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

Priscila El Kazzi[1], Nadia Rabah[1,2], Célia Chamontin[3], Lina Poulain[1], François Ferron[1,4], Françoise Debart[5], Bruno Canard[1], Dorothée Missé[6], Bruno Coutard[7], Sébastien Nisole[3] & 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 (etienne.decroly@univ-amu.fr)**

SUPPORTING MATERIAL

Table S1: Primers used for PCRs

Primers used for pDEST14_ISG20 design (5’ to 3’)		
Forward	ggggacaagtttgtacaaaaaagcaggcttctaaggaggtagaaccatgaaacatcaccatcac	
Reverse	gggaccactttgtacaagaaagctgggtcttaatctgacacagccaggcggggca	
Primers used for ISG20 mutation (5’ to 3’)		
D11A	Forward	gtggtggccatggcctgcgagatggtg
	Reverse	caccatctcgcaggccatggccaccac
E13A	Forward	gccatggactgcgcgatggtggggctg
	Reverse	cagccccaccatcgcgcagtccatggc
M14A	Forward	catggactgcgaggcggtggggctgggg
	Reverse	ccccagccccaccgcctcgcagtccatg
R53A	Forward	ggagagatcaccgattacgcaacccgggtcagcg
	Reverse	cgctgacccgggttgcgtaatcggtgatctctcc
H89A	Forward	ggcaagctggtggtgggtgctgacctgaagcac
	Reverse	gtgcttcaggtcagcacccaccaccagcttgcc
D90A	Forward	ggtggtgggtcatgccctgaagcacgact
	Reverse	agtcgtgcttcagggcatgacccaccacc
D94A	Forward	catgacctgaagcacgccttccaggcactgaaa
	Reverse	tttcagtgcctggaaggcgtgcttcaggtcatg
H93A	Forward	gggtcatgacctgaaggccgacttccaggcactg
	Reverse	cagtgcctggaagtcggccttcaggtcatgaccc
R126A	Forward	gctggaccactgcgcgcgtgtctccctg
	Reverse	cagggagacacgcgcgcagtggtccagc
R127A	Forward	ggaccactgcagggctgtctccctgcgg
	Reverse	ccgcagggagacagccctgcagtggtcc

V128A	Forward	ctgcaggcgtgcctccctgcggg
	Reverse	cccgcagggaggcacgcctgcag
H149A	Forward	gaacagcctgcttggagccagctcggtggaagat
	Reverse	atcttccaccgagctggctccaagcaggctgttc
D154A	Forward	agctcggtggaagctgcgagggcaacg
	Reverse	cgttgccctcgcagcttccaccgagct
Primers used for RT and qPCR of HIV-1 genome (5’ to 3’)		
M661	CCTGCGTCGAGAGAGCTCCTCTGG	
M667	*GGCTAACTAGGGAACCCACTG*	
AA55	*GCTAGAGATTTTCCACACTGACTAA*	
GAPDH	Forward	ACTTCAACAGCGACACCCACT
	Reverse	GTGGTCCAGGGGTCTTACTCC
ISG20	Forward	GAGCGCCTCCTACACAAGAG
	Reverse	TAGAGCTCCATCGTTGCCCT
ISG20co	Forward	TTTCAGAACGGCTGCTCCAT
	Reverse	TGGTACAGTTCCATGGTGGC
RPL13A	Forward	AACAGCTCATGAGGCTACGG
	Reverse	TGGGTCTTGAGGACCTCTGT

Table S2: Substrates used for ISG20 characterization.

Name	**Sequence from 5’ to 3’**	**ΔG (kcal/mol)**	**Modifications**
A_{27}	AAAAAAAAAAAAAAAAAAAAAAAAAAA	-0.00	No
U_{27}	UUUUUUUUUUUUUUUUUUUUUUUUUUU	-0.00	No
C_{27}	CCCCCCCCCCCCCCCCCCCCCCCCCCC	-0.00	No
$A_{20}A_mA_6$	AAAAAAAAAAAAAAAAAAA**A$_m$**AAAAAA	-0.00	2’O-methylation

$A_{20}A^{N6m}A_6$	AAAAAAAAAAAAAAAAAAAA**A^{N6m}**AAAAAA	-0.00	N6-methylation
RNA_1	AGACGAUGCGGAAAACUCUAACAAGAU	-0.72	No
RNA_2	UGACGGCCCGGAAAACCGGGCC	-14.82	No
DNA_1	AAAAAAAAAAAAAAAAAAAAAAAAAA		No
RNA_3	CUAUCCCCAUGUGAUUUUAAUAGCUUCUUA GGAGAAUGAC	-3.38	No
RNA_{3RC}	GUCAUUCUCCUAAGAAGCUA	-1.00	No
DNA_2	CTATCCCCATGTGATTTTAATAGCTTCTTAGGA GAATGAC		No
RNA_4-A_{21}	AGAUCGUGAACUUACAAACUAUACAAA	-0.43	No
RNA_4-A_{21m}	AGAUCGUGAACUUACAAACU**A$_m$**UACAAA	-0.43	2’O-methylation
RNA_5-C_{21}	AGAUCGUGAACUUACAAACUCUACAAA	-0.78	No
RNA_5-C_{21m}	AGAUCGUGAACUUACAAACU**C$_m$**UACAAA	-0.78	2’O-methylation

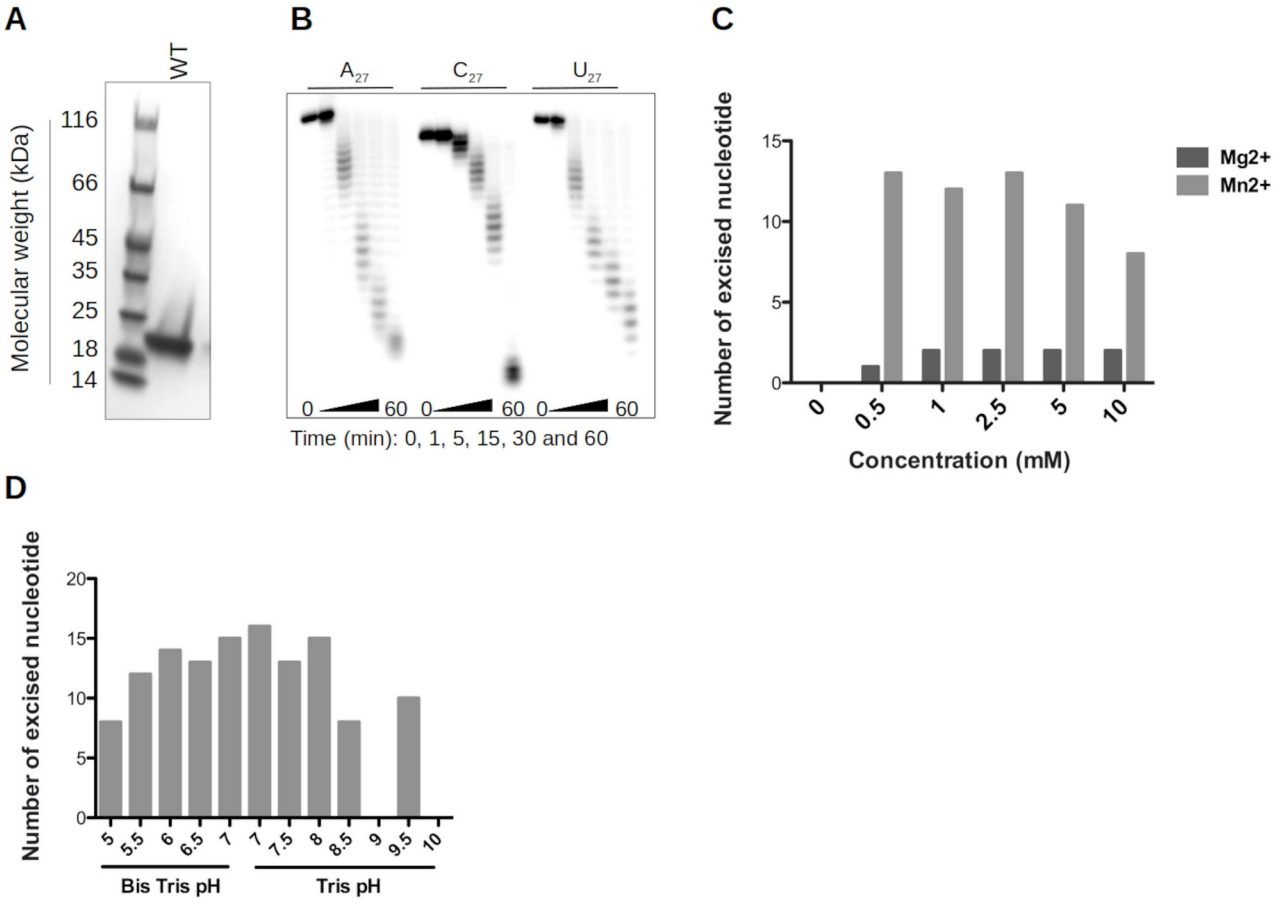

Figure S1: ISG20 purification and optimization of the exonuclease assay. Recombinant human wild-type (WT) ISG20 was expressed in *E. coli* and purified on a NTa column. A) Purified proteins were separated on 15% SDS-PAGE gels followed by Coomassie Blue staining. B) ISG20 was incubated with different 5'-radiolabeled RNAs (A_{27}, U_{27}, and C_{27}) 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 A_{27} 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 A_{27}, U_{27}, and $C_{27,}$ respectively (mean values). ISG20 seems to have a low affinity toward C_{27} because the reaction progressed more slowly compared with A_{27} and $U_{27;}$ 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 $MnCl_2$-dependent, with an optimal activation effect observed at approximately 2.5 mM (Figure S1C). Conversely, $MgCl_2$ only slightly activates ISG20 exonuclease activity. ISG20 remains active at a broad spectrum of pH (pH 5.5 to 8, Figure S1D).

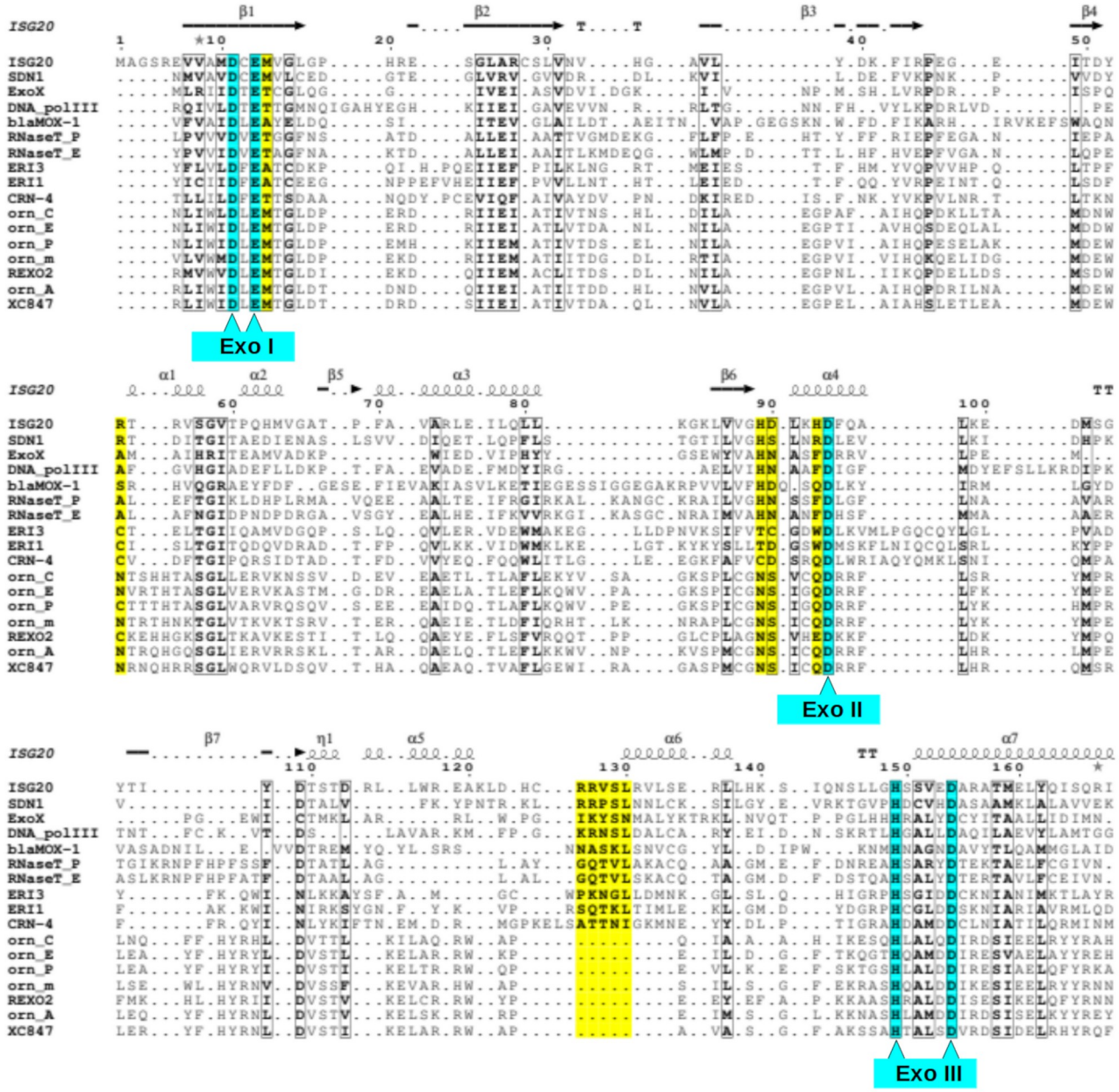

Figure S2: Multiple sequence alignment of ISG20 and other nucleases/hydrolases. Sequence and structure homologies were performed with HHpred (https://toolkit.tuebingen.mpg.de/tools/hhpred), hits were recovered and curated using TCoffee alignment. The alignment was processed with ESPript3 server. Sequences are annotated according to 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 are highlighted in cyan and those of the RBD in yellow. Informations about the proteins used for the alignement 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

ISG20	interferon stimulated gene 20kDa	H.sapien	1WLJ	A/1-168
SDN1	Small RNA degrading nuclease 1	Arabidopsis	5Z9X	A/139-297
ExoX	Exodeoxyribonuclease 10	E.coli	4FZX	C/1-152
DNA_polIII	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

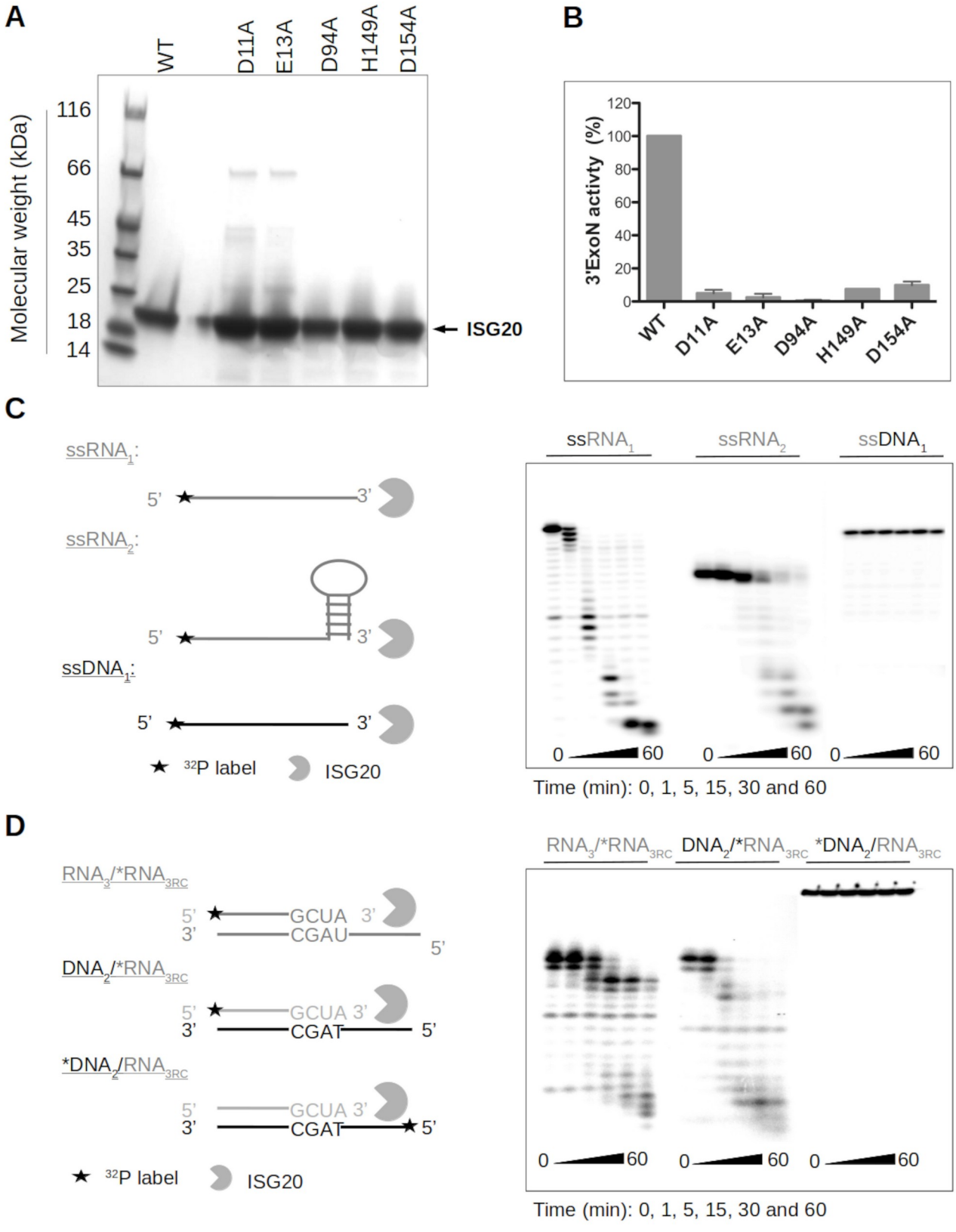

Figure S3: Characterization of the ISG20 exonuclease activity. WT and ISG20 mutants (residues within the exonuclease domain) were expressed in *E. coli* and purified on NTa columns as in S1. A) Purified proteins were analyzed on 15% SDS-PAGE gels followed by Coomassie Blue staining. A major band was 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 $ssRNA_1$, $ssRNA_2$ that forms a 3'hairpin structure, and $ssDNA_1$. D) Degradation profile of RNA_{3RC} annealed to a RNA_3 template or to a DNA_2 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 $ssRNA_2$, which is assumed to form stable hairpin secondary structures (Mfold prediction, ΔG =-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 $ssRNA_2$ 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 ($ssDNA_1$-A_{27}). Figure S3C shows that ISG20 is not active on $ssDNA_1$, 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 DNA_2 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 DNA_2 compared to the RNA_3/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 DNA_2/RNA_{3RC} heteroduplexes than when it forms dsRNA structures (RNA_3/RNA_{3RC}).

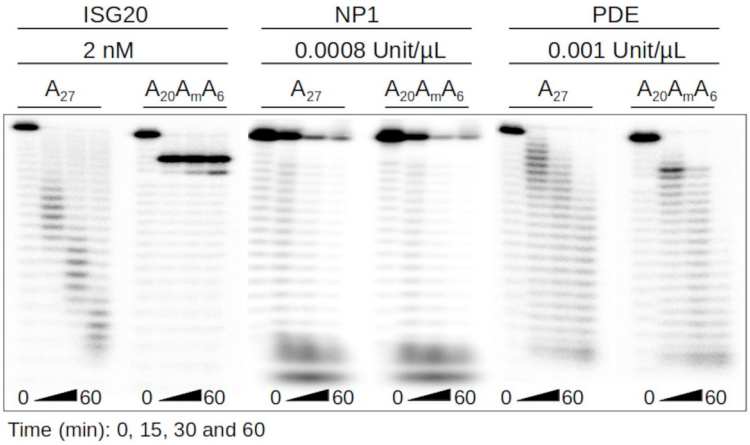

Figure S4: Comparative hydrolysis of 2'O-methylated RNA by different 3' exonucleases. The two 5'end radiolabeled RNA substrates (A_{27} and $A_{20}A_mA_6$) were incubated with ISG20, NP1, or PDE for different times, and the reaction products were analyzed by PAGE and autoradiography.

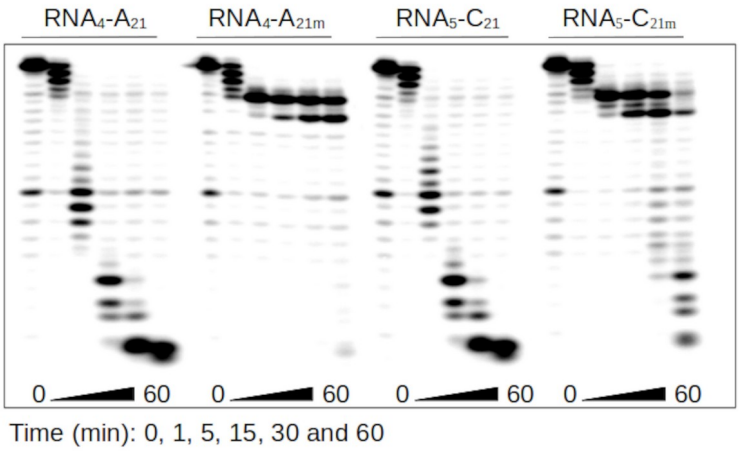

Figure S5: Effect of heteropolymeric RNA 2'O-methylation on ISG20 exonuclease activity. Evaluation of the ISG20 exonuclease activity on methylated and non-methylated RNA_4-A_{21} and RNA_4-C_{21} monitored by PAGE and autoradiography.

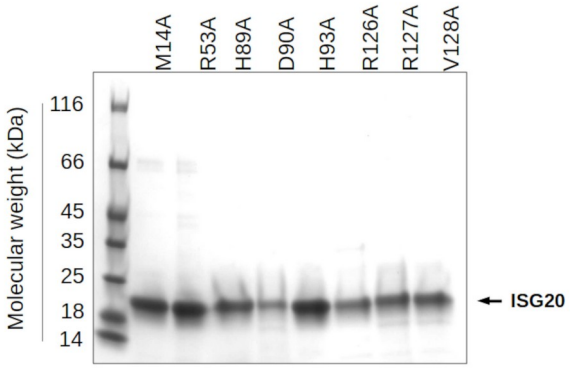

Figure S6: SDS-page of ISG20 RBD mutants. Mutants of recombinant ISG20 RBD were engineered, expressed in *E. coli*, and purified on NTa columns. Purified proteins were analyzed on 15% SDS-PAGE gels followed by Coomassie Blue staining.

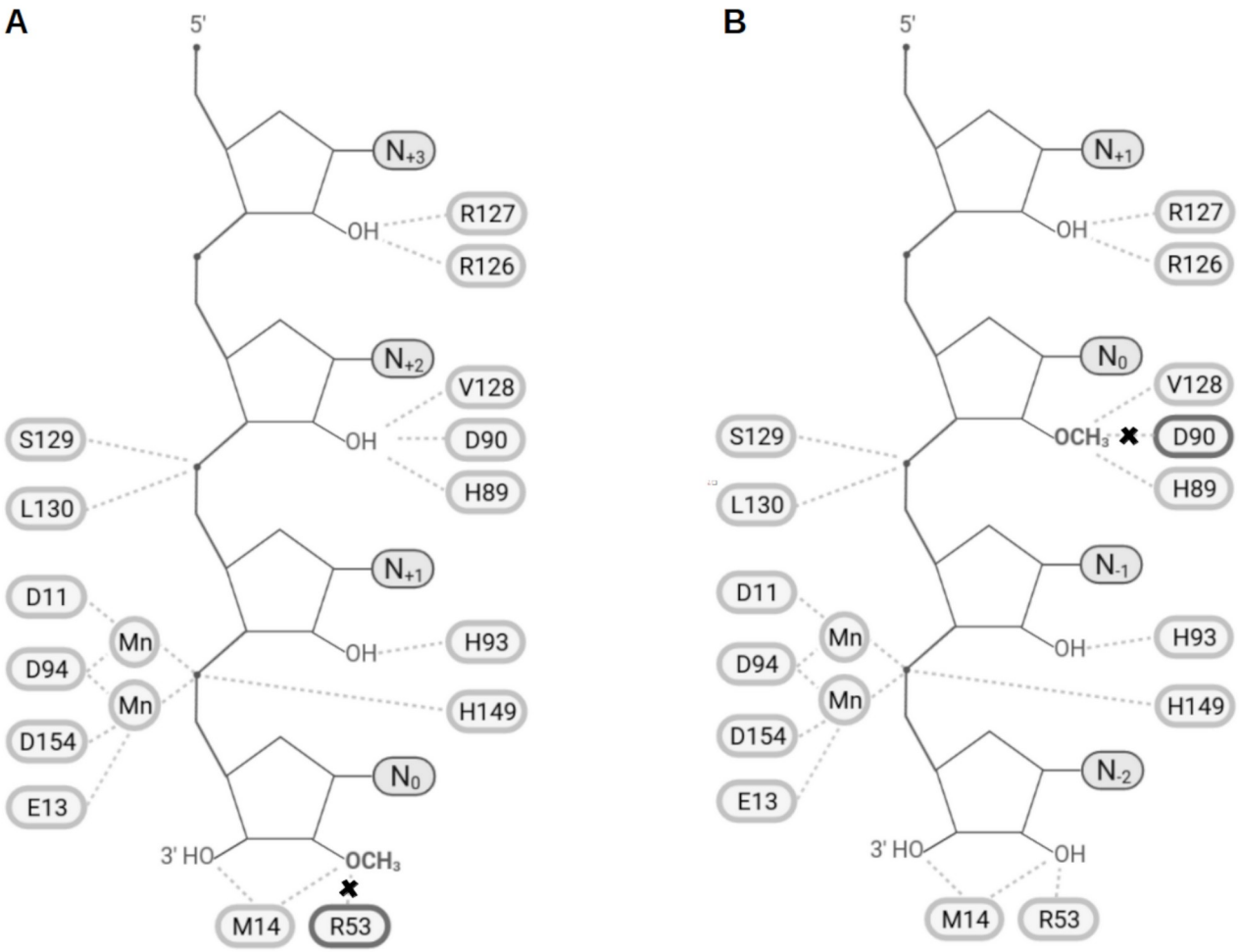

Figure S7: Representation of the N_0 and N_{-2} stops induced by RNA 2'O-mthylation on internal residues. Cartoon model of ISG20 with an RNA carrying 2'O-methylation marks performed on Biorender. A) Depiction of the N_0 stop resulting from the destruction of the interaction (black cross) between the methylated nucleotide (at N_0) 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 at N_0 and the D90 residue of ISG20.

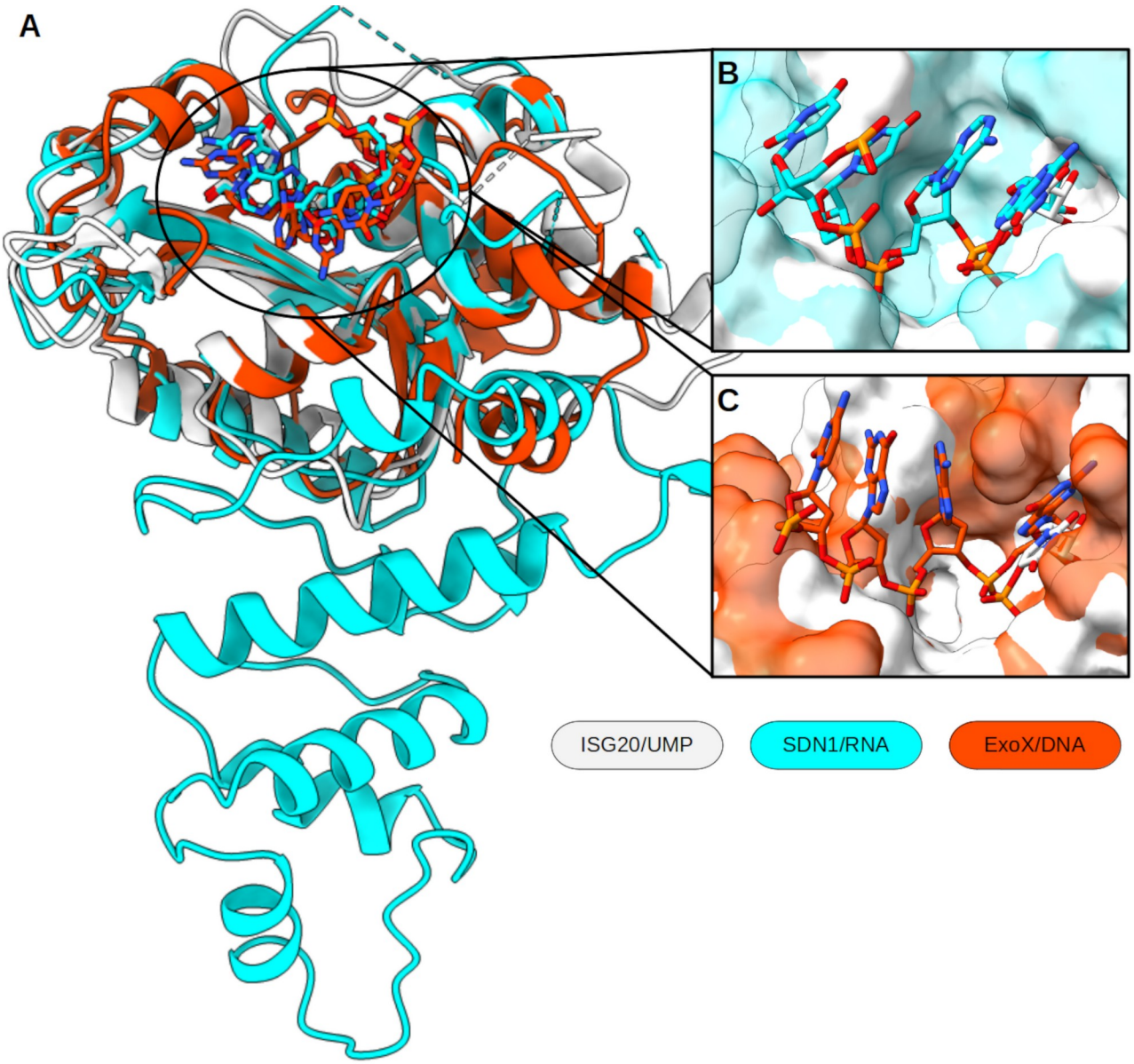

Figure S8: Structural alignment of ISG20 with its exonuclease homologues. Superimposition of ISG20 (PDB: 1WLJ, in off white), SDN1 in complex with RNA (PDB: 5Z9X, sea green), and ExoX in complex with DNA (PDB: 4FZX, orange) represented in ribbons (A), zoom on the catalytic cavity containing the RNA (B)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

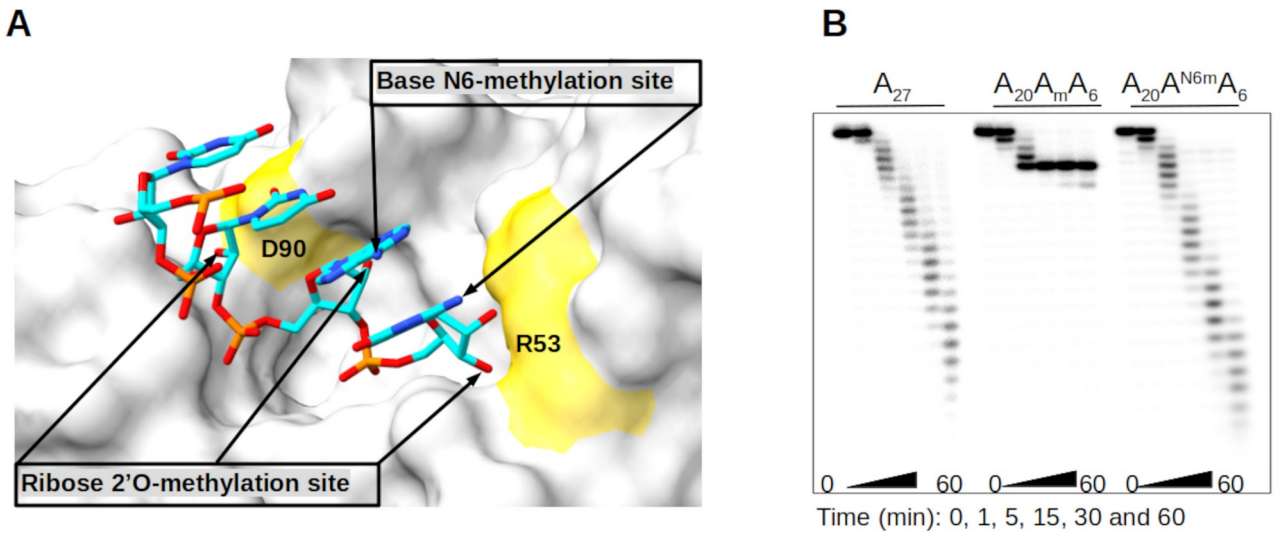

Figure S9: ISG20 exonuclease activity is impaired by RNA 2'O-methylation, but not by N6 methylation. A) Model of ISG20 in interaction with an RNA substrate built based on the superimposition of ISG20 (PDB: 1WLJ) and SDN1 in complex with RNA (PDB: 5Z9X). Surface representation of ISG20 (off-white) containing 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 N_0. RNA 2'O and N6-methylation sites are indicated. B) Comparative degradation profiles of non-methylated RNA (A_{27}), 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.