Internal RNA 2'O-methylation in the HIV-1 genome counteracts ISG20 nuclease-mediated antiviral effect.

Priscila El Kazzi¹, Nadia Rabah^{1,2}, Célia Chamontin³, Lina Poulain¹, François Ferron^{1,4}, Françoise Debart⁵, Bruno Canard¹, Dorothée Missé⁶, Bruno Coutard⁷, Sébastien Nisole³ & Etienne Decroly^{1*}

¹ AFMB, CNRS, Aix-Marseille University, UMR 7257, Case 925, 163 Avenue de Luminy, 13288 Marseille Cedex 09, France.

² Université de Toulon, 83130 La Garde, France.

³ IRIM, CNRS UMR9004, Université de Montpellier, Montpellier, France.

⁴ European Virus Bioinformatics Center, Leutragraben 1, 07743 Jena, Germany

⁵ IBMM, UMR 5247 CNRS, Université de Montpellier, ENSCM, Montpellier, France.

⁶ MIVEGEC, Univ. Montpellier, CNRS, IRD Montpellier, France.

⁷ Unité des Virus Émergents (UVE: Aix-Marseille Univ-IRD 190-Inserm 1207), Marseille, France.

*Corresponding author (<u>etienne.decroly@univ-amu.fr</u>)

SUPPORTING MATERIAL

Table S1: Primers used for PCRs

Primers used	for pDEST14_I	(SG20 design (5' to 3')	
Forward	ggggacaagtttgtacaaaaaagcaggcttctaaggaggtagaaccatgaaacatcaccatcac		
Reverse	gggaccactttgtac	aagaaagctgggtcttaatctgacacagccaggcggggca	
Primers used	for ISG20 muta	ation (5' to 3')	
D11 A	Forward	gtggtggccatggcctgcgagatggtg	
DIIA	Reverse	caccatctcgcaggccatggccaccac	
E12A	Forward	gccatggactgcgcgatggtggggctg	
EISA	Reverse	cagccccaccatcgcgcagtccatggc	
M14A	Forward	catggactgcgaggcggtggggctgggg	
11147	Reverse	ccccagccccaccgcctcgcagtccatg	
D53V	Forward	ggagagatcaccgattacgcaacccgggtcagcg	
KJJA	Reverse	cgctgacccgggttgcgtaatcggtgatctctcc	
1160 V	Forward	ggcaagctggtggtggtgctgacctgaagcac	
IIOJA	Reverse	gtgcttcaggtcagcacccaccagcttgcc	
	Forward	ggtggtgggtcatgccctgaagcacgact	
DJOA	Reverse	agtcgtgcttcagggcatgacccaccacc	
	Forward	catgacctgaagcacgccttccaggcactgaaa	
DJ4A	Reverse	tttcagtgcctggaaggcgtgcttcaggtcatg	
нозл	Forward	gggtcatgacctgaaggccgacttccaggcactg	
1155A	Reverse	cagtgcctggaagtcggccttcaggtcatgaccc	
D126A	Forward	gctggaccactgcgcgcgtgtctccctg	
K120A	Reverse	cagggagacacgcgcgcagtggtccagc	
R127A	Forward	ggaccactgcagggctgtctccctgcgg	
	Reverse	ccgcagggagacagccctgcagtggtcc	

V128A	Forward	ctgcaggcgtgcctccctgcggg		
12011	Reverse	cccgcagggaggcacgcctgcag		
H1/9Δ	Forward	gaacagcctgcttggagccagctcggtggaagat		
11143/1	Reverse	atcttccaccgagctggctccaagcaggctgttc		
D154A	Forward	agctcggtggaagctgcgagggcaacg		
D134/1	Reverse	cgttgccctcgcagcttccaccgagct		
Primers used	for RT and qPO	CR of HIV-1 genome (5' to 3')		
M661	CCTGCGTCGAGAGAGCTCCTCTGG			
M667	GGCTAACTAGGGAACCCACTG			
AA55	GCTAGAGATTTTCCACACTGACTAA			
GAPDH	Forward	ACTTCAACAGCGACACCCACT		
	Reverse	GTGGTCCAGGGGTCTTACTCC		
ISC20	Forward	GAGCGCCTCCTACACAAGAG		
15620	Reverse	TAGAGCTCCATCGTTGCCCT		
ISC-20co	Forward	TTTCAGAACGGCTGCTCCAT		
1562000	Reverse	TGGTACAGTTCCATGGTGGC		
ΡDI 13Δ	Forward	AACAGCTCATGAGGCTACGG		
	Reverse	TGGGTCTTGAGGACCTCTGT		

Table S2: Substrates used for ISG20 characterization.

Name	Sequence from 5' to 3'	ΔG	Modifications
		(kcal/mol)	
A ₂₇	АААААААААААААААААААААААААА	-0.00	No
U ₂₇	υυυυυυυυυυυυυυυυυυυυυυ	-0.00	No
C ₂₇	ссссссссссссссссссссссссс	-0.00	No
$A_{20}A_mA_6$	ΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑΑ	-0.00	2'O-methylation

$A_{20}A^{N6m}A_6$	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	-0.00	N6-methylation
RNA ₁	AGACGAUGCGGAAAACUCUAACAAGAU	-0.72	No
RNA ₂	UGACGGCCCGGAAAACCGGGCC	-14.82	No
DNA ₁	ААААААААААААААААААААААААА		No
RNA ₃	CUAUCCCCAUGUGAUUUUAAUAGCUUCUUA	-3.38	No
	GGAGAAUGAC		
RNA _{3RC}	GUCAUUCUCCUAAGAAGCUA	-1.00	No
DNA ₂	CTATCCCCATGTGATTTTAATAGCTTCTTAGGA		No
	GAATGAC		
RNA ₄ -A ₂₁	AGAUCGUGAACUUACAAACUAUACAAA	-0.43	No
RNA ₄ -A _{21m}	AGAUCGUGAACUUACAAACUAmUACAAA	-0.43	2'O-methylation
RNA ₅ -C ₂₁	AGAUCGUGAACUUACAAACUCUACAAA	-0.78	No
RNA ₅ -C _{21m}	AGAUCGUGAACUUACAAACU C mUACAAA	-0.78	2'O-methylation



Figure S1: ISG20 purification and optimization of the exonuclease assay. Recombinant human wild-type (WT) ISG20 was expressed in *E. coli* and purified on a NTa column. A) Purified proteins were separated on 15% SDS-PAGE gels followed by Coomassie Blue staining. B) ISG20 was incubated with different 5'-radiolabeled RNAs (A₂₇, U₂₇, and C₂₇) for different times and the substrate hydrolysis was followed by PAGE analysis and autoradiography. (C-D) Optimal conditions for ISG20 nuclease activity were determined by incubating 5 nM of ISG20 with 5' fluorescent-labeled A₂₇ for 30 min. Exonuclease activity was analyzed by denaturing gel electrophoresis and quantified using the FujiImager and Image Gauge software. C) ISG20 exonuclease assay in the presence of different Mn or Mg ion concentration (0, 0.5, 1, 2.5, 5, 10 mM). D) Optimal pH range for ISG20 nuclease activity.

Results of Figure S1: At 5 min post-incubation (pi), we observed the excision of 2, 7, and 8 nucleotides for A₂₇, U₂₇, and C₂₇, respectively (mean values). ISG20 seems to have a low affinity toward C₂₇ because the reaction progressed more slowly compared with A₂₇ and U₂₇; however, at 60 min pi, most RNAs were degraded. The laddering degradation profile indicates that ISG20 exonuclease activity occurs from 3' to 5' in a distributive manner. Moreover, ISG20 hydrolytic activity is MnCl₂-dependent, with an optimal activation effect observed at approximately 2.5 mM (Figure S1C). Conversely, MgCl₂ only slightly activates ISG20 exonuclease activity. ISG20 remains active at a broad spectrum of pH (pH 5.5 to 8, Figure S1D).

10020	β1	β2		β3	β4
19020	1 *10	20 3	o	40	50
ISG20 SDN1 ExcX DNA_polIII blaMOX-1 NaseT_P RNaseT_E ERI3 ERI3 CRN-4 orn_C orn_C orn_E orn_P orn_m REX02 orn_A XC847	MAGSREVVAMDCEMVGLGP. NMVAVDCEMVUCED MURILDTETCGLQG. RQIVLDTETCGLQG. UFVAIDLEAYELDQ. LFVVVDVETGGENS. YFVVIDVETGGENS. YFVVIDVETGGENS. YFVVIDVETGGENS. YFVVIDVETGGENS. YFVVIDVETGGENS. UFVVVDEATCOEG. TLLILDFEATCOEG. NLIWIDLEMTGLDP. NLIWIDLEMTGLDP. NLIWIDLEMTGLDP. RMVWVDLEMTGLDP. RMVWVDLEMTGLDT. RLIWIDLEMTGLDT.	HRE. SGLARCS GTE. CLVRV.G G. CLVRV.G G. S	LVNVHG.J VVDRDL.H SVDVI.DGK AVEVVN.RF LAILDT.AEITN ATTVGMDEKG.F AITVGMDEKG.F ILILM.G.RT VVLLNT.HT.J IVAYDV.PN.II IVAYDV.PN.II TLVTDS.HL.II TLVTDS.HL.II TLVTDS.LL.F TIVTDS.DL.F TITDG.DL.F TITDD.HL.F	NUL. Y DK.F VY L DE.F I.V. NP.M.SHL L LIG. NN.FH.VY LIFP.E LIFP.E LIFP.E NUMP.D. MEIES IIIA NILA NILA NILA NILA NILA VILA NILA NILA NULA NULA NULA	IR PEG. E. TTDY VKPNK. P. VVDY VKPDR.VD. P. VVDY VKPDR.VD. P. SPQ IKARH. IRVKEFSWAQN IEPFEGA.N. IQPE VEPFVGA.N. IQPE VKPVINR.T. ITKN HQPDKLITA. MDNW HQPDELLTA. MDNW HQPDELLDG. MDEW KQPDELLDS. MSDW HQPDELLDA. MDEW AHSLETLEA. MDEW
	Exol				
ISG20	α1 α2 β5	α3 ► 0000000.00000		β6 α4	TT
ISG20 SDN1 ExoX DNA_polIIII blaMOX-1 RNaseT_P RNaseT_E ERI3 ERI1 CRN-4 orn_C orn_C orn_E orn_M REXO2 orn_A XC847	60 RT. RVSGVTPQHMVGAT. RT. DITGITAEDIENAS. AM. AIHRITEAMVADEP. AF. GVHGIADEFLLDKP. SR. HVQGRAEYFDF. GES AL. EFTGIKLDHPLENAA. AL. EFTGIKLDHPLENAA. AL. EFTGIQDAVDEGP. CT. ELTGIQDAVDEGP. CT. SLTGIQQVDRAD. CV. DFTGIPQCSIDTAD. NTSHHTASGLLERVKNSSV. NVRHTASGLVERVKASTM. TTSHHTASGLVERVKNSSV. NVRHTASGLVERVKSSTI. KTEHHGKSGLTKAVKESTI. NTRQHRRSGLWQRVLDSQV.	70 80 P.FA. VARLE.ILQIT LSVV.DTQET.LQPFL WIED.VIPHY T.FA.EVADE.FMDYI E.FIEVAKIASVLKTI VQEE.AALTE.IFRGI VSGY.EALHE.IFRVV S.LQ.OVLER.VDEWM T.FD.VVYEQ.FQWL D.EV.EAETL.TLAFL G.DR.EAETL.TLAFL S.EE.EAIDQ.TLAFL T.ER.OAEIE.TLDFI T.AR.OAEIQ.TVAFL T.HA.QAEAQ.TVAFL	K (K (K (K (K (K (K (K (90 SKLVVGHD.LKHDFQA STILVGHS.LNRDLEV AELVIHN.AAFDIGF RAILVGHN.SSFDLGF RAILVGHN.SSFDLGF RAILVGHN.SFDLGF RAILVGHN.SFDLGF RAILVGHN.SFDLGF RAIMVAHN.ANFDHSF KSFDCGNS.UCQDRF KSPLCGNS.UCQDRF KSPLCGNS.UCQDRF RAPLCGNS.UCQDRF SPMCGNS.UCQDRF SPMCGNS.UCQDRF SPMCGNS.UCQDRF SPMCGNS.UCQDRF	100 EKE.DMSG LKI.DHPK .LPE.M.M.L MDYEFSLLKRDIPK .TRM.LGYD .LNA.AVAR GQCQYLGL.PVAD GQCQYLGL.PVAD IQCQLSRL.KYPP YQMKLSNI.QMPA LSR.YMPE .LYK.HMPR .LYK.HMPR .LYK.HMPR .LYK.YMPE .LYK.YMPE .LYK.YMPE .LYK.YMPE .LYK.YMPE .LYK.YMPE .LYK.YMPE .LYK.YMPE .LYK.YMPE
ISG20	β7η1	α5 22220.00 120	α6 <u>000000</u> . 130		α7 000000000000000000000000000000000000
ISG20 SDN1 ExoX DNA_polIIII blaMOX-1 RNaseT_F RNaseT_E ERI1 CRN-4 orn_C orn_E orn_P orn_M REXO2 orn_A XC847	YTI	. RL LWR . EAKLD . HC 	RRVSLRVLSE. RRPSINNLCK. IKYSNMALYKI KRNSLDALCA. NNASKLSNVCG. GQTVLSKACQ. GQTVLSKACQ. WPKNGLLDMNK RSQTKLTIMLE. ELSATTNIGKMNE E. S S S S S S S S S S S S S	RILLK.S.IQNSLLG SILGY.E.VRKTGVPH RKI.NVQT.P.PGLHHH RY.EI.D.N.SKRTLH YI.O.IPW.KNMH AA.GM.E.F.DNREAH TA.GM.D.F.DSTQAH GI.SL.Q.HIGRPH YY.DL.P.TIGRAH IA.A.A.H.IKESOH II.A.A.H.IKESOH II.S.G.F.TKQGTH VI.K.E.F.SKTGSH II.S.G.F.EKRASH EY.EF.A.P.KKAASH IM.S.G.F.AKSASH	SVEDARATHELYQISQRI DCVHDASAANKLALAVVEK RALYDCYITAALLIDIMN GALLDAQILAEVYLAMTGG NAGNDAVYTLQAMMGLAID SARYDTERTALFCGIVN. SALYDTERTAVLFCEIVN. SGLDDCNNIARIAVRLQD CGLDDSNIARIAVRLQD DAMDDCLNIATILQRMINM LALQDIRESTAELAYYREH LALDDIRESTAELAYYREH LALDDIRESTKELQFYRNA RALDDISESTKELQFYRNA RALDDISESTKELQFYRNA RALDDISESTKELQFYRNA
				F	

Figure S2: Multiple sequence alignment of ISG20 and other nucleases/hydrolases. SequenceandstructurehomologieswereperformedwithHHpred(https://toolkit.tuebingen.mpg.de/tools/hhpred), hitswererecovered and curated usingTCoffeealignment.The alignment was processed with ESPript3 server.Sequences are annotated accordingto their gene name.The secondary structural representation based on ISG20 crystal structure (PDB:1WLJ) is shown on the top of the alignment.Residues of the Exo I, Exo II, and Exo III domains arehighlighted in cyan and those of the RBD in yellow.Informations about the proteins used for thealignement are liste in Table S3.

Table S3: List of nuclease and their PDB accession numbers used for alignments (Figure S2) and structural studies (Figure S2).

Protein	Name	Organism	PDB	chain
		0		

ISG20	interferon stimulated gene 20kDa	H.sapien	1WLJ	A/1-168
SDN1	Small RNA degrading nuclease	Arabidopsis	5Z9X	A/139-297
ExoX	Exodeoxyribonuclease 10	E.coli	4FZX	C/1-152
DNA_polII I	DNA polymerase III subunit alpha	E.col	5M1S	D/7-181
BlaMOX-1	Beta-lactamase	Klebsiella pneumoniae	4WBG	B/9-199
RnaseT_P	Ribonuclease T	Pseudomonas aeruginosa	2F96	A/30-211
RnaseT_E	Ribonuclease T	E.coli	3V9W	C/38-219
ERI3	ERI1 exoribonuclease 3	H.Sapiens	7K05	B/24-205
ERI1	3'-5' EXONUCLEASE ERI1	H.Sapiens	1W0H	A/11-194
CRN-4	Cell death-related nuclease 4	C.elegans	5DK5	A/20-208
orn_C	Oligoribonuclease	Coxiella burnetii	3TR8	A/10-180
orn_E	Oligoribonuclease	E.coli	2IGI	A/6-176
orn_P	Oligoribonuclease	Pseudomonas aeruginosa	6N6A	A/7-177
orn_m	Oligoribonuclease	>orn_metagenom	6RK6	C/7-177
REXO2	RNA exonuclease 2 homolog,Small fragment nuclease	H.Sapiens	6N6J	B/10-181
orn_A	Oligoribonuclease	Acinetobacter baumannii	5CY4	D/16-186
XC847	Oligoribonuclease	Xanthomonas campestris	2GBZ	A/10-180



Time (min): 0, 1, 5, 15, 30 and 60

Figure S3: Characterization of the ISG20 exonuclease activity. WT and ISG20 mutants (residues within the exonuclease domain) were expressed in *E. coli* and purified on NTa columns as in S1. A) Purified proteins were analyzed on 15% SDS-PAGE gels followed by Coomassie Blue staining. A major band was detected at 20-kDa which corresponds to the expected molecular weight of ISGS20 and its mutants. A weak band around 66-kDa was also detected for the D11A and E13A mutants and

represents 1% of the total proteins. B) Mutagenesis study of ISG20 DEDDh motif: 20 nM of each purified mutant was incubated with 5'fluorescent A27 for 60 min. Exonuclease activity was analyzed by denaturing gel electrophoresis and quantified using the FujiImager and Image Gauge software. Results are the mean and standard deviation of 3 independent experiments. (C-D) The exonuclease activity of ISG20 was evaluated using different 5'-radiolabeled substrates described in the cartoon (left, the star represent the radiolabeled position) and their degradation profiles upon PAGE analysis are shown on the right panel. C) Assessment of ISG20 nuclease activity using ssRNA₁, ssRNA₂ that forms a 3'hairpin structure, and ssDNA₁. D) Degradation profile of RNA_{3RC} annealed to a RNA₃ template or to a DNA₂ substrate upon ISG20 digestion.

Results of Figure S3: Recombinant ISG20 mutants were detected at their expected molecular weight of 21-kDa (Figure S3A). Endpoint assay with 20 nM of the recombinant proteins showed a drastic reduction of RNA degradation by the DEDDh mutants (Figure S3B), which confirms the specific exonuclease activity observed with WT ISG20 and the important role of the conserved residues in the catalytic domain. To further decipher how the RNA structure and compositions regulate its exonuclease activity, ISG20 was incubated with a set of different heteropolymeric substrates (Table S2). Figure S3C shows the efficient degradation of linear single-stranded RNA (ssRNA1), and the slower hydrolysis of ssRNA₂, which is assumed to form stable hairpin secondary structures (Mfold prediction, $\Delta G = -14.80$ kcal/mol). The analysis of RNA degradation products indicates an accumulation of intermediate degradation products, suggesting that ISG20 pauses when it encounters stable double-stranded RNA structures (dsRNA). The ssRNA₂ substrate was almost completely degraded after a longer incubation period (1h). As ISG20 belongs to the DEDDh exonuclease superfamily, which contains both RNases and DNases, it was asked whether it also hydrolyzes 5'end radiolabeled single-stranded DNA (ssDNA₁-A₂₇). Figure S3C shows that ISG20 is not active on ssDNA₁, and the absence of exonuclease activity on this substrate was confirmed using a 100-fold higher concentration of the nuclease (200 nM, not shown).

To determine whether ISG20 degrades RNA involved in RNA/DNA heteroduplexes that mimic viral retrotranscription intermediates, the DNA₂ oligonucleotide was synthesized and annealed with its complementary 5'radiolabeled RNA_{3RC}. The hydrolysis of the RNA substrates by ISG20 showed an improved degradation of RNA_{3RC} paired to DNA₂ compared to the RNA₃/RNA_{3rc} duplex (Figure S3D). Overall, these results indicate that ISG20 preferentially degrades ssRNA compared with RNA containing hairpin structures or dsRNA, and that RNA is more sensitive to ISG20 degradation in DNA₂/RNA_{3RC} heteroduplexes than when it forms dsRNA structures (RNA₃/RNA_{3RC}).

ISG	20	NP1		P	DE
2 nM		0.0008 Unit/µL		0.001 Unit/µL	
A ₂₇	$A_{20}A_mA_6$	A ₂₇	$A_{20}A_mA_6$	A ₂₇	$A_{20}A_mA_6$
0 60	0 60	0 - 60	060	0 - 60	0 60
Time (min): (0, 15, 30 and	60			

Figure S4: Comparative hydrolysis of 2'O-methylated RNA by different 3' exonucleases. The two 5'end radiolabeled RNA substrates (A₂₇ and A₂₀A_mA₆) were incubated with ISG20, NP1, or PDE for different times, and the reaction products were analyzed by PAGE and autoradiography.



Figure S5: Effect of heteropolymeric RNA 2'O-methylation on ISG20 exonuclease activity. Evaluation of the ISG20 exonuclease activity on methylated and non-methylated RNA₄-A₂₁ and RNA₄-C₂₁ monitored by PAGE and autoradiography.



Figure S6: SDS-page of ISG20 RBD mutants. Mutants of recombinant ISG20 RBD were engineered, expressed in *E. coli*, and purified on NTa columns. Purified proteins were analyzed on 15% SDS-PAGE gels followed by Coomassie Blue staining.



Figure S7: Representation of the N⁰ **and N**₋₂ **stops induced by RNA 2'O-mthylation on internal residues.** Cartoon model of ISG20 with an RNA carrying 2'O-methylation marks performed on Biorender. A) Depiction of the N⁰ stop resulting from the destruction of the interaction (black cross) between the methylated nucleotide (at N₀) and the R53 residue of ISG20. B) Illustration of the N₋₂ blockage resulting from the destruction of the interaction (black cross) between the methylated nucleotide of ISG20.



Figure S8: Structural alignment of ISG20 with its exonuclease homologues. Superimposition of ISG20 (PDB: 1WLJ, in off white), SDN1 in complex with RNA (PDB: 5Z9X, sea green), and ExoX in complex with DNA (PDB: 4FZX, orange) represented in ribbons (A), zoom on the catalytic cavity containing the RNA (B) or DNA (C) substrates.



Figure S9: ISG20 exonuclease activity is impaired by RNA 2'O-methylation, but not by N6 methylation. A) Model of ISG20 in interaction with an RNA substrate built based on the superimposition of ISG20 (PDB: 1WLJ) and SDN1 in complex with RNA (PDB: 5Z9X). Surface representation of ISG20 (off-white) containing an RNA from SDN1 structure (sticks). The residues highlighted in yellow correspond to D90 and R53 that interact with the 2'O-methylated residue at N₋₂ and N₀. RNA 2'O and N6-methylation sites are indicated. B) Comparative degradation profiles of non-methylated RNA (A₂₇), 2'-O and N6-methylated RNAs at position 21 ($A_{20}A_mA_6$ and $A_{20}A^{N6m}A_6$) by ISG20 visualized by autoradiography after PAGE analysis.