Supplementary information

HIV-1 infection reduces NAD capping of host cell snRNA and snoRNA

Barbora Benoni,[1,2] Roberto Benoni,[1] Jana Trylcova,[1] Katarzyna Grab, [3] Jan Pačes,[4] Jan Weber,[1] David Stanek,[4] Joanna Kowalska, [3] Lucie Bednarova,[1] Zuzana Keckesova,[1] Lenka Gahurova[5] and Hana Cahova[1]

[1] Institute of Organic Chemistry and Biochemistry of the CAS, Flemingovo náměstí 2, Prague 6, Czechia
[2] Charles University, First Faculty of Medicine, Prague, Czechia
[3] Division of Biophysics, Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland
[4] Institute of Molecular Genetics of the Czech Academy of Sciences, Vídeňská 1083, Prague 4, Czechia,
[5] Department of Molecular Biology and Genetics, Faculty of Science, University of South Bohemia, Branišovská 1760, 37005 České Budějovice, Czechia

Supplementary Table 1: Sequences of used oligonucleotides.

Name	**Sequence** (starting at 5´ end)
RNA 3'adaptor	rApp-CNNNNNNAGATCGGAAGAGCACACGTCTG-(C3)
Reverse transcription primer	CAGACGTGTGCTCTTCCGAT
cDNA anchor sense	p-CAGATCGGAAGAGCGTCGTGT-(C3)
cDNA anchor antisense	ACACGACGCTCTTCCGATCTGGG
NAD Seq Fwd PCR primer	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
NAD Seq Rev PCR primer 1	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGCTCTTCCG ATCT
NAD Seq Rev PCR primer 2	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCG ATCT
NAD Seq Rev PCR primer 3	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCT
NAD Seq Rev PCR primer 4	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCG ATCT
NAD Seq Rev PCR primer 5	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCG ATCT
NAD Seq Rev PCR primer 6	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCG ATCT
NAD Seq Rev PCR primer 7	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGA TCT

NAD Seq Rev PCR primer 8	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGCTCTTCCG ATCT
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 primer for U1 RNA template preparation	TAATACGACTCACTATTAGGCTTACCTGGCAGGGGAG
Rev primer for U1 RNA template preparation	CAGGGGAAAGCGCGAACGCAG
U1 RNA template 20 nt	TAATACGACTCACTATTAGGCTTACCTGGCAGGGGAG
HIV-1 mRNA 20 nt	TAATACGACTCACTATAGGGTAAGCCGTACATGTAAT
Biotinylated U1 probe	GCAATGGATAAGCCTCGCCC

Supplementary Table 2: RNA-seq bioinformatic analysis of sRNAs in noninfected and HIV-1 infected cells.

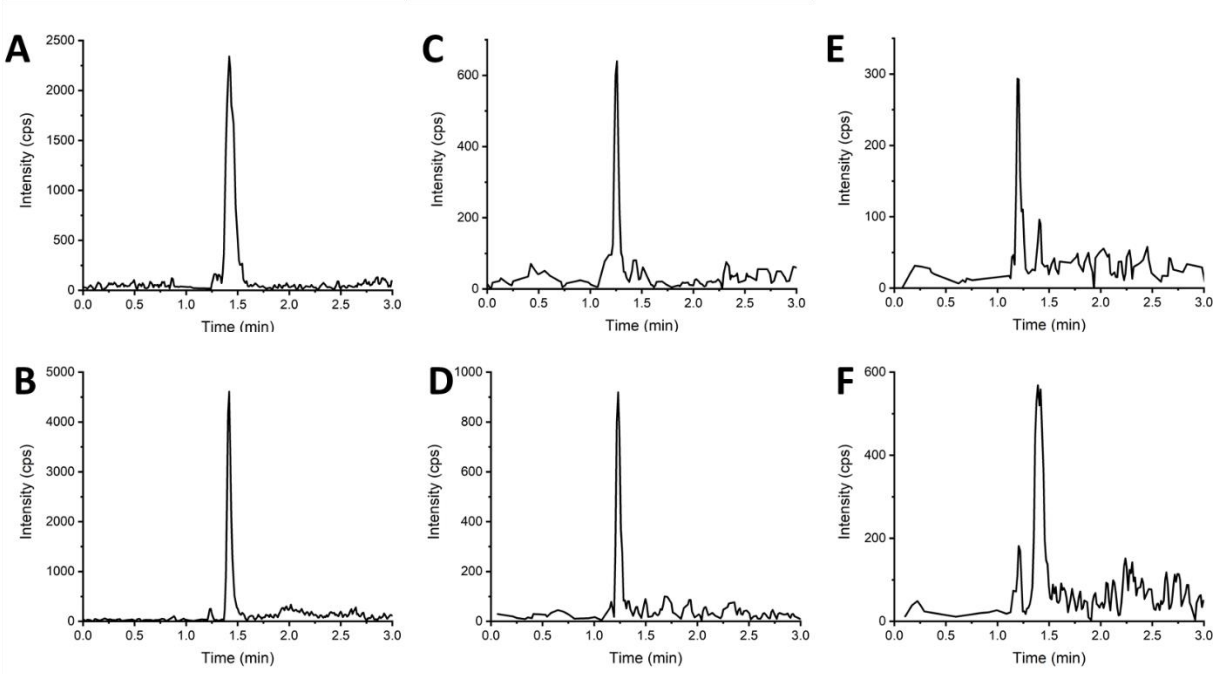
RNA	Chromosome	Start	End	Probe Strand	P-value (DESeq stats p<0.05 after correction)	FDR (DESeq stats p<0.05 after correction)	Log2 Fold Change (DESeq stats p<0.05 after correction)	Shrunk Log2 Fold Change (DESeq stats p<0.05 after correction)	noninfected	infected
Lys	6	27576067	27576139	+	2.35E-09	5.05E-07	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
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	9.562119
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	0.034462	-0.60285	-0.58191	12.73658	13.57901

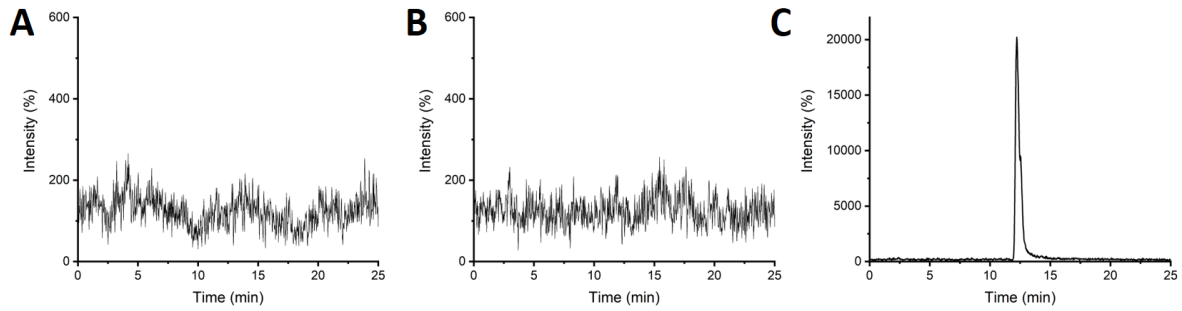
SNORD83B	22	39313819	39313911	-	7.78E-04	0.023228	-0.83373	-0.78387	9.227509	10.25942
SNORD13	8	33513475	33513578	+	0.00186	0.047545	-0.93499	-0.85516	10.00742	11.0756
SNORD45A	1	75787889	75787972	+	9.82E-04	0.028186	-1.03765	-0.94119	9.279798	10.50702
Gly	21	17454789	17454859	-	3.59E-05	0.001884	-1.02155	-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
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
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

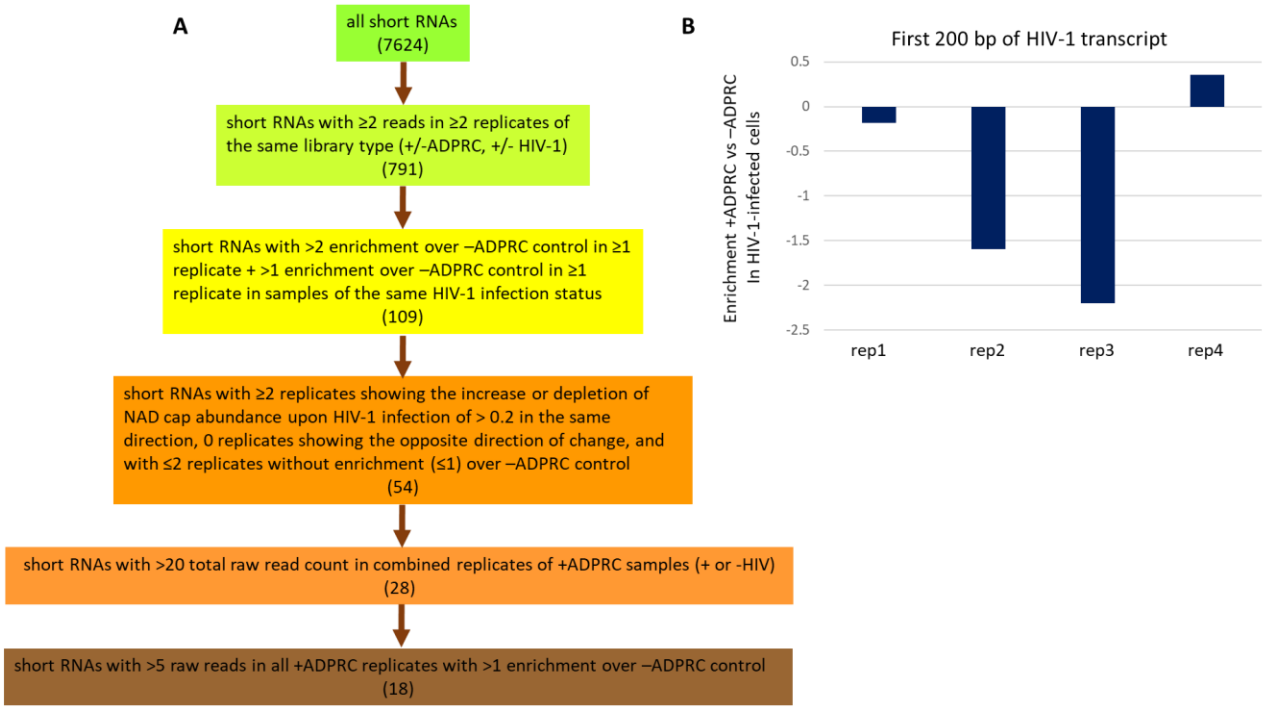
Telomerase-vert.1	3	1.7E+08	1.7E+08	-	1.24E-07	1.66E-05	-2.71506	-2.13033	3.144856	5.92663
SNORD7	17	35573657	35573753	+	1.53E-08	2.19E-06	-2.80182	-2.23003	7.432362	10.33509
U8.4	11	1.23E+08	1.23E+08	+	1.14E-09	2.72E-07	-2.69111	-2.23639	6.508183	9.262581
RNY4P9	13	49908634	49908730	-	4.73E-10	1.27E-07	-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



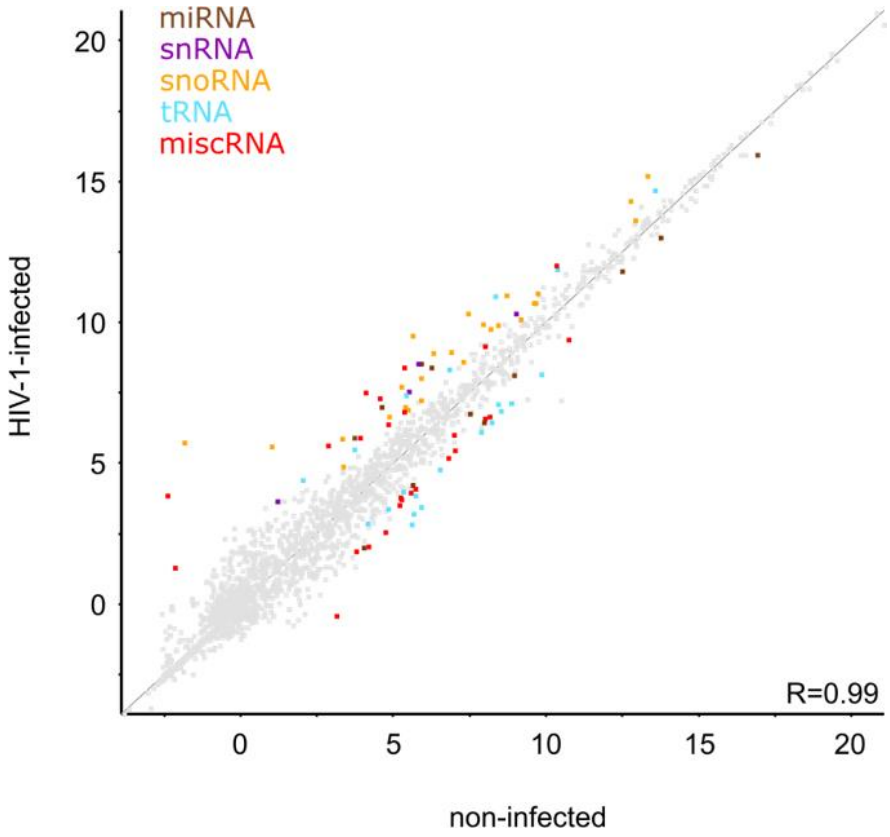
Supplementary Figure 1: MRM Chromatograms of NAD (fragment 523.90) from sRNA of A) MT-4 cells not infected, B) MT-4 cells not infected supplied with NAm, C) HIV-1 infected, D) HIV-1 infected supplied with NAm, E) MT-4 cells with DXO overexpression, and F) MT-4 cells with DXO overexpression supplied with Nam.



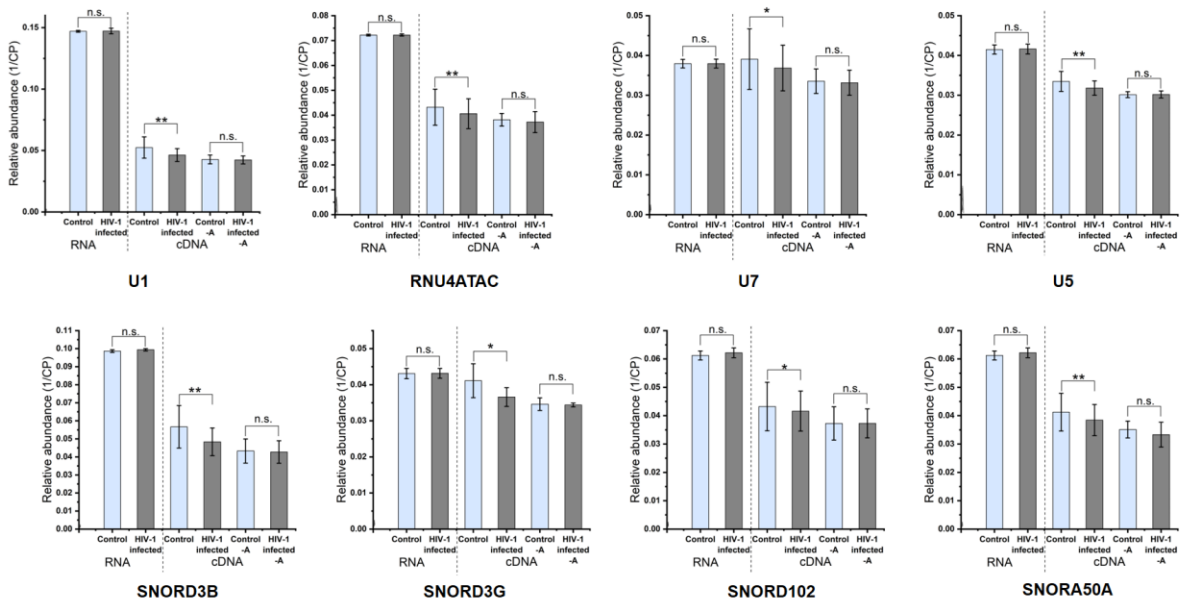
Supplementary Figure 2: EIC of NAD parent mass 664.11 from RNA isolated from A) HIV-1 virions, B) HIV-1 virions negative control (without NuP1 digestion) and C) HIV-1 virions with spiked NAD spiked.



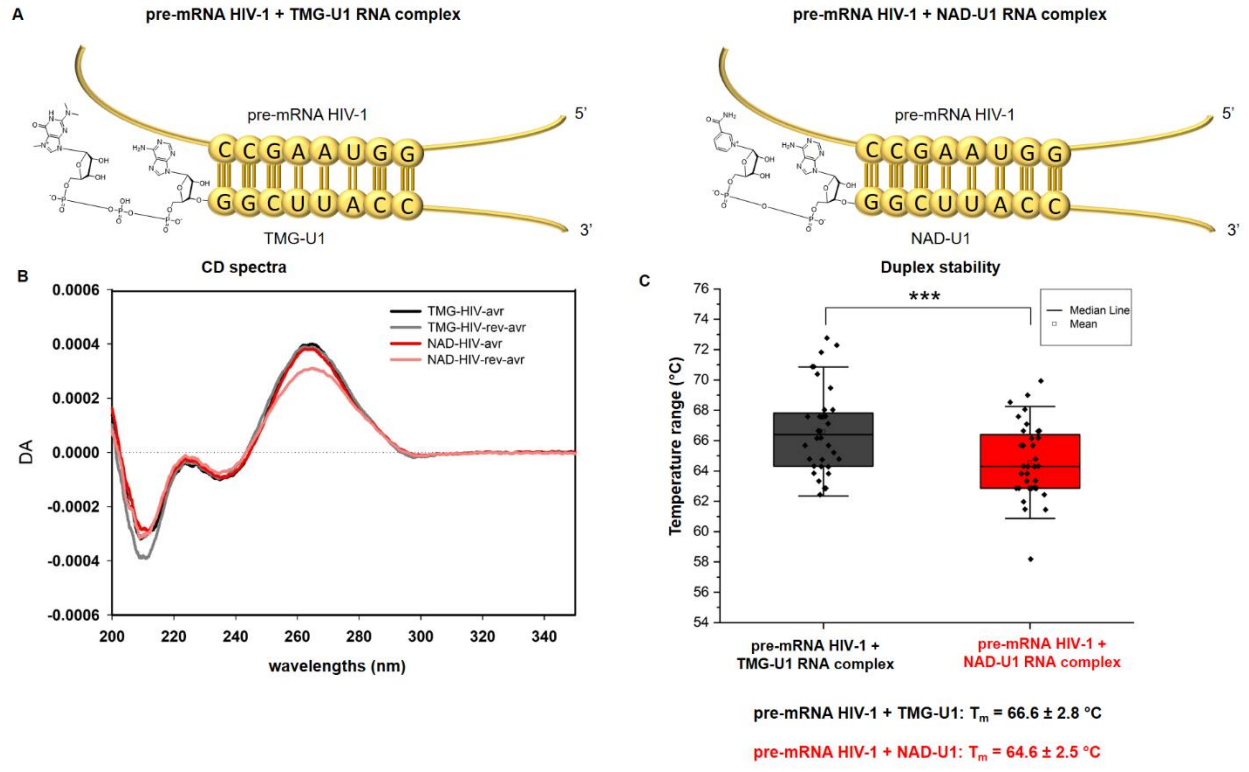
Supplementary Figure 3: A) Bioinformatical analysis workflow with the number of sRNA candidates after each step. B) The enrichment (+ADPRC vs -ADPRC) of HIV-1 transcripts.



Supplementary Figure 4: All sRNAs analysis (miRNA, miscRNA, tRNA, snRNA, snoRNA, rRNA). Total sRNAs: 7624; Expressed (in sRNA datasets): 2014 (cut-off 0.5 in at least one condition); Significantly changing noninfected vs infected (in sRNA datasets): 89 (upregulated after infection: 52, downregulated after infection: 37) – highlighted by colour coding in the scatter plot. In general, sRNAs expression does not appear to be substantially altered after HIV infection. In scatter plot, dots below and to the right from the line are more expressed in non-infected, above and to the left more in HIV-1-infected MT-4 cells. snRNAs and snoRNAs that we identified as having decreased NAD cap after HIV infection do not have decreased expression after infection (no snRNAs and snoRNAs have decreased expression following HIV-1 infection).

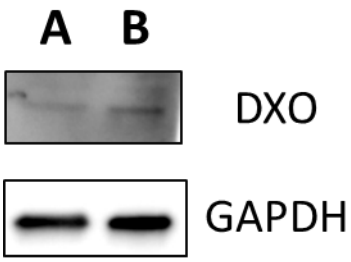


Supplementary Figure 5: Measurement of relative abundance of each candidate sRNA in control and HIV-1 infected cells by RT-PCR and relative abundance of the corresponding cDNA in samples after NAD captureSeq protocol of control and HIV-1 infected cells.

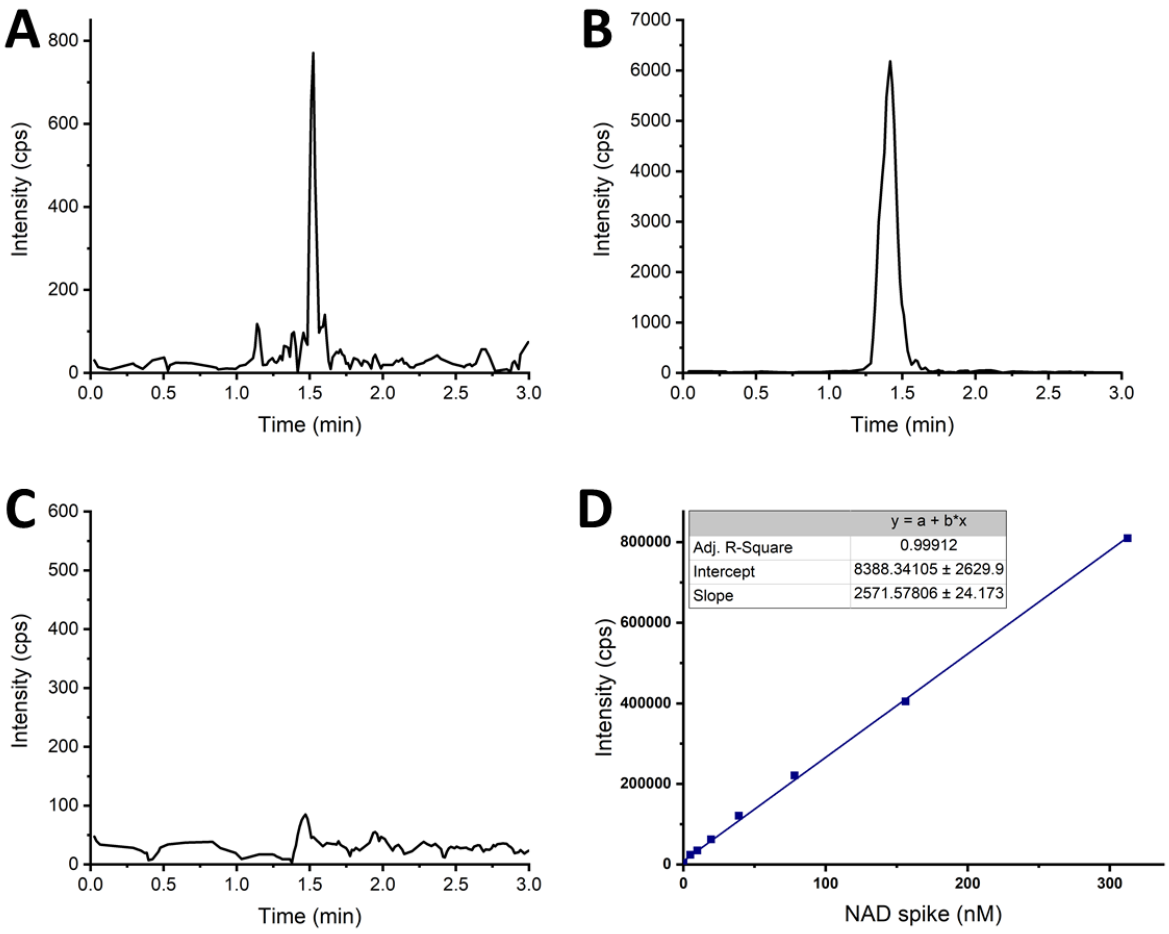


Supplementary Figure 6: NAD cap of U1 snRNA destabilizes the complex with HIV-1 pre-mRNA. A) Scheme of complex formed by complemental regions of pre-mRNA HIV-1 with either TMG-U1 (left) or NAD-U1 (right). B) CD spectra of HIV-1 pre-mRNA with

TMG-U1 (black or dark grey) and HIV-1 pre-mRNA with NAD-U1 (red and pink). C) Duplex stability measurement of pre-mRNA HIV-1 with TMG-U1 (dark grey) or NAD-U1 (red) by light cycler.



*Supplementary Figure 7: Western blot analysis of DXO and GADPH (as control) in control cells (**A**) and HIV-1 infected cells (**B**).*



Supplementary Figure 8: MRM Chromatograms of NAD (fragment 523.90) from sRNA of A) example of U1 RNA pull down from MT-4 cells with DXO overexpression. B) MT-4 cells with 1.6 nM spike of NAD, and C) negative control (without NuP1 digestion). D) Example of calibration curve used for calculation of NAD amount.

Supplementary protocol.

Viral particles purification via Sucrose cushion isolation

Virus particles were concentrated from cleared culture medium by centrifugation through a cushion of 20% (wt./wt.) sucrose in phosphate-buffered saline (PBS) (90 000 × g, 90 min, 4 °C). The pellet was resuspended in RNase/DNase buffer (Tris HCl 100 mM, MgCl2 25 mM, CaCl2 25 mM) with DNase I (10 U/mL, New England BioLabs - NEB), RNase I (200 U/mL) and RNase A (20 mg/mL, both ThermoFisher Scientific) and incubated 2 h at 37 °C. RNase/DNase treatment was stopped by adding RNAzol (Sigma-Aldrich).

RNA digestion and LC/MS analysis

HIV-1 RNA was digested by Nuclease P1 (1 U/µg of RNA, Sigma-Aldrich) 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 (Merck) 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 µL of LC-MS water. For each sample, 8 µL 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 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 ionization (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.

Synthesis of $m_3^{2,2,7}$GpppApG.

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 µm) was used to analyse the reactions progress with a flow rate of 1 mL/min. The solution of 50 mM ammonium acetate (CH_3COONH_4) pH 5.9, and the mixture of 50 mM CH_3COONH_4, pH 5.9 and ACN (1/1, V/V) were used as buffers. HiCHROM C18, 150 mm x10 mm, 5 µm, with a flow rate of 4.7 mL/min was used as semi-preparative column. Buffers for RP-HPLC were as follows: A: 50 mM CH_3COONH_4, pH 5.9, B: 50 mM CH_3COONH_4, 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 modifications. [1,2] 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 $m_3^{2,2,7}$GDP 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 (CH_3COOH, 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 ($NaBH_3CN$, 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 $NaBH_3CN$ 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 $NaHCO_3$ and the aqueous layers were extracted again with DCM / pyridine solution (2/1, v/v). The organic layer was collected and dried over Na_2SO_4 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 H_2O / 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 ($PO(OCH_3)_3$, 4.4 mL, 37.5 mmol, 18 eq., Sigma-Aldrich), then pre-distilled phosphoryl chloride ($POCl_3$, 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 PPh_3 (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 $LiClO_4$ (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 P_4O_{10} 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. $ZnCl_2$ (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 $NaHCO_3$ (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 CH_3COOH, 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, $(CH_3O)_2SO_2$ (0.72 mL, 7.6 µmol) 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 µM. 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 $[M-H]^-$ $C_{20}H_{25}N_{10}O_{14}P_2^-$: 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 µL, 0.23 mmol, 6 eq., Sigma-Aldrich) and PPh_3 (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 $NaClO_4$ (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 P_4O_{10} 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 $MgCl_2$ (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 $NaHCO_3$ (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.

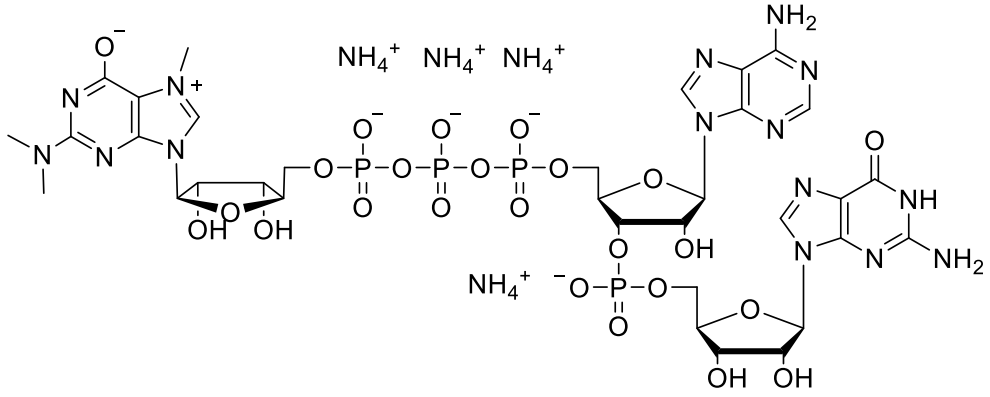
HRMS ESI (-) calcd. m/z $[M-H]^-$ $C_{33}H_{44}N_{15}O_{24}P_4^-$: 1158.16396, found: 1158.16535.

[1] Honcharenko, M. et al. Synthesis and evaluation of stability of m3G-CAP analogues in serum-supplemented medium and cytosolic extract. Bioorganic Med. Chem. ***21****, 7921–7928 (2013)*

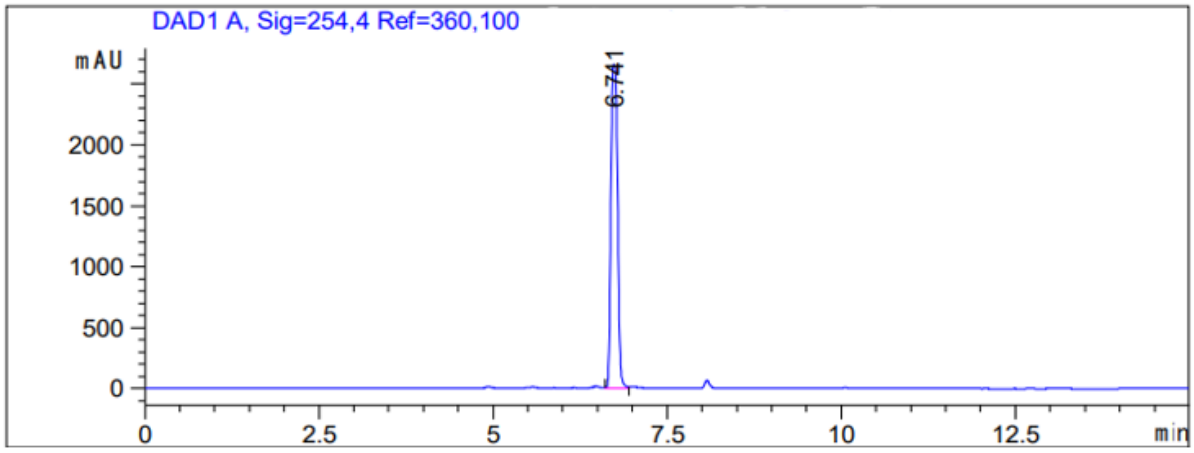
[2] Mlynarska-Cieslak, A. et al. Nicotinamide-Containing Di- and Trinucleotides as Chemical Tools for Studies of NAD-Capped RNAs. Org. Lett. ***20****, 7650–7655 (2018).*

Chemical structure

$m_3^{2,2,7}$GpppApG (TMG)



RP HPLC



HRMS (-) ESI

(Calc. $[M-H]^-$ $C_{33}H_{44}N_{15}O_{24}P_4^-$ 1158.16396)