h at 20 °C, with shaking at 1,000 r.p.m. The beads were washed five times with 200 µL of streptavidin wash buffer (50 mM Tris-HCl and 8 M urea, pH 7.4), then equilibrated by washing three times with 200 µL of 1x standard ligation buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, pH 7.4), blocked with acetylated BSA, and washed three more times with 1x standard ligation buffer as described previously. 30 µL of ligation mixture containing adenylated RNA 3' adaptor (5 µM, Tab. 1), 1x standard ligation buffer, DMSO (15%), acetylated BSA (1.5 µg), 2-mercaptoethanol (50 mM), T4 RNA ligase (15 U), and T4 RNA ligase 2, truncated K227Q (300 U) was added to the beads and the mixture was incubated at 4 °C overnight (16 h). The biotinylated RNA was rebounded by adding 7.5 µL 5 M NaCl and incubated for 1 hour at 20 °C, with shaking at 1,000 r.p.m. After washing the beads five times with the streptavidin wash buffer, they were equilibrated by washing three times with 1× first-strand buffer (50 mM Tris-HCl, 75 mM KCl and 5 mM $MgCl_2$, pH 8.3), blocked with acetylated BSA, and washed another

three times with 1× first-strand buffer. The beads were covered with 30 µL of reverse transcription mixture containing dNTPs (0.5 mM each), acetylated BSA (1.5 µg), DTT (5 mM), 1x first strand buffer, reverse primer (5 µM, Tab. 1), and Superscript III reverse transcriptase (300 U) and incubated for 1 hour at 37 °C. The rebinding procedure was repeated and samples were washed five times with streptavidin wash buffer and three times with 1x ExoI buffer. The beads were blocked with acetylated BSA in ExoI buffer and again washed three times with 1x ExoI buffer. The free primers were digested by adding 30 µL of 1x ExoI buffer and 30 U of ExoI, and the mixture was incubated for 30 min at 37 °C. The beads were washed five times with streptavidin wash buffer and three times with immobilisation buffer. The RNA was digested by adding 100 µL of 0.15 M NaOH solution and incubating for 25 min at 55 °C. The columns were centrifuged and washed with 100 µL of H_2O. The flow-through was ethanol precipitated, and the cDNA pellet was dissolved in 19 µL of Terminal deoxynucleotidyl transferase (TdT) tailing mixture containing 1x TdT buffer and CTP (1.25 mM). 20 U of TdT was added, and the mixture was incubated in a thermocycler for 30 min at 37 °C. The enzyme was thermally denatured by heating to 70 °C for 10 min. The second adaptor was ligated on the cDNA by adding 60 µL of reaction mixture containing 1x standard ligation buffer, cDNA anchor (5 µM each strand, Tab. 1), ATP (10 µM), and T4 DNA ligase (120 Weiss U) and incubating for 16 h at 4 °C. 25 cycle PCR amplification was performed in 1x Taq DNA Pol reaction buffer with 1 µM of each reverse indexed primer and forward primer (6 nt long indexes are underlined, Tab. 1), 0.2 mM dNTPs each and 2.5 U of DreamTaq DNA Polymerase in 50 µL of total reaction mixture. Initial denaturation was performed at 95 °C for 3 min, followed by annealing for 30 s at 68 °C, elongation for 30 s at 72 °C and denaturation for 30 s at 95 °C. The final extension was performed at 72 °C for 5 min. The PCR reaction mixture was loaded on 2.5% agarose gel (140 V for 2 h) and post-stained with SYBR gold. Fractions between 200–400 nt were cut out of the gel, and DNA was extracted from the gel by Monarch DNA Gel Extraction Kit (NEB). Multiplexed samples were submitted to the IMG Genomics and Bioinformatics facility for library synthesis using the NextSeq® 500/550 High Output Kit v2, 75 cycles (Illumina) and sequencing on a NextSeq 500/550, Illumina.

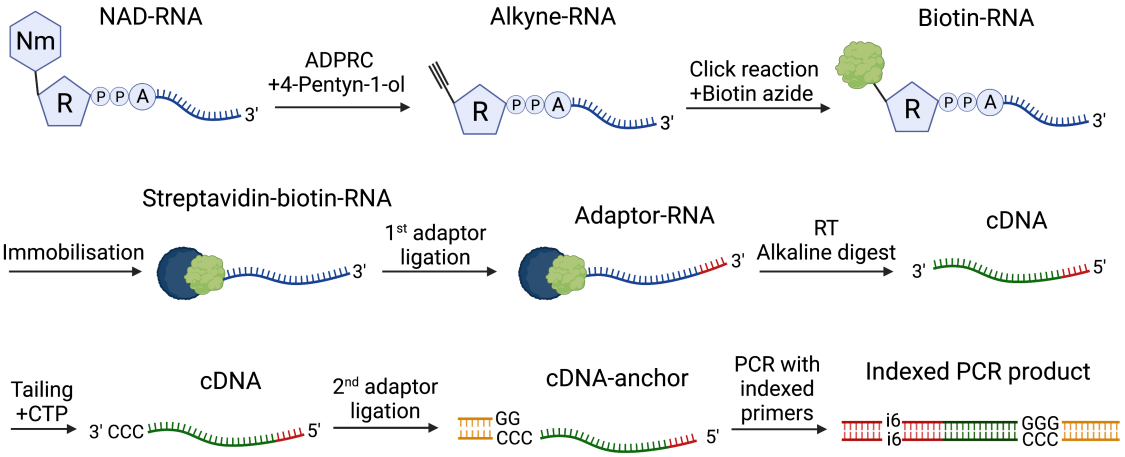


Fig. 24: Workflow of the NAD captureSeq protocol. MT-4 sRNA samples are treated with ADPRC to produce an alkyne moiety only on NAD caps. Alkyne-RNA (previously NAD-RNA) is reacted with biotin-azide in a copper-catalyzed azide-alkyne cycloaddition (CuAAC), and the biotinylated transcripts are immobilised on streptavidin beads (the RNA structure is simplified for better clarity). After ligation of the adaptor and reverse transcription (RT), the RNA is digested, and complementary DNA (cDNA) is released. cDNA is then tailed with CTP, which is a ligated double-stranded adaptor with a complementary G-overhang. cDNA with anchor is amplified using indexed primers (i6: 6 nt long indexes) by PCR and submitted for sequencing.

3.2.11 sRNA-seq library preparation

To measure the amounts of specific sRNAs in samples, three biological replicates of sRNA from MT-4 cells and MT-4 cells infected by HIV-1 were submitted for miRNA (LP-170) library preparation to SEQme. The library synthesis was done using the NEBNext®Small RNA Library Prep Set for Illumina with NEBNext Multiplex Oligos for Illumina (Index Primers Set 1-3) and sequenced on a NovaSeq 6000 System (SE 50 bp, 350-400 millions of reads, DS-270).

3.2.12 Long RNA-seq library preparation

To evaluate the splicing changes in MT-4 cells under various conditions, three biological replicates of lRNA from MT-4 cells, MT-4 cells infected by HIV-1 and MT-4 cells with DXO OE, all mentioned conditions with or without NAD repletion were submitted for library preparation to the IMG Genomics and Bioinformatics facility. KAPA RNA HyperPrep Kit with RiboErase was used to prepare the libraries, and the sequencing was performed using the NextSeq® 500/550 High Output Kit v2, 75 cycles (Illumina) and on a NextSeq 500/550, Illumina.

3.2.13 Data processing and bioinformatic analysis

Adapters and low-quality bases were trimmed using TrimGalore! v.0.4.1. Because of a bias at the ends of the reads, an additional 5 bases were trimmed from both the 5' and 3' end of the reads. Reads were mapped to the human GRCh38 genome and HIV-1 HIVNL43 NC_001802.1 genome using Hisat2 v.2.0.1. To analyse NAD presence at HIV RNA cap, we quantified RPM in the first 200 bp of HIV transcribed region and quantified the enrichment of ADPC-treated samples over controls using Seqmonk v1.47.2. Enrichment of human shortRNAs was quantified as RPM in ADPRC-treated over control samples using Seqmonk v1.47.2. To select the best candidate shortRNAs with depletion or increase of NAD-cap abundance after HIV infection, we applied a set of criteria removing shortRNAs with low read count and/or low enrichment over control (-ADPRC): (I) raw read count has to be 2 or higher for at least 2 replicates of at least one sample type (+/-ADPRC +/- HIV infection), (II) enrichment over control higher than 2 in at least one replicate and enrichment above 1 in at least one additional replicate in either HIV infected or uninfected samples, (III) of all 4 replicates, 2 or more have to show depletion/increase in the same direction (difference between HIV uninfected and infected above 0.2), and neither in the opposite direction and maximum 2 replicates the lack of enrichment (enrichment <1), (IV) total raw read count in combined replicates of +ADPRC samples (either uninfected or infected) above 20, (V) the number of reads in +ADPRC samples above 5 in the replicates showing enrichment above 1. Differential expression of sRNAs of all sRNA classes annotated within Seqmonk v1.47.2 (GRCh38v100 genome annotation) was analysed using DESeq2 implemented within Seqmonk v1.47.2. Splicing efficiency of HIV-1 transcript was analysed as the ratio of read count in random regions within introns to read count in size matched random regions in exon 3. Sequences of main donor and acceptor splice sites in HIV-1 main splicing variants (according to 44) were obtained from 45 and matched against HIV-1 HIVNL43 NC_001802.1 genome to generate genomic coordinates of exons and introns. Splicing efficiency of cellular transcripts was quantified as the ratio of reads in introns to reads in exons for 1000 genes with the highest expression in combined replicates of untreated MT-4 cells quantified using RNA-seq quantitation pipeline within Seqmonk v1.47.2 (Fig. 25). The bioinformatic analysis was done by M.Sc. Lenka Gahurová, Ph.D.

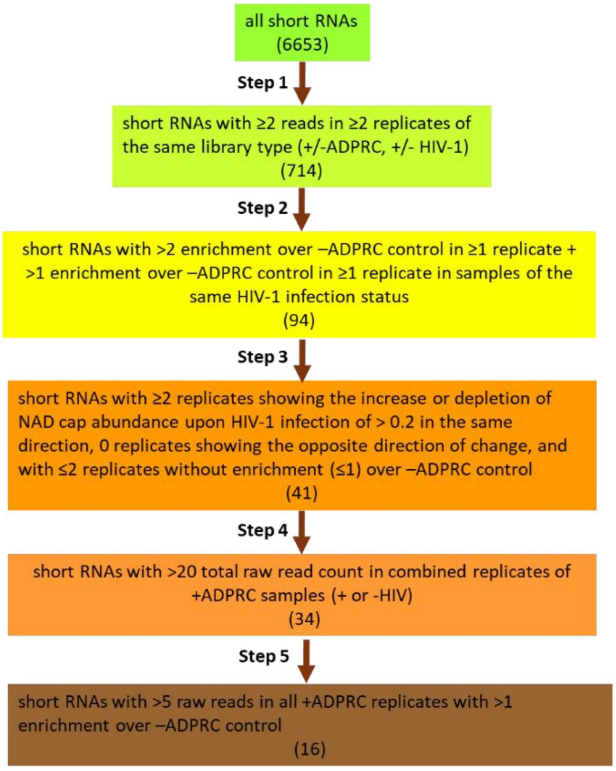


Fig. 25: Bioinformatical analysis workflow describing individual filtering steps with the number of sRNA candidates after each step.

3.2.14 Deep sequencing data validation through RT-PCR

Real-time PCR was performed to verify the NAD captureSeq data. Three biological replicates of sRNA from infected and control MT-4 cells and cDNA samples after NAD captureSeq were measured in two technical repeats on LightCycler 480 II (Roche) by Luna Universal One-Step RT-qPCR Kit according to the manufacturer protocol. 20 µL reaction mix consisted of buffer, enzyme mix, forward and reverse primer (0.4 µM, Tab. 1) and 50-100 ng template. PCR was conducted with the following program: reverse transcription 55 °C for 10 min, initial denaturation at 95 °C for 60 s, followed by 40 cycles of 95 °C for 10 s and 53 °C for 30 s. Finally, a melting curve was performed from 37 °C to 95 °C . The Cp values were calculated by LightCycler 480 Software, and the reciprocal values were plotted.

3.2.15 Chemical synthesis of TMG cap (TMGpppApG)

To enable synthesis of highly homogenous 5'-TMG-capped RNA by *in vitro* transcription we designed a trinucleotide TMG cap analog (TMGpppApG or $m_3{}^{2,2,7}$GpppApG) following the synthetic pathway described previously for

NAD-derived trinucleotides (Mlynarska-Cieslak et al., 2018). Briefly, pApG dinucleotide was synthesised by automated solid-phase synthesis, imidazole-activated, and coupled with 2,2,7-trimethylguanosine diphosphate ($m_3^{2,2,7}$GDP) in DMSO in the presence of anhydrous magnesium chloride ($MgCl_2$). The final compound was purified by ion-exchange chromatography (DEAE Sephadex A-25) using a linear gradient of triethylammonium bicarbonate buffer, followed by semi-preparative reversed-phase HPLC. The final product of the reaction was lyophilised three times, and its structure was confirmed by high-resolution MS. For structure of TMGpppApG and detailed synthetic procedure, see Appendix. Chemical synthesis of TMG cap was done by M. Sc. Katarzyna Grab.

3.2.16 RNA *in vitro* transcription

In vitro transcription was performed as described (Huang, 2003) in a 50 µL mixture (0.2 µM of template DNA (Tab. 1), 1 mM of each NTP, 5% dimethyl sulfoxide (DMSO), 0.12% triton X-100, 10 mM dithiothreitol (DTT), 4.8 mM $MgCl_2$, 1× reaction buffer for T7 RNAP, and 125 U of T7 RNAP. For capped RNA, 8 mM of NAD or TMG cap was added. To produce 20 nt long U1 snRNA containing pseudouridine, the UTP was replaced with pseudo-UTP. The mixture was incubated for 2 h at 37 °C. For higher yield, T7 RNAP was added again into the mixture and incubated for an additional two hours. The DNA template was digested by DNAse I at 37 °C for 45 min and the enzyme was heat inactivated at 75 °C for 10 min. To obtain only capped RNA, the samples were treated with 5'-polyphosphatase for 30 min at 37 °C and then with Terminator™ 5'-phosphate-dependent exonuclease for 1 h at 30 °C. Between each step, samples were purified using Clean and concentrator (Zymo).

3.2.17 U1 snRNA and HIV-1 mRNA complex stability assay

20 nt long U1 snRNA (capped either with TMG or NAD cap) and HIV-1 mRNA (Tab. 1) were prepared by *in vitro* transcription. 150 ng of U1 snRNA (TMG or NAD cap) and 150 ng of HIV-1 mRNA were mixed with 1x ResoLight dye and annealing buffer (10 mM Tris, 50 mM NaCl, 1 mM EDTA pH 7.8). The technical triplicates were prepared three times and measured on LightCycler 480 II. High-resolution melting (HRM) curves were obtained by measuring the complex stability by temperature increase to 80 °C and then decrease to 20 °C with a ramp rate 0.01 °C/s in three cycles and thus obtaining a total of

27 melting curves per sample. HRM curve analysis was performed using the LightCycler 480 Software.

3.2.18 Assessment of melting temperature by circular dichroism

The ECD spectra were measured on a Jasco 815 spectropolarimeter equipped with the Peltier type temperature control system PTC-423S/L at r.t. in the spectral range from 200 nm to 350 nm, using a 0.2 cm path length with standard instrument sensitivity, with the scanning speed of 10nm/min, response time of 8 s, and 2 spectra accumulations. For measurements, 26,000 ng of RNA duplex with TMG-U1 snRNA and/or RNA duplex with NAD-U1 was dissolved in an annealing buffer (10 mM Tris, 50 mM NaCl, 1 mM EDTA pH 7.8). The final spectra were expressed as differential absorption ΔA. Melting temperature was obtained as a result of measurement of CD signal at 264 nm in temperature range from 5 °C to 95 °C with temperature slope 24 °C/h, step 1 °C and time constant 8 sec. Each measurement was repeated 3 times. Melting temperature was calculated using sigmoid fitting by the program Sigmaplot 12.5 (Systat software). Circular dichroism measurements were done by M.Sc. Lucie Bednárová, Ph.D.

3.2.19 Radioactively labelled U1 snRNA preparation

DNA template for U1 snRNA (164 nt full length) was prepared from a plasmid (U1-SP65 plasmid containing the main U1 variant RNA-U1-1) via PCR in 50 µL reaction mix consisting of 1x DreamTaq buffer, DreamTaq DNA polymerase (1.25 U), forward and reverse primer (0.5 µM, Tab. 1), dNTPS (2mM each) and 400 ng template. PCR was conducted with the following program: initial denaturation at 95 °C for 2 min, followed by 30 cycles of 95 °C for 30 s, 60 °C for 30 s and 72 °C 1 min, the final extension was performed at 72 °C for 5 min. The PCR reaction mixture was loaded on 1% agarose gel (140 V for 2 h) and post-stained with SYBR gold. The product bend was cut out of the gel and DNA was extracted from the gel by Monarch DNA Gel Extraction Kit (NEB). ^{32}P-GTP labelled U1 snRNA was prepared by *in vitro* transcription as described previously with the addition of 0.5 µL of α-^{32}P-GTP (3.3 µM, 10 µCi/µL; Hartmann analytic), with either TMG or NAD cap.

3.2.20 U1 snRNA decapping by DXO

150 ng of 32P-GTP labelled U1 snRNA was incubated with different concentrations of DXO decapping enzyme (0 nM, 5 nM, 10 nM, 25 nM, 50 nM) in decapping buffer (10 mM Tris pH 7.5, 100 mM KCl, 2 mM DTT, 2 mM $MgCl_2$ and 2 mM $MnCl_2$) at 37 °C for 15 min. Reactions were stopped by the addition of RNA loading dye and loaded into the gel wells. Denaturing polyacrylamide gel (12.5%) was prepared from Rotiphorese gel in 1x TBE buffer and supplemented with acryloylaminophenylboronic acid (APB, 0.4%, prepared as published in Nübel et al. (2017)). The polymerised gel was pre-run at 600 V for 30 min in 1x TBE buffer and then run at 600 V for 5 hours. The gel was incubated with a phosphor imaging plate (GE Healthcare), scanned using Typhoon FLA 9500 and analysed using ImageJ.

3.2.21 Western blot

MT-4 cells with and without lentiviral plasmids were collected by centrifugation (225 × g, 5 min, at 20 °C), and proteins were extracted using lysis buffer (RIPA, cOmplete™, EDTA-free Protease Inhibitor Cocktail). The protein concentration was measured on Tecan, and approximately 20 µg of protein lysate was loaded on the 10-well protein gels (Mini-PROTEAN TGX, Bio-Rad) and run at 120 V. Gels were wet transferred on PVDF membranes at 30 V for 2 hours. Primary antibody DOM3Z Rabbit PolyAb was diluted 1:500 and incubated overnight at 4 °C. Washes were performed with phosphate-buffered saline with 0.1% Tween-20 before the addition of secondary antibody (1:10,000, peroxidase conjugate AffiniPure Goat Anti-rabbit IgG) for 1 hour at r.t. Blots were developed using SuperSignal™ West Femto or Dura enhanced chemiluminescent (Femto or Dura) and imaged on the Azure Biosystems.

3.2.22 HIV-1 infectivity determination in DXO-transduced cells

The HIV-1 infectivity in DXO-transduced cells was determined by measuring the HIV-1-induced cytopathic effect by XTT cell proliferation assay. Briefly, wild-type MT-4 and DXO-transduced MT-4 cells were seeded at $3x10^4$ cells per well in a 96-well plate in RPMI medium (phenol red-free RPMI, supplemented with 10% FBS, penicillin (100 U/mL), streptomycin (100 µg/mL), 2 mM L-glutamine and 4 mM HEPES), with and without 10 mM NAm, infected in triplicate by two-fold serially diluted HIV-1 (except for controls) and cultured at 37 °C and 5% CO_2. After 4 days, a 50:1 mixture of XTT salt and PMS

(N-methyl dibenzopyrazine methyl sulfate) electron-coupling reagent was added to the wells and incubated for 4 h at 37 °C in 5% CO_2. The formation of orange formazan dye was measured in the EnVision plate reader at 560 nm. The threshold of infectivity (TI) was calculated as:

$$TI = C_{\text{Avg}} - 5 * C_{\text{StDev}}, \tag{1}$$

where C_{Avg} is average of control absorbance and C_{StDev} is standard deviation of control absorbance. HIV-1 titer was then calculated by the Reed-Muench method and expressed as 50% tissue culture infectious dose ($TCID_{50}$).

3.2.23 Other resources

Writeful and Grammarly were variously used as tools to find/change English expressions and to proofread. Some figures were created with BioRender.com.

3.2.24 Data source and availability

The majority of data was published in Benoni et al. (2023) and in Šimonová et al. (2019).

4 Results

4.1 Preparation of control and infected MT-4 cells, and HIV-1 virions

The challenging part was to prepare sufficient quantities of HIV-1 virions and MT-4 cells under all tested conditions. For most experiments, we prepared 2 liters of the cells in 50 flasks. Each flask can contain a maximum of 40 mL of the cells for optimal conditions. Then, we divided the medium with cells into two groups: one serving as a control, and the other gradually infected with HIV-1 (Fig. 26).

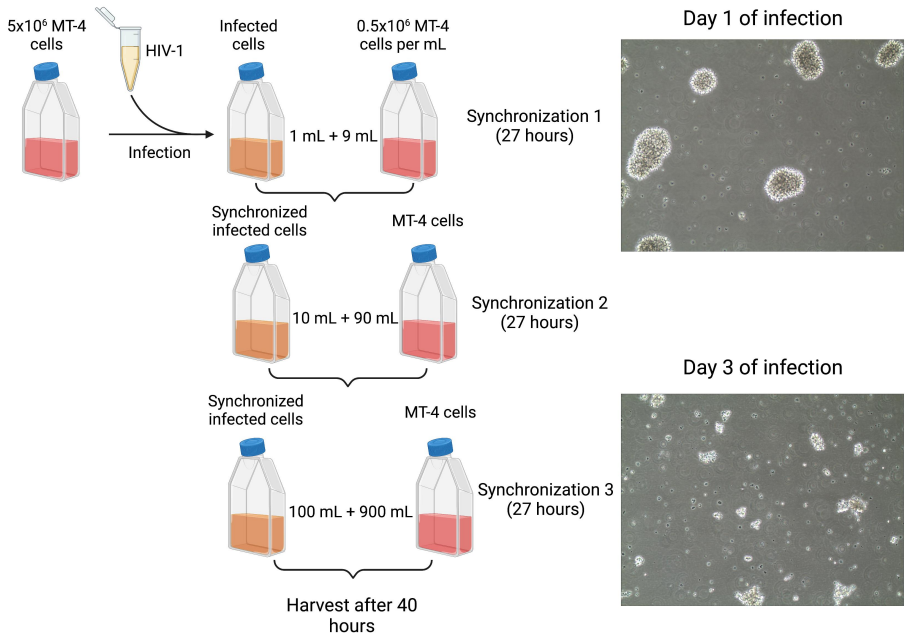


Fig. 26: HIV-1 infected cells preparation. RPMI contains a phenol red pH indicator that changes colour as pH changes due to the metabolites released by the cells. Infected cells have a higher metabolism rate and the RPMI colour is generally more orange than in non-infected cells. Infected cells are expanded by three successive additions to non-infected MT-4 cells at a ratio of 1:9. On the first day of infection, MT-4 cells create bigger clusters. With infection progress, the MT-4 cells have different shapes and do not form clusters, as visible on day 3 of infection.

The quality of the isolated RNA was assessed with TapeStation, and the presence of HIV-1 RNA in the samples was confirmed by RT-PCR (Fig. 27).

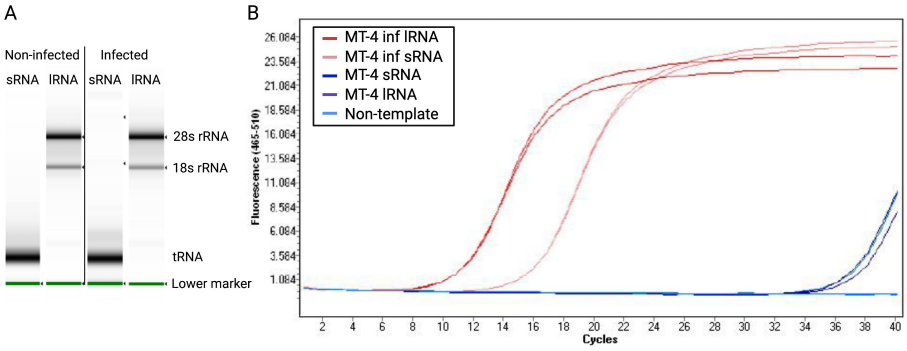
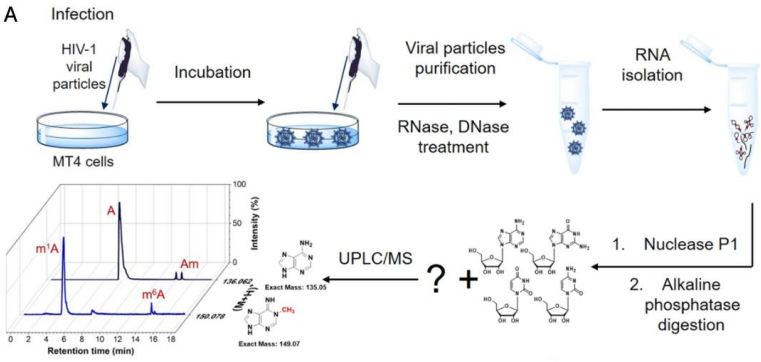


Fig. 27: Analysis of RNA samples after sRNA and lRNA isolation from MT-4 and MT-4 infected cells. A) TapeStation analysis. The most prevalent band in the sRNA sample is tRNA and the smear represents other non-coding regulatory sRNAs. B) RT-PCR detection of HIV-1 with each sample measured in technical duplicates. HIV-1 RNA was confirmed in the infected samples. The product visible after 35 cycles results from excessive cycling, which creates nonspecific amplification and errors.

4.2 LC-MS analysis of RNA in the HIV-1 packageome

RNA isolated from HIV-1 virions was used for this study, and for analysis of internal RNA modifications (Šimonová et al., 2019). We prepared four biological samples for LC-MS analysis of RNA isolated from HIV-1 virions (Fig. 28A) and found several adenosine modifications (Fig. 28B). We calculated that 4.1% of all adenosines in the packageome of the HIV-1 virus are m^1A. We prepared the deep sequencing libraries of RNA isolated from HIV-1 virions and found that m^1A is present only on tRNAs but not in the genomic HIV-1 RNA or the abundant 7SL RNA. Based on these results, we calculated the accurate tRNA composition in the HIV-1 virion: in addition to 2 copies of gRNA and 14 copies of 7SL RNA, approximately 770 copies of tRNA (Fig. 28C). LC-MS analysis and the deep sequencing libraries were done by M.Sc. Anna Šimonová, Ph.D.



B

Nucleoside*	%
A	89.3 ± 2.0
m^1A	4.1 ± 1.0
m^6A	1.1 ± 0.2
Am	0.9 ± 1.0
t^6A	3.2 ± 0.8
mS^2t^6A	1.5 ± 0.9*

* with presumption that mS^2t^6A has similar ionization properties as t^6A

C

RNA in viral particle	Copies of RNA based on literature*	LC/MS estimated copies
gRNA HIV-1	2	2
7 SL RNA	14	14
tRNA $^{Lys\ TTT-3.1}$	8	770
tRNA $^{Lys\ CTT-2.1}$	12	
tRNA $^{Lys\ CTT-1.1}$		
tRNA $^{Asp\ GTC-2.1}$	50	
tRNA $^{Asn\ GTT-2.1}$		
Rest of tRNA		

Fig. 28: A) Scheme showing the preparation of HIV-1 samples for LC-MS analysis. B) Ratios of nucleosides in the HIV-1 samples as analysed by LC-MS. C) Table showing the extrapolated numbers of copies of RNAs in HIV-1 virion based on literature (Keene et al., 2010; Telesnitsky and Wolin, 2016) and on our LC-MS data.

4.3 Changes in NAD concentration upon HIV-1 infection

HIV-1 infection causes a decrease in cellular NAD and NADH levels, a condition known as pellagra (Murray et al., 1995). We examined whether we observe the same trend under our conditions in MT-4 cells. The NAD/NADH Quantification Kit is a sensitive colorimetric assay used to detect and quantify intracellular NAD and NADH without purifying them from the samples. It is a specific assay that does not detect NADP nor NADPH.

The challenge of this procedure was to pellet a small number of cells precisely every time. The cell count was performed several times with the diluted samples to ensure the correct number of pelleted cells. Each sample was measured in two replicates, and three biological samples were prepared. As in previous work, we observed changes in NAD concentration upon HIV-1 infection. We calculated that NAD levels dropped almost fourfold, from 4.64 pmol/1000 cells (SD = 0.138) in control MT-4 cells to 1.28 pmol/1000 cells (SD = 0.025) in HIV-1 infected cells (Fig. 29).

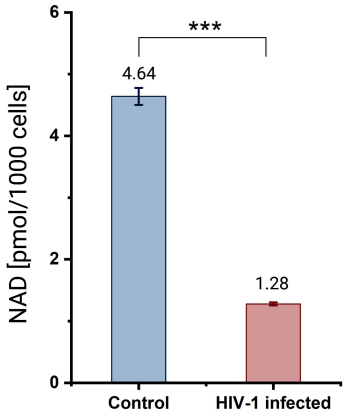


Fig. 29: Levels of free total NAD (NAD and NADH) in control and HIV-1 infected MT-4 cells. The difference between these samples is statistically significant (p-value <0.001).

4.4 Sequencing of NAD-RNAs from control and HIV-1 infected MT-4 cells

sRNAs are noncoding regulatory RNAs that bind to their targets and non-canonical cap might affect their binding efficiency. We prepared four NAD captureSeq libraries from the sRNA fraction of both control and HIV-1 infected cells. To eliminate nonspecifically interacting RNA, ADPRC was omitted in negative control samples (-ADPRC) (Fig. 24). The final products of PCR were purified using agarose gel. The area cut from the agarose gel was above the visible bands of dimers of the primers up to 400 nt (Fig. 30).

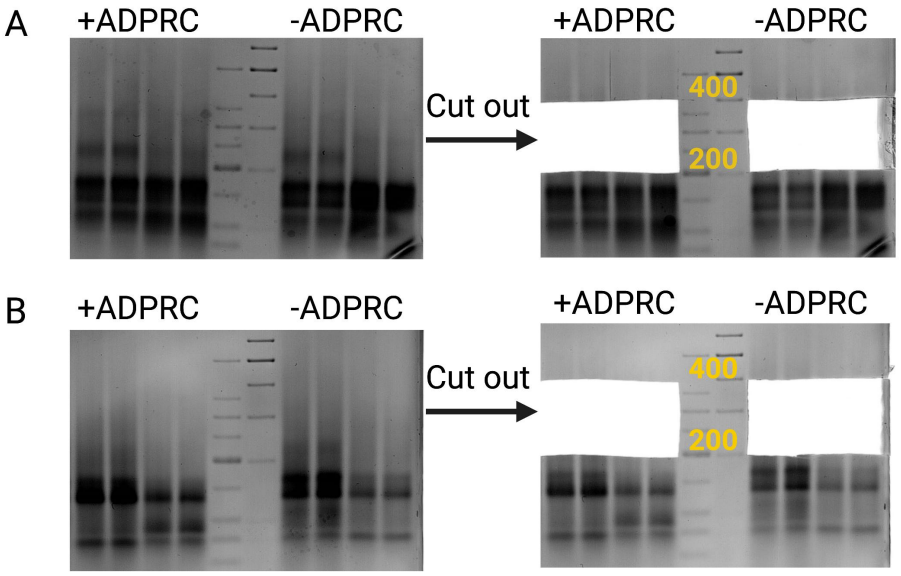


Fig. 30: Agarose gels after cDNA amplification. Each sample was loaded into four wells. Two DNA ladders were used to establish the length of DNA correctly cut from the gel: 200 nt - 400 nt. There are visible bands of primers and primer dimers. A) cDNA obtained by NAD captureSeq from healthy MT-4 cells. B) cDNA obtained by NAD captureSeq from MT-4 cells infected by HIV-1.

NAD captureSeq differs from the conventional library preparations. It was necessary

to remove the reads from the ADPRC negative control samples, the adapters had to be trimmed, the abundance of captured sRNAs in healthy and infected cells had to be compared, and several other steps in the data processing pipeline had to be optimised. The bioinformatic analysis of our libraries resulted in 94 sRNAs (2 miRNAs, 6 miscRNAs, 62 tRNAs, 12 snoRNAs, 12 snRNAs, Tab. 2) that are NAD capped in infected and control MT-4 cells (Fig. 31 A). The libraries were also mapped to the HIV-1 genome but any viral encoded RNAs were found to bear NAD cap (Fig. 31 B). This excludes the possibility that the loss of host NAD is caused by HIV-1 transcripts being preferentially NAD capped. Data analysis was done by M.Sc. Lenka Gahurová, Ph.D.

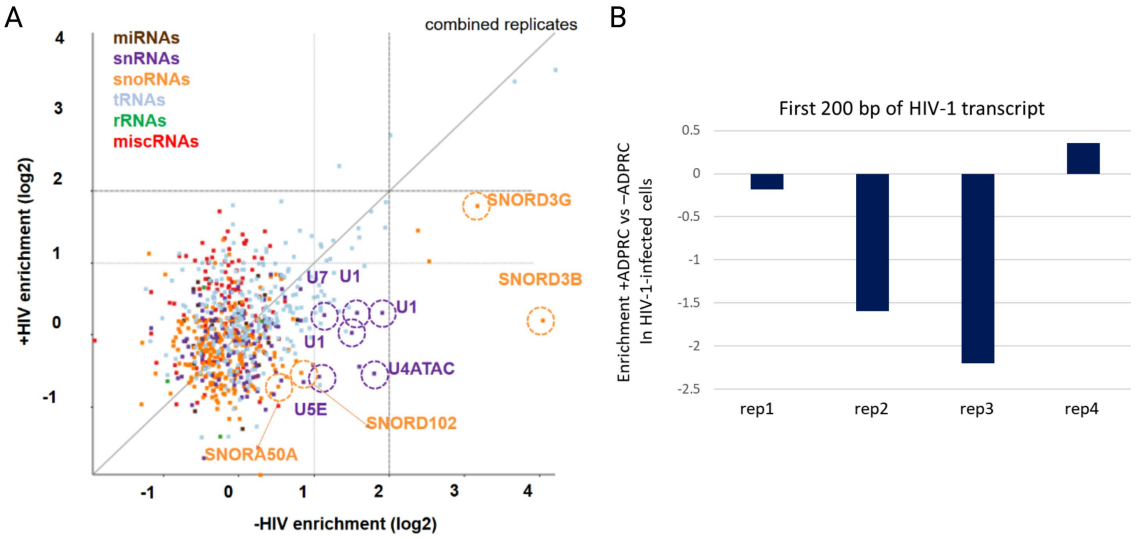


Fig. 31: Data analysis of RNA-seq. A) Scatter plot showing the enrichment (+ADPRC vs -ADPRC) in control and HIV-1 infected samples. A different colour represents each type of sRNA, and the most important sRNAs are highlighted. B) Mapping to the HIV-1 genome did not significantly enrich the transcript (+ADPRC vs. -ADPRC) in any library replicate.

NAD-capping in some sRNAs increased upon HIV-1 infection. After closer inspection, we found that these sRNAs are known to be packed in HIV-1 virions or to have increased expression due to HIV-1 infection, particularly Y RNA and miR19B (Eckwahl et al., 2016; Yin et al., 2019). NAD cap was also detected in some tRNAs but there was no clear trend due to their abundance. We observed that levels of 42 of these NAD-RNAs are significantly lower in HIV-1 infected cells, and the decrease was the most evident in two classes of sRNAs: snRNAs and snoRNAs. After we applied set of criteria regarding the number of mapped reads and consistency across replicates (Fig. 25), we identified four snoRNAs (SNORD3G, SNORD102, SNORA50A, and SNORD3B) and four snRNAs (U1, U4ATAC, U5E, and U7) which lose NAD cap upon HIV-1 infection (Fig. 32). The

highest enrichment (+ADPRC versus -ADPRC) of NAD-RNAs was observed for two sRNAs: SNORD3B and U1 snRNA.

We employed sRNA-seq to examine the expression of sRNAs in control and HIV-1 infected cells. Over 2000 sRNAs (miRNAs, miscRNAs, tRNAs, snRNAs, snoRNAs and rRNAs) were detected, and 89 were expressed differently in healthy and infected cells (Tab. 3). 52 sRNAs were upregulated and 37 downregulated after infection but neither of these 89 sRNAs was enriched in libraries prepared by NAD captureSeq. The comparison of the expression levels of identified snoRNAs and snRNAs before and after infection revealed that the amount of these sRNAs does not depend on HIV-1 infection (Fig. 32).

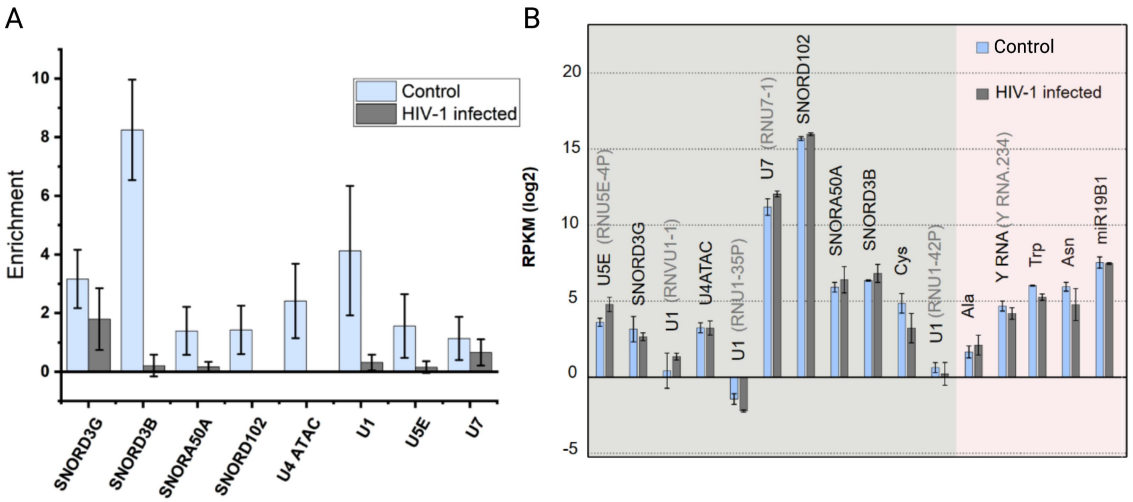


Fig. 32: Enrichment of particular sRNAs. A) Comparison of sRNAs enriched in control vs HIV-1 infected cells detected by NAD CaptureSeq. B) sRNA-seq analysis of expression of sRNAs in control and HIV-1 infected cells.

4.5 Evaluation of sRNAs levels by RT-PCR

The relative abundance of the identified snoRNAs and snRNAs was measured by RT-PCR. For each sRNA was designed a reverse and a forward primer (Tab. 1). The target sRNA sequences are very short (up to 200 nt) and contain repetitive areas, so the primer sequences were designed shorter than recommended. We measured the relative abundance of four snoRNAs (SNORD3G, SNORD102, SNORA50A, and SNORD3B) and four snRNAs (U1, U4ATAC, U5E, and U7) in biological triplicates and each measurement in technical duplicates. Already from amplification curves was visible that there is no difference in abundance of sRNAs between healthy and infected cells (Fig. 33).

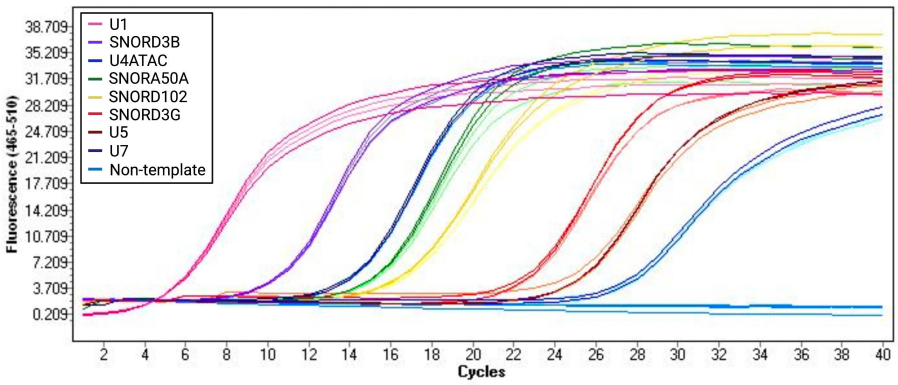


Fig. 33: Amplification curves of sRNAs isolated from infected (darker shades) and non-infected (lighter shades) MT-4 cells. Each sample was measured in technical duplicates and for each sRNA there are four curves. From the left side: U1 (the most abundant), SNORD3B, U4ATAC, SNORA50A, SNORD102, SNORD3G, U5, U7 (the least abundant). The blue linear line is a non-template control.

The relative abundance of cDNA prepared by NAD CaptureSeq was also evaluated by RT-PCR. The same set of primers used for RNA measurement was used for the set of four cDNA samples prepared from MT-4 sRNA treated with ADPRC, MT-4 sRNA not treated with ADPRC, HIV-1 infected MT-4 sRNA treated with ADPRC, and HIV-1 infected MT-4 sRNA not treated with ADPRC. The samples were measured after every library preparation, and each sample was measured in technical duplicates (Fig. 34).

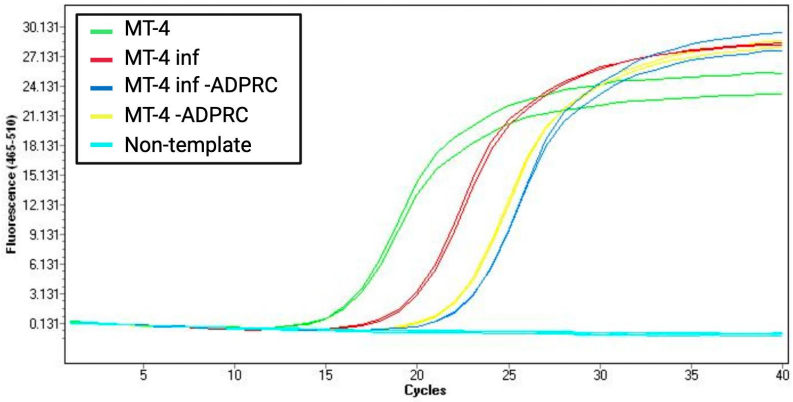


Fig. 34: U1 amplification curve from cDNA sample obtained by NAD captureSeq. Each sample was measured in technical duplicates.

RT-PCR data were analysed by LightCycler software and p-values were calculated. Statistical analysis confirmed that the number of these sRNAs does not change with infection. There was also no difference between infected and non-infected samples in -ADPRC cDNAs (ADPRC not treated - expecting only non-specific capturing). While in ADPRC-treated samples there was a significantly higher amount of captured cDNA from non-infected (control) cells than from infected cells (Fig. 35).

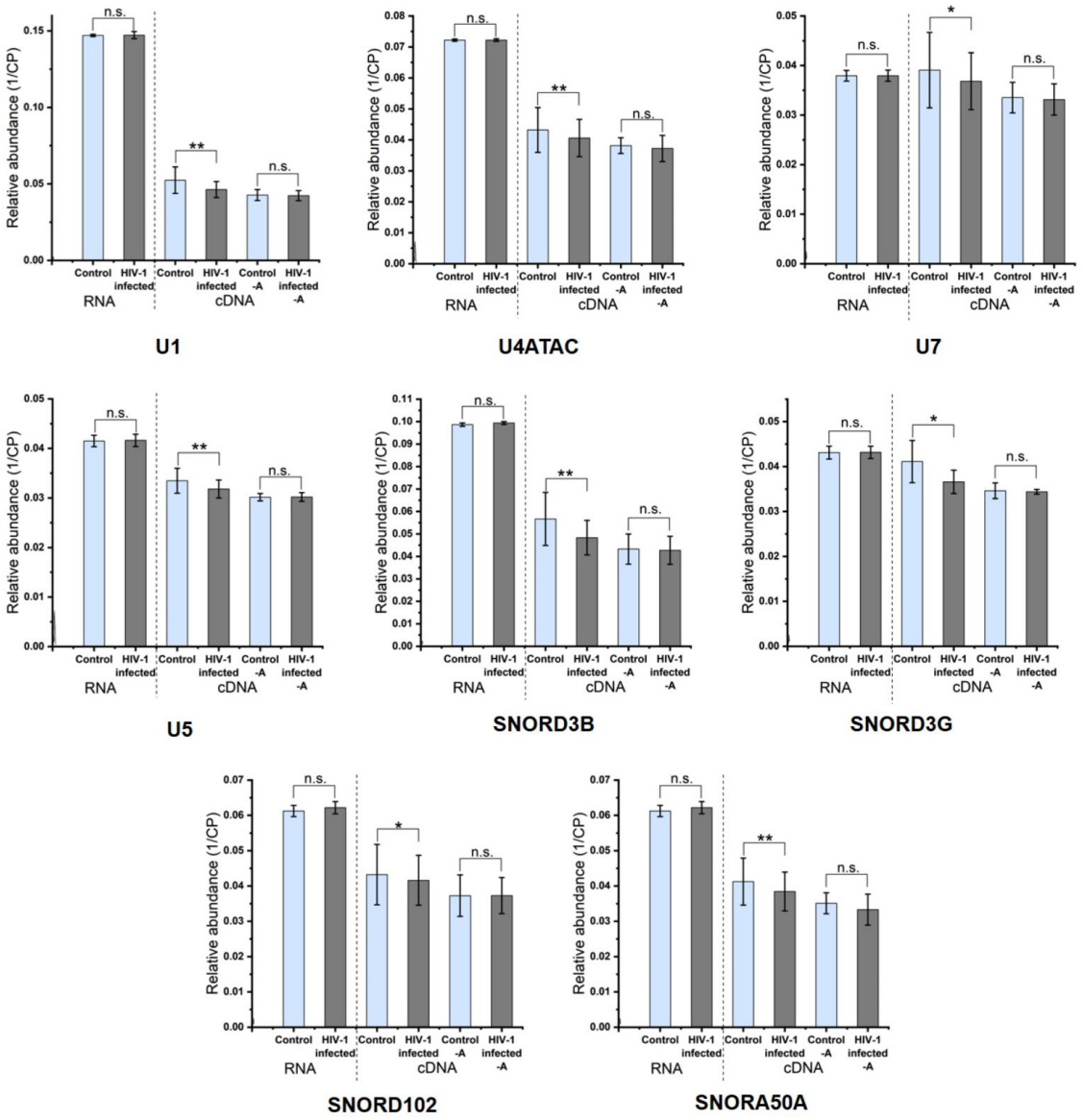


Fig. 35: Quantified RT-PCR measurements of sRNAs. The first part of each graph compares each identified sRNA abundance in total sRNA isolated from control and HIV-1 infected cells. The second part (after the dashed line) compares the relative abundance of cDNA obtained from NAD captureSeq control and infected samples. n.s. = not significant, * = p-value < 0.05, ** = p-value < 0.01.

4.6 Changes in NAD capping upon HIV-1 infection

We investigated potential differences in NAD capping between control and HIV-1 infected cells. We did not observe any statistical difference between infected and control cells in either sRNA or lRNA (Fig. 36, Fig. 54). New samples were prepared by M.Sc. Jiří František Potužník and LC-MS measurement was done by M.Sc. Anton Škríba, Ph.D.

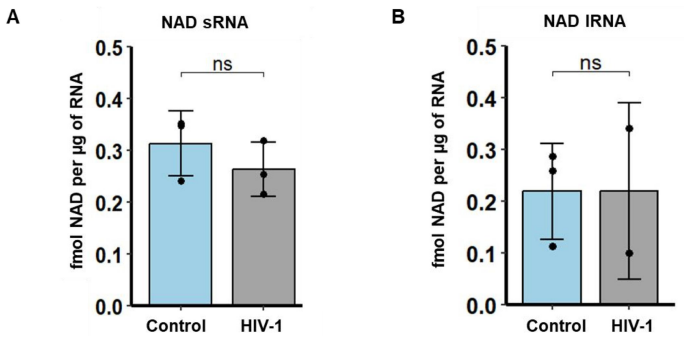


Fig. 36: Levels of NAD caps in fractions of sRNA and lRNA in control MT-4 cells and HIV-1 infected MT-4 cells. The difference between the groups is not statistically significant.

We examined whether NAD-RNAs are selectively co-packed into HIV-1 virions. Total HIV-1 RNA was isolated from the virions and measured by LC-MS. No peaks corresponding to NAD caps were observed confirming NAD captureSeq data, that HIV-1 virions do not contain any NAD-RNAs (Fig. 37).

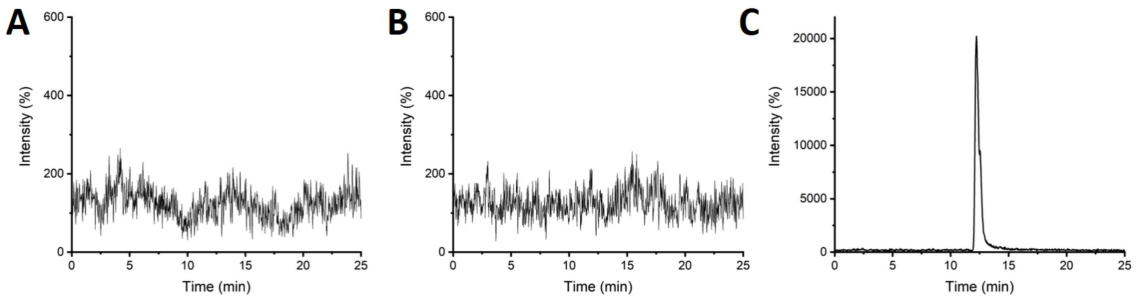


Fig. 37: EIC of NAD parent mass 664.11 from RNA isolated from A) HIV-1 virions, B) HIV-1 virions negative control (without NuP1 digestion) and C) HIV-1 virions with spiked NAD spiked.

4.7 LC-MS measurement of NAD caps on U1 snRNA

Using NAD captureSeq, we detected several sRNAs that lose NAD cap upon HIV-1 infection, and we focused our attention on U1 snRNA, which is known to be employed by HIV-1 during its replication cycle. U1 snRNA was captured and analysed by LC-MS. We observed a significant decrease in NAD capping in U1 snRNA from HIV-1 infected cells. This result supports the finding that HIV-1 infection leads to reduced NAD capping of U1 snRNA (Fig. 38, Fig. 55). New samples were prepared by M.Sc. Jiří František Potužník and LC-MS measurement was done by M.Sc. Anton Škríba, Ph.D.

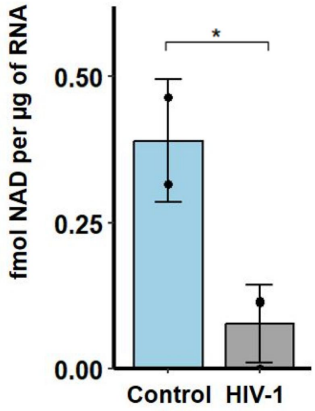


Fig. 38: Absolute quantification of NAD RNA cap in U1 pulled down from control and HIV-1 infected cells. * = p-value < 0.05.

4.8 Complex stability of U1 snRNA and HIV-1 pre-mRNA

U1 snRNA binds with a complementary region of 8 nucleotides at its 5' end to the 5' splice site of unspliced HIV-1 pre-mRNA. The canonical cap of U1 snRNA is TMG cap (Hamm et al., 1990). It was found that the nicotinamide moiety of NAD cap can form a "pseudo-base pair" with T (Bird et al., 2016) and we speculated that NAD cap in U1 snRNA can form additional interaction with HIV-1 pre-mRNA. We measured melting temperatures (Tm) of the duplexes TMG-U1 snRNA with HIV-1 pre-mRNA and NAD-U1 snRNA with HIV-1 pre-mRNA. For double verification of the results, the Tm was measured by circular dichroism and LightCycler. U1 snRNA bears Ψ at positions 4 and 5. Therefore, the UTP in the *in vitro* transcription mixture was replaced with ΨTP. To not alternate other UTPs in the U1 snRNA sequence with ΨTP and to concentrate only on the changes in interaction resulting from TMG or NAD cap, a 20 nt long 5' region of U1 snRNA was prepared. The derivative melting curves of the duplexes measured in LightCycler were created by High-Resolution Melting analysis in LightCycler software (Fig. 39).

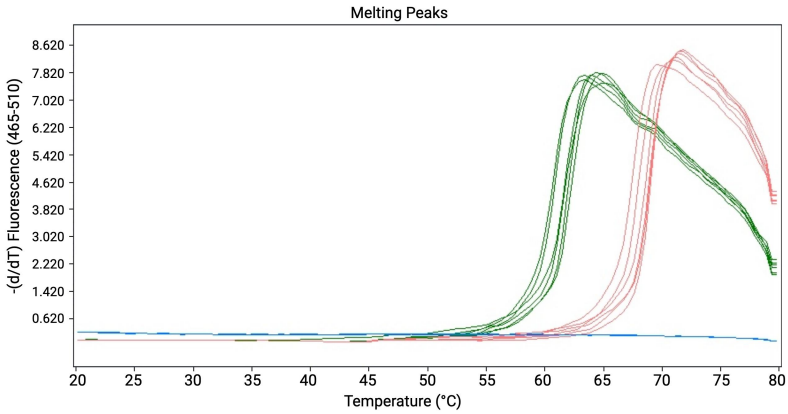


Fig. 39: Derivative melting curves obtained from measurement of RNA duplexes by LightCycler. Green curves represent a duplex of NAD-U1 snRNA and HIV-1 pre-mRNA, and red curves represent a duplex of TMG-U1 snRNA and HIV-1 pre-mRNA.

The HRM curve analysis revealed differences between the duplexes. The NAD cap in U1 snRNA lowers the duplex stability (Fig. 40). The Tm of TMG-U1 snRNA with HIV-1 pre-mRNA duplex was 70.5 °C (SD=0.6 °C) while the Tm of NAD snRNA with HIV-1 pre-mRNA duplex was 64.6 °C (SD=0.8 °C).

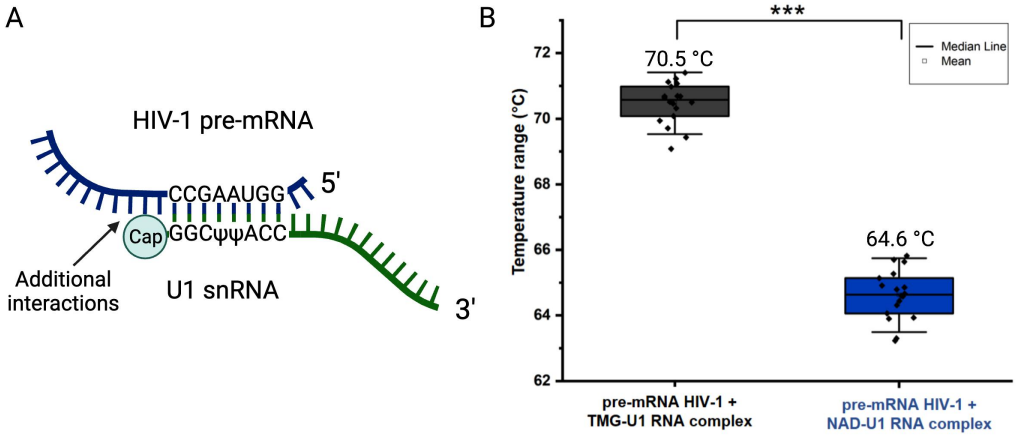


Fig. 40: Duplex stability of HIV-1 pre-mRNA and U1 snRNA (with TMG or NAD cap) measured by LightCycler. A) Schema of the prepared and annealed duplex with U1 snRNA. B) Melting temperatures measured by LightCycler. *** = p-value < 0.001.

For circular dichroism measurement had to be prepared a larger amount of RNA (13,000 ng each strand). *In vitro* transcription with ΨTP was less efficient and did not lead to a sufficient amount of RNA. Therefore, we decided to use UTP in the *in vitro* transcription mixture. The CD spectra analysis confirmed the LightCycler results and NAD cap resulted in 2.7 °C less strong interaction of U1 snRNA with HIV-1 pre-mRNA (Fig. 41, Fig. 58). CD spectra measurement was done by M.Sc. Lucie Bednárová, Ph.D.

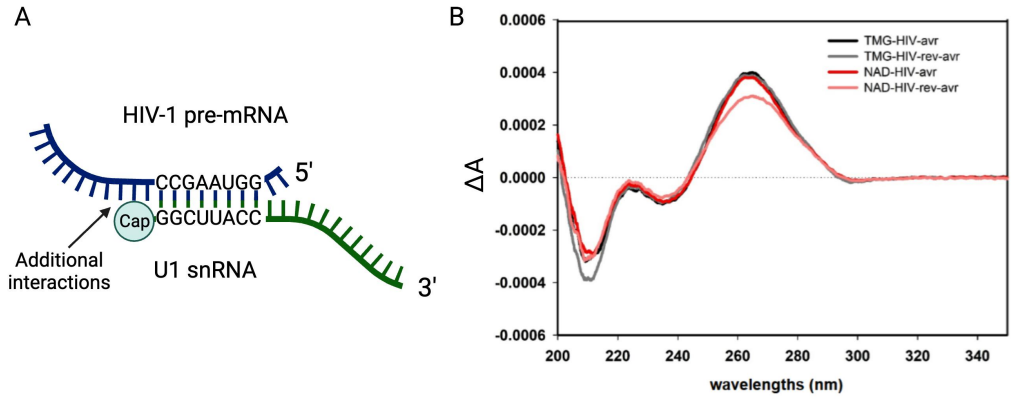


Fig. 41: Duplex stability of HIV-1 pre-mRNA and U1 snRNA (with TMG or NAD cap) measured by circular dichroism. A) Schema of the prepared and annealed duplex without pseudouridines. B) Red/pink curves represent a duplex of NAD-U1 snRNA and HIV-1 pre-mRNA, and black/grey curves represent a duplex of TMG-U1 snRNA and HIV-1 pre-mRNA. The curves are the average (avg) of three CD spectra measurements: increasing temperature from 5 °C to 95 °C (darker shades) and reverse (rev) back to 5 °C (lighter shades) expressed as differential absorption ΔA. The CD spectra of both duplexes have a positive maximum at 264 nm, which is typical for the right-handed A-type double helix of RNA.

4.9 *In vitro* assessment of DXO decapping enzyme

There are several known NAD decapping enzymes in human cells. Enzymes from the Nudix family cleave the pyrophosphate backbone of NAD cap, as well as free NAD, reducing overall concentrations of NAD. DXO removes the entire cap structure but does not hydrolyse free NAD. DXO preferentially targets subsets of NAD-RNAs among which are NAD-snoRNAs and NAD-scaRNAs (Jiao et al., 2017). Therefore, we focused on the DXO decapping enzyme.

DXO is known to have low decapping activity also on TMG cap (Jiao et al., 2013) and we examined its substrate specificity. Radioactivelly labeled U1 snRNA with TMG or NAD cap was incubated with increasing concentrations of DXO decapping enzyme (5 nM, 10 nM, 25 nM, 50 nM) or mock-treated. The data analysed by ImageJ shows that DXO preferentially decaps NAD-U1 snRNA (Fig. 42). Almost 50% of NAD-U1 snRNA was decapped by DXO after 15 minutes while only 5% of TMG-U1 snRNA was decapped.

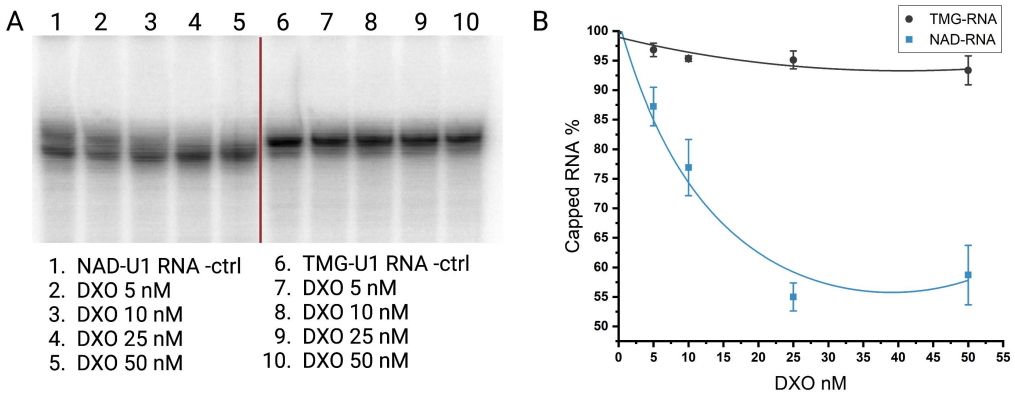


Fig. 42: Decapping of U1 snRNA by DXO enzyme. A) The scan of radioactive gel with NAD-U1 snRNA without DXO in lane 1 (-ctrl) and TMG-U1 snRNA without DXO in lane 6 (-ctrl). The upper band is the capped RNA, and the lower band represents the RNA without the cap. The other lanes contain U1 snRNA treated with increasing concentrations of DXO. The disappearance of the upper bands corresponds to the susceptibility of the substrate capped RNA to the DXO decapping enzyme. B) Analysed DXO cleavage of NAD-U1 and TMG-U1 at various DXO concentrations.

4.10 Comparison of DXO amounts in healthy and infected cells

We investigated whether the amount of DXO changes in infected and non-infected cells. We used the western blot technique to assess if the depletion of NAD caps in HIV-1 infected cells is caused by the up-regulation of DXO. The analysis revealed that the amounts of DXO in MT-4 infected cells are not elevated (Fig. 43).

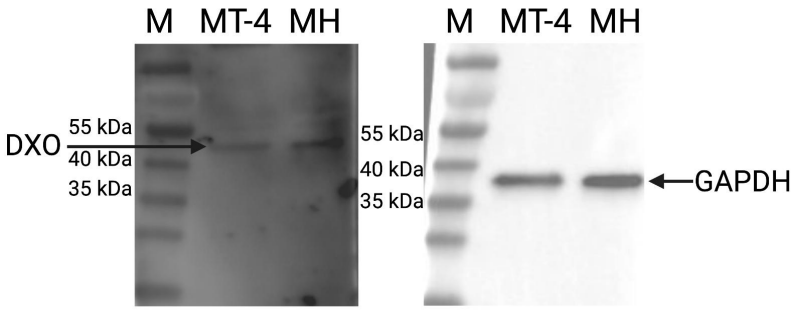


Fig. 43: Western blot analysis of DXO (44.9 kDa) and GADPH (36 kDa). The lanes were plotted, and the analysed areas of the bands were the following: DXO MT-4: 54960.170, DXO MH: 55908.877, GAPDH MT-4: 97662.437, GAPDH MH: 99765.487. M: marker, MT-4: Control MT-4 cells, MH: MT-4 cells infected by HIV-1.

4.11 Preparation of DXO transduced cells

We prepared DXO transduced cells. For DXO KD, three lentiviral vectors (each with a different target sequence) were purchased from VectorBuilder. The functionality of the KD vectors was evaluated by western blot (Fig. 44). The best results were obtained from the lentiviral vector shRNA 3 (Fig. 45), which was used in this study.

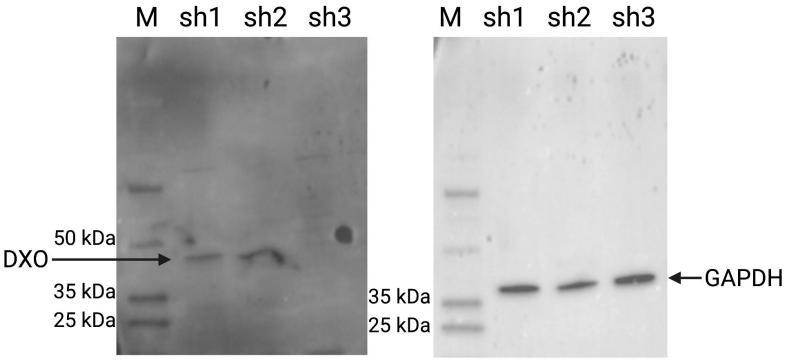


Fig. 44: Western blot evaluation of DXO (44.9 kDa) KD and GADPH (36 kDa). M: marker, sh1-sh3: shRNA1-3.

A

Vector ID	VB900039-9959kyv
Vector Name	pLV[shRNA]-EGFP:T2A:Puro-U6>hDXO[shRNA#3]
Vector Size	8347 bp
Viral Genome Size	4872 bp
Vector Type	Mammalian shRNA Knockdown Lentiviral Vector
Inserted shRNA	hDXO[shRNA#3]
Target Sequence	GTTGTTGCTGGCTTCCGTAAC
Inserted Marker	EGFP:T2A:Puro

B

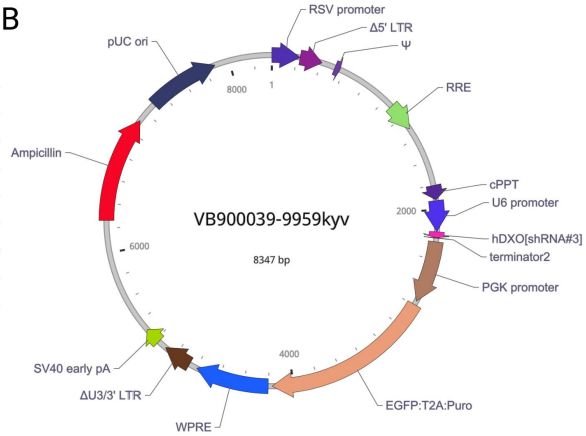


Fig. 45: Mammalian shRNA 3 DXO KD Lentiviral Vector. A) Vector summary with its target sequence. B) Vector map showing promoter areas, HIV-1 packaging signal (Ψ), and other elements.

The overexpression of DXO by a lentiviral vector is a highly efficient technique, and VectorBuilder offers only one expression system (Fig. 45).

A

Vector ID	VB900001-2728fsj
Vector Name	pLV[Exp]-EGFP:T2A:Puro-EF1A>hDXO[NM_005510.4]
Vector Size	10565 bp
Viral Genome Size	7090 bp
Vector Type	Mammalian Gene Expression Lentiviral Vector
Inserted Promoter	EF1A
Inserted ORF	hDXO[NM_005510.4]
Inserted Marker	EGFP:T2A:Puro

B

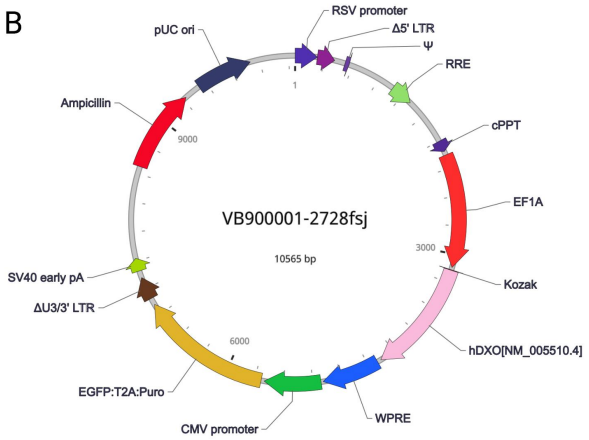


Fig. 46: Mammalian DXO overexpression lentiviral vector. A) Vector summary. B) Vector map showing promoter areas, HIV-1 packaging signal (Ψ), and other elements.

The efficacy of both systems was evaluated via western blot and compared to the wild type (WT) MT-4 cells (Fig. 47). We did not detect any DXO in the DXO KD MT-4 cells. The DXO OE led to an immense increase of the enzyme in the transduced cells. The band for DXO OE was so prevalent that the band corresponding to the DXO enzyme in WT cells was scarcely visible. Over-exposure of the membrane did not reveal any DXO in DXO KD samples, while DXO in MT-4 became more visible.

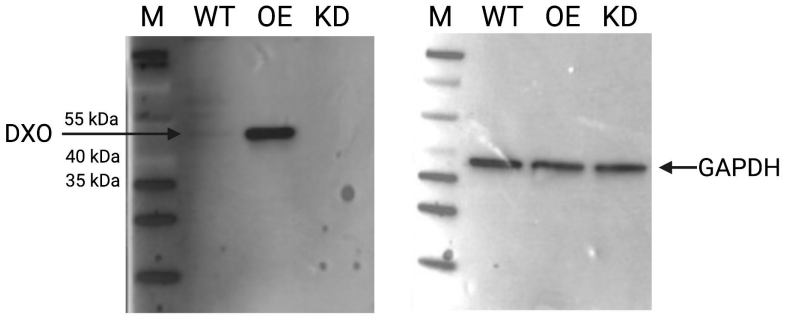


Fig. 47: Comparison of DXO (44.9 kDa) amounts in protein lysate samples from WT, DXO OE, and DXO KD MT-4 cells by western blot. GADPH (36 kDa) is a control, M: marker.

4.12 *In vivo* assessment of DXO decapping enzyme

We investigated the effect of DXO transduction on the level of NAD RNA capping in the sRNA fraction. DXO transduced cells had slower proliferation than control cells, and lower viability, especially in DXO KD cells. We measured the level of NAD caps in DXO OE and DXO KD cells by LC-MS and compared it to the control and HIV-1 infected MT-4 cells (Fig. 48, Fig. 56). Surprisingly, the level of NAD capping in RNA isolated from DXO KD cells was not elevated in comparison to the control cells. Instead, the NAD cap levels were decreased. However, the amount of NAD caps dropped drastically, more than 6-fold, in the DXO OE cells. This result confirmed that DXO functions as a deNADing enzyme in MT-4 cells, and its overexpression leads to increased NAD decapping. New samples were prepared by Ing. Pavel Vopálenský, Ph.D. and LC-MS measurement was done by M.Sc. Anton Škríba, Ph.D.

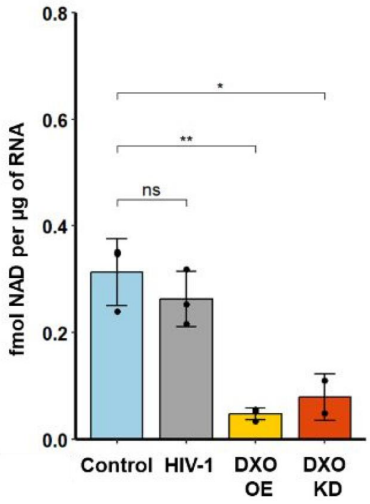


Fig. 48: Comparison of NAD cap levels in fraction of sRNA from control, HIV-1 infected, DXO OE, and DXO KD MT-4 cells. * = p-value < 0.05, ** = p-value < 0.01.

Based on our observation that HIV-1 infection leads to decreased NAD capping in U1 snRNA, we examined whether NAD capping of U1 snRNA is affected also by DXO transduction. U1 snRNA pulled down from DXO OE and DXO KD cells was measured by LC-MS and the NAD cap concentration was compared to the control and HIV-1 infected MT-4 cells (Fig. 49, Fig. 57). DXO OE leads to statistically non-significant decrease in NAD capping of U1 snRNA, while DXO KD leads to non-significant increase in NAD capping of U1 snRNA.

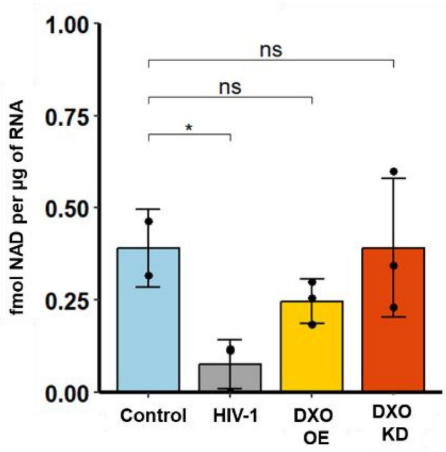


Fig. 49: Comparison of NAD cap levels on U1 snRNA in control, HIV-1 infected, DXO OE, and DXO KD MT-4 cells. * = p-value < 0.05.

4.13 Effect of DXO decapping enzyme on HIV-1 infectivity

After evaluation of the lentiviral vectors and their effects on MT-4 cells, the impact of DXO OE and DXO KD on HIV-1 infectivity was assessed. We measured the changes in cellular proliferation and viability upon HIV-1 infection via XTT colourimetric assay. The experiments were performed in four biological replicates, and the HIV-1 infectivity was expressed as $TCID_{50}$. The viral production in DXO OE MT-4 cells was the highest ($10^{4.734}$, SD=$10^{0.135}$), while in control MT-4 cells was the lowest ($10^{3.617}$, SD=$10^{0.139}$). The viral production in DXO KD cells was slightly higher than in control cells ($10^{3.955}$, SD=$10^{0.248}$) but the difference was not statistically significant (Fig. 50).

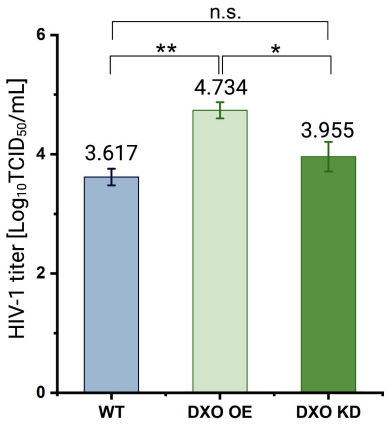


Fig. 50: Comparison of HIV-1 infectivity in WT, DXO OE, and DXO KD MT-4 cells. DXO OE significantly increased the propagation of the HIV-1 virus. n.s. = not significant, * = p-value < 0.05, ** = p-value < 0.01.

4.14 Effect of nicotinamide on HIV-1 infectivity

NAm is NAD precursor that increases intracellular NAD levels (Murray et al., 1995) and weakens the HIV-1 infection (Murray and Srinivasan, 1995). We assessed whether the NAm supplement increases intracellular NAD levels in our conditions. Since the changes in DXO KD cells are too complex, and the viability of these cells is very low, we focused on DXO OE cells. DXO OE cells, DXO OE HIV-1 infected cells, control cells, and HIV-1 infected MT-4 cells were cultured in normal media and with NAm supplemented media. The intracellular NAD levels were measured by NAD/NADH Quantification Kit. There was at least a 1.5-fold increase in the intracellular concentration of NAD, and in DXO OE infected NAm treated cells even more than a 2-fold increase in comparison to the NAm non-treated cells (Fig. 51).

We measured HIV-1 infectivity by XTT colourimetric assay. DXO OE cells and MT-4

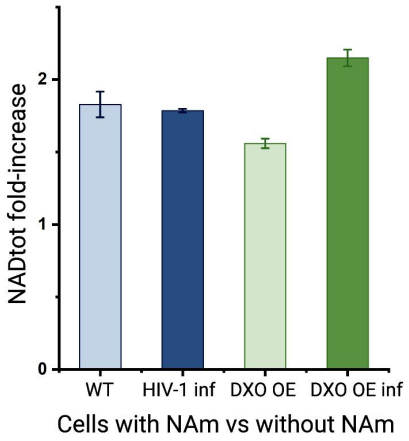


Fig. 51: Increase of cellular NAD after NAm supply expressed as fold-increase of NAD. The NAD increase was measured in control cells, HIV-1 infected MT-4 cells (HIV-1 inf), DXO OE MT-4 cells (DXO OE), and DXO OE HIV-1 infected MT-4 cells (DXO OE inf), all with and without NAm repletion.

control cells were seeded in plates in RPMI medium with or without NAm and infected by two-fold serially diluted HIV-1. As in previous experiments, the viral production in DXO OE MT-4 cells was the highest ($10^{4.51}$, SD=$10^{0.23}$). In WT MT-4 cells the $TCID_{50}$ was $10^{3.83}$ with SD=$10^{0.10}$. The viral production in MT-4 cells with NAm supply was significantly lower than in WT cells ($10^{3.37}$, SD=$10^{0.14}$) and it significantly decreased also in DXO OE cells with NAm supply where it was comparable to the WT cells ($10^{3.89}$, SD=$10^{0.09}$) (Fig. 52).

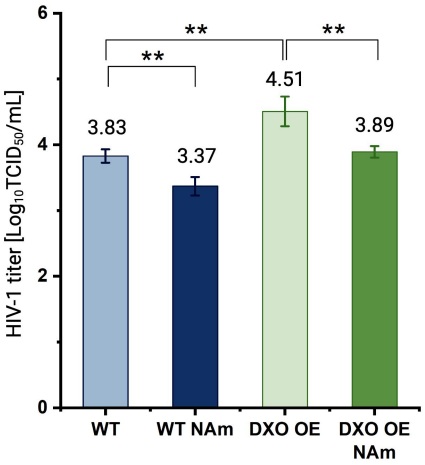


Fig. 52: Comparison of HIV-1 infectivity in WT and DXO OE MT-4 cells with or without NAm supply. NAm addition significantly decreased the propagation of the HIV-1 virus. ** = p-value < 0.01.

4.15 Changes of RNA splicing

We observed that NAD-U1 snRNA has a lower affinity to the 5' splice site of HIV-1 pre-mRNA than TMG-U1 snRNA, and we speculated that this alteration might affect the HIV-1 pre-mRNA splicing. Furthermore, we wanted to compare the changes in splicing between DXO OE, HIV-1 infected, and control MT-4 cells. To reveal the splicing patterns, lRNA was isolated from MT-4 cells, MT-4 cells infected by HIV-1, and DXO OE MT-4 cells, all with or without NAm supply, ribosomal-depleted, and the samples were submitted for RNA sequencing. The efficiency of HIV-1 pre-mRNA splicing in two main splicing regions (intron 1 and intron 2) was expressed as the ratio of reads within introns compared to exons. Fewer intron reads corresponds to more splicing events. There was statistically non-significant decrease of intron/exon ratio upon NAm addition (Fig. 53A).

We investigated the general splicing pattern in MT-4 cells under the described conditions. The top 1000 expressed genes were analysed and the results were expressed as a log of the ratio of introns and exons (Fig. 53B). The yellow colour on the heat map corresponds to higher amounts of introns and, therefore, worse splicing. The blue colour represents fewer introns and better splicing. HIV-1 infection alternated the cellular mRNA splicing. In control and HIV-1 infected MT-4 cells, the addition of NAm led to increased splicing efficiency. NAm addition in DXO OE cells did not affect the splicing pattern. Data analysis was done by M.Sc. Lenka Gahurová, Ph.D.

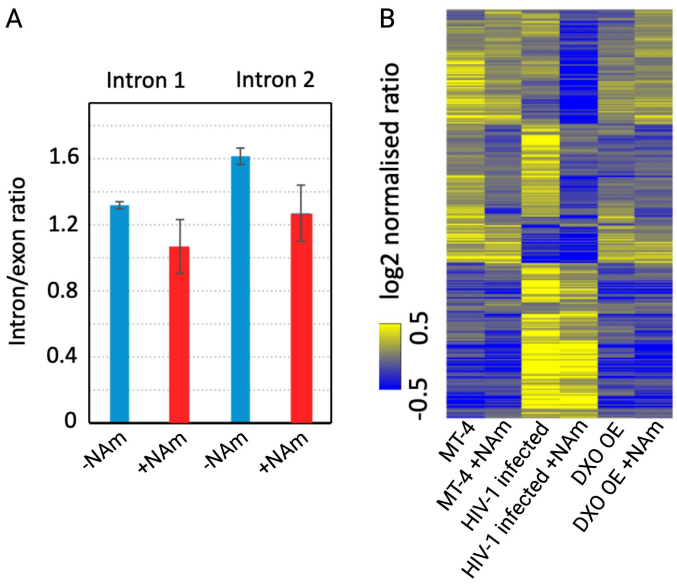


Fig. 53: Changes of RNA splicing. A) Effect of NAm addition on HIV-1 pre-mRNA splicing. B) Heat map of 1000 most expressed mRNAs from MT-4 cells, MT-4 cells infected by HIV-1, and DXO OE MT-4 cells, all with or without NAm supply.

5 Discussion

To date, more than 170 different types of RNA modifications have been identified (Boccaletto et al., 2022). HIV-1 uses RNA modifications, such as m^1A, to regulate its gene expression and replication (Burnett and McHenry, 1997). HIV-1 infection reduces free NAD levels in human cells (Murray et al., 1995). NAD was recently discovered to serve as a 5' non-canonical RNA cap, opening new research area in RNA biology (Chen et al., 2009). Several methods for capturing and sequencing NAD-RNAs have been developed, providing information on the types of modified RNAs. It was found that in bacteria, the NAD cap protects mRNA from degradation, while in mammals and plants, it may promote mRNA decay (Cahová et al., 2015; Hu et al., 2021; Zhang et al., 2019). However, little is known about NAD cap role in host-pathogen interactions. Since HIV-1 is a well-studied and characterized retrovirus, we focused our research on NAD capping in the context of HIV-1 infection. We confirmed previous findings by Murray et al. (1995) that HIV-1 infection leads to decreased NAD and NADH levels in human cells. The total NAD levels dropped almost four times upon infection (Fig. 29) and we confirmed that it is not caused by selective packing of NAD-RNAs into HIV-1 virions (Fig. 37).

Using NAD captureSeq on sRNA isolated from control and HIV-1 infected cells, we identified four snoRNAs (SNORD3G, SNORD102, SNORA50A, and SNORD3B) and four snRNAs (U1, U4ATAC, U5E, and U7) whose NAD cap levels decreased significantly upon infection (Fig. 32). NAD captureSeq was developed by Cahová et al. (2015) and it is based on the selective reaction of ADPRC and the nicotinamide of NAD on NAD-RNAs. Since then, NAD captureSeq was applied to several organisms to investigate the NAD capping pattern (Frindert et al., 2018; Wang et al., 2019b; Yu et al., 2021). The results of these studies revealed a drawback of this method: the copper used in the click reaction causes RNA fragmentation leading to a biased 5' end analyses and underestimating of the NAD-RNA number (Zhang et al., 2020). Therefore, a new method based on copper-free Strain-Promoted Azide-Alkyne Cycloaddition was developed and named SPAAC-NAD-Seq. This technique allows the identification of full-length and higher numbers of NAD-RNAs. Furthermore, the authors point out that ADPRC is a promiscuous enzyme and can interact with m^7G-RNA to a certain degree. Therefore, they employ antibodies for m^7G cap depletion (Hu et al., 2021). Recently,

another method, NADcapPro Seq, was introduced that further optimises SPAAC-NAD-Seq. In particular, the authors replaced the antibodies with an m^7G decapping enzyme to prevent m^7G-RNA contamination (Sharma et al., 2023).

Nevertheless, we aimed to explore the regulatory snRNAs and snoRNAs whose length is under 200 nt, and their fragmentation is much less probable than that of longer mRNA transcripts. Moreover, m^7G cap is on eukaryotic mRNAs while snRNAs have TMG as canonical cap and snoRNAs are intron-derived. Most importantly, SPAAC-NAD-Seq and NADcapPro Seq were developed after we started our sequencing experiments, and at that time, NAD captureSeq was the most frequently used and the most reliable method.

We provided evidence that the amounts of NAD caps are changing while the number of the identified NAD-RNAs are stable by two methods: sRNA-seq (Fig. 32) and RT-PCR (Fig. 35). Interestingly, snoRNAs have processed 5' ends as they are intron-derived and bear monophosphates on their 5' ends. It was initially thought that NAD cap could be only incorporated into RNA during transcription initiation by RNA polymerase (Bird et al., 2016). However, several recent works discovered NAD caps on snoRNAs, scaRNAs, and mtRNAs, which have exonucleolytically processed 5' ends (Bird et al., 2018; Jiao et al., 2017; Sharma et al., 2023; Zhang et al., 2020). Our study brings additional proof that RNAs can be capped with NAD after transcription initiation by yet unknown mechanism.

We investigated potential differences in NAD capping between control and HIV-1 infected cells by LC-MS. Initially, RNA samples were washed several times with RNAse-free water on columns to remove non-covalently bound cellular NAD. Nevertheless, it was shown that the samples should be washed several times with urea to ensure complete removal of cellular NAD (Frindert et al., 2018; Zhang et al., 2020). This step was included into the sample preparation. There was no statistical difference in NAD-lRNA levels from infected and control cells. We observed only statistically nonsignificant decrease in NAD-sRNA upon HIV-1 infection (Fig. 36). This could be due to similar quantities of sRNAs being enriched/depleted in the NAD captureSeq data for both control and HIV-1 infected MT-4 cells. Other possibility is that HIV-1 modulates NAD capping only in a small subset of RNAs.

U1 snRNA was the most enriched sRNA in the NAD captureSeq dataset and is involved in the splicing of HIV-1 pre-mRNAs. U1 snRNA was pulled down from control and HIV-1 infected cells and measured by LC-MS. The analysis revealed a significant decrease in the NAD cap in U1 snRNA upon HIV-1 infection (Fig. 38). These results support the findings of NAD captureSeq and formed the basis of a search for the NAD cap function.

U1 snRNA binds to the 5' splice site to define the intron sequence of HIV-1 pre-mRNA with a complementary region of 8 nucleotides. Mutations in the 5' splice site for env mRNA inhibit its expression and co-transfection with U1 snRNA containing a compensatory mutation re-establishes env mRNA expression (Lu et al., 1990). Correspondingly, mutations in the sequence of U1 snRNA inhibit viral protein expression and lead to reduced HIV-1 replication (Knoepfel et al., 2012; Mandal et al., 2010; Sajic et al., 2007). These studies prove that U1 snRNA plays an important role in the replication cycle of HIV-1 and that correct base pairing between U1 snRNA and HIV-1 pre-mRNA is essential. Therefore, we studied the stability of the HIV-1 pre-mRNA and U1 snRNA duplex by two different assay systems. Both methods revealed that NAD cap in U1 snRNA causes lower stability of the complex with HIV-1 pre-mRNA (Fig. 40, Fig. 41).

DXO is a deNADding enzyme that modulates cellular levels of NAD-RNAs (Jiao et al., 2017). We confirmed the ability of the DXO to remove the NAD cap from U1 snRNA, and we proved that the NAD cap is its preferred substrate in comparison to the TMG cap (Fig. 42). Furthermore, we showed that the amounts of DXO do not change upon HIV-1 infection (Fig. 43).

We prepared stable MT4 cell lines with DXO transduction (Fig. 47) and measured NAD cap levels in fraction of sRNA from DXO OE and DXO KD MT-4 cells by LC-MS. Surprisingly, knock down of DXO led to lower number of NAD caps in comparison to the control (Fig. 48). This discrepancy can be explained by the activity of other NAD decapping enzymes in human cells: Nudt12, Nudt16, Xrn1 or other unknown (Grudzien-Nogalska et al., 2019; Sharma et al., 2020, 2022), which could become more prominent in the absence of DXO. At the time of the experiments, only DXO and Nudt12 were known as NAD decapping enzymes. We wanted to focus only on DXO as it removes NAD

molecules from RNAs but does not hydrolyse free NAD like Nudix enzymes. Nevertheless, we tried to prepare the MT-4 cells also with DXO and Nudt12 double knock down. DXO KD led to worst viability of the cells, and double knock down suppressed the growth of the cells almost completely. Therefore, we could not obtain RNA material for LC-MS measurements. These findings suggest that NAD decapping enzymes and the NAD-RNA turn-over are important for the cells. On the other hand, overexpression of DXO led to significantly lower amounts of NAD caps in the sRNA fraction. We were interested in whether altering DXO levels influence NAD capping specifically of U1 snRNA. DXO KD led to statistically nonsignificant increase, while DXO OE to statistically nonsignificant decrease in NAD capping of U1 snRNA, suggesting that U1 snRNA is not its predominant substrate (Fig. 49).

We assessed HIV-1 infectivity in DXO transduced cells. We hypothesised that NAD caps are undesirable for HIV-1 replication and that excessive NAD cap removal might lead to increased proliferation of HIV-1 virus. Indeed, our data showed that DXO OE causes greater HIV-1 infectivity in MT-4 cells (Fig. 50). HIV-1 infectivity in DXO KD cells was slightly higher than in control cells, although not significantly. We showed that the addition of NAm to the control, HIV-1 infected, DXO OE, and DXO OE infected MT-4 cells increases the cellular levels of NAD (Fig. 51). We demonstrated that increased NAD cellular levels decrease HIV-1 infectivity. In DXO OE cells, the addition of NAm reduced HIV-1 infectivity to levels comparable to those observed in the control cells (Fig. 52).

Our last findings revealed that NAm repletion causes statistically non-significant over-splicing of HIV-1 pre-mRNA (Fig. 53). HIV-1 uses the splicing machinery of the host cells to produce its fully spliced mRNAs. Nevertheless, most viral pre-mRNAs are protected from splicing to remain at full length. Such RNA serves as the viral gRNA and encodes Gag and Gag-Pol polyproteins. Over-splicing of viral pre-mRNAs is undesirable as it leads to loss of the gRNA and decreased HIV-1 production (Bilodeau et al., 2001). Importantly, U1 snRNA is used by HIV-1 for the 5' splice site recognition, and alterations in the sequence of U1 snRNA inhibit viral protein expression and lead to reduced HIV-1 replication (Knoepfel et al., 2012; Lu et al., 1990; Mandal et al., 2010; Sajic et al., 2007).

The intron-containing transcripts retained in the nucleus are subjected to nuclear degradation or spliced to completion. U1 snRNA binds to 5' splice site of env mRNA and protects env mRNA from degradation in the nucleus. Moreover, the strength of hydrogen bonding between the splice site and U1 snRNA correlates with mRNA stability and env expression efficiency (Kammler et al., 2001; Lu et al., 1990). It may be hypothesised that NAD-U1 snRNA does not influence the splicing of HIV-1 pre-mRNA directly but rather has a destabilisation effect on the intron-containing viral mRNAs that are retained in the nucleus. Since NAD-U1 snRNA does not bind these RNAs so securely, they are not protected from degradation or over-splicing. Consequently, over-splicing results in lower amounts of gRNA, and Gag and GagPol polyproteins, which explains the lower viral production and lower HIV-1 infectivity that we observed.

Overall, our findings imply that NAD cap might play a role in antiviral defence. Possible explanation could be that the NAD cap in U1 snRNA has a destabilising effect on the complex with HIV-1 pre-mRNA, which is then not protected and cannot be exported from the nucleus by Rev protein. Instead, such RNA is retained in the nucleus and further spliced or even degraded. This leads to reduced production of viral genomic RNA, and Gag and GagPol polyproteins, and decreased assembly and proliferation of HIV-1 virions. Nevertheless, this is a very simplified explanation, and further studies would be required to determine other factors involved in the decreased HIV-1 infectivity.

6 Conclusion

To date, the studies have focused on the effect of NAD cap on mRNA half-lives and translation efficiency. In our research, we were interested in more specialised roles that NAD cap may have. Therefore, we investigated the NAD capping in a more particular model system, human cells infected by HIV-1. We focused our attention on regulatory sRNAs that lose NAD cap upon HIV-1 infection.

We demonstrate that HIV-1 infection impacts both the total cellular NAD pool and NAD capping of specific snRNAs (U1, U4ATAC, U5E, and U7) and snoRNAs (SNORD3G, SNORD102, SNORA50A, and SNORD3B). U1 snRNA is particularly interesting as it has two roles in the HIV-1 replication cycle: defining the intron sequence for splicing and stabilising HIV-1 pre-mRNA retained in the nucleus. Both processes are dependent on the strength of the bond between U1 snRNA and HIV-1 pre-mRNA. We show that NAD-U1 snRNA has a destabilising effect on this interaction, implying that NAD cap may affect HIV-1 replication and overall HIV-1 infectivity.

We provide several lines of evidence which are consistent with this hypothesis. In our DXO transducing experiments, we show that DXO overexpression leads to decreased amounts of NAD caps on sRNAs and increased HIV-1 infectivity. We demonstrate that NAD repletion in the cytoplasm leads to decreased HIV-1 proliferation. Importantly, we prove that HIV-1 infection leads to reduced NAD capping of U1 snRNA, which might be part of the virus's strategy to overcome the unfavorable effect of NAD cap.

In conclusion, this study uncovered the role of NAD cap in U1 snRNA in HIV-1 infection: affecting the stability of the complex of HIV-1 mRNA and NAD-U1 snRNA, and decreasing HIV-1 infectivity.

Contribution of the author

In the study titled "LC/MS analysis and deep sequencing reveal the accurate RNA composition in the HIV-1 virion" (Šimonová et al., 2019) I contributed by preparing the HIV-1 material at the BSL 3 laboratory. This involved producing the HIV-1 virus through infection of MT4 cells, harvesting the supernatant from infected cells, and purifying the viral particles on a sucrose cushion. The material I prepared was then used for LC-MS analysis and deep sequencing, which helped uncovering the RNA composition and modifications within the HIV-1 virions.

In the study titled: "Honeybee Iflaviruses Pack Specific tRNA Fragments from Host Cells in Their Virions" (Šimonová et al., 2022) I contributed by isolating RNA from honeybee samples and conducting Northern blot analyses using this RNA. Through this process, I learned how to perform Northern blot techniques. In this work were identified specific tRNA fragments within the virions of sacbrood virus (SBV) and deformed wing virus (DWV). The Northern blot analyses revealed that while full-length tRNAs were present in honeybee RNA, only short fragments of these tRNAs were present in viral RNA. This finding demonstrated the selective packaging of tRNA fragments by these iflaviruses.

7 References

Abdelraheim, S. R., Spiller, D. G., and McLennan, A. G. (2003): Mammalian NADH diphosphatases of the Nudix family: cloning and characterization of the human peroxisomal NUDT12 protein *Biochemical Journal*, 374(2):329–335.

Agris, P. F. (1996): The importance of being modified: roles of modified nucleosides and Mg2+ in RNA structure and function *Progress in nucleic acid research and molecular biology*, 53:79–129.

Alexandrov, A., Chernyakov, I., Gu, W., Hiley, S. L., Hughes, T. R., Grayhack, E. J., and Phizicky, E. M. (2006): Rapid tRNA decay can result from lack of nonessential modifications *Molecular cell*, 21(1):87–96.

Anderson, J., Phan, L., and Hinnebusch, A. G. (2000): The Gcd10p/Gcd14p complex is the essential two-subunit tRNA (1-methyladenosine) methyltransferase of Saccharomyces cerevisiae *Proceedings of the National Academy of Sciences*, 97(10):5173–5178.

Ashe, M. P., Pearson, L. H., and Proudfoot, N. J. (1997): The HIV-1 5 LTR poly (A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site *The EMBO journal*, 16(18):5752–5763.

Bakheet, T., Frevel, M., Williams, B. R., Greer, W., and Khabar, K. S. (2001): ARED: human AU-rich element-containing mRNA database reveals an unexpectedly diverse functional repertoire of encoded proteins *Nucleic acids research*, 29(1):246–254.

Barreau, C., Paillard, L., and Osborne, H. B. (2005): AU-rich elements and associated factors: are there unifying principles? *Nucleic acids research*, 33(22):7138–7150.

Barrett, S. P., Wang, P. L., and Salzman, J. (2015): Circular RNA biogenesis can proceed through an exon-containing lariat precursor *eLife*, 4:e07540.

Baßler, J. and Hurt, E. (2019): Eukaryotic ribosome assembly *Annu. Rev. Biochem.*, 88(1):281–306.

Bennasser, Y., Le, S.-Y., Benkirane, M., and Jeang, K.-T. (2005): Evidence that HIV-1 encodes an siRNA and a suppressor of RNA silencing *Immunity*, 22(5):607–619.

Benoni, B., Potužník, J. F., Škríba, A., Benoni, R., Trylcova, J., Tulpa, M., Spustová, K., Grab, K., Mititelu, M.-B., Pačes, J., et al. (2023): HIV-1 infection reduces NAD capping of host cell snRNA and snoRNA *ACS Chemical Biology*.

Bernstein, E., Caudy, A. A., Hammond, S. M., and Hannon, G. J. (2001): Role for a bidentate ribonuclease in the initiation step of RNA interference *Nature*, 409(6818):363–366.

Bilodeau, P. S., Domsic, J. K., Mayeda, A., Krainer, A. R., and Stoltzfus, C. M. (2001): RNA splicing at human immunodeficiency virus type 1 3 splice site A2 is regulated by binding of hnRNP A/B proteins to an exonic splicing silencer element *Journal of virology*, 75(18):8487–8497.

Bird, J. G., Basu, U., Kuster, D., Ramachandran, A., Grudzien-Nogalska, E., Towheed, A., Wallace, D. C., Kiledjian, M., Temiakov, D., Patel, S. S., Ebright, R. H., and Nickels, B. E. (2018): Highly efficient 5'capping of mitochondrial RNA with NAD+ and NADH by yeast and human mitochondrial RNA polymerase *eLife*, 7:e42179.

Bird, J. G., Zhang, Y., Tian, Y., Panova, N., Barvík, I., Greene, L., Liu, M., Buckley, B., Krásný, L., Lee, J. K., Kaplna, C. D., Ebright, R. H., and Nickels, B. E. (2016): The mechanism of RNA 5 capping with NAD+, NADH and desphospho-CoA *Nature*, 535(7612):444–447.

Birkedal, U., Christensen-Dalsgaard, M., Krogh, N., Sabarinathan, R., Gorodkin, J., and Nielsen, H. (2015): Profiling of ribose methylations in RNA by high-throughput sequencing *Angewandte Chemie International Edition*, 54(2):451–455.

Bischler, T., Hsieh, P.-k., Resch, M., Liu, Q., Tan, H. S., Foley, P. L., Hartleib, A., Sharma, C. M., and Belasco, J. G. (2017): Identification of the RNA pyrophosphohydrolase RppH of Helicobacter pylori and global analysis of its RNA targets *Journal of Biological Chemistry*, 292(5):1934–1950.

Black, D. L., Chabot, B., and Steitz, J. A. (1985): U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing *Cell*, 42(3):737–750.

Blackwood, E. M. and Kadonaga, J. T. (1998): Going the distance: a current view of enhancer action *Science*, 281(5373):60–63.

Boccaletto, P., Stefaniak, F., Ray, A., Cappannini, A., Mukherjee, S., Purta, E., Kurkowska, M., Shirvanizadeh, N., Destefanis, E., Groza, P., Avsar, G., Romitelli, Antonia annd Pir, P., Dassi, E., Conticello, S. G., Aguilo, F., and Bujnicki, J. M. (2022): MODOMICS: a database of RNA modification pathways. 2021 update *Nucleic acids research*, 50(D1):D231–D235.

Bochner, B. R., Lee, P. C., Wilson, S. W., Cutler, C. W., and Ames, B. N. (1984): AppppA and related adenylylated nucleotides are synthesized as a consequence of oxidation stress *Cell*, 37(1):225–232.

Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., and Rottman, F. M. (1997): Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA*, 3(11):1233–1247.

Borchert, G. M., Lanier, W., and Davidson, B. L. (2006): RNA polymerase III transcribes human microRNAs *Nature structural & molecular biology*, 13(12):1097–1101.

Braglia, P., Percudani, R., and Dieci, G. (2005): Sequence context effects on oligo (dT) termination signal recognition by Saccharomyces cerevisiae RNA polymerase III *Journal of Biological Chemistry*, 280(20):19551–19562.

Brengues, M., Teixeira, D., and Parker, R. (2005): Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies *Science*, 310(5747):486–489.

Brevet, A., Chen, J., Lévêque, F., Plateau, P., and Blanquet, S. (1989): In vivo synthesis of adenylylated bis (5'-nucleosidyl) tetraphosphates (Ap4N) by Escherichia coli aminoacyl-tRNA synthetases. *Proceedings of the National Academy of Sciences*, 86(21):8275–8279.

Bringmann, P., Appel, B., Rinke, J., Reuter, R., Theissen, H., and Lührmann, R. (1984): Evidence for the existence of snRNAs U4 and U6 in a single ribonucleoprotein complex and for their association by intermolecular base pairing. *The EMBO journal*, 3(6):1357–1363.

Brown, C. E. and Sachs, A. B. (1998): Poly (A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation *Molecular and cellular biology*, 18(11):6548–6559.

Burnett, B. P. and McHenry, C. S. (1997): Posttranscriptional modification of retroviral primers is required for late stages of DNA replication *Proceedings of the National Academy of Sciences*, 94(14):7210–7215.

Butler, J. S. (2002): The yin and yang of the exosome *Trends in cell biology*, 12(2):90–96.

Cahová, H., Winz, M.-L., Höfer, K., Nübel, G., and Jäschke, A. (2015): NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs *Nature*, 519(7543):374–377.

Cantó, C., Menzies, K. J., and Auwerx, J. (2015): NAD+ metabolism and the control of energy homeostasis: a balancing act between mitochondria and the nucleus *Cell metabolism*, 22(1):31–53.

Capodici, J., Karikó, K., and Weissman, D. (2002): Inhibition of HIV-1 infection by small interfering RNA-mediated RNA interference *The Journal of Immunology*, 169(9):5196–5201.

Carlile, T. M., Rojas-Duran, M. F., Zinshteyn, B., Shin, H., Bartoli, K. M., and Gilbert, W. V. (2014): Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells *Nature*, 515(7525):143–146.

Chang, C.-T., Bercovich, N., Loh, B., Jonas, S., and Izaurralde, E. (2014): The activation of the decapping enzyme DCP2 by DCP1 occurs on the EDC4 scaffold and involves a conserved loop in DCP1 *Nucleic acids research*, 42(8):5217–5233.

Chen, B. (2019): Molecular mechanism of HIV-1 entry *Trends in microbiology*, 27(10):878–891.

Chen, H., Yang, H., Zhu, X., Yadav, T., Ouyang, J., Truesdell, S. S., Tan, J., Wang, Y., Duan, M., Wei, L., Zou, L., Levine, A. S., Vasudevan, S., and Lan, L. (2020): m5C modification of mRNA serves a DNA damage code to promote homologous recombination *Nature communications*, 11(1):1–12.

Chen, Y. G., Kowtoniuk, W. E., Agarwal, I., Shen, Y., and Liu, D. R. (2009): LC/MS analysis of cellular RNA reveals NAD-linked RNA *Nature chemical biology*, 5(12):879–881.

Chiang, K., Liu, H., and Rice, A. P. (2013): miR-132 enhances HIV-1 replication *Virology*, 438(1):1–4.

Coburn, G. A. and Cullen, B. R. (2002): Potent and specific inhibition of human immunodeficiency virus type 1 replication by RNA interference *Journal of virology*, 76(18):9225–9231.

Cocquerelle, C., Mascrez, B., Hétuin, D., and Bailleul, B. (1993): Mis-splicing yields circular RNA molecules *The FASEB Journal*, 7(1):155–160.

Dahlberg, J. (1980): tRNAs as primers for reverse transcriptase *Transfer RNA: Biological Aspects*, pages 507–514.

Dai, Q., Fong, R., Saikia, M., Stephenson, D., Yu, Y.-t., Pan, T., and Piccirilli, J. A. (2007): Identification of recognition residues for ligation-based detection and quantitation of pseudouridine and N 6-methyladenosine *Nucleic acids research*, 35(18):6322–6329.

Darst, S. A. (2001): Bacterial RNA polymerase *Current opinion in structural biology*, 11(2):155–162.

Darzacq, X., Jády, B. E., Verheggen, C., Kiss, A. M., Bertrand, E., and Kiss, T. (2002): Cajal body-specific small nuclear RNAs: a novel class of 2 -O-methylation and pseudouridylation guide RNAs *The EMBO journal*, 21(11):2746–2756.

Das, A. T., Brummelkamp, T. R., Westerhout, E. M., Vink, M., Madiredjo, M., Bernards, R., and Berkhout, B. (2004): Human immunodeficiency virus type 1 escapes from RNA interference-mediated inhibition *Journal of virology*, 78(5):2601–2605.

Despic, V. and Jaffrey, S. R. (2023): mRNA ageing shapes the Cap2 methylome in mammalian mRNA *Nature*, 614(7947):358–366.

Desrosiers, R., Friderici, K., and Rottman, F. (1974): Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells *Proceedings of the National Academy of Sciences*, 71(10):3971–3975.

Dewe, J. M., Whipple, J. M., Chernyakov, I., Jaramillo, L. N., and Phizicky, E. M. (2012): The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications *Rna*, 18(10):1886–1896.

Diercks, C. S., Dik, D. A., and Schultz, P. G. (2021): Adding new chemistries to the central dogma of molecular biology *Chem*, 7(11):2883–2895.

Doamekpor, S. K., Grudzien-Nogalska, E., Mlynarska-Cieslak, A., Kowalska, J., Kiledjian, M., and Tong, L. (2020): DXO/Rai1 enzymes remove 5 -end FAD and dephospho-CoA caps on RNAs *Nucleic acids research*, 48(11):6136–6148.

Dohnalkova, M., Krasnykov, K., Mendel, M., Li, L., Panasenko, O., Fleury-Olela, F., Vågbø, C. B., Homolka, D., and Pillai, R. S. (2023): Essential roles of RNA cap-proximal ribose methylation in mammalian embryonic development and fertility *Cell reports*, 42(7).

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., Sorek, R., and Rechavi, G. (2012): Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq *Nature*, 485(7397):201–206.

Dompenciel, R. E., Garnepudi, V. R., and Schoenberg, D. R. (1995): Purification and Characterization of an Estrogen-regulated Xenopus Liver Polysomal Nuclease Involved in the Selective Destabilization of Albumin mRNA (∗) *Journal of Biological Chemistry*, 270(11):6108–6118.

Dönmez, G., Hartmuth, K., and Lührmann, R. (2004): Modified nucleotides at the 5 end of human U2 snRNA are required for spliceosomal E-complex formation *RNA*, 10(12):1925–1933.

Dunckley, T. and Parker, R. (1999): The DCP2 protein is required for mRNA decapping in Saccharomyces cerevisiae and contains a functional MutT motif *The EMBO journal*, 18(19):5411–5422.

Eaton, J. D. and West, S. (2020): Termination of transcription by RNA polymerase II: BOOM! *Trends in Genetics*, 36(9):664–675.

Eckwahl, M. J., Arnion, H., Kharytonchyk, S., Zang, T., Bieniasz, P. D., Telesnitsky, A., and Wolin, S. L. (2016): Analysis of the human immunodeficiency virus-1 RNA packageome *Rna*, 22(8):1228–1238.

Eisen, T. J., Eichhorn, S. W., Subtelny, A. O., Lin, K. S., McGeary, S. E., Gupta, S., and Bartel, D. P. (2020): The dynamics of cytoplasmic mRNA metabolism *Molecular cell*, 77(4):786–799.

Epstein, P., Reddy, R., Henning, D., and Busch, H. (1980): The nucleotide sequence of nuclear U6 (4.7 S) RNA. *Journal of Biological Chemistry*, 255(18):8901–8906.

Esteller, M. (2011): Non-coding RNAs in human disease *Nature reviews genetics*, 12(12):861–874.

Eulalio, A., Behm-Ansmant, I., Schweizer, D., and Izaurralde, E. (2007): P-body formation is a consequence, not the cause, of RNA-mediated gene silencing *Molecular and cellular biology*, 27(11):3970–3981.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998): Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans *nature*, 391(6669):806–811.

Franks, T. M. and Lykke-Andersen, J. (2007): TTP and BRF proteins nucleate processing body formation to silence mRNAs with AU-rich elements *Genes & development*, 21(6):719–735.

František Potužník, J., Nešuta, O., Škríba, A., Volenikova, B., Mititelu, M.-B., Mancini, F., Serianni, V., Fernandez, H., Spustová, K., Trylčová, J., et al. (2024): Diadenosine Tetraphosphate (Ap4A) Serves as a 5 RNA Cap in Mammalian Cells *Angewandte Chemie*, 136(6):e202314951.

Freed, E. O. (2015): HIV-1 assembly, release and maturation *Nature Reviews Microbiology*, 13(8):484–496.

Frick, D. N. and Bessman, M. J. (1995): Cloning, Purification, and Properties of a Novel NADH Pyrophosphatase: EVIDENCE FOR A NUCLEOTIDE PYROPHOSPHATASE CATALYTIC DOMAIN IN MutT-LIKE ENZYMES *Journal of Biological Chemistry*, 270(4):1529–1534.

Frindert, J., Zhang, Y., Nübel, G., Kahloon, M., Kolmar, L., Hotz-Wagenblatt, A., Burhenne, J., Haefeli, W. E., and Jäschke, A. (2018): Identification, biosynthesis, and decapping of NAD-capped RNAs in B. subtilis *Cell reports*, 24(7):1890–1901.

Fukuda, H., Chujo, T., Wei, F.-Y., Shi, S.-L., Hirayama, M., Kaitsuka, T., Yamamoto, T., Oshiumi, H., and Tomizawa, K. (2021): Cooperative methylation of human tRNA3Lys at positions A58 and U54 drives the early and late steps of HIV-1 replication *Nucleic Acids Research*, 49(20):11855–11867.

Furey, T. S., Diekhans, M., Lu, Y., Graves, T. A., Oddy, L., Randall-Maher, J., Hillier, L. W., Wilson, R. K., and Haussler, D. (2004): Analysis of human mRNAs with the reference genome sequence reveals potential errors, polymorphisms, and RNA editing *Genome research*, 14(10b):2034–2040.

Ganot, P., Caizergues-Ferrer, M., and Kiss, T. (1997): The family of box ACA small nucleolar RNAs is defined by an evolutionarily conserved secondary structure and ubiquitous sequence elements essential for RNA accumulation. *Genes & development*, 11(7):941–956.

Giegé, R. (2008): Toward a more complete view of tRNA biology *Nature structural & molecular biology*, 15(10):1007–1014.

Gomes-Filho, J. V., Breuer, R., Morales-Filloy, H. G., Pozhydaieva, N., Borst, A., Paczia, N., Soppa, J., Höfer, K., Jäschke, A., and Randau, L. (2023): Identification of NAD-RNA species and ADPR-RNA decapping in Archaea *Nature Communications*, 14(1):7597.

Gonatopoulos-Pournatzis, T. and Cowling, V. H. (2014): Cap-binding complex (CBC) *Biochemical Journal*, 457(2):231–242.

Grosjean, H., de Crécy-Lagard, V., and Marck, C. (2010): Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes *FEBS letters*, 584(2):252–264.

Grossi, E., Raimondi, I., Goñi, E., González, J., Marchese, F. P., Chapaprieta, V., Martín-Subero, J. I., Guo, S., and Huarte, M. (2020): A lncRNA-SWI/SNF complex crosstalk controls transcriptional activation at specific promoter regions *Nature Communications*, 11(1):936.

Grozhik, A. V., Olarerin-George, A. O., Sindelar, M., Li, X., Gross, S. S., and Jaffrey, S. R. (2019): Antibody cross-reactivity accounts for widespread appearance of m1A in 5'UTRs *Nature communications*, 10(1):5126.

Grudzien-Nogalska, E., Bird, J. G., Nickels, B. E., and Kiledjian, M. (2018): "NAD-capQ" detection and quantitation of NAD caps *RNA*, 24(10):1418–1425.

Grudzien-Nogalska, E., Jiao, X., Song, M.-G., Hart, R. P., and Kiledjian, M. (2016):

Nudt3 is an mRNA decapping enzyme that modulates cell migration *RNA*, 22(5):773–781.

Grudzien-Nogalska, E., Wu, Y., Jiao, X., Cui, H., Mateyak, M. K., Hart, R. P., Tong, L., and Kiledjian, M. (2019): Structural and mechanistic basis of mammalian Nudt12 RNA deNADding *Nature chemical biology*, 15(6):575–582.

Guarente, L. (2014): The many faces of sirtuins: Sirtuins and the Warburg effect *Nature medicine*, 20(1):24–25.

Gupta, S., Busch, R. K., Singh, R., and Reddy, R. (1990): Characterization of U6 small nuclear RNA cap-specific antibodies. Identification of gamma-monomethyl-GTP cap structure in 7SK and several other human small RNAs. *Journal of Biological Chemistry*, 265(31):19137–19142.

Gussakovsky, D. and McKenna, S. A. (2021): Alu RNA and their roles in human disease states *RNA biology*, 18(sup2):574–585.

Halic, M., Becker, T., Pool, M. R., Spahn, C. M., Grassucci, R. A., Frank, J., and Beckmann, R. (2004): Structure of the signal recognition particle interacting with the elongation-arrested ribosome *Nature*, 427(6977):808–814.

Hall, K. B. and McLaughlin, L. W. (1991): Properties of a U1/mRNA 5'splice site duplex containing pseudouridine as measured by thermodynamic and NMR methods *Biochemistry*, 30(7):1795–1801.

Hamm, J., Darzynkiewicz, E., Tahara, S. M., and Mattaj, I. W. (1990): The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal *Cell*, 62(3):569–577.

Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K., and Kjems, J. (2013): Natural RNA circles function as efficient microRNA sponges *Nature*, 495(7441):384–388.

Harada, F., Peters, G., and Dahlberg, J. (1979): The primer tRNA for Moloney murine leukemia virus DNA synthesis. Nucleotide sequence and aminoacylation of tRNAPro. *Journal of Biological Chemistry*, 254(21):10979–10985.

Harada, F., Sawyer, R., and Dahlberg, J. (1975): A primer ribonucleic acid for initiation of in vitro Rous sarcarcoma virus deoxyribonucleic acid synthesis. *Journal of Biological Chemistry*, 250(9):3487–3497.

Hashimoto, C. and Steitz, J. A. (1984): U4 and U6 RNAs coexist in a single small nuclear ribonucleoprotein particle *Nucleic acids research*, 12(7):3283–3293.

Hausmann, S. and Shuman, S. (2005): Specificity and mechanism of RNA cap guanine-N2 methyltransferase (Tgs1) *Journal of Biological Chemistry*, 280(6):4021–4024.

He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I., and Nogales, E. (2016): Near-atomic resolution visualization of human transcription promoter opening *Nature*, 533(7603):359–365.

Hiley, S. L., Jackman, J., Babak, T., Trochesset, M., Morris, Q. D., Phizicky, E., and Hughes, T. R. (2005): Detection and discovery of RNA modifications using microarrays *Nucleic acids research*, 33(1):e2–e2.

Höfer, K., Li, S., Abele, F., Frindert, J., Schlotthauer, J., Grawenhoff, J., Du, J., Patel, D. J., and Jäschke, A. (2016): Structure and function of the bacterial decapping enzyme NudC *Nature chemical biology*, 12(9):730–734.

Hoffman, B. E. and Grabowski, P. J. (1992): U1 snRNP targets an essential splicing factor, U2AF65, to the 3'splice site by a network of interactions spanning the exon. *Genes & development*, 6(12b):2554–2568.

Hopper, A. K. and Phizicky, E. M. (2003): tRNA transfers to the limelight *Genes & development*, 17(2):162–180.

Houzet, L., Paillart, J. C., Smagulova, F., Maurel, S., Morichaud, Z., Marquet, R., and Mougel, M. (2007): HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms *Nucleic acids research*, 35(8):2695–2704.

Hsin, J.-P. and Manley, J. L. (2012): The RNA polymerase II CTD coordinates transcription and RNA processing *Genes & development*, 26(19):2119–2137.

Hu, H., Flynn, N., Zhang, H., You, C., Hang, R., Wang, X., Zhong, H., Chan, Z., Xia, Y., and Chen, X. (2021): SPAAC-NAD-seq, a sensitive and accurate method to

profile NAD+-capped transcripts *Proceedings of the National Academy of Sciences*, 118(13):e2025595118.

Hu, J., Lutz, C. S., Wilusz, J., and Tian, B. (2005): Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation *RNA*, 11(10):1485–1493.

Hu, W.-S. and Hughes, S. H. (2012): HIV-1 reverse transcription *Cold Spring Harbor perspectives in medicine*, 2(10):a006882.

Huang, F. (2003): Efficient incorporation of CoA, NAD and FAD into RNA by in vitro transcription *Nucleic Acids Research*, 31(3):e8–e8.

Huang, J., Wang, F., Argyris, E., Chen, K., Liang, Z., Tian, H., Huang, W., Squires, K., Verlinghieri, G., and Zhang, H. (2007): Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+ T lymphocytes *Nature medicine*, 13(10):1241–1247.

Huber, J., Cronshagen, U., Kadokura, M., Marshallsay, C., Wada, T., Sekine, M., and Lührmann, R. (1998): Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure *The EMBO journal*, 17(14):4114–4126.

Hudeček, O., Benoni, R., Reyes-Gutierrez, P. E., Culka, M., Šanderová, H., Hubálek, M., Rulíšek, L., Cvačka, J., Krásný, L., and Cahová, H. (2020): Dinucleoside polyphosphates act as 5 -RNA caps in bacteria *Nature Communications*, 11(1):1052.

Isel, C., Marquet, R., Keith, G., Ehresmann, C., and Ehresmann, B. (1993): Modified nucleotides of tRNA (3Lys) modulate primer/template loop-loop interaction in the initiation complex of HIV-1 reverse transcription. *Journal of Biological Chemistry*, 268(34):25269–25272.

Iyer, M. K., Niknafs, Y. S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., Barrette, T. R., Prensner, J. R., Evans, J. R., Zhao, S., Poliakov, A., Cao, X., Dhanasekaran, S. M., Wu, Y.-M., Robinson, D. R., Beer, D. G., Feng, F. Y., Iyer, H. K., and Chinnaiyan, A. M. (2015): The landscape of long noncoding RNAs in the human transcriptome *Nature genetics*, 47(3):199–208.

Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I. W. (1994): A nuclear cap binding protein complex involved in pre-mRNA splicing *Cell*, 78(4):657–668.

Jantzen, H.-M., Admon, A., Bell, S. P., and Tjian, R. (1990): Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins *Nature*, 344(6269):830–836.

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.-G., and He, C. (2011): N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO *Nature chemical biology*, 7(12):885–887.

Jiao, X., Chang, J. H., Kilic, T., Tong, L., and Kiledjian, M. (2013): A mammalian pre-mRNA 5 end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing *Molecular cell*, 50(1):104–115.

Jiao, X., Doamekpor, S. K., Bird, J. G., Nickels, B. E., Tong, L., Hart, R. P., and Kiledjian, M. (2017): 5 end nicotinamide adenine dinucleotide cap in human cells promotes RNA decay through DXO-mediated deNADding *Cell*, 168(6):1015–1027.

Johnson, W. E. (2019): Origins and evolutionary consequences of ancient endogenous retroviruses *Nature Reviews Microbiology*, 17(6):355–370.

Jones, C. I., Zabolotskaya, M. V., and Newbury, S. F. (2012): The 5 $\rightarrow$ 3 exoribonuclease XRN1/Pacman and its functions in cellular processes and development *Wiley Interdisciplinary Reviews: RNA*, 3(4):455–468.

Julius, C. and Yuzenkova, Y. (2017): Bacterial RNA polymerase caps RNA with various cofactors and cell wall precursors *Nucleic Acids Research*, 45(14):8282–8290.

Kammler, S., Leurs, C., Freund, M., Krummheuer, J., Seidel, K., Tange, T. O., Lund, M. K., Kjems, J., Scheid, A., and Schaal, H. (2001): The sequence complementarity between HIV-1 5'splice site SD4 and U1 snRNA determines the steady-state level of an unstable env pre-mRNA. *Rna*, 7(3):421–434.

Karijolich, J. and Yu, Y.-T. (2010): Spliceosomal snRNA modifications and their function *RNA biology*, 7(2):192–204.

Kaul, D., Ahlawat, A., and Gupta, S. D. (2009): HIV-1 genome-encoded hiv1-mir-H1 impairs cellular responses to infection *Molecular and cellular biochemistry*, 323(1):143–148.

Keene, S. E., King, S. R., and Telesnitsky, A. (2010): 7SL RNA is retained in HIV-1 minimal virus-like particles as an S-domain fragment *Journal of virology*, 84(18):9070–9077.

Kirino, Y. and Mourelatos, Z. (2007): Mouse Piwi-interacting RNAs are 2 -O-methylated at their 3 termini *Nature structural & molecular biology*, 14(4):347–348.

Kiss, T. (2001): Small nucleolar RNA-guided post-transcriptional modification of cellular RNAs *The EMBO journal*, 20(14):3617–3622.

Kleiman, L. (2002): tRNALys3: The Primer tRNA for Reverse Transcription in HIV-1 *IUBMB life*, 53(2):107–114.

Knoepfel, S. A., Abad, A., Abad, X., Fortes, P., and Berkhout, B. (2012): Design of modified U1i molecules against HIV-1 RNA *Antiviral research*, 94(3):208–216.

Kowtoniuk, W. E., Shen, Y., Heemstra, J. M., Agarwal, I., and Liu, D. R. (2009): A chemical screen for biological small molecule–RNA conjugates reveals CoA-linked RNA *Proceedings of the National Academy of Sciences*, 106(19):7768–7773.

Kuehner, J. N., Pearson, E. L., and Moore, C. (2011): Unravelling the means to an end: RNA polymerase II transcription termination *Nature reviews Molecular cell biology*, 12(5):283–294.

Kuzembayeva, M., Dilley, K., Sardo, L., and Hu, W.-S. (2014): Life of psi: how full-length HIV-1 RNAs become packaged genomes in the viral particles *Virology*, 454:362–370.

Kwasnik, A., Wang, V. Y.-F., Krzyszton, M., Gozdek, A., Zakrzewska-Placzek, M., Stepniak, K., Poznanski, J., Tong, L., and Kufel, J. (2019): Arabidopsis DXO1 links RNA turnover and chloroplast function independently of its enzymatic activity *Nucleic Acids Research*, 47(9):4751–4764.

Laferté, A., Favry, E., Sentenac, A., Riva, M., Carles, C., and Chédin, S. (2006): The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components *Genes & development*, 20(15):2030–2040.

Learned, R. M., Cordes, S., and Tjian, R. (1985): Purification and characterization of a transcription factor that confers promoter specificity to human RNA polymerase I *Molecular and cellular biology*, 5(6):1358–1369.

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993): The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14 *cell*, 75(5):843–854.

Lee, S.-K., Dykxhoorn, D. M., Kumar, P., Ranjbar, S., Song, E., Maliszewski, L. E., François-Bongarçon, V., Goldfeld, A., Swamy, N. M., Lieberman, J., and Shankar, P. (2005): Lentiviral delivery of short hairpin RNAs protects CD4 T cells from multiple clades and primary isolates of HIV *Blood*, 106(3):818–826.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Rådmark, O., Kim, S., and Kim, N. V. (2003): The nuclear RNase III Drosha initiates microRNA processing *Nature*, 425(6956):415–419.

Lee, Y., Kim, M., Han, J., Yeom, K.-H., Lee, S., Baek, S. H., and Kim, V. N. (2004): MicroRNA genes are transcribed by RNA polymerase II *The EMBO journal*, 23(20):4051–4060.

Lejeune, F., Li, X., and Maquat, L. E. (2003): Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities *Molecular cell*, 12(3):675–687.

Lesbats, P., Engelman, A. N., and Cherepanov, P. (2016): Retroviral DNA integration *Chemical reviews*, 116(20):12730–12757.

Lewis, J. D., Izaurralde, E., Jarmolowski, A., McGuigan, C., and Mattaj, I. W. (1996): A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5'splice site. *Genes & development*, 10(13):1683–1698.

Li, A., Chen, Y.-S., Ping, X.-L., Yang, X., Xiao, W., Yang, Y., Sun, H.-Y., Zhu, Q., Baidya, P., Wang, X., Bhattarai, D. P., Zhao, Y.-L., Sun, B.-F., and Yang, Y.-G. (2017a): Cytoplasmic m6A reader YTHDF3 promotes mRNA translation *Cell research*, 27(3):444–447.

Li, X., Xiong, X., Zhang, M., Wang, K., Chen, Y., Zhou, J., Mao, Y., Lv, J., Yi, D., Chen, X.-W., et al. (2017b): Base-resolution mapping reveals distinct m1A methylome in nuclear-and mitochondrial-encoded transcripts *Molecular cell*, 68(5):993–1005.

Li, Y., Song, M., and Kiledjian, M. (2011): Differential utilization of decapping enzymes in mammalian mRNA decay pathways *RNA*, 17(3):419–428.

Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., Bartel, D. P., Linsley, P. S., and Johnson, J. M. (2005): Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs *Nature*, 433(7027):769–773.

Lin, H. and Spradling, A. C. (1997): A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary *Development*, 124(12):2463–2476.

Listerman, I., Sapra, A. K., and Neugebauer, K. M. (2006): Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells *Nature structural & molecular biology*, 13(9):815–822.

Liu, H., Rodgers, N. D., Jiao, X., and Kiledjian, M. (2002): The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases *The EMBO journal*, 21(17):4699–4708.

Liu, W., Jankowska-Anyszka, M., Piecyk, K., Dickson, L., Wallace, A., Niedzwiecka, A., Stepinski, J., Stolarski, R., Darzynkiewicz, E., Kieft, J., Zhao, R., Jones, D. N. M., and Davis, R. E. (2011a): Structural basis for nematode eIF4E binding an m 2, 2, 7 G-Cap and its implications for translation initiation *Nucleic acids research*, 39(20):8820–8832.

Liu, X., Bushnell, D. A., Silva, D.-A., Huang, X., and Kornberg, R. D. (2011b): Initiation complex structure and promoter proofreading *Science*, 333(6042):633–637.

Lu, X., Heimer, J., Rekosh, D., and Hammarskjöld, M. (1990): U1 small nuclear RNA plays a direct role in the formation of a rev-regulated human immunodeficiency virus env mRNA that remains unspliced. *Proceedings of the National Academy of Sciences*, 87(19):7598–7602.

Luciano, D. J. and Belasco, J. G. (2015): NAD in RNA: unconventional headgear *Trends in biochemical sciences*, 40(5):245–247.

Luciano, D. J. and Belasco, J. G. (2020): Np4A alarmones function in bacteria as precursors to RNA caps *Proceedings of the National Academy of Sciences*, 117(7):3560–3567.

Luciano, D. J., Levenson-Palmer, R., and Belasco, J. G. (2019): Stresses that raise Np4A levels induce protective nucleoside tetraphosphate capping of bacterial RNA *Molecular cell*, 75(5):957–966.

Luse, D. and Jacob, G. A. (1987): Abortive initiation by RNA polymerase II in vitro at the adenovirus 2 major late promoter. *Journal of Biological Chemistry*, 262(31):14990–14997.

Lygerou, Z., Allmang, C., Tollervey, D., and Séraphin, B. (1996): Accurate processing of a eukaryotic precursor ribosomal RNA by ribonuclease MRP in vitro *Science*, 272(5259):268–270.

Lykke-Andersen, J. (2002): Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay *Molecular and cellular biology*, 22(23):8114–8121.

Ma, S., Kong, S., Wang, F., and Ju, S. (2020): CircRNAs: biogenesis, functions, and role in drug-resistant Tumours *Molecular Cancer*, 19:1–19.

Madhani, H. D. and Guthrie, C. (1992): A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome *Cell*, 71(5):803–817.

Mandal, D., Feng, Z., and Stoltzfus, C. M. (2010): Excessive RNA splicing and inhibition of HIV-1 replication induced by modified U1 small nuclear RNAs *Journal of virology*, 84(24):12790–12800.

Maniatis, T. and Reed, R. (1987): The role of small nuclear ribonucleoprotein particles in pre-mRNA splicing *Nature*, 325(6106):673–678.

Maraia, R. J. and Lamichhane, T. N. (2011): 3 processing of eukaryotic precursor tRNAs *Wiley Interdisciplinary Reviews: RNA*, 2(3):362–375.

Marck, C. and Grosjean, H. (2002): tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features *Rna*, 8(10):1189–1232.

Marquet, R., Isel, C., Ehresmann, C., and Ehresmann, B. (1995): tRNAs as primer of reverse transcriptases *Biochimie*, 77(1-2):113–124.

Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R., and Tuschl, T. (2002): Single-stranded antisense siRNAs guide target RNA cleavage in RNAi *Cell*, 110(5):563–574.

Mattaj, I. W. (1986): Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNP protein binding *Cell*, 46(6):905–911.

Mattick, J. S. and Makunin, I. V. (2006): Non-coding RNA *Human molecular genetics*, 15(suppl_1):R17–R29.

Matzke, M. A. and Matzke, A. J. (1995): How and why do plants inactivate homologous (trans) genes? *Plant Physiology*, 107(3):679.

Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A. V., Patil, D. P., Linder, B., Pickering, B. F., Vasseur, J.-J., Chen, Q., Gross, S. S., Elemento, O., Françoise, D., Kiledjian, M., and Jaffrey, S. R. (2017): Reversible methylation of m6Am in the 5 cap controls mRNA stability *Nature*, 541(7637):371–375.

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., and Tuschl, T. (2004): Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs *Molecular cell*, 15(2):185–197.

Meyer, B. E. and Malim, M. H. (1994): The HIV-1 Rev trans-activator shuttles between the nucleus and the cytoplasm. *Genes & development*, 8(13):1538–1547.

Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., Pestova, T. V., Qian, S.-B., and Jaffrey, S. R. (2015): 5 UTR m6A promotes cap-independent translation *Cell*, 163(4):999–1010.

Miller Jr, O. and Beatty, B. R. (1969): Visualization of nucleolar genes *Science*, 164(3882):955–957.

Mlynarska-Cieslak, A., Depaix, A., Grudzien-Nogalska, E., Sikorski, P. J., Warminski, M., Kiledjian, M., Jemielity, J., and Kowalska, J. (2018): Nicotinamide-containing di- and trinucleotides as chemical tools for studies of NAD-capped RNAs *Organic Letters*, 20(23):7650–7655.

Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J., and Bordonné, R. (2002): Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus *Molecular cell*, 9(4):891–901.

Mukherjee, C., Patil, D. P., Kennedy, B. A., Bakthavachalu, B., Bundschuh, R., and Schoenberg, D. R. (2012): Identification of cytoplasmic capping targets reveals a role for cap homeostasis in translation and mRNA stability *Cell reports*, 2(3):674–684.

Müller, T. G., Zila, V., Müller, B., and Kräusslich, H.-G. (2022): Nuclear capsid uncoating and reverse transcription of HIV-1 *Annual review of virology*, 9:261–284.

Munroe, D. and Jacobson, A. (1990): mRNA poly (A) tail, a 3'enhancer of translational initiation *Molecular and cellular biology*, 10(7):3441–3455.

Murray, M. F., Nghiem, M., and Srinivasan, A. (1995): HIV infection decreases intracellular nicotinamide adenine dinucleotide [NAD] *Biochemical and biophysical research communications*, 212(1):126–131.

Murray, M. F. and Srinivasan, A. (1995): Nicotinamide inhibits HIV-1 in both acute and chronic in vitro infection *Biochemical and biophysical research communications*, 210(3):954–959.

Murthy, K. and Manley, J. L. (1995): The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation. *Genes & development*, 9(21):2672–2683.

Narayanan, M., Ramsey, K., Grebe, T., Schrauwen, I., Szelinger, S., Huentelman, M., Craig, D., Narayanan, V., and Group, C. R. (2015): Case Report: Compound heterozygous nonsense mutations in TRMT10A are associated with microcephaly, delayed development, and periventricular white matter hyperintensities *F1000Research*, 4.

Newby, M. I. and Greenbaum, N. L. (2002): Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine *Nature structural biology*, 9(12):958–965.

Ni, J., Tien, A. L., and Fournier, M. J. (1997): Small nucleolar RNAs direct site-specific synthesis of pseudouridine in ribosomal RNA *Cell*, 89(4):565–573.

Niu, K., Zhang, J., Ge, S., Li, D., Sun, K., You, Y., Qiu, J., Wang, K., Wang, X., Liu, R., et al. (2023): ONE-seq: epitranscriptome and gene-specific profiling of NAD-capped RNA *Nucleic Acids Research*, 51(2):e12–e12.

Novina, C. D., Murray, M. F., Dykxhoorn, D. M., Beresford, P. J., Riess, J., Lee, S.-K., Collman, R. G., Lieberman, J., Shankar, P., and Sharp, P. A. (2002): siRNA-directed inhibition of HIV-1 infection *Nature medicine*, 8(7):681–686.

Nübel, G., Sorgenfrei, F. A., and Jäschke, A. (2017): Boronate affinity electrophoresis for the purification and analysis of cofactor-modified RNAs *Methods*, 117:14–20.

Ohara, T., Sakaguchi, Y., Suzuki, T., Ueda, H., Miyauchi, K., and Suzuki, T. (2007): The 3 termini of mouse Piwi-interacting RNAs are 2 -O-methylated *Nature structural & molecular biology*, 14(4):349–350.

Ohira, T. and Suzuki, T. (2016): Precursors of tRNAs are stabilized by methylguanosine cap structures *Nature Chemical Biology*, 12(8):648–655.

Onafuwa-Nuga, A. A., Telesnitsky, A., and King, S. R. (2006): 7SL RNA, but not the 54-kd signal recognition particle protein, is an abundant component of both infectious HIV-1 and minimal virus-like particles *RNA*, 12(4):542–546.

Pan, S., Li, K.-e., Huang, W., Zhong, H., Wu, H., Wang, Y., Zhang, H., Cai, Z., Guo, H., Chen, X., and Xia, Y. (2020): Arabidopsis DXO1 possesses deNADding and exonuclease activities and its mutation affects defense-related and photosynthetic gene expression *Journal of integrative plant biology*, 62(7):967–983.

Pan, T. (2018): Modifications and functional genomics of human transfer RNA *Cell research*, 28(4):395–404.

Peculis, B. A., Reynolds, K., and Cleland, M. (2007): Metal determines efficiency and substrate specificity of the nuclear NUDIX decapping proteins X29 and H29K (Nudt16) *Journal of Biological Chemistry*, 282(34):24792–24805.

Pelechano, V., Wei, W., and Steinmetz, L. M. (2015): Widespread co-translational RNA decay reveals ribosome dynamics *Cell*, 161(6):1400–1412.

Pellegrini, O., Mathy, N., Condon, C., and Bénard, L. (2008): In vitro assays of 5 to 3 -Exoribonuclease activity *Methods in enzymology*, 448:167–183.

Picard-Jean, F., Brand, C., Tremblay-Létourneau, M., Allaire, A., Beaudoin, M. C., Boudreault, S., Duval, C., Rainville-Sirois, J., Robert, F., Pelletier, J., Geiss, B. J., and Bisaillon, M. (2018): 2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO *PLoS One*, 13(3):e0193804.

Purcell, D. and Martin, M. A. (1993): Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity *Journal of virology*, 67(11):6365–6378.

Ramanathan, A., Robb, G. B., and Chan, S.-H. (2016): mRNA capping: biological functions and applications *Nucleic acids research*, 44(16):7511–7526.

Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R. J., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C., and Wong-Staal, F. (1985): Complete nucleotide sequence of the AIDS virus, HTLV-III *Nature*, 313(6000):277–284.

Rehwinkel, J., Behm-Ansmant, I., Gatfield, D., and Izaurralde, E. (2005): A crucial role for GW182 and the DCP1: DCP2 decapping complex in miRNA-mediated gene silencing *RNA*, 11(11):1640–1647.

Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., and Ruvkun, G. (2000): The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans *nature*, 403(6772):901–906.

Ringeard, M., Marchand, V., Decroly, E., Motorin, Y., and Bennasser, Y. (2019): FTSJ3 is an RNA 2 -O-methyltransferase recruited by HIV to avoid innate immune sensing *Nature*, 565(7740):500–504.

Rinn, J. and Guttman, M. (2014): RNA and dynamic nuclear organization *Science*, 345(6202):1240–1241.

Robberson, B. L., Cote, G. J., and Berget, S. M. (1990): Exon definition may facilitate splice site selection in RNAs with multiple exons *Molecular and cellular biology*, 10(1):84–94.

Roca, X., Sachidanandam, R., and Krainer, A. R. (2005): Determinants of the inherent strength of human 5 splice sites *RNA*, 11(5):683–698.

Roebuck, K. A. and Saifuddin, M. (1999): Regulation of HIV-1 transcription *Gene Expression The Journal of Liver Research*, 8(2):67–84.

Roeder, R. G. (1996): The role of general initiation factors in transcription by RNA polymerase II *Trends in biochemical sciences*, 21(9):327–335.

Roeder, R. G. and Rutter, W. J. (1969): Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms *Nature*, 224(5216):234–237.

Roost, C., Lynch, S. R., Batista, P. J., Qu, K., Chang, H. Y., and Kool, E. T. (2015): Structure and thermodynamics of N6-methyladenosine in RNA: a spring-loaded base modification *Journal of the American Chemical Society*, 137(5):2107–2115.

Sajic, R., Lee, K., Asai, K., Sakac, D., Branch, D., Upton, C., and Cochrane, A. (2007): Use of modified U1 snRNAs to inhibit HIV-1 replication *Nucleic acids research*, 35(1):247–255.

Salzman, J., Gawad, C., Wang, P. L., Lacayo, N., and Brown, P. O. (2012): Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types *PloS one*, 7(2):e30733.

Samarsky, D. A., Fournier, M. J., Singer, R. H., and Bertrand, E. (1998): The snoRNA box C/D motif directs nucleolar targeting and also couples snoRNA synthesis and localization *The EMBO Journal*, 17(13):3747–3757.

Sanger, H. L., Klotz, G., Riesner, D., Gross, H. J., and Kleinschmidt, A. K. (1976): Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. *Proceedings of the National Academy of Sciences*, 73(11):3852–3856.

Saunders, A., Core, L. J., and Lis, J. T. (2006): Breaking barriers to transcription elongation *Nature Reviews Molecular Cell Biology*, 7(8):557–567.

Savarino, A., Bensi, T., Chiocchetti, A., Bottarel, F., Mesturini, R., Ferrero, E., Calosso, L., Deaglio, S., Ortolan, E., Buttò, S., Cafaro, A., Katada, T., Ensoli, B., Malavasi, F., and Dianzani, U. (2003): Human CD38 interferes with HIV-1 fusion through a sequence homologous to the V3 loop of the viral envelope glycoprotein gp120. *The FASEB journal*, 17(3):1–20.

Schaeffer, D., Tsanova, B., Barbas, A., Reis, F. P., Dastidar, E. G., Sanchez-Rotunno, M., Arraiano, C. M., and Van Hoof, A. (2009): The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities *Nature structural & molecular biology*, 16(1):56–62.

Schimmel, P. (2018): The emerging complexity of the tRNA world: mammalian tRNAs beyond protein synthesis *Nature reviews Molecular cell biology*, 19(1):45–58.

Schoenberg, D. R. and Maquat, L. E. (2012): Regulation of cytoplasmic mRNA decay *Nature Reviews Genetics*, 13(4):246–259.

Schuberth-Wagner, C., Ludwig, J., Bruder, A. K., Herzner, A.-M., Zillinger, T., Goldeck, M., Schmidt, T., Schmid-Burgk, J. L., Kerber, R., Wolter, S., Stümpel, J.-P., Roth, A., Bartok, E., Drosten, C., Coch, C., Hornung, V., Barchet, W., Kümmerer, B. M., Hartmann, G., and Schlee, M. (2015): A conserved histidine in the RNA sensor RIG-I controls immune tolerance to N1-2 O-methylated self RNA *Immunity*, 43(1):41–51.

Schwartz, S., Bernstein, D. A., Mumbach, M. R., Jovanovic, M., Herbst, R. H., León-Ricardo, B. X., Engreitz, J. M., Guttman, M., Satija, R., Lander, E. S., Fink, G., and Regev, A. (2014a): Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA *Cell*, 159(1):148–162.

Schwartz, S., Mumbach, M. R., Jovanovic, M., Wang, T., Maciag, K., Bushkin, G. G., Mertins, P., Ter-Ovanesyan, D., Habib, N., Cacchiarelli, D., Sanjana, N. E., Freinkman, E., Pacold, M. E., Satija, R., Mikkelsen, T. S., Hacohen, N., Zhang, F., Carr, S. A., Lander, E. S., and Regev, A. (2014b): Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5 sites *Cell reports*, 8(1):284–296.

Shaheen, R., Abdel-Salam, G. M., Guy, M. P., Alomar, R., Abdel-Hamid, M. S., Afifi, H. H., Ismail, S. I., Emam, B. A., Phizicky, E. M., and Alkuraya, F. S. (2015): Mutation in WDR4 impairs tRNA m7G46 methylation and causes a distinct form of microcephalic primordial dwarfism *Genome biology*, 16(1):1–11.

Shao, W. and Zeitlinger, J. (2017): Paused RNA polymerase II inhibits new transcriptional initiation *Nature genetics*, 49(7):1045–1051.

Shao, X., Zhang, H., Zhu, Z., Ji, F., He, Z., Yang, Z., Xia, Y., and Cai, Z. (2023): DpCoA tagSeq: barcoding dpCoA-Capped RNA for direct nanopore sequencing via Maleimide-thiol reaction *Analytical chemistry*, 95(29):11124–11131.

Sharma, S., Grudzien-Nogalska, E., Hamilton, K., Jiao, X., Yang, J., Tong, L., and Kiledjian, M. (2020): Mammalian Nudix proteins cleave nucleotide metabolite caps on RNAs *Nucleic acids research*, 48(12):6788–6798.

Sharma, S., Yang, J., Favate, J., Shah, P., and Kiledjian, M. (2023): NADcapPro and circNC: methods for accurate profiling of NAD and non-canonical RNA caps in eukaryotes *Communications Biology*, 6(1):406.

Sharma, S., Yang, J., Grudzien-Nogalska, E., Shivas, J., Kwan, K. Y., and Kiledjian, M. (2022): Xrn1 is a deNADding enzyme modulating mitochondrial NAD-capped RNA *Nature communications*, 13(1):889.

Sherwood, A. V., Rivera-Rangel, L. R., Ryberg, L. A., Larsen, H. S., Anker, K. M., Costa, R., Vågbø, C. B., Jakljevič, E., Pham, L. V., Fernandez-Antunez, C., et al. (2023): Hepatitis C virus RNA is 5 -capped with flavin adenine dinucleotide *Nature*, 619(7971):811–818.

Shuai, S., Suzuki, H., Diaz-Navarro, A., Nadeu, F., Kumar, S. A., Gutierrez-Fernandez, A., Delgado, J., Pinyol, M., López-Otín, C., Puente, X. S., Taylor, M. D., Campo, E., and Stein, L. D. (2019): The U1 spliceosomal RNA is recurrently mutated in multiple cancers *Nature*, 574(7780):712–716.

Shuman, S. (2000): Structure, mechanism, and evolution of the mRNA capping apparatus *Progress in Nucleic Acid Research and Molecular Biology*, 66:1–40.

Shumyatsky, G., Tillib, S., and Kramerov, D. (1990): B2 RNA and 7SK RNA, RNA polymerase III transcripts, have a cap-like structure at their 5 end *Nucleic acids research*, 18(21):6347–6351.

Šimonová, A., Romanská, V., Benoni, B., Škubník, K., Šmerdová, L., Procházková, M., Spustová, K., Moravčík, O., Gahurova, L., Pačes, J., et al. (2022): Honeybee Iflaviruses Pack Specific tRNA Fragments from Host Cells in Their Virions *ChemBioChem*, 23(17):e202200281.

Šimonová, A., Svojanovská, B., Trylčová, J., Hubálek, M., Moravčík, O., Zavřel, M., Pávová, M., Hodek, J., Weber, J., Cvačka, J., Pačes, J., and Cahova, H. (2019): LC/MS analysis and deep sequencing reveal the accurate RNA composition in the HIV-1 virion *Scientific Reports*, 9(1):8697.

Singh, R. and Reddy, R. (1989): Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. *Proceedings of the National Academy of Sciences*, 86(21):8280–8283.

Siomi, M. C., Sato, K., Pezic, D., and Aravin, A. A. (2011): PIWI-interacting small RNAs: the vanguard of genome defence *Nature reviews Molecular cell biology*, 12(4):246–258.

Slomovic, S., Laufer, D., Geiger, D., and Schuster, G. (2006): Polyadenylation of ribosomal RNA in human cells *Nucleic acids research*, 34(10):2966–2975.

Small, E. C., Leggett, S. R., Winans, A. A., and Staley, J. P. (2006): The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H box ATPase *Molecular cell*, 23(3):389–399.

Smietanski, M., Werner, M., Purta, E., Kaminska, K. H., Stepinski, J., Darzynkiewicz, E., Nowotny, M., and Bujnicki, J. M. (2014): Structural analysis of human 2 -O-ribose methyltransferases involved in mRNA cap structure formation *Nature communications*, 5(1):3004.

Smoczynski, J., Yared, M.-J., Meynier, V., Barraud, P., and Tisné, C. (2024): Advances in the Structural and Functional Understanding of m1A RNA Modification *Accounts of Chemical Research*, pages 4736–4750.

Soneoka, Y., Cannon, P. M., Ramsdale, E. E., Griffiths, J. C., Romano, G., Kingsman, S. M., and Kingsman, A. J. (1995): A transient three-plasmid expression system for the production of high titer retroviral vectors *Nucleic acids research*, 23(4):628–633.

Squires, J. E., Patel, H. R., Nousch, M., Sibbritt, T., Humphreys, D. T., Parker, B. J., Suter, C. M., and Preiss, T. (2012): Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA *Nucleic acids research*, 40(11):5023–5033.

Staley, J. P. and Guthrie, C. (1999): An RNA switch at the 5 splice site requires ATP and the DEAD box protein Prp28p *Molecular cell*, 3(1):55–64.

Statello, L., Guo, C.-J., Chen, L.-L., and Huarte, M. (2021): Gene regulation by long non-coding RNAs and its biological functions *Nature reviews Molecular cell biology*, 22(2):96–118.

Suzuki, T., Nagao, A., and Suzuki, T. (2011): Human mitochondrial tRNAs: biogenesis, function, structural aspects, and diseases *Annual review of genetics*, 45:299–329.

Swanstrom, R. and Coffin, J. (2012): HIV-1 pathogenesis: the virus *Cold Spring Harbor perspectives in medicine*, 2(12):a007443.

Taverniti, V. and Séraphin, B. (2015): Elimination of cap structures generated by mRNA decay involves the new scavenger mRNA decapping enzyme Aph1/FHIT together with DcpS *Nucleic acids research*, 43(1):482–492.

Telesnitsky, A. and Wolin, S. L. (2016): The host RNAs in retroviral particles *Viruses*, 8(8):235.

Tomecki, R. and Dziembowski, A. (2010): Novel endoribonucleases as central players in various pathways of eukaryotic RNA metabolism *Rna*, 16(9):1692–1724.

Ullu, E. and Tschudi, C. (1984): Alu sequences are processed 7SL RNA genes *Nature*, 312(5990):171–172.

Urlaub, H., Raker, V. A., Kostka, S., and Lührmann, R. (2001): Sm protein–Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure *The EMBO journal*, 20(1-2):187–196.

Vagner, S., Rüegsegger, U., Gunderson, S. I., Keller, W., and Mattaj, I. W. (2000): Position-dependent inhibition of the cleavage step of pre-mRNA 3 -end processing by U1 snRNP *Rna*, 6(2):178–188.

Valleriani, A., Zhang, G., Nagar, A., Ignatova, Z., and Lipowsky, R. (2011): Length-dependent translation of messenger RNA by ribosomes *Physical Review E*, 83(4):042903.

Van Dijk, E., Cougot, N., Meyer, S., Babajko, S., Wahle, E., and Séraphin, B. (2002): Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures *The EMBO journal*, 21(24):6915–6924.

Van Nues, R. W., Granneman, S., Kudla, G., Sloan, K. E., Chicken, M., Tollervey, D., and Watkins, N. J. (2011): Box C/D snoRNP catalysed methylation is aided by additional pre-rRNA base-pairing *The EMBO journal*, 30(12):2420–2430.

Vvedenskaya, I. O., Bird, J. G., Zhang, Y., Zhang, Y., Jiao, X., Barvík, I., Krásný, L., Kiledjian, M., Taylor, D. M., Ebright, R. H., and Nickels, B. E. (2018): CapZyme-seq comprehensively defines promoter-sequence determinants for RNA 5 capping with NAD+ *Molecular Cell*, 70(3):553–564.

Wahle, E. and Keller, W. (1992): The biochemistry of 3 -end cleavage and polyadenylation of messenger RNA precursors *Annual review of biochemistry*, 61(1):419–438.

Walters, R. W., Matheny, T., Mizoue, L. S., Rao, B. S., Muhlrad, D., and Parker, R. (2017): Identification of NAD+ capped mRNAs in Saccharomyces cerevisiae *Proceedings of the National Academy of Sciences*, 114(3):480–485.

Wang, J., Alvin Chew, B. L., Lai, Y., Dong, H., Xu, L., Balamkundu, S., Cai, W. M., Cui, L., Liu, C. F., Fu, X.-Y., Lin, Z., Shi, Pei-Yong Lu, T. K., Luo, D., Jaffrey, S. R., and Dedon, P. C. (2019a): Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA *Nucleic Acids Research*, 47(20):e130–e130.

Wang, T., Cui, Y., Jin, J., Guo, J., Wang, G., Yin, X., He, Q.-Y., and Zhang, G. (2013): Translating mRNAs strongly correlate to proteins in a multivariate manner and their translation ratios are phenotype specific *Nucleic acids research*, 41(9):4743–4754.

Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., Ren, B., Pan, T., and He, C. (2014): N6-methyladenosine-dependent regulation of messenger RNA stability *Nature*, 505(7481):117–120.

Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., and He, C. (2015): N6-methyladenosine modulates messenger RNA translation efficiency *Cell*, 161(6):1388–1399.

Wang, Y., Li, S., Zhao, Y., You, C., Le, B., Gong, Z., Mo, B., Xia, Y., and Chen, X. (2019b): NAD+-capped RNAs are widespread in the Arabidopsis transcriptome and can probably be translated *Proceedings of the National Academy of Sciences*, 116(24):12094–12102.

Wang, Z. and Kiledjian, M. (2001): Functional link between the mammalian exosome and mRNA decapping *Cell*, 107(6):751–762.

Wei, C.-M., Gershowitz, A., and Moss, B. (1975): N6, O2 -dimethyladenosine a novel methylated ribonucleoside next to the 5 terminal of animal cell and virus mRNAs *Nature*, 257(5523):251–253.

Wigington, C. P., Williams, K. R., Meers, M. P., Bassell, G. J., and Corbett, A. H. (2014): Poly (A) RNA-binding proteins and polyadenosine RNA: new members and novel functions *Wiley Interdisciplinary Reviews: RNA*, 5(5):601–622.

Wilkinson, M. E., Charenton, C., and Nagai, K. (2020): RNA Splicing by the Spliceosome *Annual review of biochemistry*, 89:359–388.

Will, C. L. and Lührmann, R. (2005): Splicing of a rare class of introns by the U12-dependent spliceosome *De Gruyter*, 386(8):713–724.

Winz, M.-L., Cahová, H., Nübel, G., Frindert, J., Höfer, K., and Jäschke, A. (2017): Capture and sequencing of NAD-capped RNA sequences with NAD captureSeq *Nature protocols*, 12(1):122–149.

Wright, D. J., Force, C. R., and Znosko, B. M. (2018): Stability of RNA duplexes containing inosine · cytosine pairs *Nucleic Acids Research*, 46(22):12099–12108.

Wurth, L., Gribling-Burrer, A.-S., Verheggen, C., Leichter, M., Takeuchi, A., Baudrey, S., Martin, F., Krol, A., Bertrand, E., and Allmang, C. (2014): Hypermethylated-capped selenoprotein mRNAs in mammals *Nucleic acids research*, 42(13):8663–8677.

Xie, N., Zhang, L., Gao, W., Huang, C., Huber, P. E., Zhou, X., Li, C., Shen, G., and Zou, B. (2020): NAD+ metabolism: pathophysiologic mechanisms and therapeutic potential *Signal transduction and targeted therapy*, 5(1):227.

Yamashita, A., Chang, T.-C., Yamashita, Y., Zhu, W., Zhong, Z., Chen, C.-Y. A., and Shyu, A.-B. (2005): Concerted action of poly (A) nucleases and decapping enzyme in mammalian mRNA turnover *Nature structural & molecular biology*, 12(12):1054–1063.

Yedavalli, V. S. and Jeang, K.-T. (2010): Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs *Proceedings of the National Academy of Sciences*, 107(33):14787–14792.

Yew, T., McCreight, L., Colclough, K., Ellard, S., and Pearson, E. (2016): tRNA methyltransferase homologue gene TRMT 10A mutation in young adult-onset diabetes with intellectual disability, microcephaly and epilepsy *Diabetic Medicine*, 33(9):e21–e25.

Yi, R., Qin, Y., Macara, I. G., and Cullen, B. R. (2003): Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs *Genes & development*, 17(24):3011–3016.

Yin, L.-B., Song, C.-B., Zheng, J.-F., Fu, Y.-J., Qian, S., Jiang, Y.-J., Xu, J.-J., Ding, H.-B., Shang, H., and Zhang, Z.-N. (2019): Elevated expression of miR-19b enhances CD8+ T cell function by targeting PTEN in HIV infected long term non-progressors with sustained viral suppression *Frontiers in Immunology*, 9:3140.

Yu, X., Willmann, M. R., Vandivier, L. E., Trefely, S., Kramer, M. C., Shapiro, J., Guo, R., Lyons, E., Snyder, N. W., and Gregory, B. D. (2021): Messenger RNA 5 NAD+ capping is a dynamic regulatory epitranscriptome mark that is required for proper response to abscisic acid in Arabidopsis *Developmental cell*, 56(1):125–140.

Zamecnik, P. G., Stephenson, M. L., Janeway, C. M., and Randerath, K. (1966): Enzymatic synthesis of diadenosine tetraphosphate and diadenosine triphosphate with a purified lysyl-sRNA synthetase *Biochemical and biophysical research communications*, 24(1):91–97.

Zhang, H., Zhong, H., Zhang, S., Shao, X., Ni, M., Cai, Z., Chen, X., and Xia, Y. (2019): NAD tagSeq reveals that NAD+-capped RNAs are mostly produced from a large number of protein-coding genes in Arabidopsis *Proceedings of the National Academy of Sciences*, 116(24):12072–12077.

Zhang, X.-O., Wang, H.-B., Zhang, Y., Lu, X., Chen, L.-L., and Yang, L. (2014): Complementary sequence-mediated exon circularization *Cell*, 159(1):134–147.

Zhang, Y., Kuster, D., Schmidt, T., Kirrmaier, D., Nübel, G., Ibberson, D., Benes, V., Hombauer, H., Knop, M., and Jäschke, A. (2020): Extensive 5 -surveillance guards against non-canonical NAD-caps of nuclear mRNAs in yeast *Nature Communications*, 11(1):5508.

Zhang, Y., Zhang, X.-O., Chen, T., Xiang, J.-F., Yin, Q.-F., Xing, Y.-H., Zhu, S., Yang, L., and Chen, L.-L. (2013): Circular intronic long noncoding RNAs *Molecular cell*, 51(6):792–806.

Zhao, G., Perilla, J. R., Yufenyuy, E. L., Meng, X., Chen, B., Ning, J., Ahn, J., Gronenborn, A. M., Schulten, K., Aiken, C., and Zhang, P. (2013): Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics *Nature*, 497(7451):643–646.

Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C.-M., Li, C. J., Vågbø, C. B., Shi, Y., Wang, W.-L., Song, S.-H., Lu, Z., Bosmans, R. P., Dai, Q., Hao, Y.-J., Yang, X., Zhao, W.-M., Tong, W.-M., Wang, X.-J., Bogdan, F., Furu, K., Fu, Y., Jia, G., Zhao, X., Liu, J., Krokan, H. E., Klungland, A., Yang, Y.-G., , and He, C. (2013): ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility *Molecular cell*, 49(1):18–29.

Zhuang, Y. and Weiner, A. M. (1986): A compensatory base change in U1 snRNA suppresses a 5 splice site mutation *Cell*, 46(6):827–835.

Zinder, J. C. and Lima, C. D. (2017): Targeting RNA for processing or destruction by the eukaryotic RNA exosome and its cofactors *Genes & development*, 31(2):88–100.

Zinshteyn, B. and Nishikura, K. (2009): Adenosine-to-inosine RNA editing *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 1(2):202–209.

Züst, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B. W., Ziebuhr, J., Szretter, K. J., Baker, S. C., Barchet, W., Diamond, M. S., Siddell, S. G., Burkhard, L., and Thiel, V. (2011): 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*, 12(2):137–143.

8 Appendix

Tab. 2: List of 94 sRNAs with enrichment in +ADPRC samples over -ADPRC control. Last three columns mark whether the sRNA passed (1) or did not pass (0) following filtering steps (St) 4, 5 and 6 as described in Fig. 25. Str: Strand, Chr: Chromosome.

sRNA	Chr	Start	End	Str	mean -HIV	mean +HIV	St 4	St 5	St 6
RNU5E–4P–201	1	11909808	11909927	-	1.06	-0.57	1	1	1
Gly	1	16678271	16678341	-	1.09	1.27	0	0	0
RNU6-40P-201	1	31497577	31497683	-	0.06	0.31	0	0	0
SNORD3G-201	1	90657750	90657964	+	3.16	1.80	1	1	1
SNORD21-201	1	92837289	92837383	+	0.66	-0.20	1	1	0
Cys	1	93516277	93516349	-	1.08	0.46	0	0	0
Gly	1	121016845	121016915	-	1.50	1.21	0	0	0
Asn	1	145287766	145287839	+	0.31	1.01	0	0	0
Gly	1	146037061	146037132	+	-0.08	1.54	0	0	0
RNVU1-1-201	1	148362370	148362533	-	1.50	0.04	1	1	1
RNVU1-30-201	1	149636766	149636929	+	0.81	0.21	1	0	0
Val	1	149708660	149708730	-	4.20	3.69	0	0	0
Met	1	153671250	153671321	+	1.25	1.32	0	0	0
Ala	2	27051214	27051286	+	0.36	1.20	1	1	1
Ile	2	42810536	42810628	+	1.09	0.61	0	0	0
RNU4ATAC-201	2	121530881	121531007	+	1.80	-0.53	1	1	1
MIR425-201	3	49020148	49020234	-	0.78	0.22	0	0	0
RNU2-64P-201	3	73110992	73111182	+	0.34	0.12	0	0	0
Cys	3	132231798	132231869	-	1.46	0.39	1	1	0
Val	3	169772230	169772302	+	1.20	0.45	0	0	0
U3.44-201	4	158700691	158700909	+	2.37	1.45	0	0	0
Leu	5	181097474	181097555	-	0.72	1.09	0	0	0
Val	5	181188416	181188488	-	1.21	1.06	0	0	0
Met	6	26286526	26286597	+	-0.07	1.28	1	1	0
Val	6	26538054	26538126	+	0.99	0.53	0	0	0
Ile	6	26554122	26554195	+	2.01	2.79	0	0	0
Ala	6	26571864	26571936	-	0.94	0.10	1	1	0
Tyr	6	26575570	26575659	+	1.02	0.32	1	1	0
Ile	6	26720992	26721065	-	0.26	-0.21	1	1	0
Ile	6	27237571	27237644	-	0.95	-0.03	0	0	0
Ser	6	27495814	27495895	+	0.60	0.82	0	0	0
Ser	6	27503039	27503120	+	0.80	1.46	0	0	0
Lys	6	27591814	27591886	-	1.95	1.50	0	0	0
Met	6	27592821	27592892	-	-0.33	0.52	1	1	0
Arg	6	27670565	27670637	-	0.78	0.68	0	0	0
Gln	6	27791356	27791427	-	0.48	0.99	0	0	0
Gln	6	27795861	27795932	-	1.75	1.73	0	0	0
Met	6	27902493	27902564	-	0.51	0.46	1	1	0
Ala	6	28607156	28607227	+	0.95	0.81	0	0	0

sRNA	Chr	Start	End	Str	mean -HIV	mean +HIV	St 4	St 5	St 6
Ala	6	28710589	28710660	+	0.81	0.79	0	0	0
Ala	6	28758364	28758435	-	3.66	3.53	0	0	0
Ala	6	28789770	28789841	-	0.59	0.36	0	0	0
Ala	6	28802800	28802870	-	0.86	0.21	1	1	0
Ala	6	28817235	28817306	-	1.26	0.49	0	0	0
Ala	6	28838444	28838515	-	-0.23	0.89	1	0	0
Leu	6	28943622	28943703	-	0.06	0.91	1	1	0
Met	6	28944575	28944647	+	0.56	0.84	0	0	0
Lys	6	28951029	28951101	+	-0.17	1.11	1	1	0
Leu	6	28989002	28989083	+	-0.48	1.01	1	1	0
Cys	7	149595214	149595285	-	-0.23	0.79	1	0	0
Y_RNA.402-201	8	99527826	99527927	+	-0.40	0.72	1	0	0
RNU1-35P-201	8	135742343	135742495	-	1.56	0.31	1	1	1
Asn	10	22229509	22229582	-	0.56	1.86	0	0	0
Ser	10	67764503	67764584	+	-0.79	1.13	1	1	0
Y_RNA.234-201	11	3663815	3663911	-	0.35	1.45	1	1	1
Lys	11	59556429	59556501	+	0.16	0.77	1	1	0
Lys	11	59560335	59560407	-	1.04	0.54	0	0	0
RNU7-1-201	12	6943816	6943878	+	1.14	0.29	1	1	1
Trp	12	98504252	98504323	+	-0.08	1.81	1	1	1
Ala	12	124921755	124921826	-	1.95	1.85	0	0	0
SNORA49-201	12	132031224	132031359	+	-0.13	1.12	1	1	0
SNORD102-201	13	27255064	27255135	+	0.82	-0.52	1	1	1
Asn	13	30673964	30674037	-	0.96	1.38	1	1	1
Glu	13	41060738	41060809	-	1.66	0.96	0	0	0
MIR19B1-201	13	91351192	91351278	+	-0.55	0.74	1	1	1
Phe	13	94549650	94549722	-	1.02	0.79	0	0	0
Pro	14	20609336	20609407	-	0.93	0.90	0	0	0
Pro	14	20684016	20684087	+	1.42	0.06	0	0	0
Lys	14	58239895	58239967	-	1.26	0.87	0	0	0
Y_RNA.351-201	14	93397139	93397234	-	0.30	1.08	0	0	0
Ile	14	102317092	102317165	+	0.23	1.18	1	1	0
Leu	16	22297140	22297221	+	1.33	0.33	0	0	0
Leu	16	57299951	57300033	+	1.09	1.23	0	0	0
SNORA50A-201	16	58559796	58559929	-	0.52	-0.71	1	1	1
Y_RNA.793-201	16	66550457	66550567	+	0.04	1.41	1	0	0
Lys	16	73478317	73478389	-	0.88	0.23	0	0	0
Gln	17	8119752	8119823	+	1.59	1.36	0	0	0
Gly	17	8221548	8221619	+	1.59	1.49	0	0	0
Ile	17	8226991	8227064	-	1.13	0.56	0	0	0
SNORD3B-2-201	17	19063346	19064136	-	4.04	0.21	1	1	1
SNORD3C-201	17	19189665	19190245	-	0.42	0.82	0	0	0
Thr	17	31550074	31550145	+	0.06	1.01	1	0	0
Cys	17	38861684	38861755	-	1.08	0.39	1	1	1

sRNA	Chr	Start	End	Str	mean -HIV	mean +HIV	St 4	St 5	St 6
RNU1-42P-201	17	48949361	48949520	-	1.90	0.32	1	1	1
U3.18-201	17	58631641	58631836	-	2.52	1.03	0	0	0
AC025362.1-201	17	64146337	64146471	+	0.40	0.50	0	0	0
SNORA50C-201	17	64146339	64146471	+	0.40	0.50	0	0	0
Met	17	82494721	82494792	-	1.33	2.35	0	0	0
RNU6-346P-201	18	76800664	76800774	-	-0.03	-0.10	0	0	0
RNU6-2-201	19	1021522	1021628	+	1.59	-0.44	1	1	0
Pseudo	19	41242237	41242309	-	0.36	1.16	0	0	0
SNORD88A-201	19	50799442	50799532	-	-0.15	0.80	1	1	0
SNORD86-201	20	2656097	2656182	+	0.56	0.78	0	0	0
RNU6-759P-201	20	35647328	35647430	-	-0.52	0.57	1	0	0

Tab. 3: RNA-seq bioinformatic analysis of sRNAs in noninfected and HIV-1 infected cells. Chr: Chromosome, Str: Probe Strand, P-value (DESeq stats p<0.05 after correction), FDR (DESeq stats p<0.05 after correction), Log2: Log2 Fold Change (DESeq stats p<0.05 after correction), ShL: Shrunk Log2 Fold Change (DESeq stats p<0.05 after correction), NI: noninfected, Inf: infected

RNA	Chr	Start	End	Str	P-value	FDR	Log2	ShL	NI	Inf
Lys	6	27576067	27576139	+	2.35E-09	5.05E-07	2.816879	2.297963	5.493564	3.064846
Y_RNA.392	12	17710934	17711046	+	1.53E-04	0.006597	3.753169	1.921711	3.249257	0.166602
Glu	13	41060738	41060809	-	3.78E-06	2.80E-04	2.506555	1.920643	5.962084	3.714684
Y_RNA.562	X	96701507	96701619	-	3.76E-06	2.80E-04	2.316154	1.840333	4.834023	2.816057
Y_RNA.83	11	93719603	93719715	+	1.61E-04	0.006656	2.307293	1.66392	4.315581	2.288709
Val	6	26538054	26538126	+	2.46E-04	0.009051	2.19048	1.596283	6.385199	4.111965
Pro	14	20633006	20633077	+	4.04E-07	4.57E-05	1.769528	1.571224	9.919452	8.327454
Lys	17	8119155	8119227	+	8.60E-07	8.40E-05	1.764935	1.557631	8.722684	7.186026
Asn	17	38751781	38751854	-	2.80E-05	0.001585	1.851575	1.540986	6.109306	4.159446
Y_RNA.17	6	99642237	99642347	-	3.89E-04	0.012873	2.063833	1.533198	3.940254	2.12569
Gly	2	1.56E+08	1.56E+08	-	2.63E-05	0.001534	1.771341	1.496111	8.240386	6.624817
Y_RNA.573	10	1.25E+08	1.25E+08	-	7.16E-05	0.003273	1.801234	1.484377	6.81489	5.317549
Pro	16	3158922	3158993	+	2.58E-05	0.001534	1.717011	1.464559	8.497743	6.896557
Y_RNA.518	20	18113467	18113579	-	4.92E-05	0.002349	1.726967	1.454194	7.173159	5.70289
Y_RNA.516	10	36626015	36626138	-	5.81E-06	4.03E-04	1.639518	1.443974	8.316263	6.900973
Y_RNA.10	1	2.41E+08	2.41E+08	+	1.74E-04	0.006938	1.752024	1.429235	5.421762	3.944697

RNA	Chr	Start	End	Str	P-value	FDR	Log2	ShL	NI	Inf
MIR4493	11	1.23E+08	1.23E+08	-	0.001641	0.043529	2.047502	1.42916	4.573041	2.746447
MIR29B1	7	1.31E+08	1.31E+08	-	1.75E-06	1.51E-04	1.585097	1.423341	8.343235	6.928911
Y_RNA.555	5	1.09E+08	1.09E+08	+	3.75E-04	0.012691	1.79408	1.41903	5.881616	4.369372
Y_RNA.476	4	37699895	37700002	-	4.74E-04	0.015445	1.816481	1.418557	5.325714	3.759005
Y_RNA.32	1	85435175	85435284	-	2.35E-04	0.009003	1.752486	1.418348	5.715662	4.190845
Gly	1	1.62E+08	1.62E+08	-	3.04E-04	0.010703	1.69115	1.377953	8.010199	6.396339
Y_RNA.435	14	63621760	63621871	+	3.49E-05	0.001875	1.517128	1.331911	10.82305	9.56219
Pro	14	20609336	20609407	-	9.84E-04	0.028186	1.620699	1.295579	6.612308	5.13773
Y_RNA.70	1	28881726	28881835	-	3.11E-04	0.010786	1.514026	1.280204	8.139081	6.858013
Y_RNA.696	17	60748940	60749046	-	5.79E-04	0.018457	1.547087	1.279207	5.337646	3.977821
Pro	14	20684016	20684087	+	8.17E-05	0.003657	1.416925	1.249693	8.480556	7.334147
MIR1468	X	63786002	63786087	-	0.001475	0.039635	1.521453	1.229377	6.145439	4.852688
Ala	12	1.25E+08	1.25E+08	-	0.001881	0.047545	1.514852	1.216813	5.429174	4.101664
Trp	17	19508181	19508252	+	6.03E-04	0.018497	1.413615	1.202712	5.418914	4.264954
U3.18	17	58631641	58631836	-	5.92E-04	0.018457	1.353132	1.16649	4.399987	3.071916
Y_RNA.653	9	70311607	70311701	-	1.93E-05	0.001221	1.0845	1.016812	7.039675	6.11551
MIR25	7	1E+08	1E+08	-	1.70E-05	0.001108	1.074476	1.009354	17.40652	16.49354
MIR942	1	1.17E+08	1.17E+08	+	0.00207	0.049986	0.885955	0.816154	9.41885	8.683841
MIR196A2	12	53991738	53991847	+	0.001445	0.039304	0.844595	0.787354	7.89752	7.222378
MIR16-2	3	1.6E+08	1.6E+08	+	2.42E-04	0.009051	0.820448	0.780155	14.21309	13.55952
MIR483	11	2134134	2134209	-	2.53E-04	0.009051	0.764608	0.731597	12.90623	12.35329
SNORD57	20	2656939	2657010	+	0.001235	0.034462ZŶ	-0.60285	-0.58191	12.73658	13.57901
SNORD83B	22	39313819	39313911	-	7.78E-04	0.023228	-0.83373	-0.78387	9.227509	10.25942
SNORD13	8	33513475	33513578	+	0.00186	0.047545	-0.93499	-0.85516	10.00742	11.0756
SNORD45A	1	75787889	75787972	+	9.82E-04	0.028186	-1.03765	-0.94119	9.279798	10.50702
Gly	21	17454789	17454859	-	3.59E-05	0.001884	-1.02155	-0.96085	13.51185	14.69965
Vault.4	5	1.36E+08	1.36E+08	-	7.14E-07	7.30E-05	-1.0333	-0.98889	7.70246	9.068642
RNU12	22	42615244	42615393	+	0.002065	0.049986	-1.26534	-1.07728	9.171878	10.57052
SNORD89	2	1.01E+08	1.01E+08	-	0.001976	0.04881	-1.27784	-1.08617	8.719049	10.17037
SNORA75	2	2.31E+08	2.31E+08	-	0.001431	0.039304	-1.28305	-1.09899	7.604465	9.045101
SNORD94	2	86135870	86136006	+	4.66E-05	0.002276	-1.23008	-1.12428	5.972054	7.395618

RNA	Chr	Start	End	Str	P-value	FDR	Log2	ShL	NI	Inf
SNORD111	16	70538005	70538098	+	2.40E-07	2.86E-05	-1.19338	-1.13099	9.531701	11.17057
SNORD67	11	46762389	46762499	-	0.001821	0.047545	-1.38906	-1.15314	5.303498	7.050073
SNORA48	17	7574713	7574847	+	7.18E-04	0.021739	-1.35028	-1.16061	3.398652	4.979776
Y_RNA.48	15	74983274	74983376	-	1.66E-04	0.006727	-1.31823	-1.17055	5.501172	6.975471
SNORD116-23	15	25091786	25091877	+	1.29E-04	0.005649	-1.34045	-1.18943	5.640678	7.250891
Lys	16	3191501	3191573	+	3.78E-04	0.012691	-1.51094	-1.27303	6.955541	8.568355
Lys	14	58239895	58239967	-	1.25E-06	1.14E-04	-1.38407	-1.27649	10.25814	11.86491
VTRNA3-1P	X	53462209	53462310	+	2.64E-05	0.001534	-1.47651	-1.30964	4.843836	6.507765
SNORD21	1	92837289	92837383	+	7.98E-04	0.023501	-1.64168	-1.31465	12.66439	14.27262
SNORD92	2	28913664	28913748	+	5.27E-05	0.002463	-1.53607	-1.33652	8.04824	9.862964
Met	8	1.23E+08	1.23E+08	-	0.001838	0.047545	-1.96982	-1.39201	3.748004	5.573608
SNORD3A	17	19188016	19188714	+	0.001028	0.029057	-2.1672	-1.48575	4.595406	6.821272
Lys	18	46089305	46089377	-	0.001934	0.048336	-2.52625	-1.4992	1.961316	4.266753
VTRNA1-1	5	1.41E+08	1.41E+08	+	3.07E-11	1.29E-08	-1.58815	-1.49956	10.43433	12.28405
SNORD15A	11	75400391	75400538	+	1.44E-05	9.70E-04	-1.78016	-1.51552	13.30564	15.23064
Glu	15	26082234	26082305	-	1.57E-04	0.006605	-1.97173	-1.53816	5.604237	7.528315
Y_RNA.266	10	88585638	88585733	-	3.19E-05	0.001755	-1.94672	-1.5885	4.233105	6.390816
MIR1246	2	1.77E+08	1.77E+08	-	4.45E-05	0.002226	-2.04627	-1.62322	6.862081	9.065045
RNU4-2	12	1.2E+08	1.2E+08	-	3.71E-06	2.80E-04	-1.92791	-1.63456	5.547839	7.616869
RNU2-29P	7	53776136	53776324	+	2.51E-04	0.009051	-2.29145	-1.63794	1.858093	4.250437
SNORD14C	11	1.23E+08	1.23E+08	-	2.16E-06	1.78E-04	-1.92397	-1.64357	6.933432	9.346354
Y_RNA.297	14	51253933	51254023	-	2.09E-04	0.008166	-2.55949	-1.70104	5.781422	8.875075
SNORD14D	11	1.23E+08	1.23E+08	-	3.61E-11	1.29E-08	-1.85959	-1.71942	8.261262	10.32275
SNORA13	5	1.12E+08	1.12E+08	+	1.09E-08	1.80E-06	-1.95549	-1.74473	6.068072	8.238724
MIR122	18	58451074	58451158	+	3.78E-05	0.001933	-2.50641	-1.80668	5.206521	7.621225
MIR215	1	2.2E+08	2.2E+08	-	1.50E-08	2.19E-06	-2.06559	-1.81535	4.276061	6.475221
SNORD14A	11	17074654	17074744	-	4.30E-06	3.08E-04	-2.28803	-1.82555	3.581273	6.048557
RN7SKP79	5	6848127	6848462	-	5.93E-04	0.018457	-3.80537	-1.85681	-2.12032	1.281619
SNORD97	11	10801467	10801608	-	4.66E-12	2.50E-09	-2.20139	-1.99241	8.977421	11.27018

RNA	Chr	Start	End	Str	P-value	FDR	Log2	ShL	NI	Inf
Glu	13	44917927	44917998	-	5.99E-07	6.43E-05	-2.54868	-2.00304	8.210685	10.91435
RNU2-3P	15	95745804	95745994	+	1.27E-06	1.14E-04	-2.6645	-2.02291	6.398067	9.188661
MIR3609	7	98881650	98881729	+	1.81E-07	2.29E-05	-2.56659	-2.05033	5.991253	8.65035
SNORD14E	11	1.23E+08	1.23E+08	-	3.23E-09	6.30E-07	-2.40956	-2.05716	5.503731	8.25771
Telomerase-vert.1	3	1.7E+08	1.7E+08	-	1.24E-07	1.66E-05	-2.71506	-2.13033	3.144856	5.92663
SNORD7	17	35573657	35573753	+	1.53E-08	2.19E-06	-2.80182	-2.23003	7.432362	10.33509
U8.4	11	1.23E+08	1.23E+08	+	1.14E-09	2.72E-07	-2.69111	-2.23639	6.508183	9.262581
RNY4P9	13	49908634	49908730	-	4.73E-10	1.27E-07	-2.66775	-2.24243	5.05719	7.913553
Y_RNA.255	9	1.33E+08	1.33E+08	-	4.55E-09	8.15E-07	-3.32347	-2.48621	4.60583	8.056031
SCARNA12	12	6967337	6967606	-	4.36E-18	4.68E-15	-3.71242	-3.11486	5.573442	9.542391
SNORD3C	17	19189665	19190245	-	2.41E-19	5.18E-16	-4.54611	-3.58749	1.042004	6.376476
RN7SL397P	3	1.2E+08	1.2E+08	+	5.52E-11	1.70E-08	-8.70295	-3.76366	-3.10515	4.363092
SNORD3D	17	19112419	19112636	-	3.17E-13	2.27E-10	-8.2338	-4.75268	-1.40851	7.710541

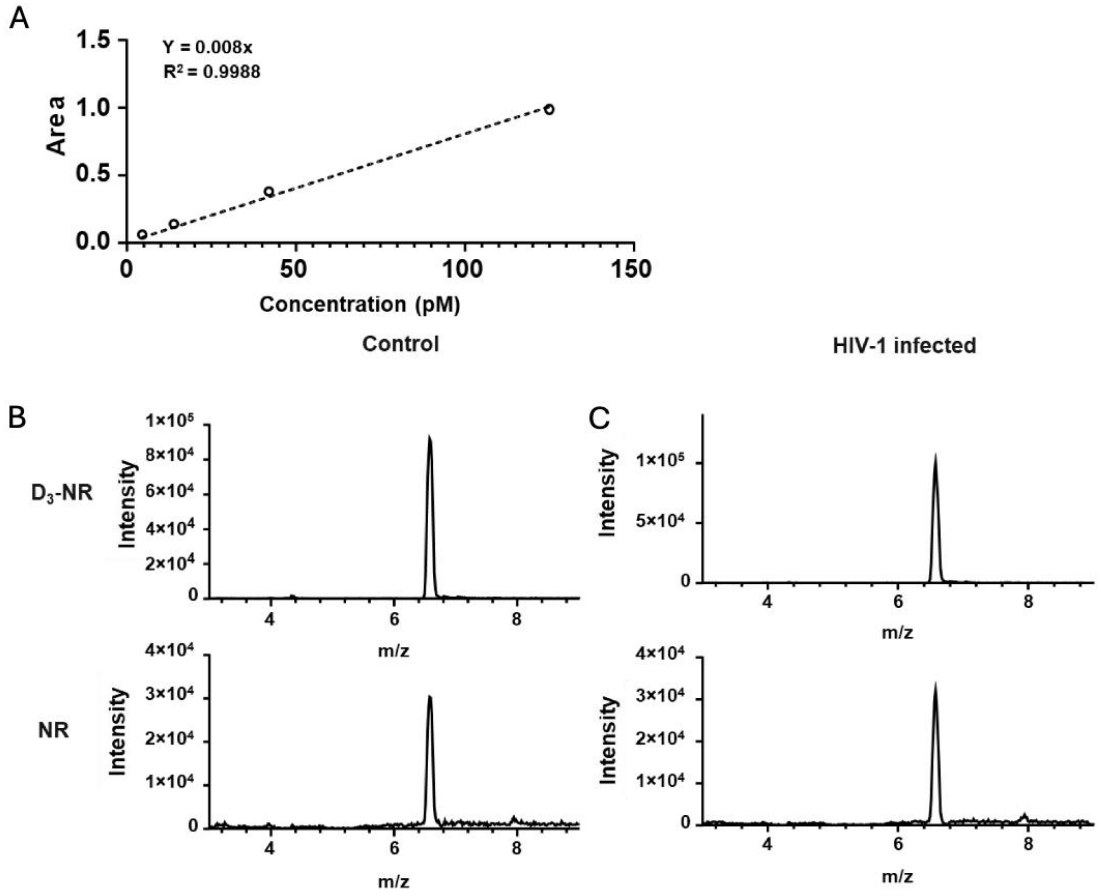


Fig. 54: LC–MS analysis of isolated sRNA and lRNA from control and HIV-1 infected cells. A) External calibration curve of NR analyte created from 4.6, 14, 42 and 125 pM concentraction. Each point is normalized to D3-NR internal standard. B) Extracted ion chromatograms of Tof-MRM transitions 258 to 126 for D3-NR used as internal standard and 255 to 123 for analysed NR in RNA from control cells. C) Extracted ion chromatograms of Tof-MRM transitions 258 to 126 for D3-NR used as internal standard and 255 to 123 for analysed NR in RNA from HIV-1 infected cells.

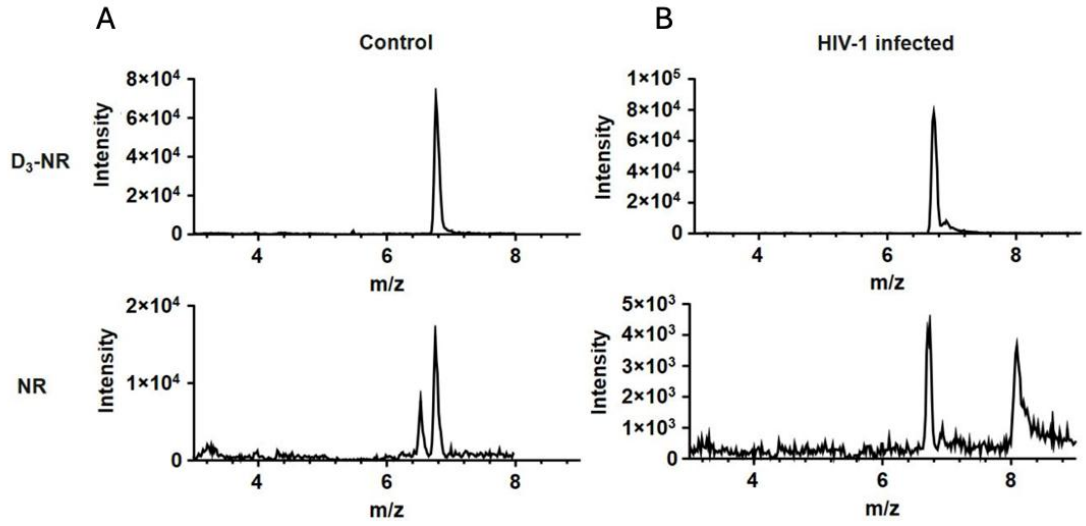


Fig. 55: LC-MS analysis of isolated U1 RNA pulled down from control and HIV-1 infected cells. A) Extracted ion chromatograms of Tof–MRM transitions 258 to 126 for D3-NR used as internal standard and 255 to 123 for analysed NR in U1 RNA from control cells. B) Extracted ion chromatograms of Tof-MRM transitions 258 to 126 for D3-NR used as internal standard and 255 to 123 for analysed NR in U1 RNA from HIV-1 infected cells.

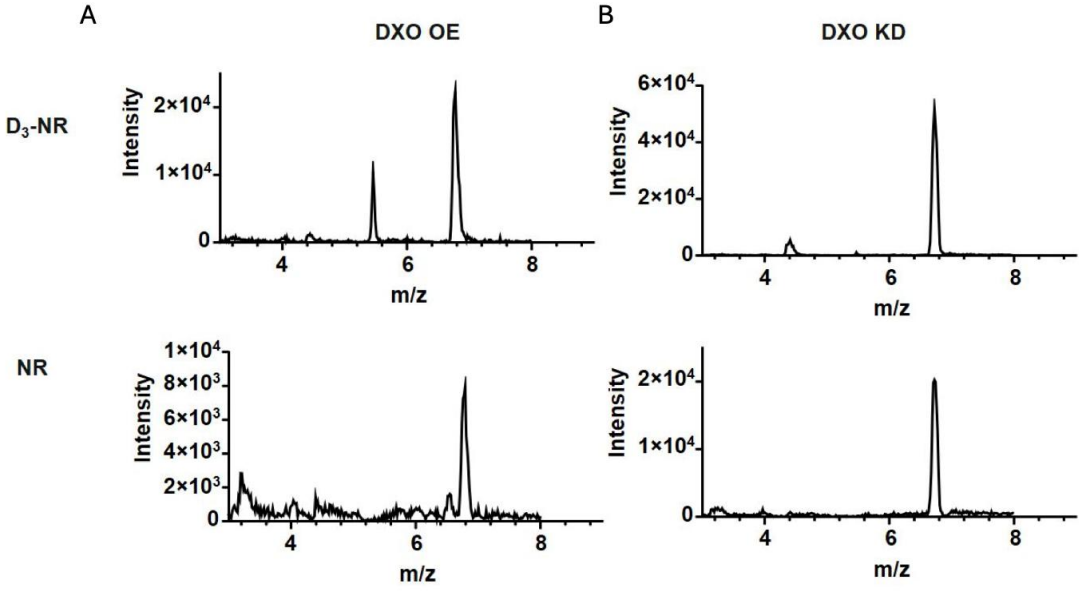


Fig. 56: LC-MS analysis of isolated sRNA from MT4-DXO OE and MT4-DXO KD cells. A) Extracted ion chromatograms of Tof-MRM transitions 258 to 126 for D3-NR used as internal standard and 255 to 123 for analysed NR in RNA from MT4-DXO OE. B) Extracted ion chromatograms of Tof-MRM transitions 258 to 126 for D3-NR used as internal standard and 255 to 123 for analysed NR in RNA from MT4-DXO KD cells.

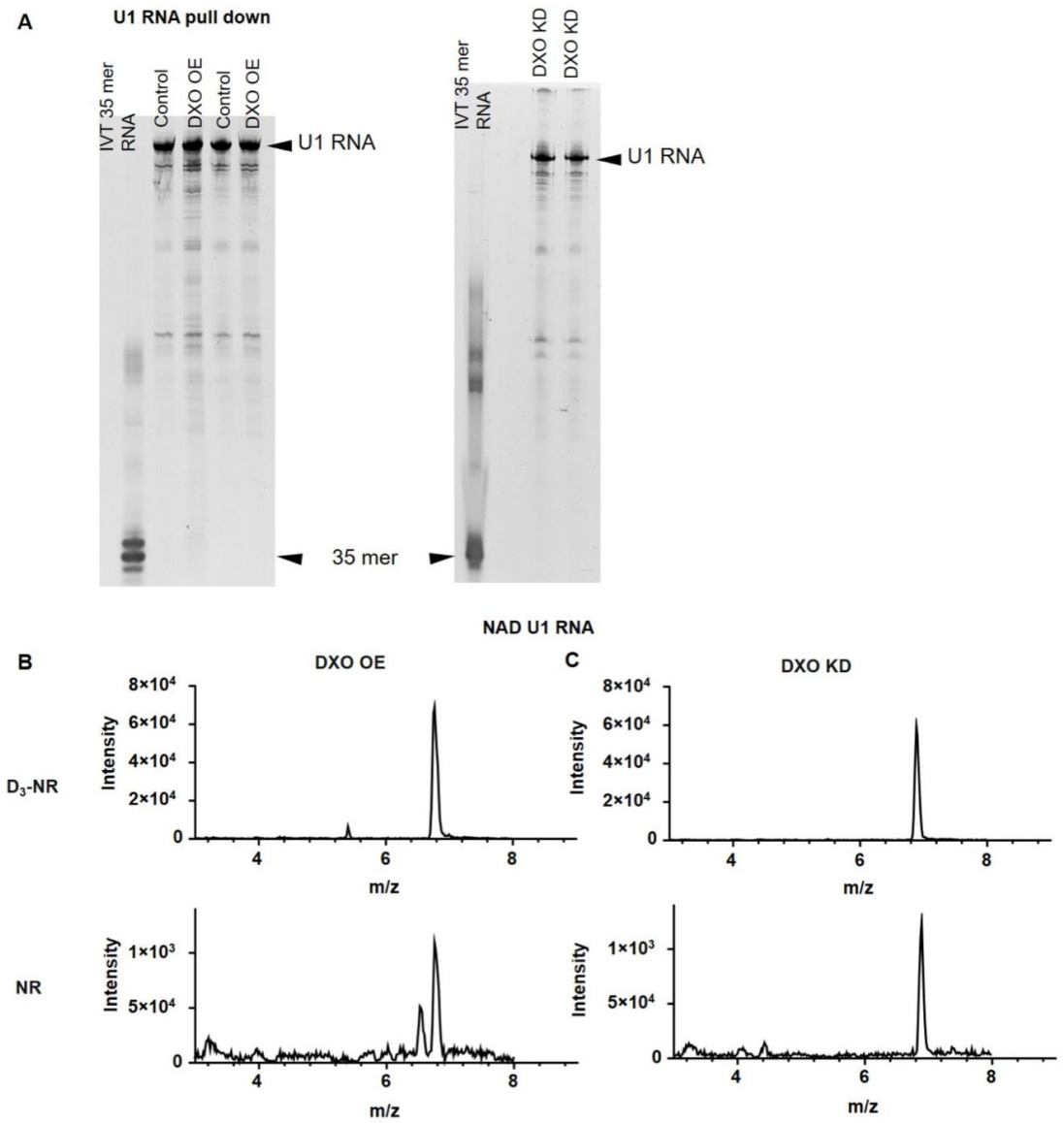


Fig. 57: LC-MS analysis of isolated U1 RNA pulled down from control and HIV-1 infected cells. A) PAGE (12.5%) analysis of U1 RNA pulled down from control, MT4-DXO OE and MT4 DXO KD cells. B) Extracted ion chromatograms of Tof-MRM transitions 258 to 126 for D3-NR used as internal standard and 255 to 123 for analysed NR in U1 RNA from MT4-DXO OE cells. C) Extracted ion chromatograms of Tof-MRM transitions 258 to 126 for D3-NR used as internal standard and 255 to 123 for analysed NR in U1 RNA from MT4-DXO KD cells.

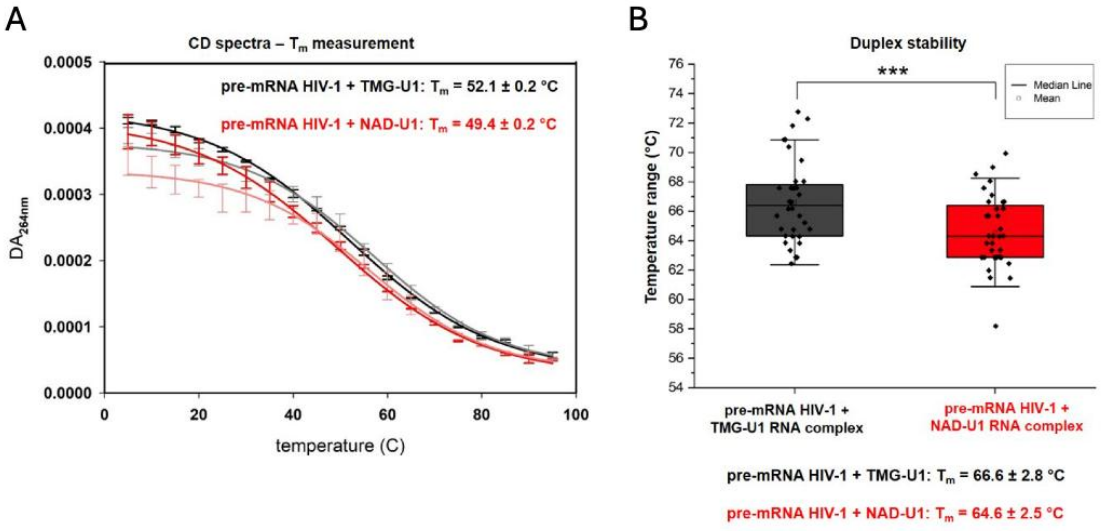


Fig. 58: NAD cap of U1 snRNA (without pseudouridines) destabilizes the complex with HIV-1 pre-mRNA. A) Tm measurement of duplex stability employing CD. B) Duplex stability measurement of pre-mRNA HIV-1 with TMG-U1 (dark grey) or NAD-U1 (red) by light cycler.

Synthesis of $m_3^{2,2,7}$GpppApG (TMGpppApG, TMG cap)

To enable synthesis of highly homogenous 5‘–TMG–capped RNA by in vitro transcription we designed a trinucleotide TMG cap analog (TMGpppApG or $m_3^{2,2,7}$GpppApG) following the synthetic pathway described previously for NAD–derived trinucleotides.8 All organic solvents that were used in the chemical syntheses under anhydrous conditions: dimethyl sulfoxide (DMSO, Honeywell, HPLC grade), N, N– dimethylformamide (DMF, anhydrous, Sigma–Aldrich) and acetonitrile (ACN, HPLC grade, J.T.Baker), were additionally dried over 4A molecular sieves. Other solvents: diethyl ether (CHEMPUR, p.a.), acetone (CHEMPUR, p.a.), methanol (MeOH, J.T. Baker, HPLC grade), acetic anhydride (CHEMPUR, p.a.) were used as received. Reversed phase HPLC (RP–HPLC) was carried out on a Agilent system for the analysis and final purification of the compounds. A Gemini column (NX–C18, 150 mm x 4.6 mm, 3 m) was used to analyse the reactions progress with a flow rate of 1 mL min-1. The solution of 50 mM ammonium acetate (CH3COONH4) pH 5.9, and the mixture of 50 mM CH3COONH4, pH 5.9 and ACN (1/1, V/V) were used as buffers. HiCHROM C18, 150 mm x10 mm, 5 m, with a flow rate of 4.7 mL min-1 was used as semi– preparative column. Buffers for RP–HPLC were as follows: A: 50 mM CH3COONH4, pH 5.9, B: 50 mM CH3COONH4, pH 5.9 / MeOH, 1/1, V/V. DEAE Sephadex A–25 was used for purification by ion exchange column chromatography. Triethylammonium bicarbonate (TEAB) buffer in deionized water was used as the

mobile phase in a linear gradient: 0 to 0.7 M TEAB for nucleoside monophosphate, 0 to 0.9 M TEAB for nucleoside diphosphate, 0 to 1.2 M TEAB for dinucleoside triphosphate. TMG cap ($m_3^{2,2,7}$GpppApG) was synthesised following the procedures formerly reported with our modifications8, 9. The synthetic pathway consisted of 9 steps. A convergent synthesis approach was used to minimize the loss of yields in each step. The starting substrate was commercially available guanosine. The scale of the individual steps as well as the techniques used to purify the intermediates were tailored to our needs. Guanosine methylation at the N7 position was carried out in the final step of the synthesis of m32,2,7GDP to avoid degradation of the compound during the 5'–phosphorylation.

a. 2',3',5'–tri–O–acetylguanosine **(1)**

Guanosine (8.83 mmol, 2.5 g, Carbosynth) was suspended in anhydrous acetonitrile (100 mL) and then 4– dimethylaminopyridine (DMAP, 0.663 mmol, 81 mg, Sigma–Aldrich) and TEA (4.87 mL) were added. The solution was stirred for 5 minutes on a magnetic stirrer at room temperature and then acetic anhydride (3 mL) was added. The reaction was carried out with continuous stirring for 50 minutes at room temperature in a closed flask. After this time, the reaction was terminated by adding 1.25 mL of methanol to the flask and stirring for a while more. The solvents were evaporated to dryness on a rotary evaporator and then cold acetone (50 mL) was added and refrigerated for 0.5h. A white precipitate precipitated on the flask, which was filtered on a Buchner funnel and washed with an additional portion of cold acetone (50 mL). The product was dried in a desiccator to give 2.18 g of the compound as a white powder. The reaction progress was controlled by RP–HPLC and low resolution MS.

b. 2',3',5'–Tri–O–acetyl–2–N,2–N–dimethylguanosine **(2)**

To a solution of compound (1) (5.0 g, 12 mmol, 1 eq.) in 99.9% acetic acid (CH3COOH, 150 mL, J.T.Baker) paraformaldehyde (1.1 g, 37 mmol, 3 eq., Sigma–Aldrich) was added and stirred for 3h at 45°C. After this time, sodium cyanoborohydride (NaBH3CN, 2.3 g, 37 mmol, 3 eq., Sigma–Aldrich) was added to the mixture and again stirred vigorously for 3 h at 45°C. The sequence of adding paraformaldehyde and NaBH3CN at 3 h intervals was repeated two more times, adding a total of 9 eq. of each. The progress of the reaction was monitored by RP–HPLC. The solvent was evaporated on a rotary evaporator and the precipitate was dissolved in a mixture of dichloromethane (DCM / pyridine, 1/1, v/v). The solution was washed 3 times with saturated NaHCO3 and the aqueous layers were

extracted again with DCM / pyridine solution (2/1, v/v). The organic layer was collected and dried over Na2SO4 and then filtered and evaporated on a rotary evaporator with toluene until residual pyridine was removed. The purity of the compound was confirmed by low resolution MS. The product was obtained as a pale grey solid.

c. 2–N,2–N–dimethylguanosine **(3)**

Compound (2) (5 g, 12.21 mmol, 1 eq.) was dissolved in H2O / methanol solution (1/1, v/v, 45 mL) and heated up to 71°C. After reaching this temperature, TEA (11.2 mL, 80 mmol, 7 eq., Sigma–Aldrich) was added and the whole mixture was stirred for another 4 h. The course of the reaction was controlled by RP–HPLC and low resolution MS. The solvent was evaporated under reduced pressure and then dried in a desiccator. The product was obtained as a white solid.

d. 2–N,2–N–dimethylguanosine 5'–monophosphate **(4)**

Compound (3) (650 mg, 2.09 mmol, 1 eq.) was suspended in trimethyl phosphate (PO(OCH3)3, 4.4 mL, 37.5 mmol, 18 eq., Sigma–Aldrich), then pre–distilled phosphoryl chloride (POCl3, 0.58 mL, 6.26 mmol, 3 eq., Merck) was added and the flask was closed with a septa. The whole reaction was stirred for 2 h on ice and the reaction was controlled by RP–HPLC. The reaction was terminated by adding 45 mL of 0.6M TEAB solution. The compound was purified on an ion exchange column using a linear gradient of TEAB buffer (0 – 0.7 M). After evaporation of the buffer, the precipitate was dried in a desiccator. Finally, the white solid was obtained.

e. 2–N,2–N–dimethylguanosine 5'–monophosphate P–imidazolide **(5)**

Compound (4) (0.9 g, 2.31 mmol, 1 eq.), imidazole (1.6 g, 23.1 mmol, 10 eq., Merck) and DTDP (1.6 g, 6.93 mmol, 3 eq., Sigma–Aldrich) were suspended in anhydrous DMF (7 mL). TEA (6.93 mmol, 0.97 mL, 3 eq.)

and PPh3 (1.82 g, 6.93 mmol, 3 eq.), both from Sigma–Aldrich, were then added and stirred for 3 h at room temperature. After this time, the reaction was terminated by adding a solution of LiClO4 (0.74 g, 6.93 mmol, 3 eq.) in cold ACN (70 mL). After cooling the mixture for 1 h in a refrigerator, the precipitate was centrifuged, washing each time with ACN. The final product was dried in a desiccator over P4O10 to give a light grey precipitate.

f. 2–N,2–N–dimethylguanosine 5'–diphosphate **(6)**

Compound (5) (954 mg, 2.16 mmol, 1 eq.) was suspended in DMF (25 mL) followed by the addition of orthophosphoric acid (V) TEA salt (3.7 g, 9.24 mmol, 4 eq.) and stirred

until dissolution. ZnCl2 (5.04 mg, 36.96 mmol, 16 eq., Sigma–Aldrich) was then added and stirred vigorously at room temperature for 3 h. The progress of the reaction was controlled by RP–HPLC and low resolution MS. After 3 h, the reaction was terminated by adding an aqueous solution (250 mL) of disodium EDTA (13.76 g, 36.96 mmol, 8 eq.) and NaHCO3 (1/2 m EDTA) to adjust to pH 7. Compound (6) was purified on a DEAE Sephadex A–25 ion exchange column using a linear gradient of TEAB (0 – 0.9 M) and then evaporated to dryness and lyophilized. The product was obtained as a white powder.

g. 2–N,2–N,7–N–trimethylguanosine 5'–diphosphate **(7)**

Compound (6) (348.54 mg, 0.74 mmol) was dissolved in water (7.22 mL). Using glacial CH3COOH, the solution was adjusted to pH 4 and the mixture was stirred for 1 h at room temperature. The progress of the reaction was controlled by RP–HPLC. Then, (CH3O)2SO2 (0.72 mL, 7.6 mol) was added and stirred vigorously. The solution of 1M NaOH was used to maintain pH 4. The reaction was terminated after 2.5 h and then extracted 3 times with DCM. The aqueous layer was collected and concentrated on a rotary evaporator. Compound (7) was purified on a DEAE Sephadex A–25 ion exchange column using a linear gradient of TEAB (0 – 0.9 M) and then evaporated to dryness and lyophilized. The product was obtained as a white powder.

h. Solid–phase synthesis of pApG **(8)**

Automatic solid phase syntheses of dinucleotide 8 was carried out using the phosphoramidite chemistry. The previously reported procedure (Mlynarska–Cieslak et al.) was applied, increasing the reaction scale to 50 M. Modifications were introduced at the stage of cleavage of the dinucleotide from the support, increasing the time of incubation of the sample with AMA (ammonium hydroxide/methyl amine, 1/1, v/v) (1.5 h instead of 1 h) and increasing the incubation temperature (55°C instead of 37 °C). This maximised the final efficiency of this step. The further treatments were unchanged. Compound 8 was obtained as a pale grey solids with 82% of yield.

HRMS ESI (–) calcd. m/z [M–H]- C20H25N10O14P2-: 691.10324, found: 691.10381.

i. Activation of pApG (Im–pApG) **(9)**

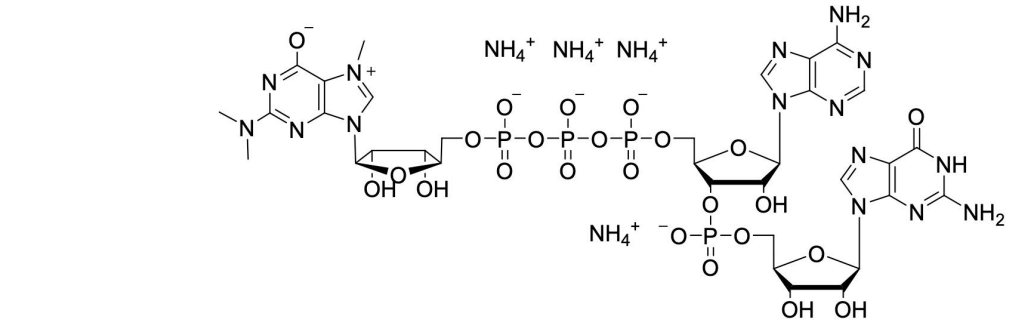
Compound (8) as a triethylammonium salt (29.95 mg, 0.038 mmol, 1 eq.), imidazole (41.13 mg, 0.6 mmol, 16 eq., Merck) and DTDP (49.90 mg, 0.23 mmol, 6 eq., Sigma–Aldrich) were suspended in anhydrous DMSO (1.15 mL). TEA (31.82 L, 0.23 mmol, 6 eq., Sigma–Aldrich) and PPh3 (59.42 mg, 0.23 mmol, 6 eq., Sigma–Aldrich) were then added

and stirred for 2 h at room temperature. After this time, the reaction was terminated by adding a solution of NaClO4 (46.31 mg, 0.38 mmol, 10 eq.) in cold ACN (11.50 mL). After cooling the mixture for 1 h in a refrigerator, the precipitate was centrifuged, washing each time with

cold ACN. The final product was dried overnight in a desiccator over P4O10 to give a light yellow precipitate in the yield of 91%.

j. $m_3{}^{2,2,7}$GpppApG **(10)**

Compound (9) as a sodium salt (141.03 mg, 0.20 mmol, 1 eq.) was dissolved in DMSO (11 mL) and compound (7) as triethyl ammonium salt (148.62 mg, 0.31 mmol, 1.5 eq.) was added. Dry MgCl2 (222.22 mg, 1.63 mmol, 8 eq., Sigma–Aldrich) was then added and the mixture was vigorously stirred at room temperature. The reaction was controlled by performing analyses on RP–HPLC every half hour. After 5 h, the reaction was terminated by adding an aqueous solution (110 mL) of disodium EDTA (606.51 mg, 1.63 mmol, 8 eq.) and NaHCO3 (1/2 m EDTA) to adjust to pH 7. The compound (10) was purified using Flash system followed by RP HPLC. The final synthesis product was lyophilized 3 times and its structure was confirmed by high resolution MS.



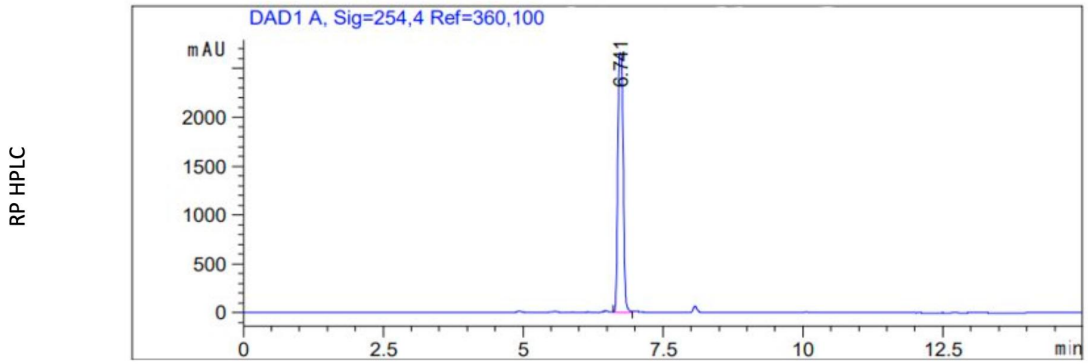


Fig. 59: Structure of $m_3{}^{2,2,7}$GpppApG and final product purification using RP HPLC.

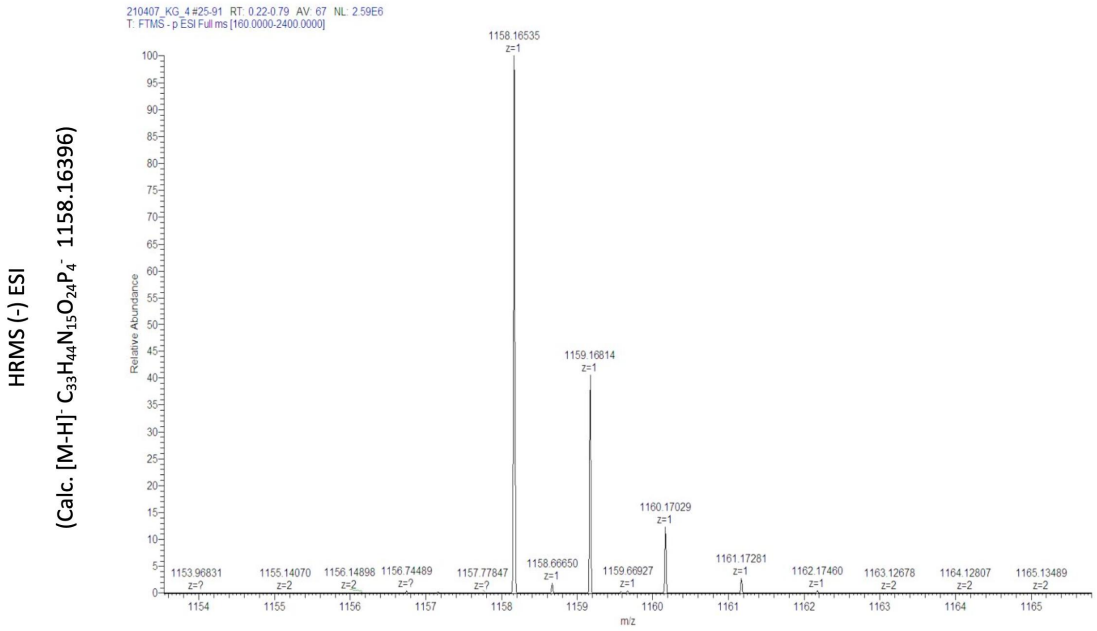


Fig. 60: HRMS ESI (-) calcd. m/z M–H- $C^{33}H^{44}N^{15}O^{24}P^{4}$-: 1158.16396, found: 1158.16535.

9 Attachments

Attachment 1

Benoni B., Potužník J. F., Škríba A., Benoni R., Trylcova J., Tulpa M., Spustová K., Grab K., Mititelu M.B., Pačes J., Weber J., Stanek D., Kowalska J., Bednarova L., Keckesova Z., Vopalensky P., Gahurova L. and Cahova H. "HIV-1 infection reduces NAD capping of host cell snRNA and snoRNA". ACS Chemical Biology 2024 19 (6), 1243-1249.

IF(2024) = 3,5

Attachment 2

Šimonová A., **Svojanovská, B.**, Tryl'cová J., Hubálek M., Moravčík O., Zavřel M., Pávová M., Hodek J., Weber J., Cvačka J., Pačes J., Cahová H. "LC/MS analysis and deep sequencing reveal the accurate RNA composition in the HIV-1 virion." Scientific Reports 2019 9(1): 8697.

IF(2019) = 3,99

Attachment 3

Šimonová A., Romanská V., **Benoni, B.**, Škubník K., Šmerdová L., Procházková M., Spustová K., Moravčík O., Gahurova L., Pačes J., Plevka P., Cahová H. "Honeybee Iflaviruses Pack Specific tRNA Fragments from Host Cells in Their Virions." ChemBioChem 2022 23(17): e202200281.

IF(2022) = 3,2