THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université

le 11 juillet 2022 par

Priscila Binta El kazzi

The interplay of RNA N7- and 2'O-methylation in viral replication

Discipline
Biologie-Santé

Spécialité
Maladies infectieuses

École doctorale
ED62

Laboratoire/Partenaires de recherche
AFMB
UMR7257 CNRS
Aix-Marseille Univ

Composition du jury
Nolwenn JOUVENET Rapporteuse
Institut Pasteur, CNRS UMR3569, Paris, France.
Andrea CIMARELLI Rapporteur
CIRI, CNRS, UMR5308, ENS de Lyon, Lyon, France.
Nathalie ARHEL Examinatrice
IRIM, CNRS UMR9004, Montpellier, France.
Bruno COUTARD Président du jury
UVE: AMU-IRD 190-Inserm 1207, Marseille, France.
Etienne DECROLY Directeur de thèse
AFMB-AMU-UMR7257 CNRS, Marseille, France.
Dorothée MISSE Co-directrice de thèse
MIVEGEC-CNRS-IRD, Montpellier, France

AFMB ARCHITECTURE ET FONCTION DES MACROMOLÉCULES BIOLOGIQUES

Affidavit

Je soussigné, Priscila Binta El kazzi, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de Etienne Decroly et Dorothée Missé, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect à la fois de la charte nationale de déontologie des métiers de la recherche et de la charte d'Aix-Marseille Université relative à la lutte contre le plagiat.

Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Marseille, le 26/4/2022.

Affidavit

I, undersigned, Priscila Binta El kazzi, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific direction of Etienne Decroly et Dorothée Missé, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the French national charter for Research Integrity and the Aix-Marseille University charter on the fight against plagiarism.

This work has not been submitted previously either in this country or in another country in the same or in a similar version to any other examination body.

Place Marseille, date 26/4/2022.

Liste de publications et participation aux conférences

1) Liste des publications réalisées dans le cadre du projet de thèse :

1. **Structure-function analysis of the nsp14 N7-guanine methyltransferase reveals an essential role in *Betacoronavirus* replication.**

   Ogando NS, **El Kazzi P**, Zevenhoven-Dobbe JC, Bontes BW, Decombe A, Posthuma CC, Thiel V, Canard B, Ferron F, Decroly E, Snijder EJ.

   Co-first author

   Proc Natl Acad Sci U S A. 2021 Dec 7,118(49):e2108709118. doi: 10.1073/pnas.2108709118.

2. **Fluoxetine targets an allosteric site in the enterovirus 2C AAA+ ATPase and stabilizes a ring-shaped hexameric complex.**

   Hurdiss DL, **El Kazzi P**, Bauer L, Papageorgiou N, Ferron FP, Donselaar T, van Vliet ALW, Shamorkina TM, Snijder J, Canard B, Decroly E, Brancale A, Zeev-Ben-Mordehai T, Förster F, van Kuppeveld FJM, Coutard B.

   Co-first author

   Sci Adv. 2022 Jan 7,8(1):eabj7615. doi: 10.1126/sciadv.abj7615. Epub 2022 Jan 5.

3. **A dual mechanism of action of AT-527 against SARS-CoV-2 polymerase.**

   Shannon A, Fattorini V, Sama B, Selisko B, Feracci M, Falcou C, Gauffre P, **El Kazzi P**, Delpal A, Decroly E, Alvarez K, Eydoux C, Guillemot JC, Moussa A, Good SS, La Colla P, Lin K, Sommadossi JP, Zhu Y, Yan X, Shi H, Ferron F, Canard B.

   Nat Commun. 2022 Feb 2,13(1):621. doi: 10.1038/s41467-022-28113-1.

4. **Identification of potent inhibitors of arenavirus and SARS-CoV-2 exoribonucleases by fluorescence polarization assay**

   Sergio Hernández, Mikael Feracci, Carolina Trajano De Jesus, **Priscila El Kazzi**, Rafik Kaci, Laura Garlatti, Clemence Mondielli, Fabrice Bailly, Philippe Cotelle, Franck Touret, Xavier de Lamballerie, Bruno Coutard, Etienne Decroly, Bruno Canard, François Ferron, karine ALVAREZ

   Submitted to Antiviral Research

5. **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 Coutard[6], Sébastien Nisole[3], Bruno Canard[1], Dorothée Missé[7] & Etienne Decroly[1*]

In revision NAR

2) Participation aux conférences et écoles d'été au cours de la période de thèse :
   1. Les conférences:
      a. La XXIe édition des JFV - Du 28 au 29.03.2019 - **Présentation orale - Obtention d'un lauréat par société francophone de virologie.**
      b. La journée scientifique de Infectiople- 03.07.2020 - **Présentation orale - Obtention d'une médaille d'argent pour le projet.**
      c. XXIIIèmes des JFV Du 28 au 29.04.2021- **Présentation de poster.**
      d. The Epitranscriptome EMBO Courses & Workshops - Du 09 au 11.02.2022 - **Présentation de poster**.
      e. XXIVème des JFV Du 11 au 12.04.2022- **Présentation orale.**
   2. Les workshops:
      a. Atelier « Cristallographie des Protéines : des données de diffraction à la carte de densité » - Du 11 au 15.06.2019
      b. Réseau National de Formation en Biologie Structurale Intégrative (RéNaFoBiS) - Du 21 au 28.06.2019
      c. MedIM - Du 07 au 11.12.2020 (07 décembre 2020)
   3. Vulgarisation scientifique:
      a. Souk des sciences - Le 26.05.2019
      b. Festival des sciences et de l'innovation Marseille 2021 - Du 09 au 10.10.2021

Abstract/Resumé

Abstract

In recent years, epitranscriptomic modifications have been detected in numerous viral RNA, but the physiological function of most of them remains barely understood. Among these modifications, RNA methylations are an important modification induced by specific viral or cellular RNA methyltransferases. The first part of this manuscript focuses on RNA 2'O-methylation. This post-transcriptional modification prevents viral RNA detection by RIG-like receptors (RIG-I and MDA5) that regulate type-1 interferon expression and is therefore considered a "self" marker. In this work, we address the possibility that RNA 2'O-methylation may interplay with the antiviral action of an interferon-induced restriction factor, namely ISG20. Briefly, we show that ISG20 exonuclease activity is strongly impaired when it encounters a 2'O-methylation mark within the RNA. ISG20 stops two nucleotides upstream (N_{-2}) and at the methylated residue (N_0). Structure-function analyses revealed that ISG20 R53 and D90 residues play a key role in the RNA hydrolysis impairment, which results from a steric clash of these residues with the 2'O-methylated nucleotide. We next extrapolated our observations to HIV-1 which is naturally 2'O-methylated by the host FTSJ3. By comparing the sensitivity to ISG20 of HIV-1 RNAs extracted from hypo-methylated viruses (produced in FTSJ3-KO cells) and from normally methylated viruses, we showed that internal 2'O-methylation protects the HIV-1 genome from ISG20 degradation. We confirmed this observation in infected cells and showed that ectopically expressed ISG20 drastically reduces the replication of hypomethylated VSV-G pseudotyped HIV-1 virus, as a consequence of impaired reverse transcription. Altogether, our results shed light on a new pro-viral role of viral RNA 2'O-methylation. Indeed, we demonstrated that HIV-1 2'O-methylation promotes viral replication by limiting ISG20-mediated restriction.

In the second part of the manuscript, we characterized the CoV nonstructural protein 14 (nsp14) which is a bifunctional protein harboring an N-terminal 3′-to-5′ ExoN domain and a C-terminal N7-MTase domain that is presumably involved in viral mRNA capping. We integrate structural, biochemical, and virological data to identify 4 conserved regions essential for nsp14's enzymatic activities and virus viability. We identified several residues involved in the formation of the N7-MTase catalytic pocket, which presents a fold distinct from the Rossmann fold observed in most known MTases and we assess their importance for in vitro enzymatic activity and for virus replication. Our results identify the N7-MTase as a critical enzyme for betacoronavirus replication and define key residues of its catalytic pocket that can be targeted to design inhibitors with a potential pan-coronaviral activity spectrum.

Keywords : ISG20, epitranscriptome, 2'O-methylations, innate immune response, interferon, 3'Exonuclease. FTSJ3, HIV-1.

Resumé

Ces dernières années, des modifications épitranscriptomiques ont été détectées dans de nombreux ARN viraux, mais la fonction physiologique de la plupart d'entre eux reste mal connue. Parmi ces modifications, les méthylation sont des modifications importantes induites par des ARN méthyltransférases spécifiques virales ou cellulaires. Dans la première partie de ce manuscrit, nous avons étudié le rôle de 2'O méthylations connues pour empêcher la détection de l'ARN viral par les récepteurs de type RIG-I et MDA5 qui régulent l'expression de l'interféron de type 1. Par conséquent, la méthylation de l'ARN 2'O est considérée comme un marqueur du sois. Dans ce travail, nous avons abordé la possibilité que la méthylation de l'ARN 2'O puisse interagir avec l'action antivirale des facteurs de restriction induits par l'interféron, et nous avons démontré que l'activité exonucléase de l'ISG20 est altérée par la 2'O-méthylation de l'ARN. Nous avons élucidé une nouvelle conséquence des modifications épitranscriptomiques de l'ARN viral et concluons que l'activité de restriction de l'ISG20 peut être contrecarrée par la 2'O-méthylation induitepar les 2'O-méthyltransférases hôtes. Nos conclusions sont étayées par une étude biochimique combinée à des expériences réalisées dans le cadre de l'infection par le VIH. En bref, nous montrons que l'activité ISG20 est fortement altérée lorsque l'exonucléase rencontre des marques épitranscriptomiques comme les 2'O-méthylations. ISG20 fait une pause deux nucléotides en amont (N-2) et au niveau du résidu méthylé (N0). Les analyses structure-fonction ont révélé que les résidus ISG20 R53 et D90 jouent un rôle clé dans l'altération de l'hydrolyse de l'ARN, qui résulte d'un clash stérique entre ces résidus et le nucléotide 2'O-méthylé. Nous avons ensuite extrapolé nos observations au VIH-1 qui est naturellement 2'O-méthylé par la MTase FTSJ3 des cellules. En comparant la sensibilité à l'ISG20 des ARN du VIH-1 extraits de virus hypo-méthylés (produits dans les cellules FTSJ3-KO) et de virus normalement méthylés, nous avons montré que la 2'O-méthylation interne protège le génome du VIH-1 de la dégradation par l'ISG20. Nous avons confirmé cette observation dans les cellules infectées et montré que ISG20 exprimé de manière ectopique réduit considérablement la réplication du virus de pseudotypes VSV-G/VIH-1 hypométhylé en conséquence d'une altération de la transcription inverse. Au final, nos résultats mettent en lumière un nouveau rôle pro-viral de la 2'O-méthylation de l'ARN viral. En effet, nous avons démontré que la 2'O-méthylation du VIH-1 favorise la réplication virale en limitant la restriction médiée par ISG20.

Dans la seconde partie du manuscrit, nous avons caractérisé la protéine non structurale 14 (nsp14) du CoV, qui est une protéine bifonctionnelle composée d'un domaine 3'-exonucléase en N-terminal et d'un domaine N7-MTase en C-terminal. Ce dernier est vraisemblablement impliqué dans la méthylation des structures coiffes ARN viraux. Nous intégrons des données structurelles, biochimiques et virologiques pour identifier 4 régions conservées essentielles aux activités enzymatiques de nsp14 et à la viabilité du virus. Nous avons identifié plusieurs résidus impliqués dans la formation de la poche catalytique de la N7-MTase, qui présente un repliement distinct de celui observé dans la plupart des MTases connues et nous évaluons leur importance pour l'activité enzymatique *in vitro* et pour la réplication du virus. Nos résultats identifient la N7-MTase comme une enzyme critique pour la réplication des betacoronavirus et définissent les résidus clés de sa poche catalytique qui peuvent être ciblés pour concevoir des inhibiteurs avec un spectre d'activité potentiel pan-coronaviral.

Mots clés : ISG20, epitranscriptome, 2'O-methylations, innate immune response, interferon, 3'Exonuclease. FTSJ3, HIV-1.

Acknowledgments

I'm sitting here and I'm thinking wow this was quite a journey! I have acquired so many memories and experiences that have changed and shaped me during these past years. Many people were involved and have accompanied me during this journey and I would like to thank them for that! Well, I'm not really good with emotional talk! But I will try my best to express my gratitude to you all.

First I would like to thank the members of the jury for accepting to evaluate and be part of this thesis committee.

I would also like to thank la Foundation Mediterranée for granting me a scholarship and Yves Bourne for having me at the AFMB.

This thesis couldn't have been done without the great supervision of Etienne. Well Etienne the day asked me to work with you, and I was so intrigued as to why? And I'm grateful that you gave me this opportunity. Your enthusiasm for science is very infectious just like the viruses we are trying to characterize. The way that you prioritize your students made me feel so privileged in terms of supervision. I really enjoyed all the chatting sessions we had, whether science-related or not. At a point, I realized that I should've never told you about the "trotinette" adventure, but since you enjoy this story so much I'm glad to leave you with this souvenir. Finally thank you Dr Etienne (I had to say it at least once :)) for all your investment in improving my skills.

The other person who also played a role in my scientific growth is you Bruno Co and I want to thank you for that. During my master's internship, you used to say "il faut toujour prendre du recul pour réévaluer ses objectifs", this sentence got engraved in my head and was very useful during my thesis. I would also like to thank you for always challenging my brain! Whenever I think I know everything about my project, you always have THE question… . I enjoyed sharing your enthusiasm for the 2C project and thank you for all your advice and physical/virtual availability.

Francois, I've learned so many stuff from you from banal to the most complex stuff. I'm so grateful for all the time you invested in teaching me how to do structural modeling, use a terminal, GIMP, and how to make some proper figures for an article…. the list is so long but

those are what I can think of. Thank you for all the constructive discussion we had about the nsp14. Any student is lucky to have you as a supervisor.

Nadia, I'm so grateful for the opportunity you gave me by providing me with some teaching hours in Toulon. It was a great experience that made me go out of my comfort zone and boosted my confidence. Thank you for all your advice, always making sure that we are doing fine, being our cheerleader during the infectiople day, the covoiturage to Toulon, the coffee time and the funny stories we shared.

I would like to express my gratitude to Barbara, you taught me how to deal with radioactivity. Your advice on how to set up my exonuclease experiments and electrophoresis gels was all appreciated. At last thank you for replying to my urgent needs for substrates during your holidays. Truthful thanks to all AFMB members who have given me valuable advice, especially Cécilia, Veronique Z, Gerlind, Christophe, and Olivier.

Bruno Ca, I'm so impressed by your path! Thank you for sharing your career stories with us, it did inspire me! And for your suggestions on nice books to read. I enjoyed all the constructive discussions we had and sharing your passion for wolves.

Sarah, Viola, Fabien, Alexi, and Gab you brought so much joy to me during this journey. All the laughter, pranks, beer pongs, karaoke parties, and trips we had together made my days. Sarsoura, my partner in crime, I'm so glad to have met you. Your friendship is one of the best things I got from this journey. Thank you for always being here and lending me your shoulders when I really needed it. Being far from home is not always easy, but Viola sharing a flat with you made it feel like home. Thank you for all your support and motivational speech. Fabien, you taught me so many lab tricks and I'm so grateful for that. Thank you for always making time to have coffee whenever I pop at your office and of course for providing the free coffee. I think I should also thank Renaud for that. Finally thank you both for all your troubleshooting tips about protein purification and for always rescuing me whenever I urgently needed an AKTA.

To the queens, Bhawna and Alice, sharing the lab with you was a great pleasure. All the singing and dancing we had in the lab, girls talking, and deep scientific discussion we had made my days in the lab so fun. Thank you both for all your help when I needed it! Miss Sama, I enjoyed sharing your passion for food, thank you for always saying yes to any spontaneous plan. Although I overreacted sometimes, I did enjoy all our discussions and they made me aware of so many things. Thank you for your support, and help during the writing of this thesis. Alice, I would like you to know how thankful I'm for all your nice messages and handmade gifts. It was fun brainstorming with you about the 2'O-methylation project.

A sprinkle of some Cuban vibes from Oney and Sandra made this journey quite exceptional. Thank you Sandra for organizing the almost-weekly to try fancy-schmancy restaurant and Oney for turning Bhawna's place into our brunch and apéro sanctuary. I would also like to thank all the “djeunzz” of the lab with whom I spent lots of pleasant moments, in particular Léa, Camille, Ash, Adran, Pierre, Valentine, Axel, Vincent, Mika, Sergio S, and H.

To Coralie, thank you for being the familiar face of polytech at the begining of this journey. It was comforting to know that you were going to be there and it allowed me to adapt quickly to the new environment. I enjoyed all of our discussions and coffee breaks. I appreciate your efforts in teaching me the right direction to start the "biz à Marseille"! I'm still confused though which side it starts , in my defense, the Covid crisis helpful at all!

In these few lines, I would like to express appreciations to the best officemates ever Vero and Priscila S. During those past three and a half years, I enjoyed sharing with you guys a a calm, fun and positive working space. Véro, I would like to thank you for your investment in correcting grammatical errors in my cover letters, abstracts and especially all the key emails that I had to send in French. I don't know what I would do without you :/ Priscila, first it’s a great pleasure sharing the same name! It was fun to receive emails addressed to you from Etinne and having to send them back to you. Thank you for all the laboratory tips you taught me, especially for passing on your knowledge on the purification of nsp14.

Fabienne, The lab feels so dead in your absence! Thank you for all the energy you bring with you in the morning and for all the technical assistance you provide for the lab.

I’m thankful to my collaborators Sébastien, Natacha, Lisa, and Daniel for their help and resources.

At last but not least, I would like to thank Christelle, one of the persons that matters the most to me. Although your plate was full your were always there to support me. I consider myself lucky to have you as a best friend. Thank you for all the free therapy and for the long list of anime that kept me busy during the weekends.

Finally, I would like to thank my dear family, “les Azzi et dérivés” for all their warm support. I’m particularly grateful to my aunts for always being there during the ups and downs of this journey, and for knowing how to lift my spirit when necessary. All of your faith and encouragements were truly motivating, they kept me going forward.

I would like to dedicate this thesis to
all the people who supported me during this journey,
especially my aunts.

“No matter how hard or impossible it is, never lose sight of your goal.”

-Monkey D Luffy

Abbreviations

2′O-MTase	: 2′O-methyltransferase
5mC	: 5-methylcytosine
ADAR	: RNA-acting adenosine deaminase
ALKBH5	: α-ketoglutarate-dependent dioxygenase AlkB homology 5
CARD	: Caspase activation and recruitment domain
circRNA	: Circular RNA
CoV	: Coronavirus
CTD	: C-terminal domain
CTD	: C-terminal domain
DC	: Dendritic cell
DENV	: Dengue virus
DNase	: Deoxyribonuclease
dsRNA	: Double-stranded RNA
ExoN	: Exonuclease
FBL	: Fibrillarin
FTO	: Fat mass and obesity associated protein
GMP	: Guanosine monophosphate
Gp-enz	: Enzyme–GMP
GTase	: Guanylyltransferase
GTP	: Guanosine triphosphate
HIV-1	: Human immunodeficiency virus type 1
hnRNPU	: heterogeneous nuclear ribonucleoprotein U
I	: Inosine
IFN	: Interferon
IFN-I	: Type I IFN
IFN-II	: Type II IFN
IFN-III	: Type I IFN
IKKε	: IκB kinase ε
IRAK	: IL-1R-associated kinases
IRF	: IFN-regulatory factor
ISG	: Interferon-stimulated gene
L protein	: Large protein

LGP2 : Laboratory of genetics and physiology 2
lncRNA : Long non-coding RNA
MAPK : Mitogen-activated protein kinases
MDA5 : Melanoma differentiation-associated protein 5
METTL : Methyltransferase-like protein
miRNAs : MicroRNA
mRNA : Messenger RNA
mRNA : Messenger RNA
MyD88 : Myeloid differentiation primary response protein 88
N6mA : N6-methyladenosine
N7-MTase : N7-methyltransferase
^{N7m}G : N7-methylguanosine
ncRNA : Non-coding RNA
NF-κB : Nuclear factor-κB
NLR : NOD-like receptor
Nm : Nucleotide 2'O-methylation
NOD : Nucleotide-binding oligomerization domain
npcRNA : Non-protein-coding RNA
NS5 : Nonstructural protein 5
nsp : Non-structural protein
PAMP : Pathogen-associated molecular pattern
PB : Processing bodies
piRNA : Small interacting RNAs with piwi
PRNTase : Polyribonucleotidyl transferase
PRR : Pattern recognition receptor
PTM : Post-translational RNA modification
PUS : Pseudo-uridine synthases
RF : Rossmann fold
RIG-I : Retinoic acid-inducible gene I
RIPK1 : Receptor-interacting protein kinase 1
RLR : RIG-I like receptors
RNA : Ribonucleic acid
RNase : Ribonuclease
rRNA : Ribosomal RNA
SARS-CoV : Severe acute respiratory syndrome coronavirus
siRNAs : Silent RNA

SMARCA5 : SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5

sncRNA : Short non-coding RNA

snoRNAs : Small nucleolar RNA

snRNAs : Small nuclear RNA

ssRNA : Single-stranded RNA

SWI/SNF : Switch/sucrose non-fermentable

TAK1 : Transforming growth factor β-activated kinase 1

TBK1 : TRAF family member-associated NF-κB activator binding kinase 1

TLR : Toll-like receptors

TOP1 : DNA topoisomerase 1

TPase : 5'-triphosphatase

TRAF : Tumor necrosis factor receptor-associated factor

tRNA : Transfer RNA

vRNA : Viral RNA

WNV : West Nile virus

YTH : YTH N6-Methyladenosine RNA Binding Protein

ZIKV : Zika virus

ZNFX1 : Zinc finger nuclear transcription factor, X-box binding 1-type containing 1

Ψ : Pseudouridine

Table of contents

Table of figures

Chapter 1: General Introduction

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

Chapter 3: Structure-function analysis of the nsp14 N7-guanine methyltransferase reveals an essential role in betacoronavirus replication

Chapter 1

General Introduction

1. The RNA: a diverse and essential molecule for life

Ribonucleic acid (RNA) is a macromolecule playing a key role in all known forms of life biology. It consists of a polymer of ribonucleotides linked by 5'-3' phosphate and is considered to be an elemental biomolecule having a vital role in the evolution of life and plays a significant function in almost all cellular processes. The RNA molecule is quite diverse in nature, which allows it to perform myriad biological functions. At least two distinct categories of RNA exist, notably, messenger RNA (mRNA) which is synthesized by the RNA pol-II and contains the genetic information that will be translated into proteins, and non-protein-coding RNA (npcRNA) or non-coding RNA (ncRNA). The ncRNAs have a subset of members holding a diversity of functions such as gene decoding, regulation, expression, and even sometimes catalytic activities known as ribozymes. Two well-known ncRNAs are transfer RNA (tRNA) and ribosomal RNA (rRNA) which assist protein synthesis. The ncRNA is further divided into long noncoding RNA (lncRNA), short non-coding RNA (sncRNA), and circular RNA (circRNA)[1]. In turn, the sncRNAs are classified into microRNAs (miRNAs), small interfering or silent RNAs (siRNAs)[2], small nucleolar RNAs (snoRNAs), and small nuclear RNAs (snRNAs), small interacting RNAs with piwi (piRNA)[2,3]. All of these varieties of RNA molecules display distinct roles in different life forms.

2. Overview of the epitranscriptome

Changes in RNA composition or conformation by nucleotide modifications are known as post-transcriptional or epitranscriptomics modifications that control the fate and function of the transcript. Post-transcriptional modifications (PTMs) were first discovered on RNAs abundantly found in the cell such as tRNAs or ribosomal RNAs. Using traditional biochemical techniques, the structural role of some modifications, their localization, and their function have been elucidated. Recently, the advent of new molecular mapping techniques, based on sequencing methods or mass spectra analysis, has allowed the identification of myriads of modified transcripts in different tissues under various physiological conditions such as viral infections. To date, hundreds of PTMs have been identified among which RNA capping, 2'O-methylation, N6-methyladenosine, N1-methyladenosine, 5-methylcytidine, N4-acetylcytidine, pseudouridine (Ψ) are the best described (Figure 1). Those PTMs participate in dynamically the fine-tuning of many RNA functions such as RNA stability, splicing, nuclear export, miRNA biogenesis, RNA localization, degradation, sensing by innate immunity, and translation. In addition to these, the

epitranscriptomic modifications also play a crucial role in host-pathogen interactions and in the regulation of virus replication mechanisms. Indeed, many studies have recently shown that viruses have evolved strategies to subvert the cellular RNA modifying enzymes[4,5]. As a result, viruses decorate their genome or mRNA in order to 1) optimize viral replication, and 2) control virus sensing by the innate immune system (Figure 1). In turn, the host cells have also been shown to modulate the PTM of viruses as a defense mechanism to limit their replication[4,6,7].

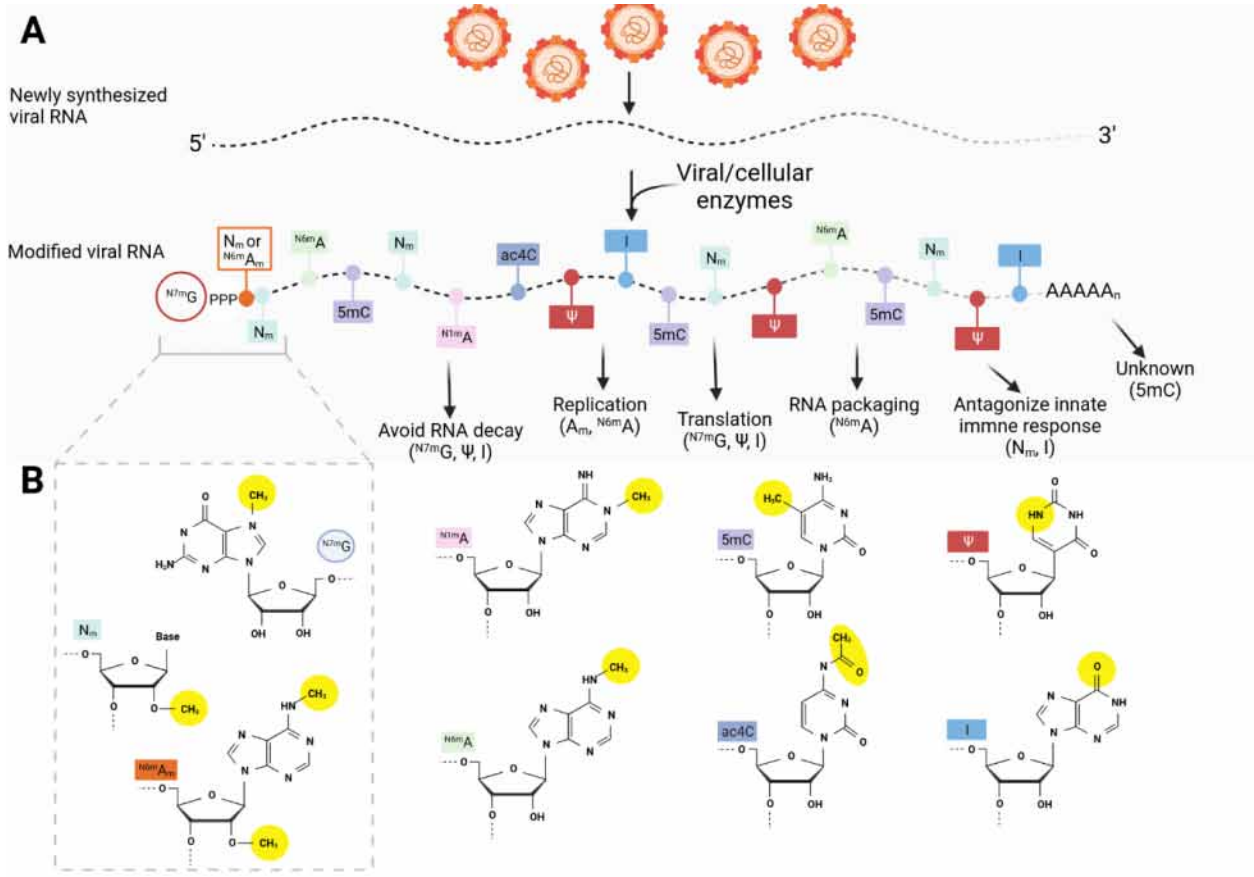

Figure 1: Overview of viral mRNA epitranscriptomic modifications and their outcomes. *A) During the viral life cycle newly synthesized viral RNA are prone to several epitranscriptomic modifications installed either by cellular or viral enzymes, among the commonly occurred modification are: RNA capping ($^{N7m}GpppN_1$), 2'O-methylation (N_m), N1-methyladenosine (^{N1m}A), N6-methyladenosine (^{N6m}A), 5-methylcytidine (5mC), N4-acetylcytidine (ac4C), pseudouridine (Ψ) and inosine (I). These epitranscriptomic marks direct various biological functions of viral RNA as indicated in the figure. B) Representation of different residue bearing epitranscriptomic marks. In yellow is highlighted the chemical change that has taken place.*

3. Host-parasite epitranscriptome: an evolutionary trajectory

Viruses are cellular parasites that infect cells and cause various diseases with serious consequences for global health. They exist as extracellular life forms that strictly need a host organism to replicate. Therefore, to promote their replication, viruses must maintain optimal cellular function and counteract some of their host physiological changes that limit their propagation. It's now evident that vRNA bear PTMs that can modify host immune responses, viral virulence, and the specificity of their interactions with host proteins. Despite the increasing knowledge about the epitranscriptome, little is known about the vRNA PTM landscape[4,5] and its outcome on viral replication. One of the best characterized PTMs in the viral world is the modification of the viral RNA by the addition of a 5'-end cap structure (Figure 1), described below). Which consists of an inverted N7-methylguanosine moiety linked to the first transcribed nucleotide (N_1) through a 5'-5' triphosphate bridge ($^{N7m}GpppN_1$)[8]. This cap structure is a feature adopted by viruses from eukaryotes due to its key role in the initiation of protein translation, protecting the RNA from cellular 5′ exoribonucleases (ExoNs), and escaping innate immune sensing[9–12].

Interestingly, during infection, the viral genome is confronted by many cellular RNA-modifying enzymes and can thus be marked. Among the possible modifications are the N6-methyladenosine (^{N6m}A), pseudouridine (Ψ), inosine (I), 5-methylcytosine (5mC), and 2'O-methylation (N_m, where N represents any nucleotide) are found inside the viral mRNA and can influence its metabolism and function[4,5] (Figure 1). The viral transcript can be edited by specific host enzymes. Two of the most common vRNA editing events are the isomerization of uridine to pseudouridine (or 5-ribosyl uracil) and the deamination of adenosine to inosine (Figure 2). The enzymes responsible for each editing event include the pseudo-uridine synthases (PUS), and RNA-acting adenosine deaminase (ADAR)[13–16]. However, until now, the enzymes required to reconvert the pseudouridilation and deamination reactions remain unknown, suggesting that these RNA modifications are irreversible (Figure 2).

The ^{N6m}A modification is probably the best-characterized PTM. This is due to the specificity and sensitivity of the mapping techniques available. Remarkably, the ^{N6m}A modification is an abundant PTM found in cellular and viral RNAs and its regulation is dynamic as proteins are available to set, remove, and detect this modification when necessary[17–21]. To date, no viral N6-methyltransferase (also known as the writer) has yet been identified and all identified ^{N6m}A within vRNA are associated with the host writers namely the METTL13 and METTL14[17,22]

(Figure 2). One of the peculiarities of this modification is that the methyl group can be erased by demethylases (erasers: FTO and ALKBH5) when necessary[23,24] (Figure 2). The various functions of ^{N6m}A modification are directed by the YTH domain family reader proteins (YTHDC1, YTHDC2, YTHDF1, YTHDF2, and YTHDF3)[17] (Figure 2). Numerous pieces of evidence highlight the importance of ^{N6m}A in the pathogenicity of some viruses.

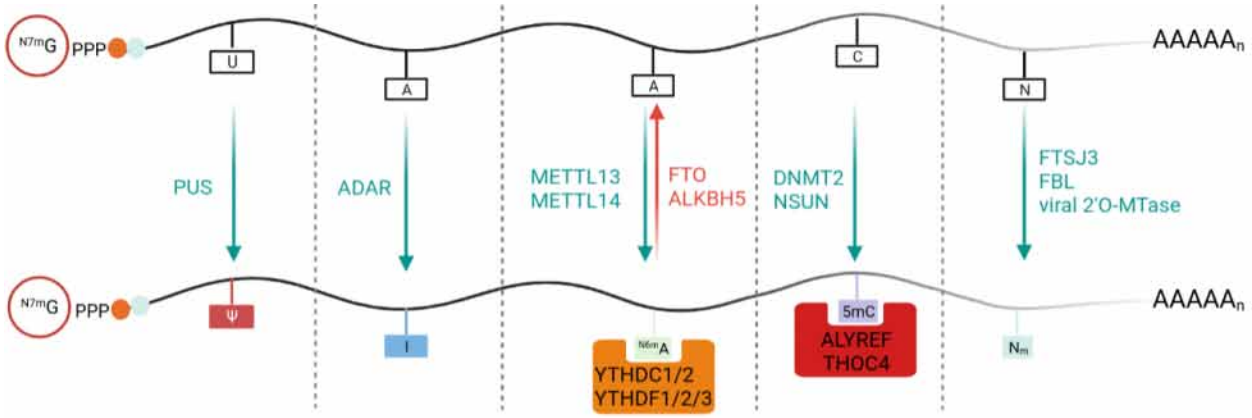

Figure 2: Regulation of viral epitranscriptome. *During the viral life cycle, viral RNAs undergo various key epitranscriptomic modifications necessary to regulate host-pathogen interactions. These modifications can be catalyzed either by cellular or viral enzymes. The cellular enzymes PUS and ADAR are responsible for the editing of uridine (U) - pseudouridine (Ψ) and adenine (A) - inosine (I) respectively. The adenine residues of viral RNA can be methylated at the N6 position (^{N6m}A) by the cellular methyltransferases METTL13 and 14. This modification can be removed by the cellular erasers FTO and ALKBH5. The cellular effector proteins YTHDC1 and 2 and YTHDF1,2 and 3 specifically recognize the ^{N6m}A with the RNA and regulate the function of the latter. Another type of viral RNA epitranscriptomic modification mediated and regulated by the cellular machinery is the 5-methylcytosine (5mC), this mark is deposited by DNMT2 and NSUN and recognized by ALYREF and THOC4 proteins. Finally, viral RNA nucleotides can be 2'O-methylated (N_m) by specific viral enzymes or by cellular enzymes, in particular FTSJ3 and FBL.*

In contrast to RNA N6-methylation and 2'O-methylation, the interplay of 5mC in the viral transcript during viral infection is barely understood. However, the 5-methylated cytosine is a common epitranscriptomic modification encountered in eukaryotic DNA and RNA. Many studies highlight the regulatory role of 5mC in RNA metabolism, namely export, ribosome assembly, translation, and RNA stability[25]. Different analytical approaches have identified the regions of concentration of 5mC marks in RNA without however revealing a clear pattern, but they are mainly found in the GC-rich regions of the 5'- and 3'-untranslated regions (UTRs), some are also found in RNA and on the poly(A) tail[26] . The enzymes responsible in particular for the 5mC of cellular and viral RNA belong to two specific subgroups of methyltransferases, namely DNMT2 or NOL/NOP2/sun (NSUN)[27,28] (Figure 2). To date, it is unknown whether the methyl group of the 5mC modification can be erased. However, the RNA transport adaptor protein,

Aly/REF export factor (ALYREF or THOC4) has been identified as a putative reader of the 5mC sites within the mRNA[26] (Figure 2).

Accumulating evidence demonstrates that some viruses contain internal 2'O-methyl marks within their genome. The enzymes involved in this methylation process can be either of viral origin or of the host[29–33]. Among the latter, the cellular FTSJ3 has been proven to actively methylate HIV-1 RNA and these epitranscriptomic marks shield from sensing by the innate immune sensor MAD5. Moreover, FBL has been shown to be important for the replication of numerous viruses, the proviral role of this cellular MTase has been linked to its ability to methylate the viral genome, but this remains, however, to be demonstrated[30,34,35] (Figure 2) (see below for more information on this epitranscriptomic mark).

4. RNA capping: a fundamental epitranscriptomic mark

In eukaryotes, mRNA cap structure displays various functions, it 1) controls mRNA synthesis and splicing, 2) protects mRNA from 5'-exonuclease induced decay, 3) allows mRNA recruitment at the ribosome for their translation into proteins, and 4) allows the discrimination of exogenous RNA as "non-self". Given the fundamental role of the cap in the biology of mRNAs, viruses have developed mechanisms to hide their RNA 5'-ends with a cap or cap-like structures indistinguishable from that of the host.

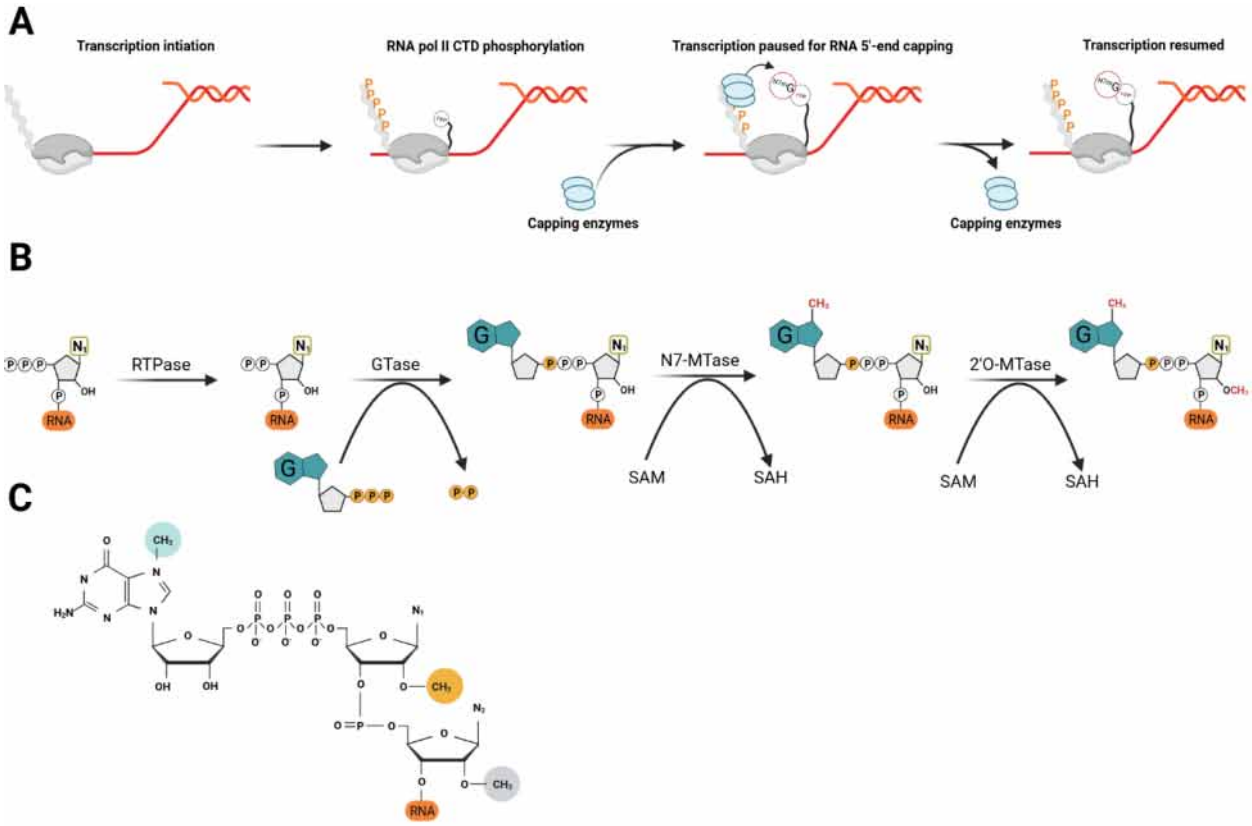

Figure 3: Eukaryotic mRNA transcription and capping. *RNA capping occurs co-transcriptionally. At transcription initiation, phosphorylation of RNA pol II CTD domain promotes the recruitment of RNA capping enzymes. Then the transcription is paused so that the capping enzymes can catalyze a cap structure on the 5'-end of the nascent transcript. After this, the capping enzymes are released from the hyperphosphorylated RNA pol II CTD domain leading to the processive transcription. B) The cap-0 structure is formed on nascent RNA by three sequential enzymatic reactions. First, the RTPase hydrolyzes the γ-phosphate of the nascent RNA to yield a diphosphate RNA and inorganic phosphate (Pi). Then, GTase then transfers the GMP (Gp) from a GTP molecule to the 5'-diphosphate RNA to generate $GpppN_1$-RNA and a molecule of PPi. In the final step, N7-MTase transfers the methyl group from SAM to the cap guanine to form the cap-0 structure ($^{N7m}GpppN_1$) and releases SAH moiety as a by-product. The cap-0 is further methylated by a 2'O-MTase reaction on the ribose-2'-O position of the first nucleotide, generating the cap-1 structure ($^{N7m}GpppN_m$). C) Structure of RNA cap-2. In bleu is highlighted the N7-methylation of the cap, while the 2'O-methylation of the first and second nucleotide of the RNA are highlighted in orange and grey respectively.*

Eukaryotic mRNA capping results from a series of enzymatic reactions catalyzing modifications that are added co-transcriptionally to the nascent RNA in the nucleus. The cap structure is not encoded by DNA but rather results from the recruitment of cellular enzymes by the C-terminal domain (CTD) of the RNA pol-II after the initiation of the RNA transcription. The enzymes that catalyze this RNA modification are the RNA 5'-triphosphatase (RTPase), guanylyltransferase (GTase), and an N7-methyltransferase (N7-MTase)[36]. *Saccharomyces cerevisiae* code for tree independent enzymes which ensure the RNA Capping (CET, CEG, and

ABD1) and the knockout of these enzymes results in a growth defect of the yeast[37]. In human cells, the capping enzyme (human capping enzymes; hCE) is a bifunctional protein with an N-terminal domain harboring the triphosphatase activity and the C-terminal domain GTase domain[38,39]. The capping occurs on the transcript of RNA polymerase II while it is only about 25-30 nucleotides (Figure 3A). As the nascent RNA has a 5′-triphosphate ($pppN_1$-RNA), the capping process begins with the removal of the γ-phosphate from the transcript by an RTPase. Then a GTase catalyzes the addition of a guanosine monophosphate (GMP) moiety on the 5'-diphosphate of the RNA leading to the formation of a cap-0 structure ($GpppN_1$). This occurs via a two-step reversible reaction: the GTase first reacts with the α-phosphate of the guanosine triphosphate (GTP), forming an enzyme–GMP (Gp-enz) intermediate, then the GMP is transferred from Gp-enz to ppN_1-RNA to generate $GpppN_1$-RNA. Finally, the guanine moiety of the cap is methylated at the N7 position by an S-adenosyl-L-methionine (SAM)-dependent N7-MTase (human Cap N7 methyltransferase, hCM) thus providing the minimal RNA cap structure, termed cap-0 ($^{N7m}GpppN_1$)[8] (Figure 3B).

After cap synthesis, capping enzymes dissociate from the CTD domain of RNA pol-II and the latter allows transcription to resume[40] (Figure 3B). The RNA transcript, which is in the form of premature RNA (pre-mRNA)), can next undergo a series of splicing events and 3 'end polyadenylation generating a mature and functional RNA (mRNA). Upon maturation, the mRNA is exported from the nucleus to the cytoplasm, where it recruits translation factors and the ribosomal subunits for translation. While being translated, the transcripts must be protected from premature degradation by 5'- end 3'-exonucleases. These processes are governed by mRNA 5′-end N7-methyl cap structure, and 3'end poly(A) tail[7,8].

5. RNA cap 2'O-methylation: "self" mimicry

Eukaryote cap-0 structure can be further methylated by 2′O-methyltransferases (2′O-MTases). It consists of an enzymatic transfer of a methyl moiety (CH_3) to the 2'O of the ribose of the first and second transcribed nucleotides thus yielding cap-1 ($^{N7m}GpppN_m$) and cap-2 ($^{N7m}GpppN_m$-N_m) structures, respectively[43]. While the cap-1 is more common in lower eukaryotes, higher eukaryotes harbor a cap-2 structure[44–46]. In humans, the cap methyltransferase (CMTR) 1 and 2 methylate respectively N_1 and N_2 of the mRNA[47,48] (Figure 3B & C). Dysregulation of CMTR1 has been associated with various diseases, particularly through mechanisms related to RNA discrimination of self versus non-self[30,49,50]. This is expected since

the 2'O-methylation of cap-0 N_1 serves as a feature of the host's innate defense mechanism[51]. Conversely, the role of N_2 methylation is still unknown.

6. Nuclear RNA cap quality control system

Pre-mRNA splicing and mRNA 3'-end polyadenylation have been associated with a cap quality control mechanism. In mammals, the multifunctional protein DXO/Dom3Z specifically recognizes and degrades unmethylated cap ($GpppN_1$-RNA). Knockdown of DXO/Dom3Z induces an accumulation of pre-mRNA in the nucleus of cells as well as defective cleavage of mRNA 3′-end for polyadenylation[52], highlighting the link between cap N7-methylation and pre-mRNA splicing and polyadenylation[52]. In contrast, *Saccharomyces cerevisiae* presents two functionally redundant sets of RNA cap quality control machinery. The first consists of a protein complex: the Rai1-Rat1 system, within which Rai1 performs pyrophosphatase and decapping activity[53] while Rat1 (homologous to mammalian XRN2) does the exoribonuclease activity[54]. The second known cap quality system is the Dxo1 protein which possesses both decapping and 5′ to 3′ exonuclease activity[55].

7. Cytoplasmic RNA turnover

The fate of mRNAs is under the control of two major processes including the transcript decapping and the removal of the poly(A) tail[7,8], which control the mRNA decay (Figure 4). Cap removal is catalyzed by cytoplasmic Decapping Scavenger (DcpS), upon stimulation by decapping activator proteins. Following this, the transcripts are prone to 5' to 3' exonucleases such as XRN1 in the processing bodies (PB) (Figure 4). It is thus uncommon to find 5′-triphosphate mRNAs in the cytoplasm. Moreover, eukaryotes have evolved mechanisms that recognize 5'-triphosphate-RNA and trigger an appropriate innate immune response[9,10]. Being the most common cellular invaders to produce cytoplasmic mRNAs, viruses have been under selective pressure to develop strategies to protect their 5'-ends from detection by innate immune sensors and decay by 5'-exonucleases.

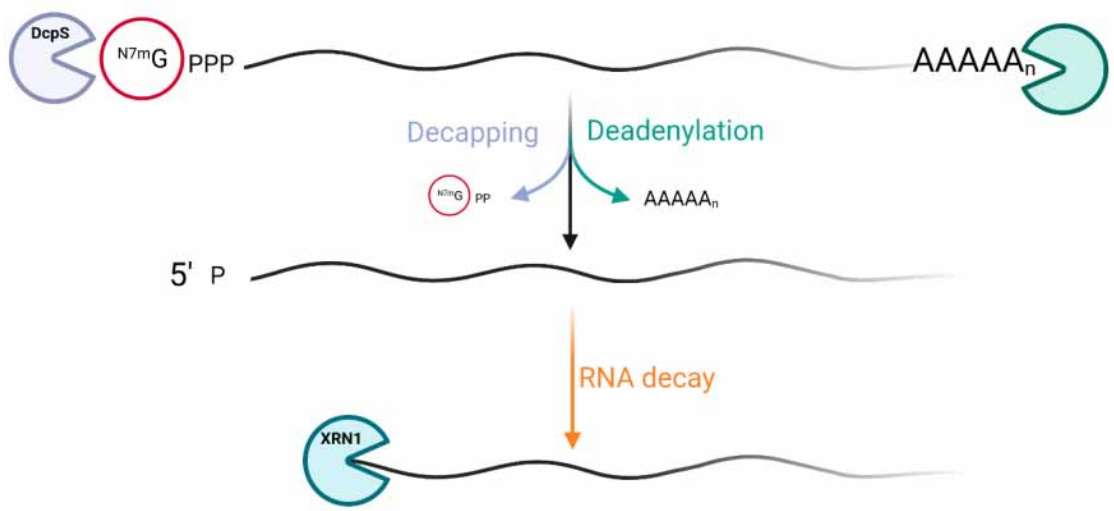

Figure 4: Cytoplasmic mRNA turnover mechanism. *Protein expression is down-regulated by the mRNA turnover process. The latter is directed by mRNA 5'end decapping (blue) and 3'-end poly(A) deadenylation (sea green). Once decapped by the DcpS (blue), mRNA is transferred to the PB where it is prone to a 5' to 3' decay by the XRN1 (5' exonuclease).*

8. Cytoplasmic RNA recapping: an emerging concept

Initially considered to exclusively occur in the nucleus, RNA capping has recently been described in the cytoplasm of mammalian cells. It appears that uncapped mRNA stored in the PB[56–58] can be re-addressed to polysomes after re-capping for translation[56,59]. Targeted substrates for cytoplasmic recapping contain a 5'-end structure resistant to decay by XRN1 and a stable poly(A) tail[60]. This process is possible because the capping metabolon necessary for cap-0 formation is found in the cytosol at a basal concentration[60]. In mammals, the cytoplasmic capping pathway occurs in the following manner: 1) first an ATP-dependent kinase phosphorylates 5′-monophosphate RNA 2) then the GTase domain of RNGTT caps the resulting diphosphate RNA 3) the cap is subsequently methylated by the cytoplasmic RNMT-RAMAC (cap N7-MTase) to generate a cap-0 structure[60–62] (Figure 5). Interestingly, studies have shown certain mRNA are dynamically de- and re-capped in the cytoplasm. This could be one of the mechanisms used by the cell to regulate protein synthesis[63,64]. Now it can be questionable whether viruses resort to the cytosolic capping machinery, evidence has it that the hepatitis B virus could use this capping pathway[65].

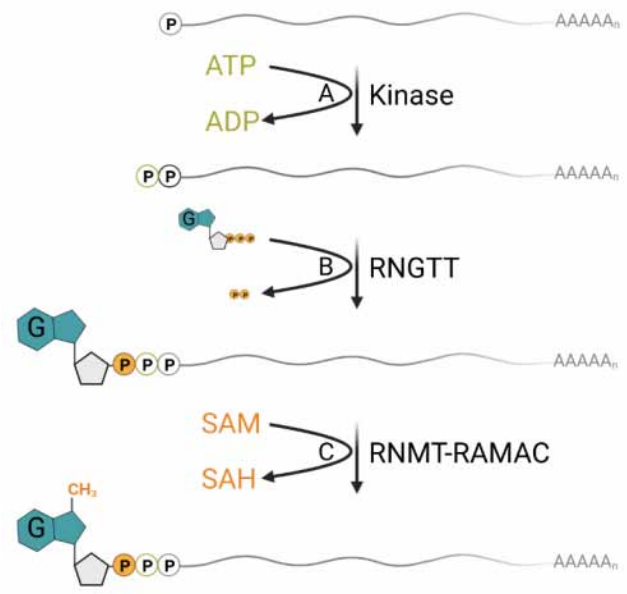

Figure 5: Model mRNA recapping pathway in mammalian cytoplasm. *A) The 5'-monophosphate RNA is phosphorylated by an ATP-dependent kinase. B) Following this, the GTase domain of RNGTT transfers a GMP on the 5'-diphosphate RNA to yield the synthesis of capped RNA (GpppN$_1$-RNA). C) The cap structure is methylated at the N7 position by the RNMT-RAMAC heterodimer to generate an mRNA with a cap-0 structure at its 5'-end.*

9. Viral RNA capping: a diverse mechanism apparatus

Given the fundamental role of the cap in mRNA turnover and expression, viruses often decorate the 5'-extremity of their RNA with a traditional cap or a cap-like structure for an efficient replication[66]. The cap-like structures consist of peptides (i.e., VPg) or pseudo cap (e.g., NAD, FAD, etc..) described to protect the 5'-end of a limited set of viral RNA (see below in the section: who needs a cap). However, even though most viral mRNAs carry a cap structure different capping mechanisms were evidenced. At least three different RNA capping strategies exist in the viral world: 1) The first is through the host capping machinery, 2) the second consists of stealing the cap of the host mRNAs (known as "cap snatching"), 3) the third is based on the use of viral encoded capping machinery. Within this diversity, viral mRNA (vRNA) capping may or may not occur canonically depending on whether or not it follows the same pathway as eukaryotes[66] (Figure 3B).

9.1. Who needs a conventional cap? Role of alternative 5'-end modification of vRNA

The cap structure of some RNA populations such as human U snRNAs has trimethylguanosine (TGM: ^{N2,2,7m}GpppRNA, Figure 6). This structure is catalyzed on ^{N7m}GpppRNA by the trimethylguanosine synthase 1 (TGS1)[67]. In this form, some TMG-capped coding RNAs can bind eIF4E and undergo translation[67]. Additionally, the cap-0 and 1 structure can be further methylated. This occurs especially when the N_1 of the cap is adenosine residue. Indeed, in this case, the adenosine can be methylated at the N6 position to generate either ^{N7m}Gppp^{Nm6}A or ^{N7m}Gppp^{Nm6}A$_m$[68] (Figure 6).

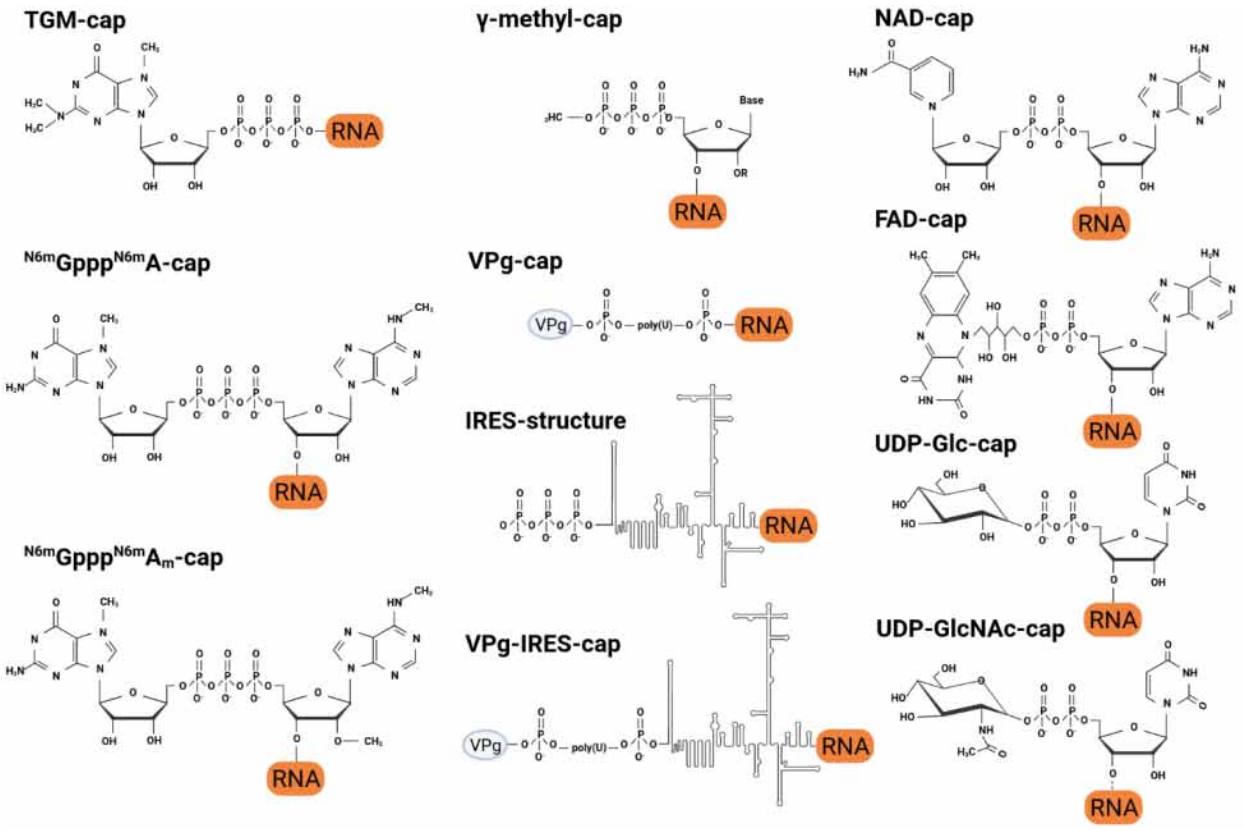

Figure 6: The diversity of the RNAs 5'-end structures. *Structural representation of eukaryotic and viral RNAs 5'end caps. Not all caps are presented here.*

The acquisition of a cap structure is by far the only mechanism used by viruses to ensure efficient translation and avoid recognition by the innate immune sensors. Unlike eukaryotic mRNA, the *Picorna-*, *Poty-*, and *Caliciviridae* do not have a cap structure at the 5′-end of their RNAs. They are rather covalently bound to a protein termed viral protein genome-linked (VPg)

(Figure 6), which ensures the recruitment of the eIF4E[69,70]. However, the translation of viral protein is accomplished by an unusual mechanism. It is mediated by an internal ribosome entry site (IRES) structure[32] found in the 5′-untranslated region of the vRNA[33] (Figure 6). The IRES have also been found in numerous cellular and viral RNAs including those of Lentivirus, Hepacivirus, and Pestivirus. In addition to the IRES, lentiviruses also decorate the 5'-end of their RNA with a cap structure. Conversely, in pestiviruses, the genomic RNA remains uncapped and harbors a 5′-triphosphate thus promoting high levels of expression of host interferon-stimulated genes[71–73].

Various eukaryotic RNAs have been recently evidenced to protect their 5'-end with other types of structures. For instance, snRNAs (small nuclear RNA) synthesized by RNA polymerase III such as U6 and 7SK have γ-methyl caps (CH_3pppRNA, Figure 6) catalyzed by a γ-MTase. Interestingly, several non-canonical structures are observed at the 5'-end of some eukaryotic and viral RNA of which the nicotinamide adenine dinucleotide (NAD), flavin adenine dinucleotide (FAD), uridine diphosphate glucose (UDP-Glc), and uridine diphosphate N-acetylglucosamine (UDP-GlcNAc)[74,75] (Figure 6). The function of these alternative cap structures remains elusive except for the NAD+ cap. The latter cannot support translation and promote mRNA decay, unlike the conventional cap structure[76]. The recent discovery of these alternative structures at the 5’-end of human and viral RNA offers an exciting area for research.

9.2. Viruses using the host capping machinery

Most viruses that use RNA pol II for RNA synthesis during their life cycle, such as *Retroviridae*, *Bornaviridae*, and DNA viruses (except *Poxviridae* for which the replication is cytoplasmic), use the capping machinery of the host cells[66]. Interestingly, some viruses such as this apply to HIV have been shown to contain unconventional cap structures (e.g. TMG)[67]. However, neither the mechanisms allowing the formation of these unusual cap structures nor their functional consequences are yet understood.

9.3. Cap snatching mechanism

Multinegaviruses such as influenza viruses (*Orthomyxoviridae*), Lassa virus (*Arenaviridae*), hantavirus (*Hantaviridae*), and rift valley fever virus (*Phenuiviridae*) do not encode for any cap-synthesizing machinery nor can employ that of their host[77,78]. They have, however, acquired the ability to steal the caps of host mRNA in a process called “cap snatching”. In this case, the enzymatic complex organized around the viral polymerase contains a protein with a cap-binding

domain (PB2), which recruits the cellular cap RNAs. The cap snatching mechanism begins with the binding of host mRNA to viral RdRp complex in a cap-1 or cap-2 specific manner[79] (Figure 7). Then, the transcript is cleaved 10-20 nucleotides downstream of the cap structure by an endonuclease domain present in the enzymatic complex associated with the viral RNA. Finally, this short, capped RNA thus generated is used as primers by the polymerase to initiate viral RNA synthesis[80,81] (Figure 7). The consequence of this mechanism is that the viral RNAs produced are chimeric and contain a 5'-end of 10 to 15 variable nucleotides "stolen" from cellular RNA. One of the important advantages of this viral strategy is that this mechanism induces the shut down of cellular RNA expression in favor of viral RNA expression. Interestingly, 'decapped' host mRNAs released upon endonucleolytic cleavage are prone to the cytosolic to nuclease action therefore in their downregulation[66].

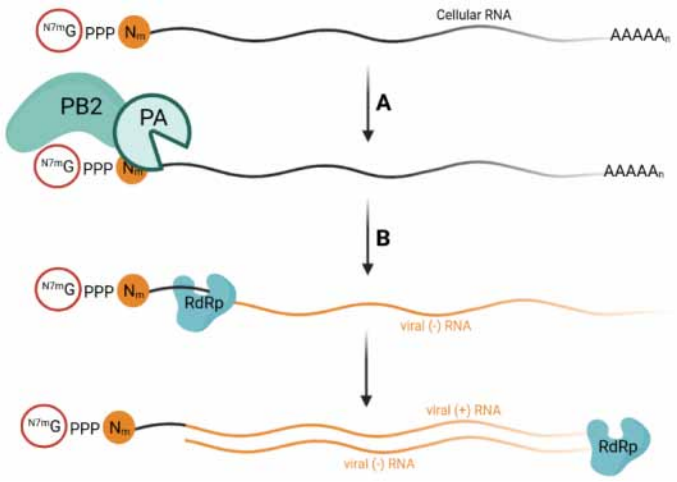

Figure 7: Snatching off the host cap. *The RNA-capping mechanism of negative-sense RNA viruses such as those of the Orthomyxoviridae family consists of stealing the cellular mRNA cap structure. A) The PB2 subunit of the viral RdRp binds to the cap structure -namely a cap-binding domain- of cellular mRNAs (which are enriched in the P-bodies), then the PA subunit releases short capped RNAs via its endonuclease activity. B) The short capped RNAs are used by the RdRp to prime viral transcription and generate viral mRNA [viral (–)RNA] using the viral negative-stranded RNA [viral (–)RNA] as a template.*

9.4. Virally encoded capping mechanism

Additionally, some cytosolic replicating viruses that encode for their own capping enzymes have been shown to follow the same capping pathway as eukaryotes, this applies to the poxviruses, reoviruses, flaviviruses, etc. It is noteworthy that although the capping involves the 3

canonical enzymatic activities (RTPase, GTase, and N7-MTase) the functional organization of these machinery and the regulation of enzymes involved in the capping process are not necessarily conserved (Figure 3B). Additionally, some of the viruses that encode for their own machinery use an alternative capping pathway. This is exemplified by the alphaviruses and *Mononegavirales* (described below), which use a PolyRiboNucleotidylTranferase (PRNRase) activity instead of a GTase to ensure their cap synthesis.

9.4.1. The canonical vaccinia virus RNA capping pathway

Viruses that encode for the capping system are characterized to follow the conventional pathway are the following: flaviviruses, vaccinia virus (a dsDNA virus from the *Poxviridae* family), and mammalian orthoreovirus (dsRNA virus from the *Reoviridae* family). Characterization of the enzymatic system of the last two viruses played a crucial role in the general understanding of the RNA capping process[66]. Vaccinia virus mRNA capping is performed by a unique heterodimer system, in which the first subunit (D1) harbors the RTPase, GTase, and N7MTase activity under the influence of the second subunit (D2)[82]. Nascent mRNA 5′-triphosphate is first hydrolyzed by RTPase to yield RNA 5′-diphosphate, which is then transferred to the other domains for a subsequent GMP addition and N7-guanine methylation[83,84]. The latter reaction is promoted by an allosteric stimulation from the D12 protein[83,84]. Vaccinia mRNA capping ends with 2'O-methylation of the cap-0 structure by viral VP39[85].

9.4.2. Non-canonical viral RNA capping pathways

9.4.2.1. Alphavirus-like togaviruses capping pathway

Viruses such as chikungunya virus, semliki forest virus, and sindbis virus only harbor the cap-0 structure and have reshuffled their capping process. Indeed, alphaviruses encode for the non-structural protein 1 (nsP1) which harbors both N7-MTase and GTase activities. The capping process starts with the methylation of a GTP before being covalently linked to the catalytic histidine of the GTase forming an ^{N7m}Gp-enz: intermediate[8,86] (Figure 8). The N7-methylated GMP is then transferred to the processed ppN_1RNA to generate a capped viral transcript (^{N7m}GpppN$_1$RNA)[8,86]. The RTPase activity is supposed to involve the nsp2[88] (Figure 8). Interestingly, the cryo-electron microscopy structure of nsP1 of the chikungunya virus, reveals that nsP1 forms a dodecamer associated with intracellular membranes containing the viral replication complexes as RNA synthesis reactors. The ring shape of the complex ensures the protection of nascent RNA from detection by the innate immune sensor and controls the exit of

properly capped viral RNA in the cellular cytoplasm. Although the vRNA exported from the intracellular vesicle is protected by a 5' cap-0 (as the 2'O-methylation is lacking), its recognition by RIG-like receptors is limited by a secondary structure found downstream of the cap structure[51,88] (see below).

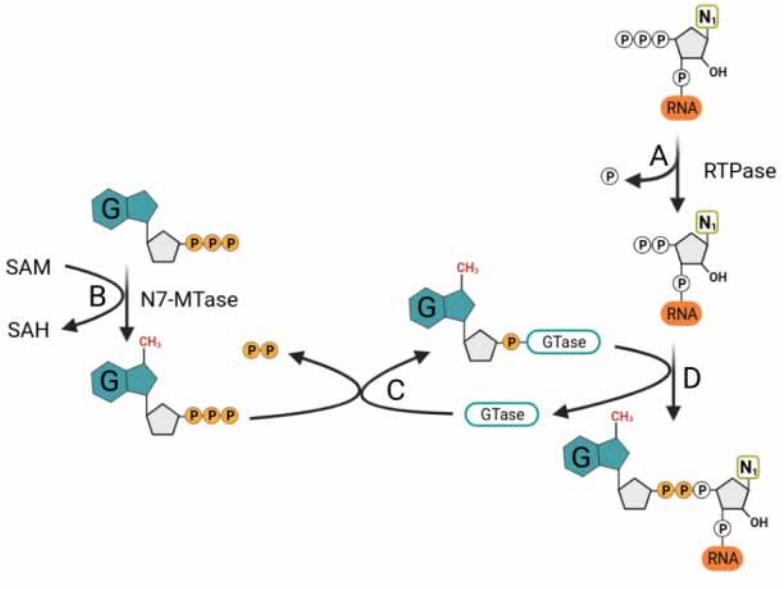

Figure 8: Alphavirus-like togaviruses capping mechanism. *A) The RTPase hydrolyses the γ-phosphate of the viral RNA to release a diphosphate RNA and Pi. B) Simultaneously, a GTP molecule is methylated at its N7 position by the SAM-dependent N7-MTase. C) The GTase forms a covalent link with the N7-methyl-GTP and releases pyrophosphate moiety (PPi). D) The GTase then transfers the m7GMP molecule (m7Gp) to the 5′-diphosphate RNA to create m7GpppNp-RNA. The box indicates viruses that acquire their cap structures using such a capping pathway.*

9.4.2.2. *Mononegavirales* fashionable capping pathway

Members of the *Mononegavirales* have evolved a cap-synthesizing apparatus independently from other known eukaryotes. They encode for a multifunctional large protein (L protein), that harbors an RNA-dependent RNA polymerase (RdRp), cap synthesis by using a polyribonucleotidyl transferase (PRNTase) activities (instead of a GTase), and the subsequent cap methylation. Unlike the conventional eukaryotic way, *Mononegavirales* L protein transfers a GDP to the RNA 5′-end[89–92]. The capping process occurs in the following manner, the PRNTase domain of the L protein first forms a covalent link between its catalytic histidine and the nascent vRNA (enzyme-pN_1-RNA intermediate) (Figure 9). Then a GDP generated from GTP[89], by a yet unidentified NTPase, is added to the 5′-monophosphorylated vRNA (Figure 9). The MTase domain of *Mononegavirales* L protein carries bifunctional activity, it harbors N7- and 2'O-MTase activities and allows the formation of a cap-1 structure. Thus, even if the reaction

pathway is extraordinary, the product (cap-1 structure) is indistinguishable from the native cap-1 structure cellular mRNA. Regardless of the high sequence similarities between the MTase domains of *Mononegavirales*, it appears that the cap methylation process occurs in various manners in the *Mononegavirales* order. Indeed, while VSV cap N7-methylation exclusively occurs if the N_1 of the transcript is 2'O-methyl guanosine[207] (Figure 9), the RSV methylation apparatus operates in the opposite direction[89] (Figure 9). Interestingly, it is likely that Ebola MTase responsible for vRNA cap 2'O-methylation also carries a cap-independent activity allowing internal adenosine 2'O-methylation[95]. The actual presence of such internal methylation inside the vRNA remains to be confirmed and their biological outcome one viral replication is still elusive but it is speculated that they might interplay with RNA sensing by RIG-Like receptors such as MDA5 (see below).

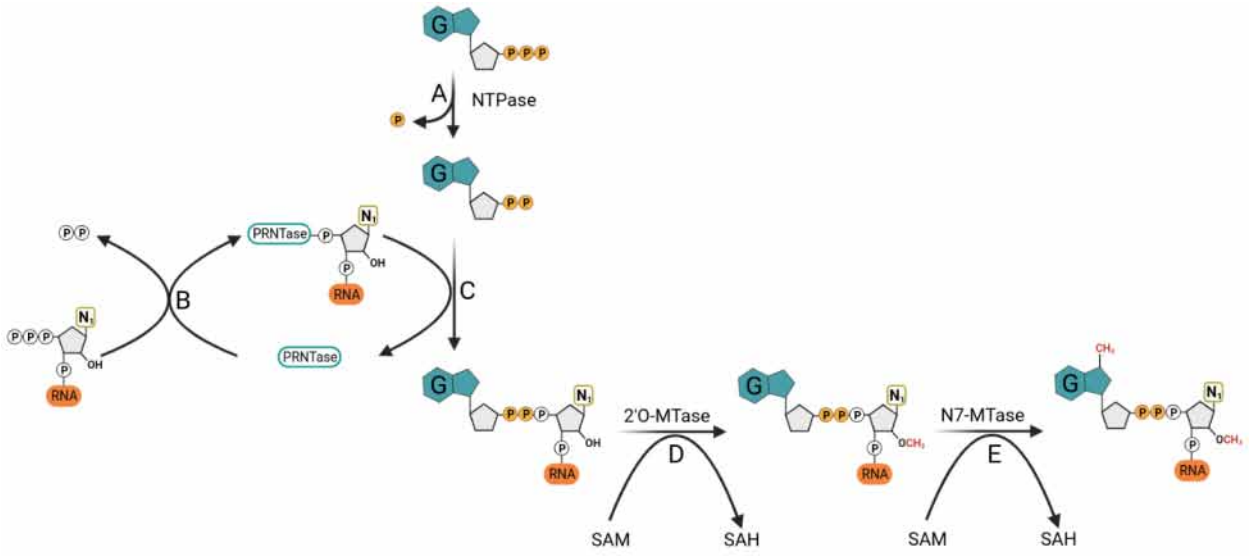

Figure 9: Mononegavirales RNA capping pathway. *A) The NTPase hydrolyses the γ-phosphate of GTP (Gppp) and generates a GDP (Gpp) and inorganic phosphate (Pi) moiety. B) On the other hand, the polyribonucleotidyl transferase (PRNTase) forms a covalent bond with the nascent viral RNA ($pppN_1$-RNA; in which N_1 is the first transcribed nucleotide) and releases a pyrophosphate moiety (PPi) and PRNTase–N_1-RNA intermediate. C) Following this, the PRNTase transfers the RNA molecule to the GDP to yield $GpppN_1$-RNA. D-E) In the case of VSV, a 2'O-MTase transfers the methyl group from SAM moiety to the N_1 (first nucleotide) of the RNA. The capping reaction ends with the methylation of the cap at the N7 position of the guanosine by a SAM-dependent N7-MTase.*

9.4.2.3. The curious case of coronavirus capping machinery.

The *Coronaviridae* family has spawned three major zoonotic introductions, namely severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome CoV, and the recently identified SARS-CoV-2. These human pathogens belong to the

Betacoronavirus genus, which contains numerous CoVs circulating in bat species[96–99]. Although their ability to cross species barriers has been demonstrated, there are no effective prophylactic solutions to prevent any CoV outbreak. Coronavirus capping machinery is attractive for the development of antiviral strategies. Indeed we can expect that the inhibition of CoV capping would result in a strong impairment in vRNA translation into protein and increase the early detection of vRNA by innate immune sensors (see below). The genetic analysis of the viral enzymes encoded by CoV has allowed others and us to identify nsp16 as a putative 2'O-MTase and its enzymatic activity was confirmed biochemically[100,101]. Next nsp14 exonuclease was reported to carry N7-MTase activity by a complementation assay in yeast devoid of N7-MTase activity[102]. Based on these results, and on biochemical data demonstrating the N7- and 2'O-MTase activity of nsp14 and nsp10/nsp16 complexes, it was inferred that the CoV capping involves the subsequent action of different viral enzymes. CoV capping pathway was first assumed to occur in a canonical way as followed: the nsp13 (RTase) starts by removing the γ-phosphate from the nascent 5′-triphosphorylated RNA[103,104], followed by a transfer of a GMP to the RNA's dephosphorylated 5′-end, a role assumed to be catalyzed by the nsp12 nucleotidyltransferase domain[105–107]. The cap is first methylated by the nsp14 (N7-MTase) to generate a cap-0 structure. The latter is converted by the nsp10/nsp16 (ribose 2′-O-MTase) complex into a cap-1 structure[100,102]. Although many *in vitro* studies have demonstrated and characterized the SARS-CoV in detail the methylation pathway on synthetic RNAs, the true mechanism of capping still remained hypothetical. This is especially true in absence of biochemical data elucidating the exact role of the rôle of the NIRAN domain present in the N-terminal part of the nsp12 RdRp. Interestingly, a recent preprint proposes that CoVs cap synthesis occurs in an unexpected noncanonical way. Indeed the study suggests that CoV capping implies the nsp12 NiRAN domain and the nsp9 protein[108]. Accordingly, the authors propose that the NiRAN domain transfers pppRNA to the amino terminus of nsp9, forming a covalent RNA-nsp9 intermediate in a process called RNAylation and realizing ppi (Figure 10). The NiRAN domain is next proposed to transfer RNA to GDP, forming the canonical cap structure (GpppA-RNA, Figure 10). This process is reminiscent of the PRNTase activity previously described for *Mononegavirales* capping (Figure 9), but in the case of the latter, at least two proteins are needed for that step. The N7-MTase (nsp14) and 2'O-MTase (nsp10/nsp16) are then supposed to transfer the methyl groups forming functional cap-1 structures (Figure 10). The replication-transcription complex (composed of nsp7 & 8 and nsp12) bound to nsp9 is thus the key player in the capping reaction. This was validated by a mutagenesis

study on nsp9 and NiRAN domain catalytic and conserved residue Lys 73 and asp 218 respectively. Thus it is now likely CoV follows an extraordinary and unconventional capping mechanism essential for successful SARS-CoV-2 replication. The discovery of this capping pathway constitutes a new target for the development of anti-SARS-CoV antivirals[108].

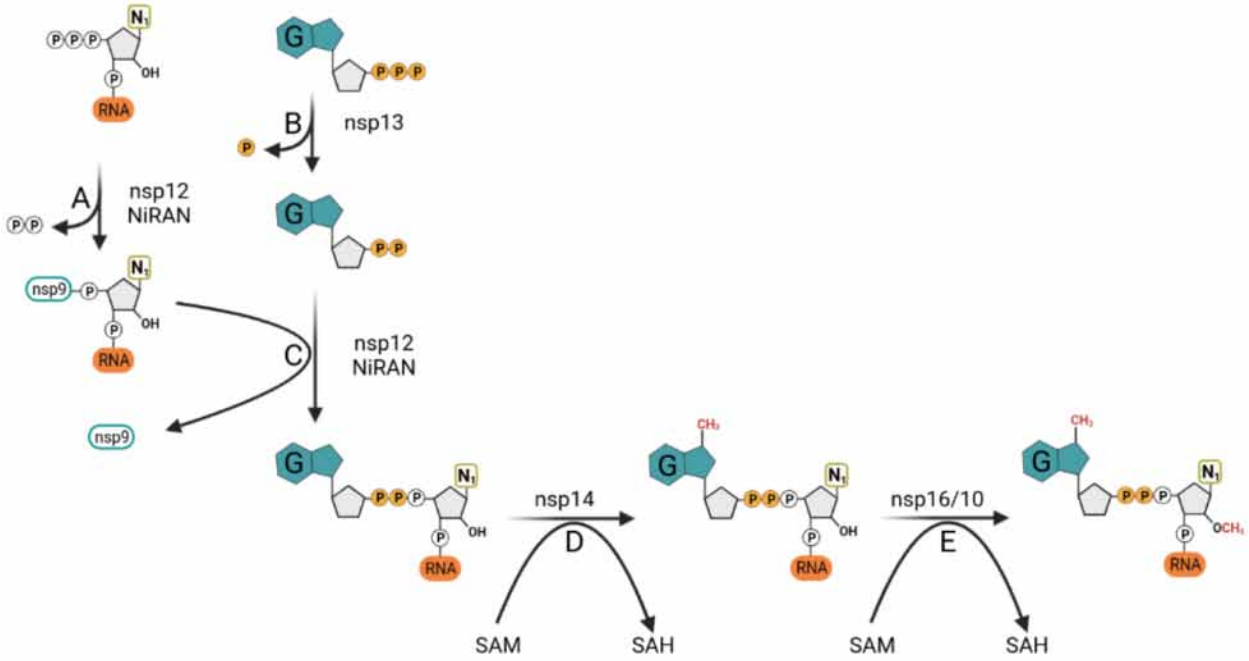

Figure 10: The outstanding coronaviruses capping mechanism. *A) The NiRAN domain of the nsp12 catalyzes the transfer of nsp9 on the nascent RNA (pppN$_1$-RNA) releasing pyrophosphate (PPi) and a covalent nsp9-N$_1$-RNA intermediate. B) The nsp13 hydrolyses the γ-phosphate of GTP (Gppp) to generate an inorganic phosphate (Pi) moiety and a GDP (Gpp). C) The latter is used by the nsp12 NiRAN to replace the nsp9 on the RNA and generate GpppN$_1$-RNA. D) Following this, the guanine moiety is methylated at position N7 by a SAM-dependent nsp14. E) CoVs capping reaction ends with the 2'O-methylation of the N1 (first nucleotide) of the RNA by the nsp16/10 complex.*

9.4.2.3.1. Coronaviruses nsp14 harbor two enzymatic activities and the MTase fold is uncanonical.

Most known MTases share a similar Rossmann fold (RF) organization[109,110] (Figure 11), a higher-order structure fairly conserved among dinucleotide-binding enzymes[109,111]. However, structural analysis of SARS-CoV nsp14 highlights an outstanding non-conserved RF within its MTase domain (characterized as a non-RF) (Figure 11). Biochemical characterization of nsp14 has revealed that this protein plays several functions in the viral replication cycle. Its function is not only limited to the cap N7-methylation, it is also involved in maintaining replication fidelity

via its 3' to 5' ExoN domain[112,113]. Although the N7-MTase and ExoN domains of nsp14 are functionally distinct[114–117], truncations and alanine substitutions in the ExoN domain can affect the N7-MTase activity[102,118]. The latter stems from the fact that both domains are structurally intertwined, so any destabilization of one could have serious repercussions on the other[114,115]. Moreover, in silico analysis highlights the NRF as a unique feature shared only among CoVs[115,119]. So far, the significance of the CoV N7-MTase domain has been assessed on murine hepatitis virus (a model of betacoronavirus). Alteration of the N7-MTase increases virus susceptibility to innate immune response and strongly reduces vRNA translation efficiency, as a consequence viral replication is drastically decreased[120]. Given the key role of N7-MTase in viral survival and its level of conservation within CoVs, it constitutes an attractive target for a broad spectrum anti-CoV compound.

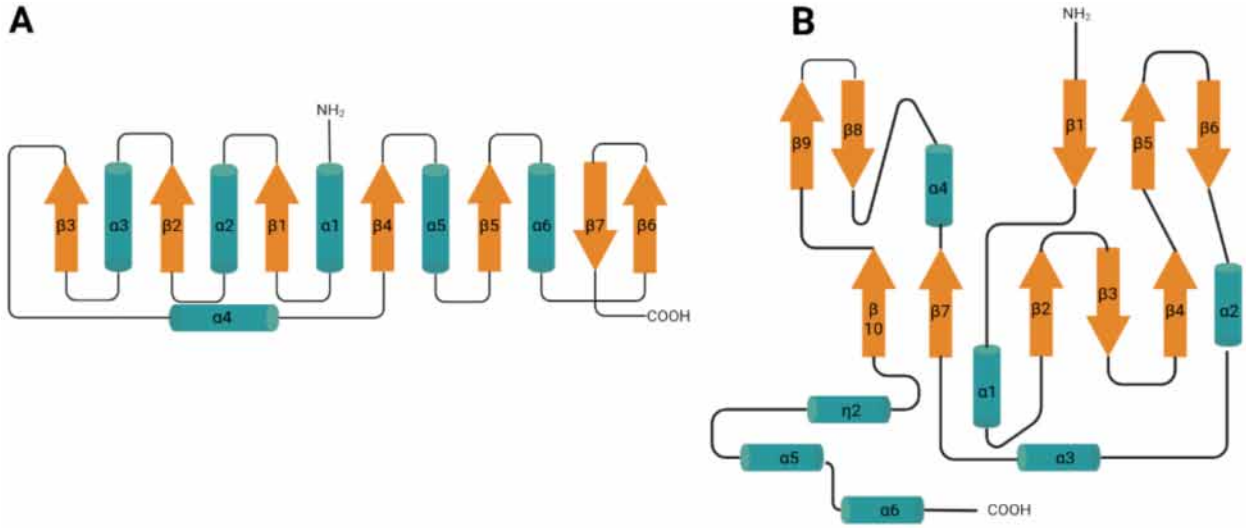

Figure 11: Topographic depiction of a conventional and unconventional methyltransferases secondary structure. *A) Representation of an MTase with a conventional Rossmann fold organization (β–α–β). B) Topology representation of SARS-CoV nsp14, with no apparent Rossmann fold organization. In orange and sea green are the beta (β) and alpha (α) sheets respectively.*

Unlike the N7-MTase domain of nsp14, the structure of SARS-CoV 2'O-MTase namely nsp16 reveals the presence of a canonical Rossmann fold[101,115]. interestingly, both proteins interact with the nsp10. However, the consequences of this interaction are very different for each protein complex. The nsp10/nsp14 interaction stimulates ExoN activity (30-fold) but does not regulate N7-MTase activity, whereas the 2'O-MTase activity of nsp16 is regulated by nsp10[101,121]. The latter acts as a molecular switch that stabilizes nsp16 structure into a confirmation prone to specifically recruit RNAs bearing a cap-0[101]. It has recently been shown that a stable and active nsp10/nsp14/nsp16 complex can be formed *in vitro*, and thus capable of ensuring both N7- and

2'O-methylation of cap structures. Finally, although the role of nsp16 in cap methylation is now clearly elucidated, the importance of its catalytic activity is regulated by nsp10 remains mysterious.

10. RNA internal 2'O-methylation: an enigmatic epitranscriptome.

The 2'O-methylation of the nucleotide ribose is not limited to the cap structure, it can also occur inside the RNA chain. Internal epitranscriptomic 2'O-methylation is a ubiquitous modification found in many RNAs. It has first been described on abundant cellular RNA and is considered a common and highly conserved modification found at multiple locations in the tRNA, rRNA, and snRNA[122–124]. With the development of new sensitive detection techniques, the 2'O-mark has recently been evidenced within the mRNA[125] and at the 3′-end of sncRNAs, such as miRNAs and siRNAs[126,127,127]. Additionally, numerous studies have recently highlighted the presence of 2'O-modified nucleotides within vRNA[29–33]. Nevertheless, their biological implication in viral replication remains poorly characterized although increasing evidence of the critical role of such epitranscriptomic modifications was recently published (see below).

10.1. The controversial role of RNA 2'O-methylation in viral replication.

The advent of new high-sensitivity "agnostic" mass spectrometry technologies has recently made it possible to better characterize the epitiranscriptomic modifications present in the genome of various viruses. These methods have shown that the most abundant modifications are the well-described N6-Adenosine methylation, 5-methylcytosine, the pseudouridine, and the 2'O methylation[29–33], etc. RNA 2'O-methylation is thus frequently detected in viral genomes by mass spectrometry technologies, including those of Flaviviruses, SARS-CoV-2, Retroviruses (human immunodeficiency virus type 1 (HIV-1), and murine leukemia virus (MLV)), Hepatitis C virus, and polioviruses[29–33]. Unfortunately, these techniques do not allow for the determination of the exact position of the 2'O-methylations inside the vRNA. This difficulty is now overcome with the advent of the Ribomet-seq technique[128], which allows the identification of abundantly methylated positions within viral genomes. Using such techniques it has been recently demonstrated that the HIV genome is highly 2'O-methylated at 17 specific positions[30]. However, the mechanism governing the setting up of these marks remains poorly characterized for most viruses and the involvement of viral or cellular viral enzymes in such RNA modification is still an open question.

10.1.1. Is viral MTase involved in the internal methylation of their own genome?

As several viruses encode for their own 2'O-MTase supposedly involved in the capping process, one possibility is that the viral 2'O-MTase participates in the internal methylation detected in the viral genome. In this context, the discovery that the purified nonstructural protein 5 (NS5) of West Nile (WNV), dengue (DENV), and Zika (ZIKV) viruses harbor an internal adenosine specific 2'O-methylation within the viral genome[33,129] (Figure 12) -in addition to its MTase activity targeting the vRNA capping structure- is an important observation. Interestingly, in the case of DENV and ZIKV, the NS5-induced internal 2'O-methylations are rather nucleotide-specific than sequence-specific, but this observation remains to be confirmed for WNV. Moreover, the ability of viral MTase to establish internal 2′O-methylation seems not unique to flaviviruses. Indeed, our laboratory recently demonstrated that the MTase domain of Ebola virus protein L also specifically 2′O-methylate internal adenosines[95] (Figure 12). Although there are now several viral enzymes that have been demonstrated to catalyze internal RNA 2'O-methylation, proof of their direct involvement in vRNA methylation during the virus life cycle is still barely documented, and further experiments to demonstrate this are needed. In addition, although an increasing number of viruses have been detected to contain 2'O-methylation inside their RNA, it is noteworthy that the function of these internal modifications is still poorly understood.

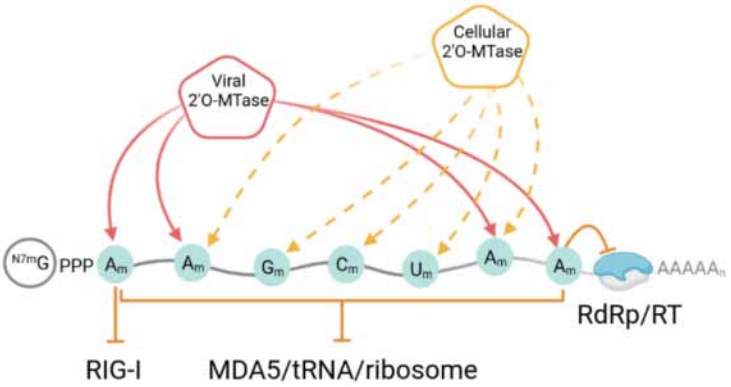

Figure 12: Biological functions of viral RNA 2'O-methylation. *Some cellular RNA 2'O-MTase can be involved in the 2'O-methylation of viral RNA; this can occur on all four nucleotides. However, to date, the identified viral 2'O-MTase only methylates adenosines within the viral genome. The 2'O-methylation of the first nucleotide at the 5'-end of the viral RNA limits innate immune sensing by the cytosolic RNA sensors RIG-I and MDA5 to prevent the expression of IFN. 2'O-methylated nucleotides (N_m) within viral RNA masks the genome from MDA5 recognition, inhibit the elongation of the viral RNA-dependent RNA polymerase (RdRp), retrotranscription by reverse transcriptase (RT), and prevent the interaction with tRNA or ribosome which downregulate translation.*

10.1.2. Evidence that cellular MTase participates in viral internal methylation.

Surprisingly, sensitive "agnostic" mass spectrometry studies have shown that in addition to A_m, DENV, and ZIKV also contain 2'O-methylated cytosine, uridine, and guanosine in their genome[29]. As *in vitro* assays have not evidenced the capacity of the viral 2'O-MTase to target these ribonucleotides, the possible involvement of cellular MTases in the 2'O-methylation process of these viral genomes is an open question (Figure 12). In line with this, several viruses that do not code for any MTase activity were reported to contain 2'O-methylated nucleotides inside their own genome[29–31]. The first direct evidence of the involvement of a cellular MTase in the internal methylation of a viral genome was provided by a study carried out on retroviruses. Members of the *Retroviridae* family such as human immunodeficiency virus type 1 (HIV-1) and murine leukemia virus (MLV), have been first shown to contain internal 2'O-methylated nucleotides[30,31]. Concerning HIV, a proteomic approach has revealed that the nascent proviral RNA recruits through the cellular RNA-binding protein TAR (TRBP) a cellular 2'O-MTase (called FTSJ3) to methylate at least 17 specific residues[30] (Figure 12). Mapping of residues modified by the TRBP/FTSJ3 complex revealed that methylated sites are highly conserved among HIV-1 strains[30]. Functional assessment of these internal methylation shows that they shield HIV RNA from innate immune recognition by the RIG-like receptor differentiation-associated protein 5 (MDA5, see below) (Figure 12). Whether MLV recruits cellular 2'O-MTases like HIV-1 to shield its genome from innate immune sensing remains unclear. However, in the case of MLV, the consequence of internal 2'O-methylation is poorly characterized and is perhaps more complicated than expected. Indeed studies have demonstrated that the reverse transcriptase (RT) activity of MLV is impaired by RNA 2'O-methylation (Figure 12). Indeed biochemical studies have demonstrated that reverse transcriptase of the virus is blocked when it encounters an N_m on the RNA template, suggesting that internal 2'O-methylation present on the viral genome might play an antiviral role. For retroviruses, it is likely that internal 2'O-methylation might have both proviral and antiviral effects, and thus the fine-tuning of these processes probably regulate the viral infection. Apart from retroviruses, the landscape of viruses functionally interacting with cellular 2'O-MTase is poorly described and will probably increase in the future. Several negative-stranded viruses have been reported to be dependent on FBL -a cellular MTase involved in the specific methylation of snoRNA- for efficient replication, particularly Hendra and Nipah viruses, measles virus, mumps virus, VSV,

and RSV[34,35]. In the case of VSV, it is likely that the cellular 2'O-MTase promotes vRNA synthesis[35], however, the molecular basis of this remains unclear and further studies are needed to determine whether the effect on viral replication is the result of FBL induced the direct 2′O-methylation of the vRNAs or if FBL regulate the viral replication by indirect effect. Further work is needed to clarify these questions.

10.1.3. The dual effect of RNA internal 2'O-methylation (pro- and anti-viral effect)?

Although it is now evident that several RNA viruses carry internal 2'O-methylation, the consequence of such modification is still puzzling as pro-viral and anti-viral effects have been linked with such modifications. It is thus likely that the fine-tuning of the methylation process is an important mechanism participating in the control of virus replication and in the host-pathogen interaction. We can postulate that there is a kind of spatial (which RNA and on which position) and temporal (when the methylation occurs during the virus life cycle) regulation of 2'O methylations. In some cases, 2'O-methylation has been reported to stabilize the mRNA and or promote virus replication[30] (e.g., HIV). RNA 2'O-methylation has also been associated with enhanced pathogenicity of diverse viruses. Indeed, Yang et al., (2021) have demonstrated that upon infection, the 2'O-methylation profile of the SARS-CoV-2 genome increases while that of the unrequired host RNA significantly declines[32]. Conversely, RNA 2'O-methylation has also been reported to limit the activity of the DENV viral RdRp[33] or the reverse transcriptase from retroviruses such as ML[130]. Finally, the 2'O-methylation landscape of diverse viruses remains poorly characterized, and the MTase involved in this process remains elusive as well as the outcome of this epitranscriptomic modification on viral replication (figure 12). In the next section, we are going to describe some possible consequences of such vRNA modifications summarized in Figure 12 and illustrated in the recent literature.

10.2. The interplay of internal 2'O-methylation with mRNA translation.

Despite its diverse proviral role, RNA 2'O-methylation has been shown to affect the dynamics of mRNA translation either by interplaying with tRNA/mRNA or mRNA/ribosome interactions. This was first demonstrated by an *in vitro* study showing that the presence of a 2'O-methylated mRNA codon decreases the efficiency of translation, the effect is particularly drastic when this modification is at the second position of the codon[131,132]. The molecular basis of

this inhibition derives from an alteration at the level of decoding of RNA following an excessive rejection of the cognate tRNA during RNA translation into proteins (Figure 12). Therefore, the presence of 2'O-methylations in some mRNAs can selectively decrease translation elongation[133]. These raise the possibility that C/D box snoRNAs guided 2'O-methylation within the coding regions potentially regulates protein expression by interfering with translation.

11. Immunostimulatory RNA orchestrated innate immune activation

Exogenous threats such as bacterial or viral infections are quickly recognized by the host's innate immune system. Loss of homeostatic control of these defense mechanisms can lead to adverse consequences such as fatal infection or autoimmunity. Thus to maintain faithful immune homeostasis, the cellular defense machinery must be able to recognize and eradicate invaders. Mechanisms allowing the discrimination of "self" from "non-self" molecules were selected as key players to limit cellular infection by pathogens in the early phase of infection. The detection is based on sensors expressed in cells and able to recognize either pathogen-associated molecular patterns (PAMPs, ie uncapped mRNA, LPS, glycoconjugates, etc..) or the presence of a molecule in an unexpected subcellular compartment (ie DNA or RNA molecule in endosome). In other words, evolution has selected different kinds of sensors -pathogens recognizing receptors (PRRs)- able to detect PAMPs. These PRRs are located in different subcellular compartments and are able to monitor foreign or abnormal molecules (non-self) from natural components of the host cell (self). The cell holds a myriad of PRRs capable of specifically recognizing different kinds of molecules such as vRNA. To date, the best-characterized families of PRRs involved in detecting exogenous or abnormal nucleic acids are the Toll-like receptors (TLR), the retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs), and the nucleotide-binding oligomerization domain-like receptor (NLRs). Viral genomes constitute an essential class of pathogen-associated molecular patterns (PAMPs) that activate these sensors. The recognition of foreign materials by these different PRRs induces a cascade of intracellular events, leading to the transduction of transcription factors stimulating the production of interferon pathways (those mechanisms are further detailed below). The main innate immune defense mechanism is either through the type I interferon (IFN), interleukin-1 (IL-1)-mediated proinflammatory responses, or the combination of both responses[134,135]. The

expression of type I IFNs and IL-1β rely on the activation of a set of innate immune sensors, known as pattern-recognition receptors (PRRs).

11.1. The Toll-like receptors (TLRs)

The TLRs are one of the best characterized PRRs in mammals. They take their name after the *Drosophila* Toll proteins, which have been primarily found to be involved in the embryogenesis and the antimicrobial response in Drosophila[136]. The TLRs have been classified into thirteen families with respect to their specific PAMPs[137]. With regard to their cellular localization, they are divided into two subfamilies namely the cell surface TLRs (includes TLR1, TLR2, TLR4, TLR5, TLR6, and TLR10) and endosomal intracellular TLRs (consisting of TLR3, TLR7, TLR8, TLR9, TLR11, TLR12, and TLR13). This family of PRRs is mainly expressed on the surface and/or endosomal compartment of innate immune cells such as dendritic cells (DCs) and macrophages. However, some of them can be found in other host cells (e.g., fibroblast cells and epithelial cells)[138]. Despite their vast variety, the TLRs share a common structural organization composed of an extracellular leucine-rich domain (LRR) responsible for the PAMPs recognition, a transmembrane hydrophobic domain, and a cytosolic signaling domain, termed the TIR domain (Toll receptor /interleukin-1)[138–140]. Key sensors in the recognition of non-self RNAs are TLR3, TLR7, TLR8, and TLR13[8–10].

Upon binding of foreign RNA to the LRR, the dedicated TLR undergoes conformational changes to form an active homodimer, which is characterized by an "m"-shaped architecture. This however is not the case for TLR8, which pre-exists as a homodimer in its inactive form[141–143]. Following activation, the dimerized cytoplasmic TIR domains recruit an adaptor protein for intracellular signal transduction. In general, the myeloid differentiation primary response protein 88 (MyD88) is the key adapter used by all TLRs except for TLR3 (Figure 13). Recruitment of MyD88 to the respective TLR triggers the assembly of the "Myddosome", a higher-order protein complex composed of MyD88 and three different IL-1R-associated kinases (IRAK) including IRAK 1, 2, and 4 (Figure 13). The recruitment of tumor necrosis factor receptor-associated factor 6 (TRAF6) by the Myddosome activates the transforming growth factor β-activated kinase 1 (TAK1) which in turn stimulates the mitogen-activated protein kinases (MAPKs), nuclear factor-κB (NF-κB), and IFN-regulatory factor 5 (IRF5)[137,144,145]. Activation of these effectors induces the expression of IFN (Figure 13). On the other hand, the transduction pathway of the cytoplasmic signal emitted by TLR3 occurs differently, first the

adapter protein involved here is the TIR domain-containing adapter inducing interferon-β (TRIF) (Figure 12). The latter can induce IFN expression through two distinct paths: 1) TRIF triggers the TRAF family member-associated NF-κB activator binding kinase 1 (TBK1) and IκB kinase ε (IKKε) to activate IRF3. 2) TRIF associates with TRAF6 through the receptor-interacting protein kinase 1 (RIPK1) to activate NF-κB and MAPK via TAK1[137] (Figure 13). Activated kinases IKKε and TBK1 respectively phosphorylate the transcription factors "Interferon regulatory factors 3 and 7" (IRF3 and IRF7) and nuclear factor κ-B (NF-κB). These factors then migrate to the nucleus to interact with the promoters of the target genes inducing the production of INF-I and pro-inflammatory cytokines (Figure 13).

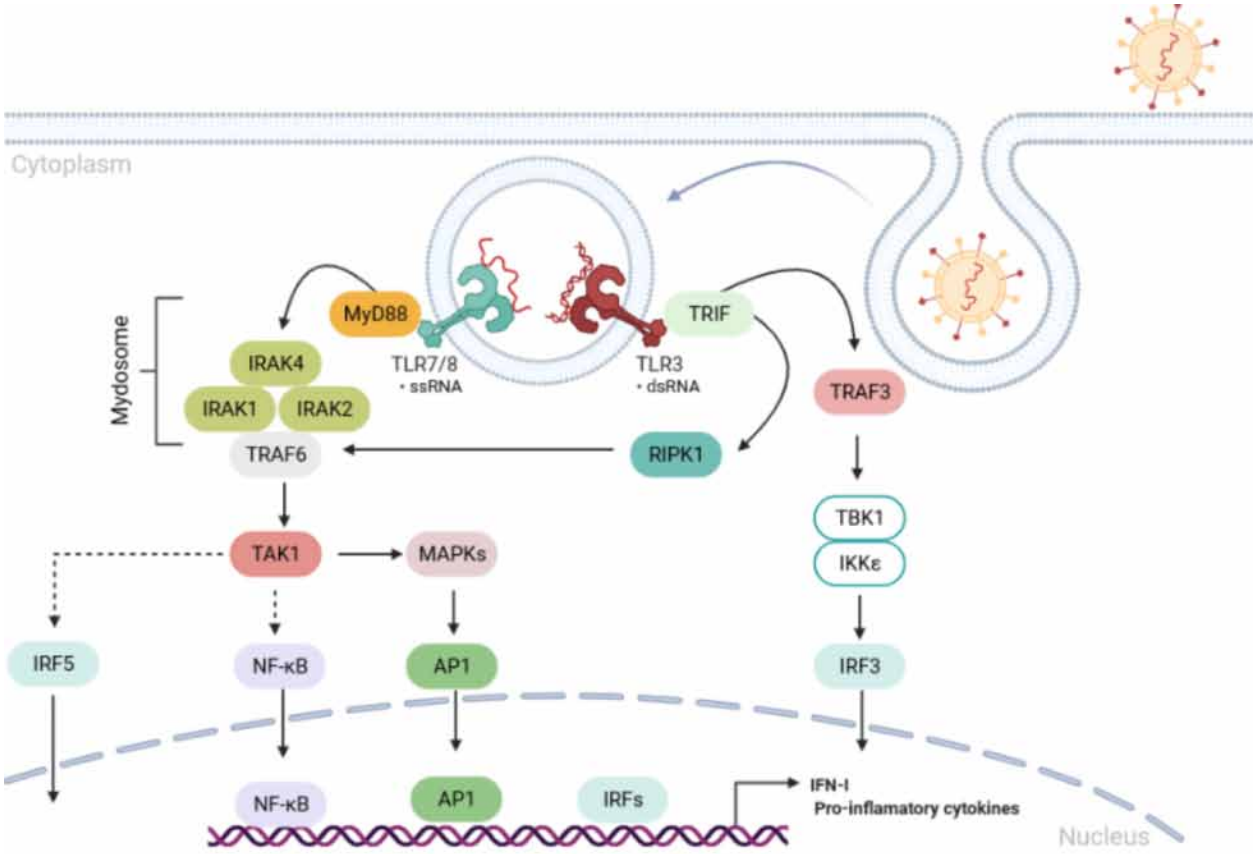

Figure 13: TLR Signaling Pathways. *RNA-sensing TLR3, 7, and 8 are localized to the endosome. TLR3 recognizes dsRNA while TLR7 and TLR8 sense ssRNA. Activated TLR7 and TLR8 engage the adaptor protein MyD88 to form the "Myddosome", which is a protein complex composed of MyD88, IRAK1, IRAK2, and IRAK4. The Myddosome recruits TRAF6 and TAK1 to activate IRF5, MAPKs, and NF-κB. MAPKs further activate AP1. TLR3 triggers through TRIF the activation IRF3 via the TRAF3-TBK1-IKKε axis or MAPKs and NF-κB via the TRAF6-RIPK1-TAK1 axis. Once activated the transcription factors IRF5, IRF5, NF-κB, and AP1 migrate to the nucleus and activate IFN-I and pro-inflammatory cytokines expression.*

11.1.1. RNA features activating TLRs and countermeasures.

The characterization of the different TLRs shows that they recognize their ligands with a well-defined specificity. For instance, TLR3 contains two RNA-binding sites in its LRR domain capable of selectively binding double-stranded RNAs (dsRNA) in a sequence-independent manner[146]. This innate immune sensor has been demonstrated to activate IFN in response to some endocytic viral pathogens including WNV[147], influenza A virus (IAV)[148], poliovirus[149], etc... . So far, the structural elements that activate TLR3 have mainly been characterized for poliovirus. Poliovirus-derived dsRNA harboring internal stem-loop structures has been shown to selectively agonize the TLR3[150].

Despite their strong structural homology and common preference for single-stranded RNA (ssRNA), the TLR7 and TLR8 are functionally distinct[151,152]. Both TLRs contain two ligand-binding sites in their LRR domain, named sites 1 and 2. Site 1 consists of a conserved nucleotide-binding site, with TLR7 having a preference for guanosine (G) while TLR8 for uridine (U)[143,153,154]. As for site 2, it harbors an oligonucleotide-binding activity. TLR8 selectively accommodates a UG dinucleotide in its site 2, however, TLR7 requires a trinucleotide motif containing a U in the second position[143,153,154]. The TLR7 has been shown to induce innate immune responses to several RNA viruses, such as IAV, VSV, HIV-1, and SARS-CoV-2[151,152,155,156]. Various characterization studies have allowed the selection of diverse GU-rich ssRNA fragments of HIV, SARS-CoV-1, and 2 that agonize TLR7 an[89,92,93–91].

Interestingly, the insertion of a 2'O-methylated residue in a short RNA has been shown to strongly limit cytokine and IFN expression due to its antagonist effect on TLR7[159]. From the study done by Ringear et al (2019), it can infer that HIV-1 particles contain 2’O-methylated vRNAs which limit the MDA5 sensing (see below). One can wonder if this PTM could also contribute to limiting vial detection during the internalization stage? Given TLR7 key role in the control of SARS-CoV-2 pathogenesis[160,161], the same question could be addressed for SARS-CoV-2, which has recently been shown to contain numerous 2’O-methylated uridines. Further studies are needed to evaluate whether internal 2′O-methylation can limit vRNA sensing by TLR during virus entry in the host cells.

11.2. RIG-I like receptors: cytosolic sensors of immunogenic RNA

Although the TLRs constitute an effective alert system during pathogen invasion, their contribution to the antimicrobial response remains limited due to their specific expression in a

limited set of immune cells. In addition, TLRs are localized at the extracellular surface or within the endosomal compartments and are thus unable to sense the cytoplasmic or nucleolar nucleic acids of invaders. The detection of cytosolic viral PAMPs is in part covered by the ubiquitous RLRs, which are specialized in RNA sensing. They contain three members: RIG-I, melanoma differentiation-associated protein 5 (MDA5), and laboratory of genetics and physiology 2 (LGP2)[162–164]. All three are NTP-dependent RNA helicases belonging to the SF2 superfamily and are mainly found in the cytoplasm, except for RIG-I which is also present in the nucleus but in a small amount[165]. Both share common domain architecture comprising a central helicase flanked by an N-terminal tandem caspase activation and recruitment domain (CARD) and a C-terminal domain (CTD). Unlike the other RLR, LGP2 lacks the signal transducer domain the N-terminal CARD[166], and was proposed to act as a regulator of RIG-I and MDA5[167].

In the absence of stimulation, RIG-I adopts an autoinhibitory conformation while MDA5 resides in an open and flexible conformation (Figure 14). Upon RNA binding to the CTD and helicase of RIG-I and MDA5, the CARDs domains are released allowing the recruitment of specific protein[168] (Figure 14). The exposed CARDs are subsequently modified by various enzymes (ie E3 ubiquitin ligases and phosphatases) rendering them fully active. Following this, RIG-I and MDA5 are addressed to their common adaptor protein: mitochondrial antiviral signaling protein (MAVS) by the 14-3-3ε and 14-3-3η proteins, respectively (Figure 14). MAVS is mainly found at the outer mitochondrial membrane and has also been localized at mitochondria-associated membranes of the ER and peroxisomes[169–171]. Upon activation, MAVS induces a cascade of signals which stimulates the TRAF3-TBK1-IKKε axis leading to IRF3 or IRF7 activation and expression of type I and III IFNs (Figure 14). Moreover, MAVS signaling also induces pro-inflammatory cytokine expression through the IRF3-NF-κB axis (Figure 14).

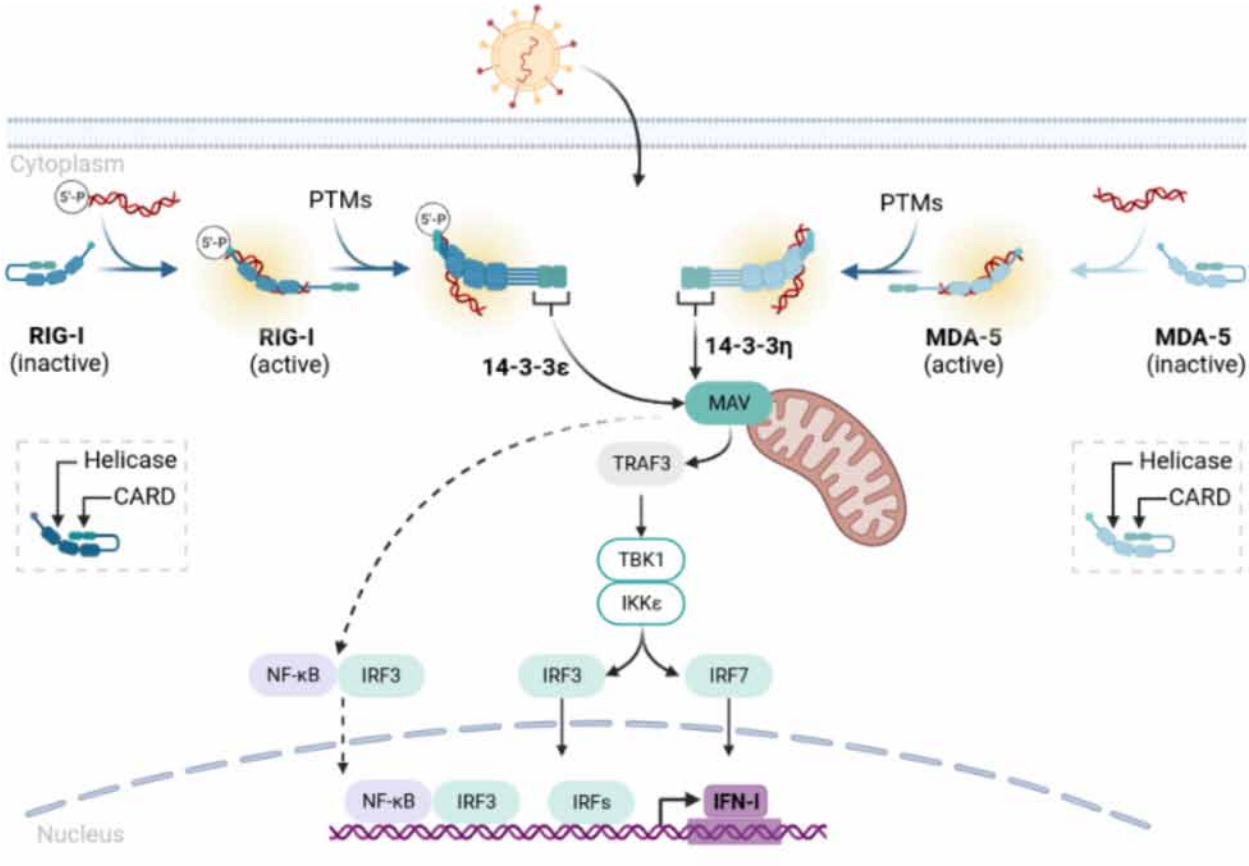

Figure 14: RLR Signaling Pathways. *The cytoplasmic RLR namely RIG-I and MDA5 preferentially recognize 5'-mon, di- or triphosphorylated dsRNA and long dsRNA, respectively. In absence of a ligand binding, RIG-I is held in an auto-repressed and closed conformation, while MDA5 is more open and flexible. Upon RNA-ligand binding to RIG-I and MDA5, both receptors undergo conformational changes that expose their N-terminal CARDs, which are then subjected to posttranslational modifications (PTMs) allowing oligomerization of the RLR along with their ligand. Following this, the CARD domain of RIG-I and MDA5 are subsequently engaged with the chaperon protein 14-3-3ε and 14-3-3η respectively for cytosol-to-mitochondrial translocation. Both RLR bind and activate MAVS on mitochondria, which subsequently activates the transcription factors IRF3 or IRF7 via TRAF3-TBK1-IKKε. Once activated, IRF3 or IRF7 migrate to the nucleus and promote IFN-I expression. The RLR can also induce the expression of pro-inflammatory cytokines via the MAVS-IRF3-NF-κB axis.*

11.2.1. How RIG-like receptors detect RNA and induce interferon.

The activation of RIG-I and MDA5 is done in three main axes: 1) Binding of the ligand, 2) post-translational modifications of the sensors and 3) multimerization of the sensors. The first step of this cascade of events is absolutely key in order to ensure the specific detection of invaders like viruses. RIG-I and MDA5 recognized thus different RNA features and it has been demonstrated that RIG-I recognized specifically the 5' end of RNA and sense among other mis-capped RNA, whereas MDA5 specifically recognized non-self dsRNA[172,173]. The molecular

basis of such specific detection is described below. Various evidence shows that these signal-competent RLRs can act separately or cooperatively to recognize viruses. Negative-strand RNA viruses, particularly members of the Paramyxoviridae, Orthomyxoviridae, Rhabdoviridae, Bunyaviridae, and Filoviridae families are mainly recognized by RIG-I, whereas positive-strand RNA viruses such as Picornaviridae family are mainly sensed by MDA5. Moreover, both RIG-I and MDA5 are involved in detecting members of the Reoviridae and Flaviviridae families, which highlights their collaborative work in inducing an antiviral response[174–177]. Major efforts have been done to characterize ligands that activate the RLRs.

The molecular basis of the RNA sensing was elucidated by combining biochemical assay with structural studies of thées RLRs (Figure 15). RIG-I specifically recognizes blunt ends of dsRNA with a 5'cap-0 structure, 5'di- or triphosphate group, which are commonly found in the (sub)genomes and replication intermediates of RNA viruses but absent in most cellular RNAs (Figure 15). In addition, RIG-I agonistes does not require to be fully complementarity, it can tolerate wobble base pairs, mismatches, and bulge elements like those found in the (sub)genomic panhandle structures of RNA viruses[178,179]. Although ligand specificity by RIG-I is believed to be sequence-independent, some RNA motifs have been proposed to promote binding. Among these: the poly(U/UC) stretch of hepatitis C virus (HCV) 3′UTR[180] and the GA-rich motifs embedded in the multi-branch loop structures of the cellular lncRNA lnc-Lsm3b[181], and AU-rich motifs found in the measles virus, Sendai virus, and Kaposi's-sarcoma-associated herpesvirus derived RNA transcripts[182–184]. Strikingly, RIG-I can also bind RNase L-cleaved cellular RNA bearing 5′-OH and 3′-monophosphates and exogenous circular RNA[185,186]. The structural basis for RIG-I activation by these unusual ligands needs further investigation. Structural analysis of RIG-I unveils a sensor (residue H830) for the 2′OH group of 5'pppN and the cap-0 structure. The binding assay shows insertion N_m at the first position of 5'pppRNA decreases RIG-I affinity (20-fold), the effect is the more drastic for an RNA with a cap-1 structure (200-fold). The characterization of the molecular basis of this inhibition shows a steric clash between the RIG-I H830 and the 2'O-methyl group of the N_1 of the transcript. Thus, the 2'O-methylation of the cap-0 is assigned as a self hallmark that prevents RIG-I recognition. Interestingly, many viruses adopt the cap-1 structure on the 5'-end of their RNA and therefore limit their detection by RIG-I (Figure 15). However, during the replication process, these viruses synthesize minus-strand intermediate non-capped RNA and can thus be detected by such receptors[51].

In contrast, ligand feature requirements for MDA5 sensing are vastly less understood,

even the structure of MDA5 reveals that it forms a kind of donut-like structure able to accommodate dsRNA[187] (Figure 15). MDA5 is preferentially activated by the presence of a long dsRNA structure (> 300 bp). Such dsRNA are often present in untranslated regions of the viral genome forming hairpin structures or are formed during the replication of the genome that needs the synthesis of the complementary strand of the genomic RNA[174,188]. Furthermore, it seems that the detection of long double-stranded RNA structure results from the formation of MDA5 concatemers covering the RNAs, necessary for the activation of the signal transduction. It has been recently suggested that RNA 2'O-methylation limits the MDA5 sensing[30], however, the molecular basis for this is not completely elucidated. It can infer from the structure of MDA5 in complex RNA that internal methylation creates a steric clash with MDA5 thereby abolishing RNA recognition.

Once bound to the RNA ligand, the CARD of RIG-I and MDA5 are dephosphorylated by PP1α or PP1γ respectively[189]. RIG-I dephosphorylation induces K63-linked polyubiquitination, which is required to promote its "head to tail" ATP-dependant oligomerization[190] (Figure 15). RIG-I undergoes covalent K63-linked ubiquitination at the CARDs and CTD[191–193] while that of MDA5 occurs only in its helicase domain[194].

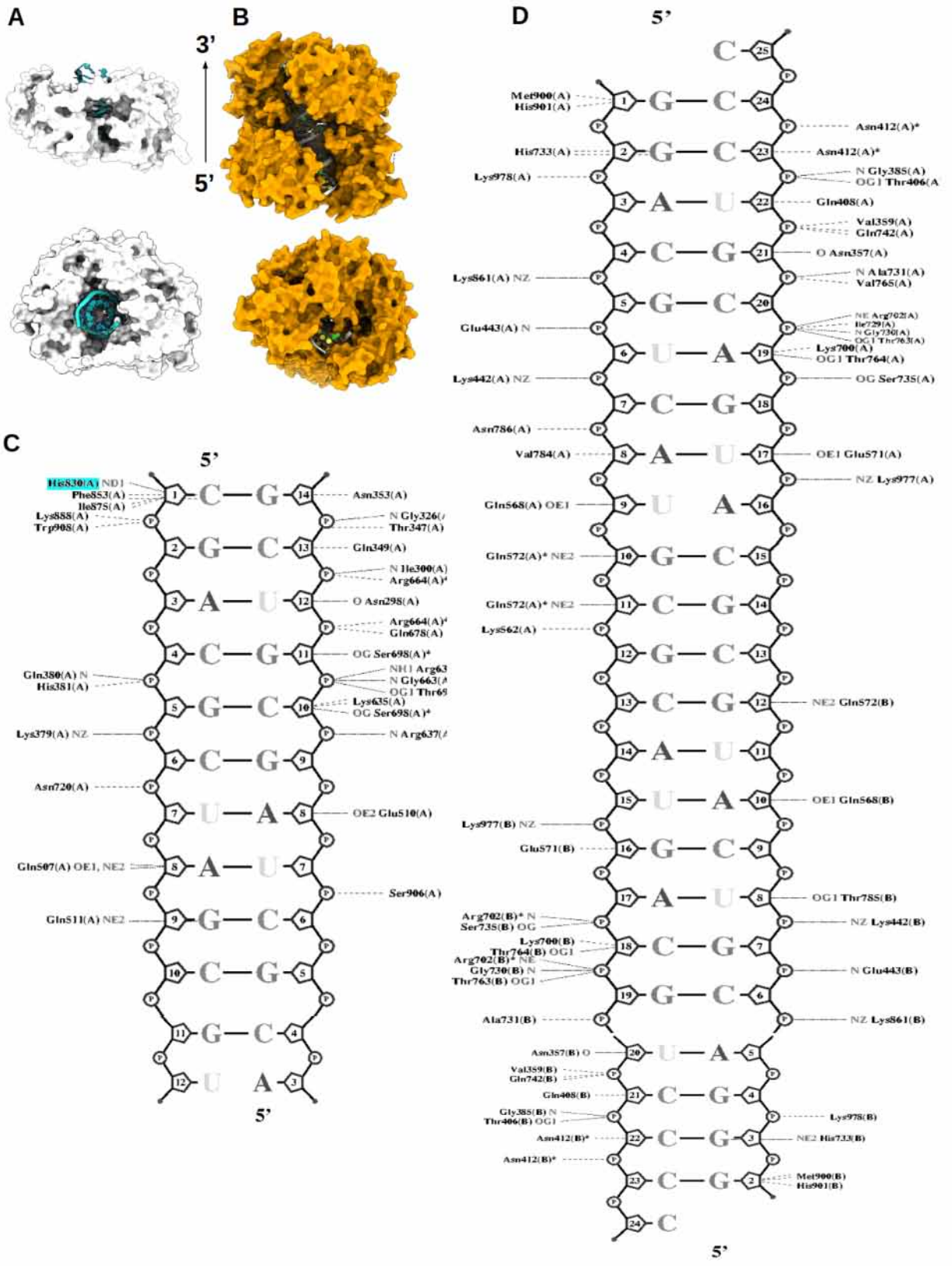

Figure 15: Structural analysis of RLRs complexed with an RNA. *A) Surface depiction of RIG-I (off-white) in complex with a dsRNA (cyan, PDB) . Once bound to its ligand and activated, RIG-I forms a filament along its substrate from the 5 'to the 3'-extremity characterized by a head to tail multimerization. B) Surface depiction of*

MDA5 (orange) in complex with a dsRNA (off-white, PDB: 5E3H). MDA5 forms a donut-like filament along its substrate from the 5 'to the 3'-extremity. C) Nucplot of RIG-I structure. The plot represents the different interactions taking place between RIG-I and its RNA ligand. Highlighted in cyan the sensor for the cap-1 2'O-methyl group. D) Nucplot of MDA5 structure. The plot represents the different interactions taking place between MDA5 and its dedicated RNA ligand.

11.2.2. LGP2 function

LGP2 has been shown to be a major regulator of RLR signaling. It can induce negative feedback on RIG-I signaling, through various mechanisms that include: 1) competition with RNA binding, 2) blockage of RIG-I multimerization 3) inhibition of RIG-I ubiquitination and 4) competition with IKKε for MAVS interaction[166,195,196]. In contrast, LGP2 promotes MDA5-mediated antiviral signals. Indeed, LGP2 depleted mice were shown to be more susceptible to some RNA viruses, in particular, EMCV which is mainly recognized by MDA5[197,198]. Digging into the mechanistic, unveiled that LGP2 first associate with MDA5-stimulatory RNA and then promotes MDA5 nucleation and oligomerization through a cooperative interaction[167,199].

11.3. NLRs and their involvement in innate immune signaling.

NLRs constitute the third large family of intracellular PRRs capable of triggering immune responses to a variety of microbe and exogenous or cell-derived danger signals[200]. They are found in lymphocytes, macrophages, dendritic and non-immune cells, for instance in epithelium[201]. Moreover, the NLRs are characterized by their NOD domain which mediates dNTPase activity and the multimerization of the NLR protein. Moreover, the NLRs display an N-terminal effector domain and a C-terminal LRR domain[202]. The architectural organization of NLRs includes, 1) an N-terminal effector domain consisting of protein-protein interaction domains such as the caspase recruitment domain (CARD), pyrin (PYD) and the repeat domain of baculovirus inhibitor (BIR), 2) a central NOD domain, and finally, 3) a C-terminal leucine-rich repeat (LRR) motif that detects a set of conserved microbial patterns and regulates NLR activity. Similar to RLRs, NLRs are present in monomeric states, the inactive conformation is maintained by an intramolecular interaction between the LRR and NOD domains preventing spontaneous NLR oligomerization and signaling. Recognition of ligands induces conformational changes which result in NLR multimerization and exposure of effector domains for specific partner scaffolds[202]. Based on their N-terminal domain, the NLRs are classified into 1) CARD-containing

NLRs: NODs, 2) PYD-containing NLRs: NALPs, and 3) BIR-containing NLRs: NAIPs. NLR activation is proposed to occur in the following manner: 1) recognition of the PAMPs by C-terminal LRRs, 2) triggers conformational changes resulting in sensor oligomerization via its NOD domain. 3) the exposed NLR effector domains in turn induce the scaffold and activation of other effector molecules (containing CARD and PYD domains) thus promoting signal transduction[202]. Given the diversity of NLR effector domains, these are involved in the activation of various signaling pathways. Most of the 22 known human NLRs function as a scaffold for the assembly of inflammasomes, which are composed of large multiprotein complexes that induce the caspase-1 and subsequent maturation of the inflammatory cytokines IL-1β and IL-18. While these latter induce the inflammatory cascade (resulting in the production of type II IFN (IFN-II)), caspase 1 is said to trigger at the same time the programmed lytic cell death (pyroptosis). On the other hand, some NLRs have been shown to induce IFN-I expression without however triggering the caspase pathway; among these, we can find NOD2 (alias: NLRC2). This NRL has been shown to recognize viral ssRNA upon RSV infection and mediate IFN-I antiviral responses via the MAVS-IRF3 axis[203]. NOD2-depleted transgenic mice were deficient in IFN-I responses and more susceptible to RSV pathogenesis[203]. Recent evidence shows that some NLRs cooperate with specific DEAH-box RNA helicases (DHX) to recognize viral dsRNA and then trigger inflammasome- and/or IFN-I-dependent antiviral responses. This has been described for DHX15, which sensitizes NLRP6 to recognize dsRNAs during EMCV infection and subsequently induces IFN-I and III expressions via MAVS[204]. The formation of the viral NLRP6/DHX15/dsRNA complex was described by an immunoprecipitation study combined with mass spectrometry. The intimate cooperation between NLRP6 and DHX15 has further been substantiated by the fact that DHX15 knockout greatly decreases NLRP6 affinity for vRNA[204]. Unfortunately, RNA features that activate the NLR are largely unknown, and whether the epitranscriptome influences its function remains to be determined.

11.4. Emerging RNA sensors that activate IFN-I

Recent advances highlight the presence of non-canonical RNA sensors. Indeed, various helicases involved in RNA metabolism have emerged to induce innate immunity upon vRNA recognition. Among which is the zinc finger nuclear transcription factor, X-Box Binding 1-type containing 1 (ZNFX1) an SF1 helicase localized at the mitochondrial outer membrane that inhibits the replication of VSV and other viruses[205]. Depiction of the antiviral mode of action shows that ZNFX1 binds VSV and poly (I:C) RNA and triggers the IFN-I response via MAV[205].

Besides the classical RLR sensors, many DEA(D/H)-box helicases of the SF2 superfamily have emerged to be involved in non-self RNA recognition and subsequent innate immune signaling. These non-classical sensors can either be the main actor in RNA detection or serve as a potentiator. Among the helicases that directly recognize non-self RNA are DHX9, DHX33, DDX1, DHX15, etc….

In response to poly(I:C), DHX9 and DHX33 have been found to individually mediate IFN induction via the MAV pathway in myeloid DCs[206,207]. This innate immunity sensor function was then validated for each of these helicases in a viral context (Figure 16). Moreover, the DEAH-box protein, DHX15 has also been identified as a poly(I:C)-binding protein and mediates IFN-I and cytokine responses to poly(I:C) or RNA virus infection in myeloid DCs. DHX15 interacts with MAVS which promotes IRF3, MAPKs, and NF-κB activation (Figure 16). RNA recognition by DHX15 is independent of its ATPase activity and is potentiated by DHX9[208].

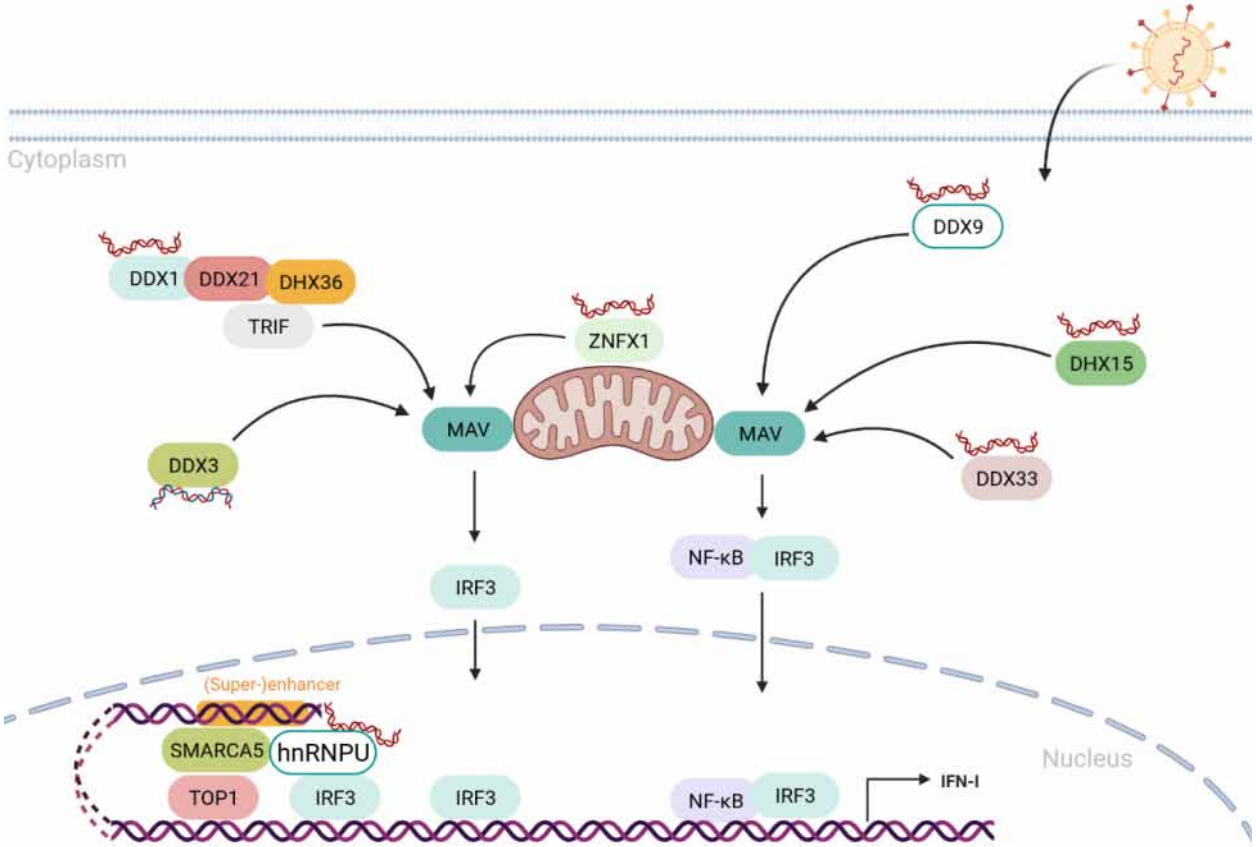

Figure 16: Emerging RNA sensors signaling pathway. *The mitochondrial helicase ZNFX1 induces MAVS-dependent IFN-I expression in response to non-self dsRNA binding. Some DEA(D/H)-box helicases have also been shown to be involved in non-self dsRNA sensing and triggering of IFN-I response. DDX9, DHX15, and DHX33 induce IFN-I and proinflammatory stimulation in response to dsRNA in myeloid DC and via the MAVS-IRF3-NF-κB axis. The DEA(D/H)-box helicases can also cooperate among themselves to induce innate immune response as exemplified by DDX1, which form complex with DDX21 and DHX36 upon dsRNA recognition to induce IFN-I and*

proinflammatory responses via TRIF, MAVS, and IRF3. The DDX3 directly binds to the 3'-end abortive cellular and HIV-1 dsRNA to induce MAVS-dependent IFN-I responses. The nucleus has recently been shown to contain an innate immune RNA sensor namely HnRNPU, which upon activation associates with SMARCA5 and TOP1 to stimulate the (super-)enhancer that drives the transcription of IFN-I and proinflammatory genes.

The cooperation of various helicases notably that of DDX1, DDX21, and DHX36 allows sensing of non-self dsRNA in the cytosol of myeloid DCs. In this complex, DDX1 is the sensor that recognizes foreign dsRNA, while DDX21 and DHX36 recruit TRIF for IFN-I and cytokine responses. Knockdown of any component of this pathway blocks IFN-I and cytokine responses to poly I:C, IAV, and reovirus[209] (Figure 16).

Most focuses were on sensors that discriminate RNA with an unusual 5'-end structure. A remarkable recent study highlights the presence of a dsRNA 3'-end sensor. Indeed, DDX3 has been shown to interact with cellular and abortive HIV-1 RNA lacking a 3'-end poly(A) tail in monocyte-derived DCs and macrophages. Upon this recognition, the DEAD-box helicase triggers IFN-I responses in a MAVS- and TRAF3-dependent manner[210] (Figure 16).

To our knowledge, all known RNA sensors are mainly extranuclear. However, a recent study highlights the role of heterogeneous nuclear ribonucleoprotein U (hnRNPU, alias: SAFA) in maintaining nuclear self RNA homeostasis. The hnRNPU is normally involved in many cellular processes including the organization of nuclear chromatin, regulation of telomere length, transcription, alternative splicing, and mRNA stability. Its ability to discriminate HSV-1-derived dsRNA in the nucleus and the subsequent induction of IFN-I assign it a nuclear sensor function[211]. To activate the IFN-I expression and response, hnRNPU uses a fairly sophisticated signaling pathway from that of extranuclear PRRs. Upon sensing viral RNA, hnRNPU oligomerizes and initiates the formation of enhanceosome at the distal enhancer of IFN. The enhanceosome consists of oligomerized hnRNPU, SMARCA5 (SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 5), and TOP1 (DNA topoisomerase 1), two components of the SWI/SNF (SWItch/Sucrose Non-Fermentable) nucleosome remodeling complex that activate IFN-I distal enhancers and super-enhancers. Moreover, recruitment of IRF3 and 7 by hnRNPU are crucial for IFN expression[211] (Figure 16).

The ability of all these sensors to discriminate self from non-self indicates that there are some molecular signatures that direct this process such as epitranscriptomic marks and this needs to be further elaborated in future studies. Interestingly, some of these helicases interact and assist cellular MTases during the RNA 2'O-methylation process. Such interaction might interplay with the RNA sensing process, as it interferes with does with RLRs recognition. One example is the helicase DHX15 has been shown to be CMTR1 partner during the capping process[212]. Strikingly,

two studies show contradictory effects of DHX15 on CMTR1 2'O-methylation activity[212,213]. However, the in-depth analysis of the data presented in both papers suggests that this cap-0 2'O-MTase cooperates with DHX15 to induce internal methylations which is a still undemonstrated function of CMTR1! That needs to be addressed. Moreover, DHX9 has been shown to assist FTSJ3 during the methylation of HIV-1 RNA, in this case, DHX9 rather plays a proviral role during infection[30]. These open up numerous questions, such as whether the non-canonical RNA sensors interact with cellular 2'O-MTases? can they recruit cellular 2'O-MTases and promote vRNA methylation? And whether the outcome of these 2'O-methylations of the viral genome has a pro- or antiviral effect?

12. Interferon Induced genes

In the late 50s, Isaacs and Lindenmann discovered the presence of a protective substance capable of interfering with the spread of inactive influenza virus, which was subsequently named "interferon" (IFN)[214]. Since that discovery, many IFNs have been identified and are now classified into three distinct families: IFN-I, II, and III. Types I and III are involved in antiviral innate immunity in most cells of the body, while type II plays a major role as a communication molecule between specialized cells of the immune system. The IFN-I are ubiquitous and conserved in vertebrate cytokines. They are divided into several classes, including IFN-α, IFN-β, IFN-κ, IFN-ε, and IFN-ω, with IFN-α/β being the major key players. These cytokines are structurally and genetically related given their common ancestor[215–217]. There are 13 different IFN-α genes in humans, while all the others are encoded by a single gene. The IFN-I do not have an antiviral property, but they rather are mediators that bind to interferon receptors (INFNAR) which transmit intracellular signals to induce the expression of antiviral effectors[216]. Thus, when an infected cell produces interferon, the proinflammatory cytokine binds to the IFNAR leading to the induction of hundreds of interferon-stimulated genes (ISG) involved in establishing an antiviral state. The response to interferon is autocrine and paracrine as ISGs are induced in the interferon-producing cell as well as in nearby cells following the diffusion of interferon around the infection site. The interferons also display antiproliferative, pro-apoptotic properties and are likely to orchestrate the adaptive immune response. They are involved in the maturation of DC, the cytotoxicity of natural killer cells, the differentiation of cytotoxic T lymphocytes, and the production of high antibody titers during viral infections[218]. Due to all these properties, IFNs are considered crucial cytokines to circumvent viral spread early in infection and to achieve the

switch between innate and adaptive host responses. The IFN signaling pathway consists thus of two distinct steps: 1) an induction phase triggered by the detection of viral molecules by specialized cellular sensors leading to the first wave of IFN synthesis, followed by 2) a response activated upon binding of IFN to its cognate receptor (IFNAR) and characterized by the expression of hundreds of ISG.

Upon viral PAMPs recognition by innate immune sensors and subsequent IFN-I expression, the latter act in a paracrine and autocrine way to activate the IFNAR. This promotes the recruitment of tyrosine kinases, namely Janus kinases (JAK1) and Tyrosine kinase 2 (TYK2) (Figure 17). Following this, the kinases phosphorylate the signal and activator transducers in transcription factors 1 and 2 (STAT1 and STAT2), which in turn form a heterodimer and associate with the IRF9 to generate the interferon-stimulated gene factor 3 (ISGF3) complex (Figure 17). The ISGF3 complex is then translocated into the nucleus, where it binds the Interferon sensitive response elements (ISRE), and stimulates the transcription and expression of IFN-stimulated genes (ISGs) (Figure 17). These proteins orchestrate the early response against different pathogens including viruses. Briefly among the hundred or so genes induced by interferon, there are proteins capable of limiting viral replication by acting on different cellular levers: Certain innate immunity sensors are overexpressed in order to optimize pathogen detection (RIG-I, MDA5), and proteins acting on major cellular metabolic pathways are produced in order to control RNA decay (2-5-oligoadenylate synthetase (OAS), protein kinase R (PKR), ribonuclease L (RNase), ISG20), translation (IFN-induced protein with tetratricopeptide repeats, (IFIT*)), or pyroptosis, and restriction factors (ie APOBEC family members viperin, tethering, Trim5 alpha, MxA) specifically targeting viral replication steps are produced.

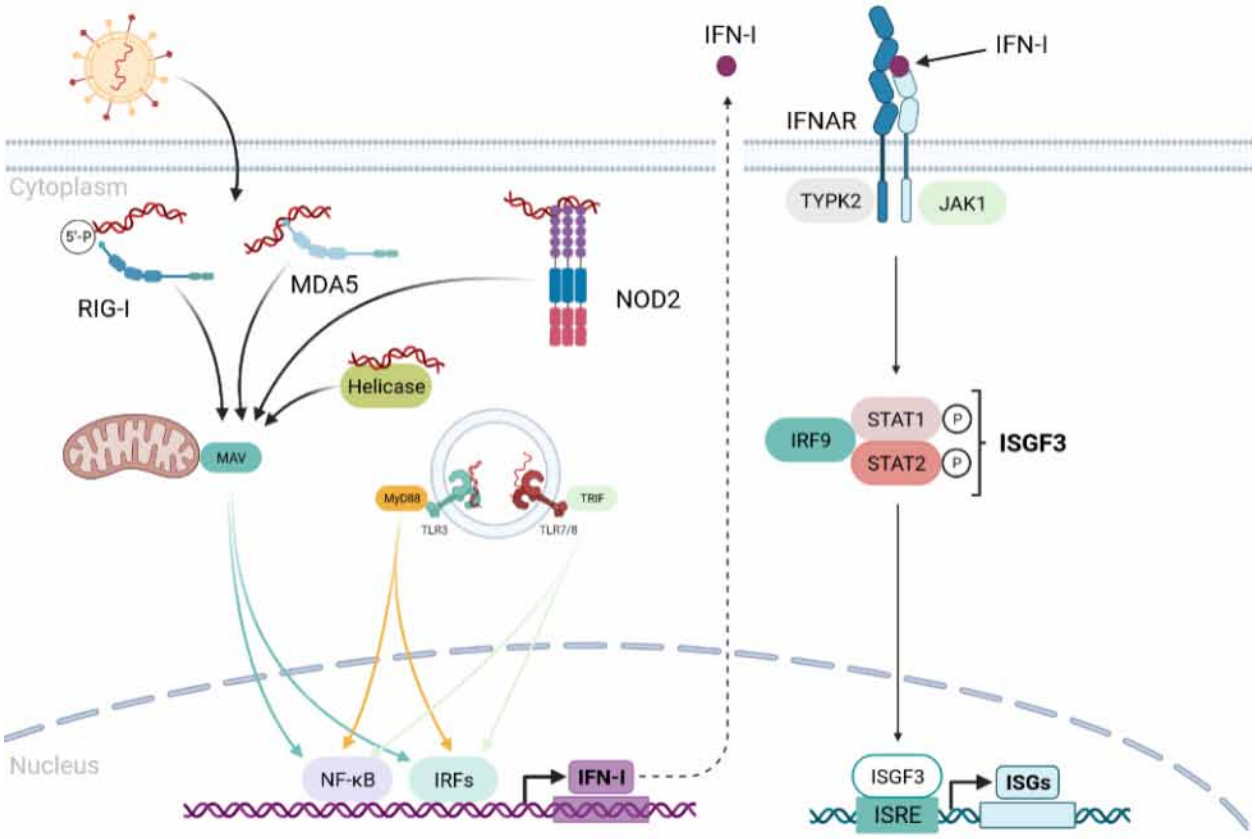

Figure 17: Expression of ISGs through the JAK and STAT pathway. IFN-I induced by RLR (RIG-I and MDA5), TLR (TLR3, 7, and 8), NLR (NOD2) or non-canonical RNA sensor (Helicase) upon specific RNA ligand recognition binds to IFNAR and activate the TYPK2 and JAK which in turn phosphorylate STAT1 and 2. Upon activation, STAT1 and 2 interact with IRF9 to form the ISGF3 complex, which then migrates to the nucleus and upregulates the expression of many ISGs.

*The IFITs are a family of IRF-inducible genes involved in many antiviral processes against several RNA viruses. Thus many viruses can escape the IFITs antiviral action either by suppressing its expression or protecting their own RNA with a cap structure

It has been demonstrated that the antiviral activity of some ISGs can be controlled by RNA 2'O-methylation. Indeed, the 2'O methylation of the VRNA cap structure was evidence to limit RIG-I sensing[50] and also limit the capture of vRNA by IFIT members[49], and the sensing of vRNA by MDA5 was reported to depend on the vRNA internal 2'O-methylation level[30]. In this work, we address the possibility that the antiviral activity of ISG20: an Exonuclease induced by interferon, is overcome by the internal methylation of vRNA. We thus described in more detail ISG20 in the next paragraph.

12.1. The interferon-stimulated gene 20 (ISG20)

The interferon-stimulated gene 20 kDa protein (ISG20), is the second known human RNase to be induced by the IFN in response to viral invasion[219]. It was first discovered in the late 90s in Daudi cells as a new ISG[220]. ISG20 expression can be upregulated by type I (IFN-α/β) and type II (IFN-γ) IFNs[221,222]. However, in absence of IFN stimulation, the transcription factors specificity protein 1 (SP-1) or upstream stimulating factor 1 (USF-1) ensure a constitutive and basal expression of ISG20, thus allowing the protein to perform its physiological function, among which maintaining homeostasis of self RNAs[221]. Interestingly, ISG20 was also identified in the late 90s as a human estrogen-regulated transcript (HEM45) in breast cancer cell lines, thus HEM45 was assigned as a second name[223]. Adding to this, ISG20 has been linked to numerous tumorigenic processes and was highlighted as a potential biomarker, however, the molecular basis of ISG20 involvement in this process is yet to be described[82,222,224–226]. Due to its small molecular size, ISG20 is quite dispersed in the cell. However, a study specifically addressed to its distribution reveals that it is quite concentrated in the nucleoli and Cajal bodies of promyelocytic leukemia (PML), which are well-known intracellular factories for RNA processing. Interestingly, its identification in the Cajal body suggests that its activity is not limited to an immune response, it could be involved in other cellular functions such as biogenesis and/or snRNA and rRNA maturation.

12.2. Structural and functional insight into ISG20 exonuclease activity.

ISG20 belongs to the DEDDh superfamily and members of this subgroup share high structural homology, notably at the level of the exonuclease (ExoN) domain. Within which the similarities are distributed over the Exo I, Exo II, and Exo III motifs defined by four conserved amino acids: three aspartates (D), one glutamate (E), and a histidine residue (h). Despite their highly conserved ExoN domain, DEDDh exonucleases can harbor RNase or/and DNase activity depending on their affinity, thus highlighting divergent and not necessarily conserved substrate-binding sites[227,228].

Structural analysis of ISG20 in complex with a uridine 5'-monophosphate (UMP), reveals an ExoN domain quite similar to other members of the DEDDh superfamily, notably that of exonuclease I and the ε DNA subunit polymerase III and exonuclease X (ExoX) from *Escherichia coli* and *Arabidopsis* small RNA degrading nucleases 1 (SDN1), suggesting that they may follow the same catalytic mechanism[229,230] (Figure 18). Moreover, the ISG20 crystal

structure shows that within its active site, the ExoN domain motifs including D11, E13, D94, and D154 coordinate two manganese (Mn) ions[229]. Which is reminiscent of its activity[231]. In contrast, the conserved H149 of the Exo III domain interacts with the phosphate group of the UMP, suggesting that its the key residue that directs the nucleophilic attack of the phosphodiester bridge during substrate decay[229,232]. Moreover, the structure also shows that the Methionine 14 (M14) and Arginine 53 (R53) residues are involved in a hydrogen bonding with the 2′OH of the UMP ribose underlining ISG20 affinity for RNA[229,231] (Figure 18). Although the 3D structure of ISG20 gives crucial information on the catalytic and interaction mechanism of a nucleotide in the catalytic pocket, it fails to give more details on the substrate-binding domain.

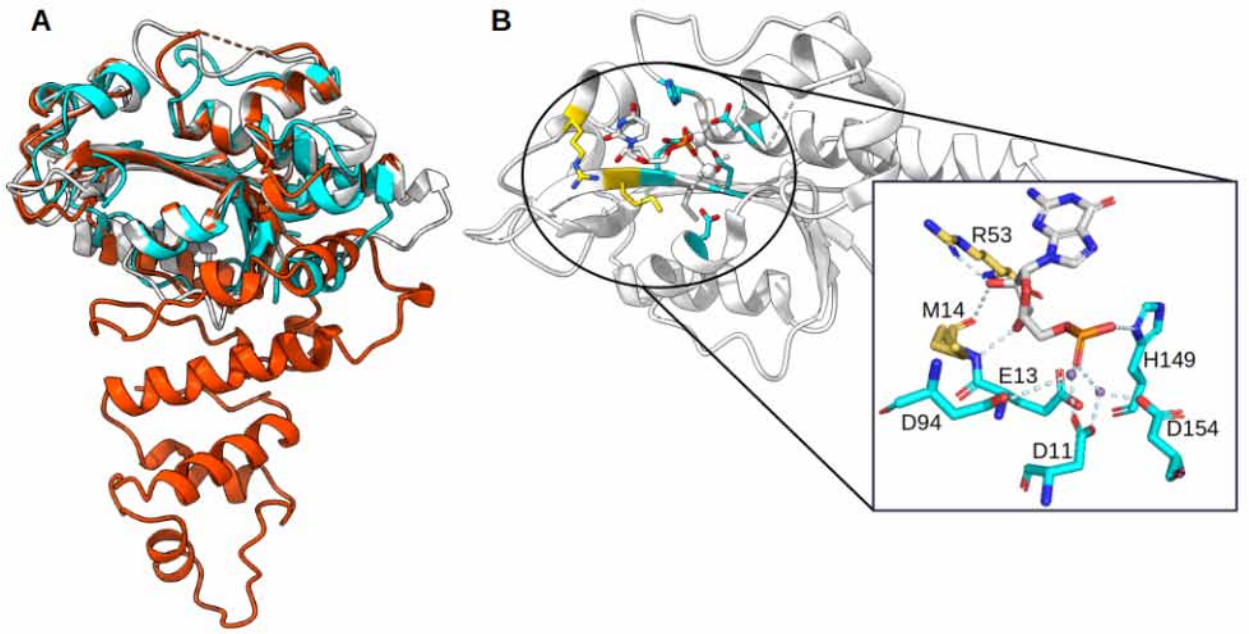

Figure 18: Structural analysis of ISG20. *A) Superimposition of ISG20 (PDB: 1WLJ, in off-white), SDN1 in complex (PDB: 5Z9X, orange), and ExoX (PDB: 4FZX, in cyan) represented in ribbons showing the high coverage of the catalytic pocket of the three homologs. B) Ribbon depiction of ISG20 complexed to a UMP (PDB: 1WLJ, in off-white) Highlighted in yellow are the residues that engage with the 2'OH of the ribose of the UMP and in cyan the conserved residues of the DEDD exonuclease domain. Zoom on the ISG20 catalytic pocket showing the interaction (blue) between the residues of the ExoN motifs domain (cyan), manganese (purple), and the UMP (grey).*

Regarding ISG20 enzymatic mechanism, it displays a 3' to 5' non-specific ssRNA decay. It can also degrade RNA substrates with a stem-loop structure at the 3′end, with, however, a lower efficiency than ssRNA implying that ISG20 operates poorly in the presence of a secondary structure[231,233]. Strikingly, ISG20 has been reported to harbor a weak DNAse activity however, recent evidence highlights its efficient ExoN activity on deoxyuridine-containing ssDNA. The latter is the result of the nucleotide editing done by a specific antiviral factor, the human apolipoprotein B mRNA-editing catalytic polypeptide-like 3A protein (APOBEC3A) that

deaminates cytosines in viral genomes[234]. Perhaps some structural information of ISG20 in complex with DNA and/or deoxyuridine containing ssDNA could shed light on the molecular mechanism for this difference in affinity between these two types of DNA. The fact that the substitution of a single conserved aspartic acid (D) by an uncharged residue (alanine or glycine) is sufficient to abolish its ExoN activity demonstrates that ISG20 is functionally related to the DEDD superfamily[227,228]. By analyzing more closely ISG20 structure, it appears to be solely composed of a unique ExoN domain without any apparent regulatory domain. This suggests that regulation of its activity may require interactions with cellular partners leading to local and specific activation and preventing cell toxicity. Another possibility is that the epitranscriptomic mark present on the self RNAs would limit off-target effects or its subcellular localization.

12.3. ISG20: a multitasking antiviral

ISG20 is an antiviral factor that inhibits the replication of a broad range of RNA viruses including Flaviviruses (Yellow fever, West Nile, DENV, bovine viral diarrhea virus, hepatitis C)[235], Togaviruses (Sindbis, Chikungunya, Venezuelan equine encephalitis viruses), Rhabdoviruses (vesicular stomatitis virus), Orthomyxoviruses (influenza virus), Bunyaviruses, Hepadnaviruses (hepatitis B)[236] and Retroviruses (HIV-1)[237]. Due to its 3' ExoN activity, ISG20 has long been thought to restrict viral propagation via vRNA decay[235–237], however, recent evidence, highlights that for some viruses it mediates antiviral activity resulting from a translational inhibition[6,238]. The molecular mechanism of the switch between its two modes of action for a proper antiviral response remains unknown and opens up some interesting perspectives for future discoveries. It is noteworthy that the human ISG20 exhibits two cellular isoforms, namely ISG20L1 and ISG20L2. Despite their sequence similarity, to date, none of the isoforms have been proposed to display exhibit any antiviral activity[239,240].

Given that ISG20 does not have any apparent regulatory domains, it has long been questioned how it selectively spares self RNA from decay? A captivating recent study shows that ISG20 has been shown to borrow the ^{N6m}A regulatory pathway (previously described) for antiviral activity in the context of HBV infection[241]. The HBV transcripts actually contain two ^{N6m}A marks specifically distributed on 5'- and 3'-end epsilon stem-loop structure (ϵ)[241]. This epitranscriptomic mark is a critical signal for viral packaging. ISG20 was first shown to specifically target HBV ϵ stem-loop and inhibit viral replication by RNA decay[233]. However, the molecular mechanism of this recognition has just been elucidated, it appears that the epitranscriptomic ^{N6m}A mark reader, particularly YTHDF2 (promote RNA destabilization and

decay) recruits ISG20 for antiviral activity. The interaction between YTHDF2 and ISG20 occurs in an HBV-independent manner and depletion of this ^{N6m}A reader abolishes IFN-dependent HBV RNA decay[241]. Therefore confirms that ISG20-directed antiviral activity can be regulated by a co-factor, of which to date only YTHDF2 is identified.

Two different studies have highlighted ISG20 alternative antiviral mode of action, notably through the downregulation of viral protein translation, however, both studies propose different mechanisms[6,238]. The first shows that ISG20 mediates an indirect translational inhibition by upregulating the expression of IFIT1. this was confirmed by showing that an IFIT1-susceptible alphavirus mutant displayed the same virulence profile as WT in ISG20 deficient mice[238]. In contrast, the second study shows that ISG20 selectively targets non-self vRNA for inhibition of viral protein translation in an IFIT1-independent manner. Through a refined study, the authors show that ISG20 inhibits the translation of any exogenous DNA of viral origin and the host. However, when a CMV-GFP cassette is inserted into the host genome by a retroviral vector, ISG20 fails to restrict GFP expression implying that the transcript produced by host cells bears molecular signatures that allows ISG20 to distinguish self from non-self RNA. Although the molecular basis for translation inhibition is still elusive, the authors show that upon VSV infection, ISG20 addresses the viral genome to the PB, where RNA translation is repressed[6].

12.4. The putative interplay of RNA 2'O-methylation with ISG20 activity

As previously mentioned, ISG20 belongs to the large DEDD superfamily. The latter includes numerous 3'-end RNA trimming enzymes supporting various biological functions such as RNA proofreading functions, maturation, and turnover complexes[242]. RNA 2'O-methylation processes have been proven to play a key role in the stability of miRNAs, siRNA, and piRNAs and notably by protecting these RNA transcripts from decay. This has well been established in the case of *Arabidopsis,* which encodes for a 2'O-MTase (the small RNA MTase HEN1) responsible for 3'-end methylation of miRNAs and siRNA, mutation of this MTase has been proven to promote RNA decay by the SDN1 and 2[243]. Biochemistry assays showed that SDN1 specifically targets small ssRNA, and its activity is impeded by the 2′O-methyl mark on the 3′-termini of small RNAs[243]. Given the high structural coverage between SDN1 and ISG20 along with their functional similarities, it is conceivable that RNA 2'O-methylation also affects the function of this IFN-induced restriction factor.

Aim of the thesis

Post-transcriptional modifications such as RNA methylations and particularly the N7-methylation of the cap and RNA 2'O-methylation are important regulators of the structure and biological function of the cellular and viral transcripts. RNA 2'O-methylation is now considered a "self" marker that allows the discrimination between cellular and viral RNA by RIG-I-like receptors which regulates IFN production. Besides the selective detection by RIG-I-like receptors, 2'O methylations may play an additional role in the intracellular survival of viruses. In chapter 2, we addressed the possibility that RNA 2'O-methylation affects the anti-viral activity of ISG20.

Study approaches:

In vitro:

- Produce and purify WT and mutants recombinant ISG20 for characterization.
- Evaluate the effect of RNA 2'O-methylation on ISG20 exonuclease activity.

In silico:

- Assess the impact of RNA 2'O-methylation on ISG20 structure.

In cellula:

- Uncover the pro/antiviral effects of vRNA 2'O-methylation (done by our collaborators).

As some cytosolic replicating viruses have acquired their own capping machinery to ensure the efficient translation of their mRNAs, protect them from degradation by cellular 5′ exoribonucleases, and escape innate immune sensing. Given the essential role of methylation in the capping process, I was involved in a side project consisting of evaluating the biological relevance of the different coronaviruses cap N7-MTase (chapter 3).

Study approaches:

In silico:

- Annotate coronaviruses nsp14 signature sequence based on SARS-CoV structure (done by François Ferron).

In vitro:

- Produce and purify SARS-CoV and MERS-CoV nsp14 WT and mutants of the MTase.
- Evaluate MTase and exonuclease activity of the nsp14s.

In cellula:

- Assess the impact of nsp14 N7-MTase mutant in SARS-CoV-1 and 2, MERS-CoV, and MHV viability (done by our collaborators).

Chapter 2

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 Coutard[6], Sébastien Nisole[3], Bruno Canard[1], Dorothée Missé[7] & 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] Unité des Virus Émergents (UVE: Aix-Marseille Univ-IRD 190-Inserm 1207), Marseille, France.

[7] MIVEGEC, Univ. Montpellier, CNRS, IRD Montpellier, France.

*Corresponding author (etienne.decroly@univ-amu.fr)

In revision

Abstract

RNA 2'O-methylation is a "self" epitranscriptomic modification that allows discriminating between host cell and pathogen mRNAs. Indeed, human immunodeficiency virus 1 (HIV-1) induces the 2'O-methylation of its genome by recruiting the cellular FTSJ3 methyltransferase, thereby impairing its detection by RIG-like receptors. Here, we show that RNA 2'O-methylation also limits the anti-viral activity of interferon-stimulated gene 20-kDa protein (ISG20). Using biochemical experiments we found that ISG20-mediated degradation of 2'O-methylated RNA pauses two nucleotides upstream of and at the methylated residue. Structure-function analysis indicated that the nuclease activity was inhibited through a steric clash between the ISG20 R53 and D90 residues and the 2'O-methylated nucleotide. Moreover, hypo-methylated HIV-1 genomes produced in FTSJ3-KO cells were more prone to degradation by ISG20 than those produced in cells that express wild type FTSJ3. Consequently, the retrotranscription and production of hypomethylated viruses was impaired, demonstrating the direct antagonist effect of 2’O-methylation on ISG20-mediated viral degradation.

Keywords: ISG20, epitranscriptome, 2’O-methylation, interferon, 3’exonuclease, FTSJ3, HIV-1.

Introduction

RNA structure and functions are regulated by post-transcriptional modifications (PTMs). To date, more than 140 chemical epitranscriptomic modifications have been described and they control most RNA-mediated cellular processes (e.g. splicing, nuclear export, microRNA biogenesis, localization, translation mechanisms, degradation and sensing by innate immunity) (1). Despite the increasing knowledge on PTM roles in RNA biological functions, the viral RNA (vRNA) epitranscriptomic modification landscape is still poorly characterized (2,3), and the role of these chemical modifications in virus replication remains misunderstood. RNA capping is one of the best characterized PTMs in the viral world, which consists of a guanosine moiety linked via a 5'-5' triphosphate to the first nucleotide (N_1) at the 5'end of most vRNAs ($GpppN_1$) (4). This cap structure is methylated at the N7 position of the guanosine residue by an S-adenosyl-L-methionine-dependent N7-methyltransferase (MTase), forming the cap-0 ($^{m}GpppN_1$) structure. In some cases, the first nucleotide (N_1) of the cap-0 structure is methylated on the ribose 2'OH by a virally encoded 2'O-MTase, leading to the formation of a cap-1 structure ($^{m}GpppN_m$) (5). The cap 2'O-methylation is sensed by the host cytosolic retinoic acid-inducible gene I protein (RIG-I) as a hallmark of the self (6,7). The detection of "non-self" RNAs (*i.e.* vRNAs without cap-1 structure) by RIG-like receptors initiates a cascade of events, leading to the secretion of type I interferon (IFN) and proinflammatory cytokines (7). In addition, it has been shown that some IFN-stimulated genes (ISG), such as IFN-induced proteins with tetratricopeptide repeats (IFIT) 1/3, can specifically discriminate capped RNA without the N_1 2'O-methylation mark (8). The modality of recognition of this RNA structure as "non-self" by these restriction factors has been established by X-ray crystallography (9). By sequestering "non-self" RNAs, IFIT1/3 block their translation into proteins (10,11).

In the viral world, several MTases involved in cap methylation have been characterized biochemically and structurally (12–16). The activity of these MTases depends on their ability to recruit and accommodate specific substrates in their catalytic pocket. Besides their functional role in RNA capping, the MTases of Flaviviruses and Filoviruses can methylate internal adenosines at the 2'O position within the vRNA (17–20). These internal epitranscriptomic marks have been mapped in the genome of some RNA viruses using RiboMethSeq (21), Nm-seq (22) and sensitive "agnostic" mass spectrometry techniques (2). The presence of 2'O-methylated nucleotides (N_m) has been demonstrated within the genome of Zika virus (ZIKV), Dengue virus (DENV) (2), and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (22)

Strikingly, internal N_m have been identified also in the genomes of polioviruses and hepatitis C viruses (2), although they do not encode any known MTase. This suggests that host 2'O-MTases might be recruited to promote these epitranscriptomic vRNA modifications. In line with these findings, type 1 human immunodeficiency virus (HIV-1) recruits the complex composed of a transactivation-responsive RNA-binding protein (TRBP) and the cellular 2'O-MTase FTSJ3 to catalyze the 2'O-methylation of 17 internal nucleotides in its genome (23). These modifications have a "pro-viral" effects at least by limiting the sensing by the RIG-like receptor MDA5 (23). Although the exact mechanism whereby MDA5 discriminates self and non-self has not been fully characterized, it has been suggested that the internal 2'O-methylation might prevent MDA5 interaction with the vRNA (24). Moreover, internal RNA 2'O-methylation might also confer resistance towards cellular restriction factors, including the IFN-induced exonucleases RNase L (25) and ISG20 (26,27).

In this study, we focused on ISG20, a 3' to 5' exonuclease that inhibits the replication of a broad range of RNA viruses, including Flaviviruses (Yellow fever, West Nile, DENV, bovine viral diarrhea virus, hepatitis C) (28), Togaviruses (Sindbis, Chikungunya, Venezuelan equine encephalitis viruses), Rhabdoviruses (vesicular stomatitis virus), Orthomyxoviruses (influenza virus), Bunyaviruses, Hepadnaviruses (hepatitis B) (29) and Retroviruses (HIV-1) (30). ISG20 is a 20-kDa protein that is enriched in the nucleolus and Cajal bodies (31). It contains three conserved exonuclease motifs (Exo I, II and III) and belongs to the DEDDh box 3'-5'exonucleases. Members of this superfamily are usually involved in ssRNA and ssDNA degradation, but ISG20 might preferentially targets RNA substrates (32,33). This hypothesis is supported by the crystal structure of ISG20 in complex with uridine 5'monophosphate (UMP) showing a specific hydrogen bond between residues M14 and R53 of ISG20 and the 2'OH moiety of the nucleotide ribose (32) that might stabilize the RNA in the catalytic pocket. Here, we produced and purified human ISG20 to precisely characterize its exonuclease activity on RNA substrates, and 2'O-methylation role in ISG20 RNase activity. We confirmed our biochemical results in an infectious cell model and demonstrated that the HIV-1 genome, which is naturally 2'O-methylated by the host FTSJ3, is more resistant to ISG20-induced degradation than hypomethylated vRNA. In addition, using pseudotyped HIV-1 particles, we demonstrated that the vRNA stability and luciferase reporter gene expression are reduced in the absence of 2'O-methylation. Together, our data shed light on a new pro-viral role of HIV-1 RNA 2'O-methylation by FTSJ3 in protecting the viral transcripts against ISG20-mediated degradation during the early infection stages.

Materials and methods

Cloning of ISG20 coding sequence

The coding sequence of human ISG20 (amino acids 1 to 181) followed by a C-terminal hexa-histidine tag was cloned into the pDEST14 vector (Thermo Fisher Scientific) using the Gateway cloning system, as indicated by the manufacturer. Mutations were introduced in the expression plasmid by PCR using the QuickChange Site-Directed Mutagenesis Kit (Stratagene) according to the manufacturer's instructions. Primers are listed in table S1.

ISG20 expression and purification

ISG20 was produced in *Escherichia coli* BL21-CodonPlus (DE3)-RIPL bacteria (Agilent). Cells were grown in LB medium containing 1% glucose, 100 µg/mL of ampicillin, and 34 µg/mL chloramphenicol at 37°C. Expression was induced at OD_{600nm}= 0.7 with 0.2 mM isopropyl-thiogalactopyranoside (IPTG) and stopped after 4h. Cells were harvested by centrifugation at 8,000 g at 4 °C for 20 min, and pellets were stored at -80 °C until purification. Pellets were thawed on ice and resuspended in lysis buffer (50 mM Tris pH 7, 50 mM NaCl, 1 mM β-mercaptoethanol, 1 mM PMSF, and 10% glycerol). The suspension was sonicated and centrifuged at 30,000 g 4 °C for 30 min. After clarification, 0.3% (w/v) of polyethyleneimine was slowly added to the lysate with gentle shaking at 4 °C for 30 min, followed by another centrifugation step. The supernatant was collected and supplemented with 5 mM imidazole pH 7.9 and 500 mM NaCl. Soluble ISG20 was purified on NiNTA resin, the column was washed with buffer W1 (50 mM Tris pH 7.9, 500 mM NaCl, 1 mM β-mercaptoethanol, 25 mM imidazole), W2 (50 mM Tris pH 7.9, 1 M NaCl, 25 mM imidazole, 1 mM β-mercaptoethanol), W3 (50 mM Tris pH 7.9, 100 mM NaCl, 25 mM imidazole, 1 mM β-mercaptoethanol), and W1. Proteins were eluted with buffer E (50 mM Tris pH 7.9, 250 mM NaCl, 250 mM imidazole and 1 mM β-mercaptoethanol), then dialyzed against 50 mM Tris pH 7, 150 mM NaCl, 1 mM β-mercaptoethanol and 0.1 mM dithiothreitol (DTT). Finally, proteins were concentrated on Amicon Ultra 5 kDa, and stored in 50% glycerol at -20 °C.

Synthesis of RNA and DNA substrates

The synthetic RNAs (Biomers) and DNAs (Eurofins Genomics) used in this study are listed in Table 1 and S1. Synthetic RNAs and DNAs (HPLC grade) were purchased. A_{27} with an N6-methyl adenosine at position 7 from the 3'-end ($A_{20}A^{N6m}A_6$) was synthesized as previously described (34). Synthesis was performed at the 1 mmol scale on an ABI 394 DNA synthesizer using commercially available 5'-O-DMTr-2'-O-pivaloyloxymethyl-3'-O-(O-cyanoethyl-N,N-diisopropylphosphoramidite N-6-phenoxyacetyl adenosine (Chemgenes Corp). The N6-methyl adenosine (A^{N6m}) was introduced in the sequence using the 5'-O-DMTr-2'-O-pivaloyloxymethyl-3'-O-(O-cyanoethyl-N,N-diisopropylphosphoramidite N6-methyl adenosine synthesized by us. After deprotection and release from the solid support upon basic conditions (DBU then aqueous ammonia treatment at 37 °C for 4 h), $A_{20}A_{N6m}A_6$ was purified by IEX-HPLC to high purity.

^{32}P radiolabeling of the 5' of substrates:

The non-fluorescent substrates were radiolabeled at their 5' end using T4 polynucleotide kinase (New England) and [γ^{32}P]ATP (Perkin Elmer) according to the manufacturer's instructions.

Exonuclease assays.

Nuclease activity assays were performed by mixing 2 nM of recombinant ISG20 with 500 nM of radiolabeled substrate in optimized buffer (50 mM Tris-HCl pH 7, 2.5 mM $MnCl_2$, 1 mM β-mercaptoethanol, 0.1 mM DTT and 0.1% Triton X-100). Reaction was performed at 37 °C and at each time point (0, 1, 5, 15, 30 and 60 min), 5 µL of suspension was taken and a double volume of loading buffer containing 96% formamide and 10 mM EDTA was added to stop the reaction. Then, the digested products were loaded on 7 M urea-containing 14% (wt/vol) polyacrylamide gels (acrylamide/bisacrylamide ratio, 19:1) buffered with 0.5xTris-NH2-taurine-EDTA and run at 65 W. Results were visualized by phosphorimaging using a Typhoon-9410 variable-mode scanner (GE Healthcare). ISG20 exonuclease activity was quantified using the FujiImager and Image Gauge analysis software. The nuclease activity of NP1 (New England Biolabs) and PDE (Sigma-Aldrich) was tested according to manufacturer recommendation at 37 °C. All assays were done at least twice.

Structural modeling

2'O-methylated UMP was designed using PubChem Sketcher V2.4 (https://pubchem.ncbi.nlm.nih.gov//edit3/index.html) then aligned with the UMP of the ISG20 structure (1WLJ) *via* UCSF Chimera (35) to visualize the molecular basis of ISG20 inhibition at N_0. Modeling of ISG20 with an RNA was done by superimposing the previously described structure of this nuclease (PDB:1WLJ) with that of SDN1 (PDB: 5Z9X) in ChimeraX (36). Methyl groups were added to the RNA of this model at the 2'O of riboses at positions N_{-2} and N_0 using the chimera build structure option that guaranties proper angles and bond distances.

Cells and viruses

HEK-FTSJ3-KO and parental HEK293T (HEK-WT) cells were cultured at 37 °C and 5% CO_2 in Dulbecco's modified Eagle's medium containing 10% fetal calf serum, supplemented with 1% penicillin-streptomycin (Thermo Fisher Scientific) as described by Ringeard et al., 2019. VSV-G pseudotyped NL4-3.Luc.R-E (pNL4-3.Luc.R-E) (37) was produced by transient transfection of HEK-WT or HEK-FTSJ3-KO using calcium phosphate, to produce WT HIV-1 particles or particles containing hypomethylated RNA genomes (FTSJ3-KO) (23), respectively. HIV-1 particles were concentrated at 22,000 g (Beckman Coulter) at 4 °C for 1h. Viral RNA was extracted using the RNeasy Mini Kit (Qiagen) and incubated with DNase, following the manufacturer's instructions. RNA concentration and purity were evaluated by spectrophotometry (NanoDrop 2000c, Thermo Fisher Scientific).

Western blot analyses

Cells were lysed in a buffer containing 50 mM Tris-HCl (pH7.4), 150 mM NaCl, 1 mM EDTA, 1% Triton X-100 and EDTA-free protease inhibitor cocktail (Roche). Protein extracts were analyzed on 10% SDS-PAGE gels, and transferred onto nitrocellulose membranes. Membranes were incubated with primary antibodies against FTSJ3 (rabbit; Bethyl Laboratories), Flag (clone M2; Sigma-Aldrich), and ß-actin (mouse) (makers), followed by anti-mouse and anti-rabbit HRP-conjugated antibodies. Chemiluminescent bands were imaged with a ChemidocTM MP Imager and analyzed for quantification using the Image LabTM desktop software (Bio-Rad Laboratories).

RT-qPCR analysis

HIV-1 RNA (50 ng) was reverse-transcribed using an oligo(dT) primer (located at the 3' end of HIV-1 RNA) or the M661 primer (in the Gag gene) (Zack et al., 1990) and the PrimeScript RT Reagent Kit (Perfect Real Time, Takara) in a 20 µl reaction. Real-time PCR amplifications were performed in triplicate using Takyon Rox SYBRMasterMix dTTP Blue (Eurogentec) on an Applied Biosystems QuantStudio 5 instrument. RT products were quantified using the M667/AA55 primers that amplify the R/U5 region of HIV-1, and the following program: 3 min at 95 °C followed by 40 cycles of 15 s at 95 °C, 20 s at 60 °C, and 20 s at 72 °C. Primers are listed in table S1.

Quantification of HIV-1 RT products by qPCR

HEK293T cells were transduced with VSV-G-pseudotyped HIV-1 (produced in HEK-WT or HEK-FTSJ3-KO cells) pre-treated with 250 U/mL benzonase (Merck) at 37°C for 20 min. Cells were collected 6 h post-infection and total DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen). Real-time PCR reactions were as described above. RT products were quantified using M667/M661 primers and results were normalized to *GAPDH* expression level. Primers are listed in table S1.

Results

Internal 2'O-methylation of RNA limits ISG20 activity at N_{-2} and N_0 of the methylated residue

Recombinant human ISG20 was purified (Figure S1A) and its exonuclease activity was assessed using synthetic 27-nucleotide RNAs (A_{27}, C_{27}, U_{27}) (Figure S1B) after optimization of the reaction conditions (Figure S1C-D). The laddering degradation profile confirmed ISG20 distributive 3' exonuclease activity. Sequence analysis showed that ISG20 contains conserved "DEDDh " motifs (Figure S2). After alanine substitution of residues in the Exo I (D11A, E13A), Exo II (D94A) and Exo III (H149A, D154A) motifs, the five recombinant ISG20 mutants were purified (Figure S3A). The exonuclease activity of wild type (WT) ISG20 was confirmed using synthetic A_{27} RNA, whereas that of the five mutants was strongly impaired (Figure S3B). Moreover, assessment of ISG20 nuclease activity using $ssRNA_1$, $ssRNA_2$ that forms a 3'hairpin structure, and $ssDNA_1$ showed that the presence of an RNA secondary structure stalled ISG20 exonuclease activity and that the enzyme was inactive on ssDNA (Figure S3C). Lastly, ISG20

also degraded RNA involved in RNA/DNA heteroduplexes (Figure S3D).

Next, to investigate whether the internal 2'O-methylation mark conferred resistance to ISG20-mediated degradation, radiolabeled A_{27} fragments carrying a 2'O-methyl group at different positions were incubated with ISG20, and their decay analyzed on urea-PAGE. A_{27} that contained one N_m ($A_{26}A_m$) at the 3'end or two nucleotides upstream of its 3'end ($A_{24}A_mA_2$) was almost completely protected from ISG20-mediated degradation (Figure 1A). Conversely, if the methylation was 7 nucleotides upstream of the 3'end ($A_{20}A_mA_6$), the RNA remained sensitive to ISG20 exonuclease activity, but hydrolysis paused before further degradation of the substrate. The stalling occurred at two distinct positions. The first main pause was detected two nucleotides upstream the N_m (N_{-2}) and the second at the N_m (N_0). This inhibition is likely to be ISG20-specific because the 3'exonuclease activity of two other 3'-5' exonucleases, namely nuclease P1 (NP1) and phosphodiesterase I (PDE), was barely affected by the presence of N_m (Figure S4). Then, the same experiment was repeated to evaluate whether 2'O-methylation on other ribonucleotides than "A" led to the same degradation pattern. $A_{20}U_mA_6$ and $A_{20}G_mA_6$ displayed a similar degradation profile as $A_{20}A_mA_6$ (Figure 1A). However, the pause observed at N_{-2} and N_0 seemed to be less pronounced with $A_{20}C_mA_6$, and ISG20 overcame the N_m after a longer incubation period (2h). The 2'O-methylation mark also protected heteropolymeric RNAs from degradation (Figure S5), but less efficiently than the homopolymeric A_{27} (Figure 1). These findings indicated that ISG20 exonuclease activity pauses two nucleotides upstream of the N_m, and at the N_m and that the length of these pauses depends on the N_m nature and the environment in which it is located.

Table 1: *List of the RNA substrates used for the exonuclease assays with ISG20*

Name	**Sequence from 5' to 3'**	**Modifications**
A_{27}	AAAAAAAAAAAAAAAAAAAAAAAAAAA	No
$A_{26}A_m$	AAAAAAAAAAAAAAAAAAAAAAAAAA**A_m**	2'O-methylation
$A_{26}U_m$	AAAAAAAAAAAAAAAAAAAAAAAAAA**U_m**	2'O-methylation
$A_{26}C_m$	AAAAAAAAAAAAAAAAAAAAAAAAAA**C_m**	2'O-methylation
$A_{26}G_m$	AAAAAAAAAAAAAAAAAAAAAAAAAA**G_m**	2'O-methylation

$A_{24}A_mA_2$	AAAAAAAAAAAAAAAAAAAAAAAA**A$_m$**AA	2'O-methylation
$A_{20}A_mA_6$	AAAAAAAAAAAAAAAAAAAA**A$_m$**AAAAAA	2'O-methylation
$A_{20}U_mA_6$	AAAAAAAAAAAAAAAAAAAA**U$_m$**AAAAAA	2'O-methylation
$A_{20}C_mA_6$	AAAAAAAAAAAAAAAAAAAA**C$_m$**AAAAAA	2'O-methylation
$A_{20}G_mA_6$	AAAAAAAAAAAAAAAAAAAA**G$_m$**AAAAAA	2'O-methylation
RNA_4-A_{27}	AGAUCGUGAACUUACAAACUAUACAAA	No
RNA_4-A_{27m}	AGAUCGUGAACUUACAAACUAUACAA**A$_m$**	2'O-methylation
RNA_4-C_{27}	AGAUCGUGAACUUACAAACUAUACAAC	No
RNA_4-C_{27m}	AGAUCGUGAACUUACAAACUAUACAA**C$_m$**	2'O-methylation

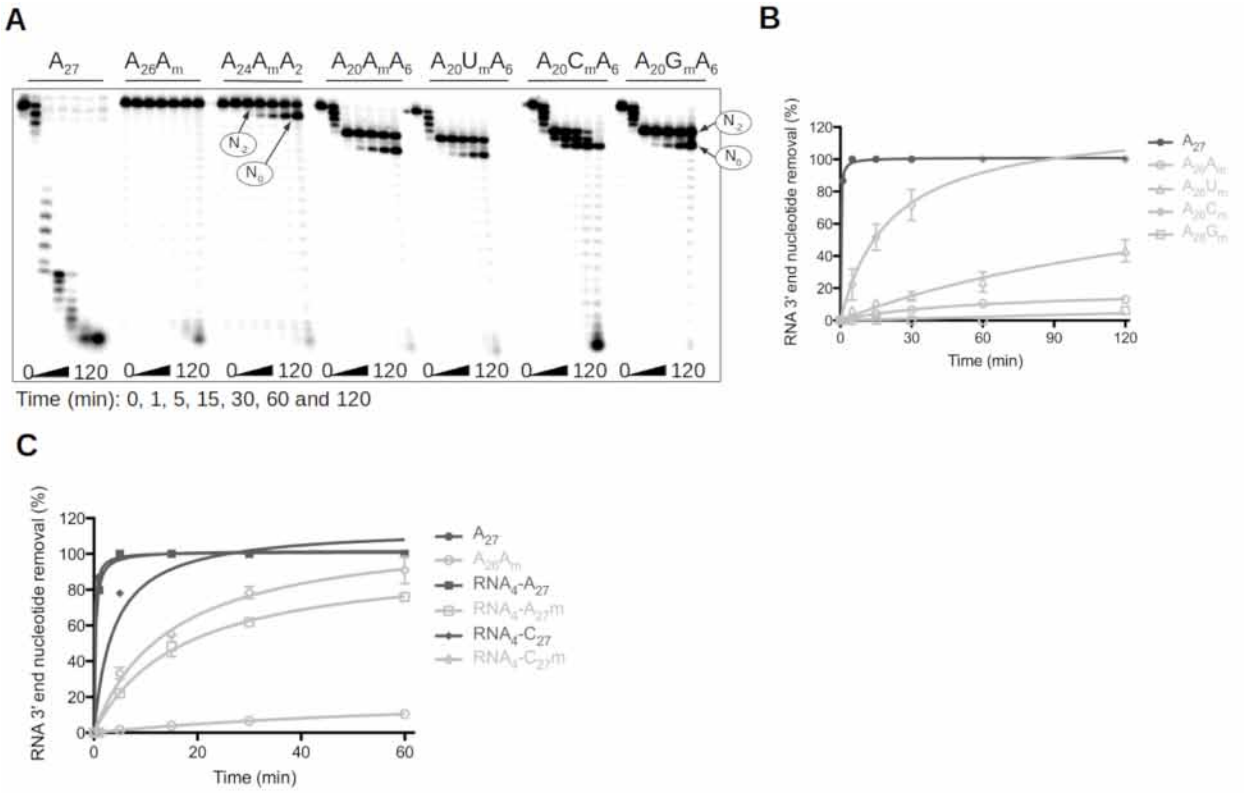

Figure 1: Impact of RNA 2'O-methylation on ISG20 nuclease activity. *(A-C) Recombinant human ISG20 was incubated with different 5'-radiolabeled RNAs carrying 2'-O-methylated residues at various positions in a time-course experiment, and the substrate hydrolysis was followed by PAGE. The degradation kinetics of various substrates in (B and C) were monitored by PAGE separation. Quantification of nucleotide removal, relative to the*

total RNA length, was done using the FujiImager and Image Gauge analysis software. A) Effect of RNA 2'O-methylation on ISG20 exonuclease activity. The assay was performed using non-methylated A_{27}, and 3'-end 2'O-methylated $A_{26}A_m$, $A_{24}A_mA_2$, $A_{20}A_mA_6$, $A_{20}U_mA_6$, $A_{20}C_mA6$, and $A_{20}G_mA_6$. B) Quantification of non-methylated (A_{27}) and 3'end 2'O-methylated ($A_{26}A_m$, $A_{26}U_m$, $A_{26}C_m$, $A_{26}G_m$) RNA degradation by ISG20. C) Quantification of non-methylated RNA_4-A_{27} and RNA_4-C_{27}, and 3'-end 2'O-methylated RNA_4-A_{27m}, RNA_4-C_{27m} degradation by ISG20. The results in panel B and C correspond to the mean values and standard deviation of three independent experiments.

2'O-methylation of poly-A at its 3'end suppresses degradation by ISG20

To further characterize the stalling observed at the N_m, ISG20 was incubated with a poly-A carrying A_m, G_m, C_m or U_m at its 3'end for different times. Following PAGE separation, the hydrolysis of RNA substrates was quantified by densitometry. This showed that the 3'end excision kinetics of $A_{26}N_m$ varied in function of N_m nature. Indeed, ISG20 fully hydrolyzed unmethylated A_{27} in ~5 minutes (Figure 1B). After 120 min ~100% of $A_{26}C_m$, ~43% of $A_{26}U_m$, and 13% of A_{27m} were degraded, conversely $A_{26}G_m$ was almost completely protected from ISG20 degradation (6%). These results confirm that the presence of a methyl group on the 2'O of the last nucleotide strongly reduces RNA decay and increases its lifespan to a different extent in function of the methylated residue (G>A>>U>>C). Again, ISG20 was more active on heteropolymeric RNAs that contained an N_m at their 3'end, as indicated by the degradation of ~75% and 95% of RNA_4-A_{27m} and RNA_4-C_{27m} RNA, respectively, after 60 min of incubation (Figure 1C).

2'O-methylated ribonucleotides limit ISG20 RNase activity through steric hindrance

Next, the previously described 3D structure of ISG20 in complex with UMP (Horio et al., 2004) was used to determine whether ISG20 exonuclease activity inhibition resulted from a direct inhibition of the catalytic reaction or from altered substrate recognition. A methyl group was modeled on the 2'O of the UMP ribose to determine its impact on the catalytic/binding pocket of ISG20. This modification led to a potential steric clash of the methyl group with the M14 and R53 residues of ISG20 (Figure 2A), suggesting the UMP repositioning to accommodate the methyl group. To test this hypothesis, these two ISG20 residues were substituted with the less bulky alanine residue (Figure S6) and the activity of each mutant was evaluated using A_{27} and $A_{26}A_m$ (as substrate to analyze the pause at N_0). The exonuclease activity of these two ISG20 mutants toward A_{27} was slightly reduced compared with WT ISG20, confirming the role of these

residues in RNA stabilization in the ISG20 catalytic pocket (Figure 2B). When incubated with $A_{26}A_m$, WT ISG20 and the M14A mutant did not show any nuclease activity. Conversely, the R53A mutant could degrade $A_{26}A_m$, indicating that this mutant can overcome the stop at N_0 (Figure 2B). Therefore, it can be inferred that the stop at N_0 results from a steric hindrance mechanism concerning residue R53 of ISG20 (Figure S7A).

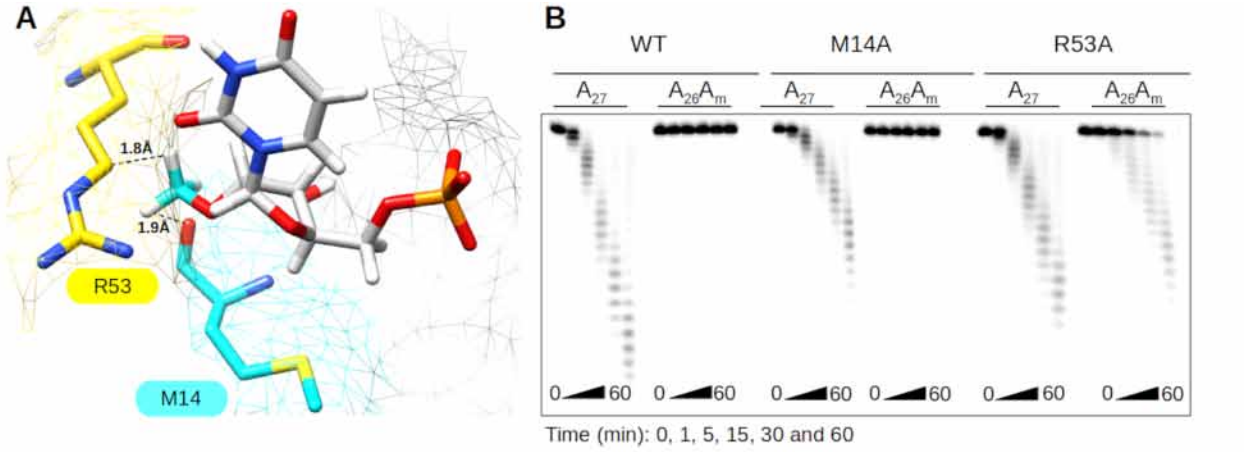

Figure 2: Molecular basis of ISG20 inhibition by 2'O-methylation at N_0. *A) Structural model of ISG20 showing the steric clash between residues R53 (yellow surface) and M14 (cyan surface), and 2'O-methylated UMP (UMP is in grey and 2'O-methyl in cyan). The methylated UMP was built using PubChem Sketcher V2.4 and superimposed on the UMP present in the catalytic pocket of the ISG20 structure (PDB: 1WLJ). B) The exonuclease activity of wild-type (WT) and M14A and R53A mutant ISG20 on methylated and non-methylated A27 RNAs was monitored by PAGE analysis as in Figure 1.*

Then, the molecular mechanism of the N_{-2} stop was investigated using the structural information of two ISG20 homologs (SDN1 and ExoX) in complex with a nucleic acid chain (Chen et al., 2018; Wang et al., 2013). Superimposition of ISG20 with SDN1 and ExoX showed a good fit of their catalytic sites with a Root-Mean-Square Deviation of 0.836 Å and 1.148 Å, respectively, and the superposition of the nucleic acid chains with the UMP of ISG20 (Figure S8). Based on the structural alignment, putative candidate residues of the RNA binding domain (RBD) were selected: M14, R53, H89, D90, H93, R126, R127, V128, S129 and L130 (Figure 3A, B). To experimentally test the structural model, alanine scanning was carried out on residues that were most likely to affect the binding (H89A, D90A, H93A, R126A, R127A, V128A) and the different proteins were produced and purified (Figure S6) for RNA nuclease assays. Although most ISG20 mutants conserved their exonuclease activity on A_{27} (Figure 3C), the D90A, R126A and R127A mutants showed reduced hydrolytic capacity that may result from disruption of the

interaction between ISG20 and the RNA substrate (Figure 3B). According to our structural model, the stop at N_{-2} could implicate residue(s) H89, D90 and/or V128 that are supposed to interact with the 2'OH of N_0 in the RNA and could be struggling to accommodate the methylated nucleotide (Figure S7B). To confirm this hypothesis, the effect of 2'O-methylation on the ISG20 H89A, D90A, and V128A mutant activity on $A_{20}A_mA_6$ was assessed at a concentration of 20 nM instead of 2 nM to overcome their decreased enzymatic activity. All mutants showed the same activity profile as WT ISG20, including the two intermediate degradation products detected at N_{-2} and N_0, with the exception of the D90A mutant that only stopped at N_0 (Figure 4A). These results confirmed that D90 is the key player in the N_{-2} stop, which is likely due to a steric clash with the N_m (Figure S7B).

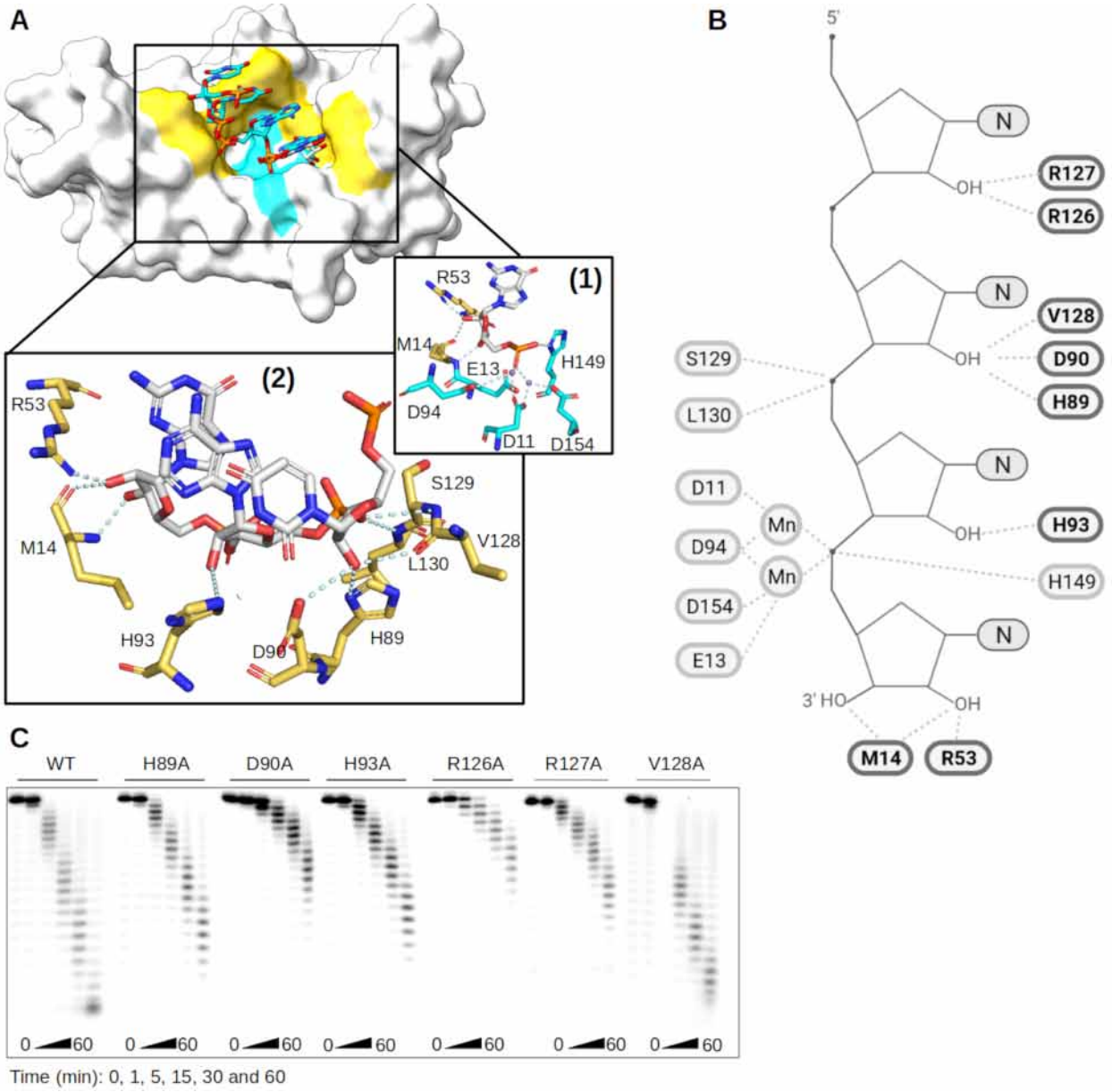

Figure 3: Structural activity relationship of ISG20 RNA binding Domain. *A) Model of ISG20 in interaction with an RNA substrate was built based on the superimposition of ISG20 (PDB: 1WLJ) and SDN1 in complex with RNA (PDB: 5Z9X). Surface representation of ISG20 (off-white) containing an RNA from SDN1 structure (sticks). Highlighted in yellow are the residues of the RNA binding domain (RBD) and in cyan the conserved residues of the exonuclease domain. Zoom on the ISG20 catalytic pocket showing the interaction (cyan) between the (1) residues of its catalytic domain (cyan), manganese (purple), and the nucleotide to be excised (grey), (2) residues of its RBD (yellow) and the RNA (grey). B) Cartoon representing the ISG20 model in interaction with an RNA substrate performed using Biorender in which the residues of ISG20 RBD interacting with the RNA 2'OH are highlighted in bold. C) Mutagenesis analysis of the residues highlighted in panel (B). The different ISG20 mutants were produced and their exonuclease activity was followed as in Figure 1 using non-methylated* A_{27}.

A double mutant R53A/D90A was next generated. Despite partial loss of its nuclease activity on substrates at 2 nM, the double mutant digested both methylated ($A_{26}A_m$ and $A_{24}A_mA_2$) and unmethylated A_{27} in a time-dependent manner at 20 nM (Figure 4B). Although the mutant still stalled during hydrolysis of methylated RNA ($A_{24}A_mA_2$), it could bypass the N_{-2} and N_0 stops, unlike WT ISG20. Altogether, these results give insights into the molecular basis of the hydrolysis blockage that is caused by two consecutive steric hindrances between the D90 and R53 residues of ISG20 and the methylated nucleotide (Figure 4C & S7). The specificity of the 2'O-methylation-linked inhibition was confirmed by introducing another PTM, adenosine N6-methylation (A_{N6m}), in the RNA substrate. This methylation did not interact with the ISG20 RBD in the structural model (Figure S9A) and consequently did not affect ISG20 exonuclease activity (Figure S9B).

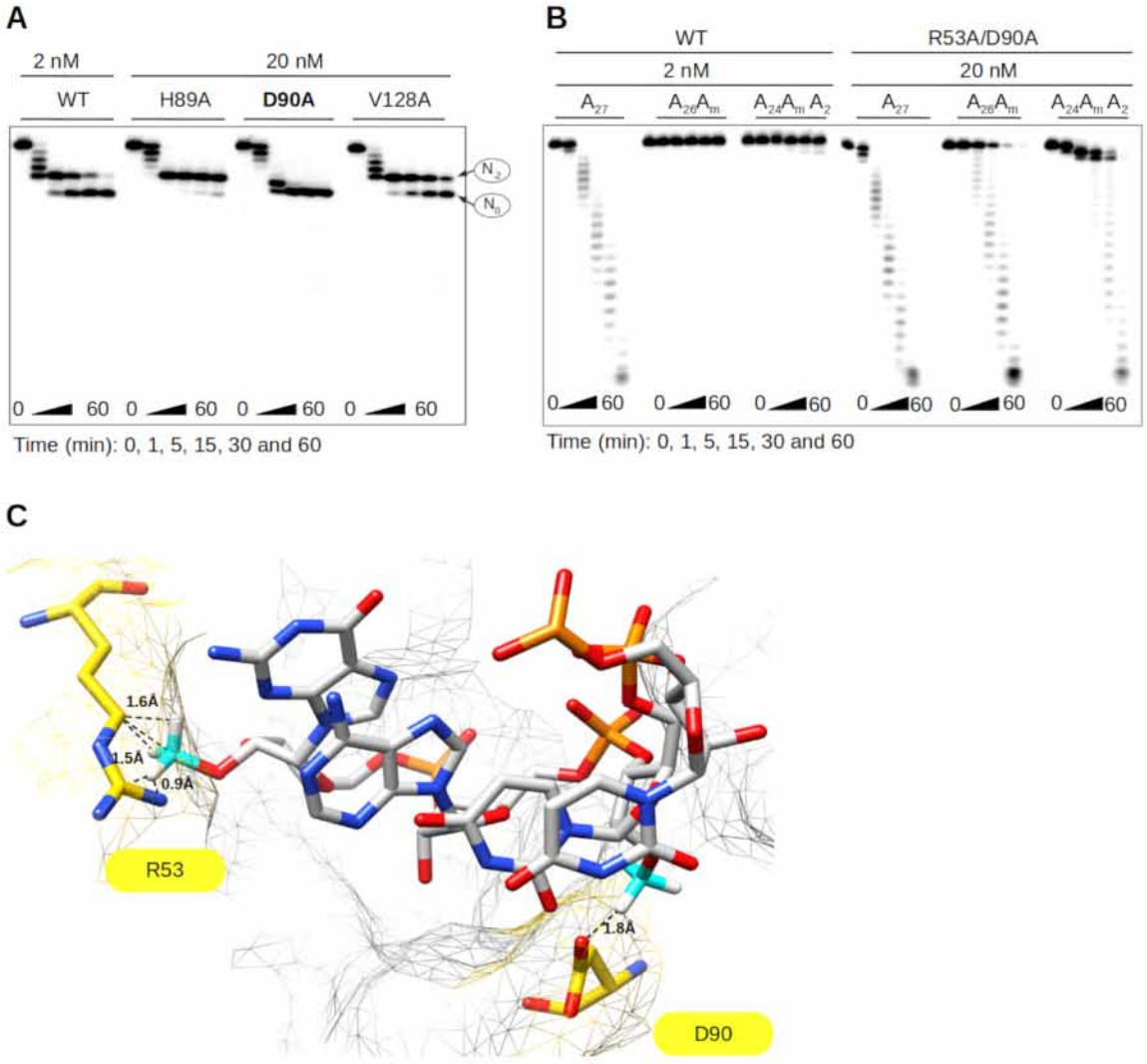

Figure 4: The D90 and R53 residues play a key role in ISG20 inhibition by 2'O-methylation. *A) Exonuclease activity of WT ISG20 and the H89A, D90A, and V128A mutants was assessed on $A_{20}A_mA_6$, and the RNA degradation*

products were separated on PAGE and analyzed by autoradiography, as in Figure 1. B) Exonuclease activity of WT ISG20 and the double mutant R53A/D90A was assessed using the A_{27}, $A_{26}A_m$ and $A_{24}A_mA_2$ RNA substrates. Degradation was monitored by PAGE and analyzed by autoradiography, as described in Figure 1. C) Structural model of ISG20 catalytic site (PDB: 1WLJ) with a methylated RNA (the methyl groups were added to the RNA using the chimera build structure option). Residues D90 and R53 are in yellow, the RNA is in grey, the RNA 2'O-methyl moieties in cyan. The model shows the steric hindrance between the distinct methylated nucleotides and the R53 and D90 residues of ISG20.

Internal 2'O-methylation protects HIV-1 RNA from ISG20-mediated degradation

In the HIV-1 genome, internal 2′O-methylation is catalyzed by the cellular 2'O-MTase FTSJ3 (23). Therefore, parental and FTSJ3-KO HEK293T cells (Figure 5A) were used to produce HIV-1 virions. After extraction, methylated (from parental cells) and hypomethylated (from FTSJ3-KO cells) full-length vRNAs were incubated with increasing concentrations of recombinant human ISG20. After 1h incubation, the exonuclease reaction was stopped by heating, and the remaining HIV-1 vRNAs were reverse transcribed using primers that anneal to the 5' (M661, in the Gag gene) or the 3' end (oligodT) of the HIV-1 genome. Quantification of the RT products by qPCR (Figure 5B, C) showed that HV-1 RNA could be degraded by ISG20 and that the 5' end of the genome was less prone to RNA degradation, as expected for 3' to 5' exonuclease activity. Moreover, hypomethylated HIV-1 RNA was more sensitive to ISG20-mediated degradation than methylated RNA. Then, methylated and hypomethylated HIV-1 RNAs were incubated with the double ISG20 R53A/D90A mutant, the activity of which is barely affected by RNA 2'O-methylation (Figure 4B). Methylated and hypomethylated HIV-1 RNAs were degraded to a similar extent (Figure 5D), confirming that N_m presence in the HIV-1 genome does not alter the activity of the double mutant.

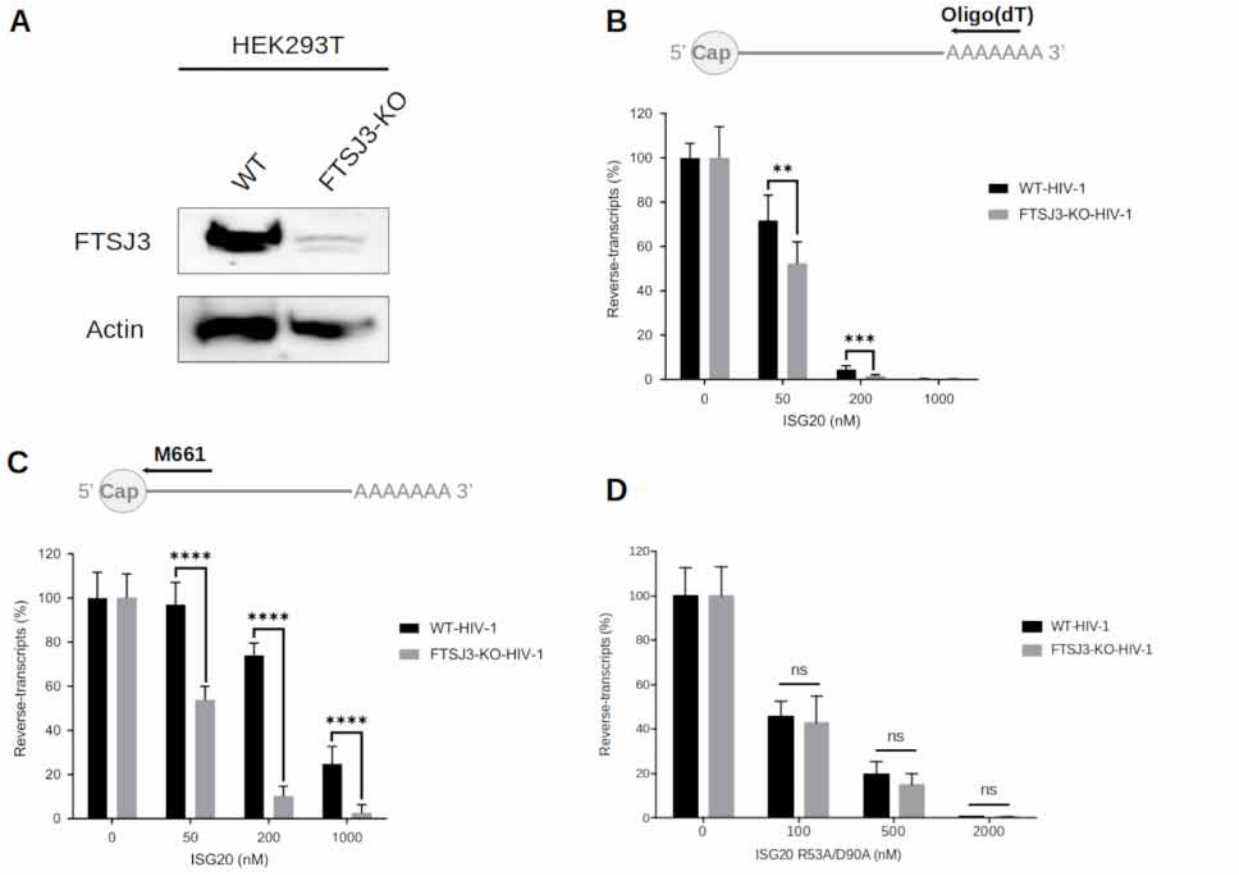

Figure 5: Hypomethylated HIV-1 RNA is sensitive to in vitro ISG20-mediated degradation. *A) FTSJ3 expression in parental (WT) and FTSJ3-KO HEK293T cells was assessed by western blotting before transfection of the HIV-1 molecular clone pNL4.3. B-C) 50ng of RNA extracted from HIV-1 particles produced in parental (WT) or FTSJ3-KO HEK29T cells was incubated with ISG20 at the indicated concentrations at 37 °C for 1h. After ISG20 inactivation (20 min at 70 °C), RNA products were reverse transcribed using a B) oligo(dT) or C) the M661 primer, and the reverse transcripts quantified by qPCR. The relative amount of HIV-1 reverse transcripts in each condition is shown. D) The exonuclease activity of ISG20 R53A/D90A was evaluated using HIV-1 RNA extracted from viral particles produced in parental (WT) or FTSJ3-KO HEK29T cells. RNA degradation products were reverse-transcribed using the M661 primer, and the reverse transcripts quantified by qPCR as in C. Data are the mean ± standard deviation of 3 independent experiments performed in triplicate. **p <0.01, ***p <0.001 and **** p <0.0001 (Student's t-test).*

These results prompted us to evaluate the antiviral effect in ISG20-expressing cells during infection. To this aim, luciferase-encoding VSV-G pseudotyped viruses were produced in parental and FTSJ3-KO HEK29T cells to generate HIV-1 pseudoparticles that contained methylated and hypomethylated vRNAs, respectively. Then, HEK293T cells that overexpress or not Flag-tagged human ISG20 (Figure 6A) were infected with these viruses and reverse transcription of HIV-1 RNA in infected cells was monitored by qPCR. Quantification of HIV-1 DNA products at 6h post-infection showed a decreased reverse transcription efficiency of hypomethylated viruses in ISG20-expressing cells (Figure 6B). This confirmed the increased

sensitivity of hypomethylated vRNA to ISG20-mediated degradation in infected cells. This observation was supported by the lower luciferase expression observed in ISG20-expressing cells infected with hypomethylated compared with methylated pseudotyped viruses (Figure 6C, FTSJ3-KO). Collectively, these results indicate that RNA 2'O-methylation induced by FTSJ3 plays a key role in protecting the HIV genome against ISG20 antiviral activity.

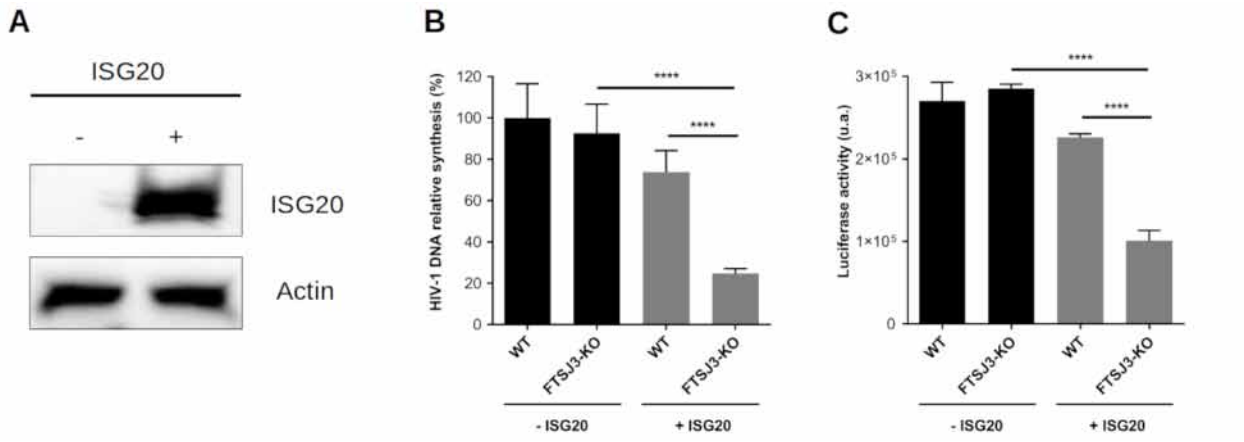

Figure 6: ISG20 restricts the reverse transcription of hypomethylated VSV-G pseudotyped HIV-1. *A) HEK293T cells were transfected with an empty plasmid or with a plasmid encoding Flag-tagged human ISG20. The expression of ISG20 was analyzed by western blotting using an anti-Flag antibody. 48h post-transfection, cells were infected with VSV-G-pseudotyped HIV-1 particles encoding luciferase, produced in parental (WT) or in FTSJ3-KO HEK293T cells. B) 6h post-infection, total cell DNA was extracted and HIV-1 reverse transcripts were quantified by qPCR. C) HIV-1 infection was estimated by measuring luciferase activity in cell extracts 24h post-infection. Data are the mean ± standard deviation of 3 independent experiments performed in triplicate; **p <0.01, ***p <0.001 and **** p <0.0001 (Student's t-test).*

Discussion

ISG20 is an IFN-regulated protein with antiviral activity against a broad range of RNA viruses (28–30,38). Although it has been suggested that its antiviral activity is linked to its exonuclease activity, ISG20 also inhibits the translation of vRNA (39). Several recent studies have shown that the genome of many RNA viruses contains PTM, including N_m (1–3,22,23). Here, we found that this epitranscriptomic modification can inhibit ISG20 exonuclease activity and thereby counteract the host antiviral response mediated by ISG20. Using synthetic RNAs carrying 2'O-methylation at different positions, we demonstrated that ISG20 nuclease activity is specifically impaired by this epitranscriptomic mark. We also observed that this inhibition is modulated by several factors, such as the nature of the methylated nucleotide moiety (G_m>A_m>U_m>C_m>>A), its position in the RNA sequence, and the substrate sequence (Figure 1 &

Figure S5). We found that ISG20 activity was more strongly impaired when the 2'O-methylation occurred in the context of homopolymeric RNA, such as poly-A, and at the 3'end of the sequences (Figure 1). Together, these data indicate that N_m can regulate RNA decay and that synthetic poly-A is quite resistant to ISG20-mediated degradation when modified at its 3' end. However, the proportion of methylated A residues in viral poly-A tails has not been investigated yet, to our knowledge, leaving the biological relevance of this finding unclear.

In addition, we showed that the degradation of RNA carrying internal 2'O-methylation ($A_{20}N_mA_6$) by ISG20 was paused at N_{-2} and N_0 from the N_m (Figure 1A). Through a structure-guided analysis, we selected key residues of ISG20 RBD that may be involved in this inhibition (M14, R58, H89, D90, and V128; Figure 3A & B). Alanine mutagenesis of these residues revealed that the pauses at N_{-2} and N_0 were overcome by ISG20 harboring the D90A and R53A mutation, respectively (Figure 2 & 4A & S7), indicating that this inhibition occurs through steric hindrance. This was further confirmed by the finding that the double mutant R53A/D90A can hydrolyze methylated RNA without pausing at N_{-2} and N_0 (Figure 4B). Moreover, this inhibition seems to be specific to RNA 2'O-methylation because N6-methylation, another important epitranscriptomic mark present on vRNA, barely affected ISG20 exonuclease activity (Figure S9B). This observation is supported also by the structural model of ISG20 in the presence of an RNA substrate showing that N6-methylation is outside the RNA binding pocket (Figure S9A). This is consistent with the study by Imam et al. demonstrating that ISG20 targets HBV transcripts containing A^{N6m} (40). Therefore, it is likely that ISG20 inhibition is specifically related to the presence of RNA 2'O-methylation, suggesting that viruses that induce such internal PTM, using their own (18–20) or the host MTase (23), might limit ISG20 antiviral activity. This possible role of internal 2'O-methylation is reminiscent of other viral strategies to overcome the cell defenses, such as cap 2'O-methylation that prevents vRNA sensing by RIG-I (6,7) and its subsequent sequestration by IFIT1/3 restriction factors (8).

Although the viral epitranscriptome is still poorly characterized in most viruses, the recent development of sequencing-based methods to identify modified nucleotides and their location has boosted the knowledge on the mechanisms and functions of 2'O-methylation in RNA viruses (1–3,22,23). For instance, HIV-1 exploits the cell machinery to catalyze 2'O-methylation of its own genome through a tightly regulated process. Indeed, HIV-1 recruits FTSJ3, a cellular MTase, to catalyze the 2'O-methylation of 17 specific residues of its own RNA. This prevents the sensing of vRNA by MDA5, which in turn limits the production of type 1 IFN (23). Therefore, we asked whether internal methylation of HIV-1 RNA might also influence its

resistance to ISG20 restriction. We addressed this question by comparing the sensitivity to ISG20 of HIV-1 RNAs extracted from viruses produced in parental HEK293T cells or cells that lack FTSJ3. We found that internal 2'O-methylation protects the HIV-1 genome from ISG20 degradation (Figure 5B & C). We validated this result in cells infected with normally methylated or hypomethylated VSV-G pseudotyped HIV-1. We also showed that ectopically expressed ISG20 drastically reduced luciferase expression in cells infected with hypomethylated VSV-G pseudotyped HIV-1 as a consequence of RT impairment (Figure 6B & C). Collectively, these results indicate the importance of internal 2'O-methylation in the early phase of HIV-1 infection by counteracting ISG20 antiviral activity. Indeed, 2'O-methylation is a viral countermeasure to the cellular antiviral response originally identified as the subversion of "self" sensing to impair type 1 IFN expression. Our study also demonstrates that at least for HIV-1, 2'O-methylation also directly prevents the antiviral effect of ISG20 by protecting the vRNA from its exonuclease activity. The importance of 2'O-methylation in viral replication has been shown in various RNA viruses, including Nipah and Hendra viruses (41), but the dual function observed in HIV-1 (limiting vRNA sensing and degradation) must be investigated in other viruses to determine whether this dual mechanism is universal.

The analysis of SARS-CoV-2 epitranscriptome also showed the presence of internal 2'O-methylation marks (22). However, the implicated MTase remains unknown. These N_m marks were mainly detected in the 5' and 3' untranslated regions (UTR) and in the non-coding regions of the viral RNA (22) that are enriched in hairpin structures. Therefore, these methylation marks might prevent the detection of vRNA by MDA5, as observed for HIV-1. Their impact on ISGs that target double stranded RNA, such as 2′-5′-oligoadenylate synthetase 1 (42) and the RNase L, is questionable. Moreover, as a recent study revealed a transient upregulation of ISG20 expression upon SARS-CoV-2 infection (43), the resistance to ISG20-mediated restriction could result from the adoption of countermeasures (e.g. hyper-2'O-methylation at the 3′UTR) (22) evolved by the virus. More studies are needed to confirm these hypotheses in the context of SARS-CoV-2 infection. Internal 2'O-methylation has also been described in Flaviviruses, and *in vitro* studies showed that the NS5 MTases of DENV (19) and ZIKV (18) specifically methylate adenosines. However, the epitranscriptome analysis of both viruses highlighted the presence of 2'O-methylation marks on all four ribonucleotides (2). This implies that the host 2'O-MTases might also participate in these methylation processes. This is consistent with the binding of fibrillarin, a nucleolar 2'O-MTase, in the DENV and ZIKV genomes (44). These observations

suggest that there is synergy and fine-tuning in the 2'O-methylation performed by viral and cellular MTases to shield vRNAs from innate immune censors and the IFN response.

The identification of the functions associated with internal 2'O-methylation on MDA5 sensing and ISG20 exonuclease activity may also be considered in the context of improving mRNA-based therapeutics. Indeed treatment with exogenous RNA can be hampered by the rapid degradation of RNA molecules and/or its capacity to induce inappropriate innate immunity stimulation. Various PTMs are used to overcome these drawbacks. For instance, addition of an RNA cap structure (45–47), a 5′ and 3′ untranslated region (48), or Poly-A (47,49) promotes mRNA stabilization and increases protein translation. In addition, modified nucleosides, such as A_{N6m}, 5-methylcytidine, and pseudo-uridine, can be used to avoid innate immunity overstimulation and to increase RNA translation (50,51). In the light of the present (this study) and past (23) results showing 2'O-methylation role in RNA sensing, the evaluation of 2'O-methylated nucleotides could contribute to improve the transcript half-life by protecting from ISG20-mediated degradation and the cellular response.

Acknowledgment

We thank Yamina Bennasser (IGH, Montpellier, France) for kindly providing HEK293T cells knocked-out for FTSJ3 (HEK-FTSJ3-KO). The luciferase-encoding HIV-1 plasmid (pNL4-3.Luc.R-E) was kindly provided by Nathaniel R Landau (Aaron Diamond AIDS Research Center, The Rockefeller University, New-York, USA). We thank Coralie Valle for her careful critical reading of this manuscript.

Funding

P.E.K was supported by the Foundation Méditerranée infection. The research was part of the project supported by the Agence nationale de la recherche [ANR-20-CE11-0024-02 and ANR-17-CE15-0029]; the Foundation pour la recherche médicale [FRM-REPLI80C/U160]; and the Agence Nationale de la Recherche sur le SIDA et les Hépatites virales [ANRS-ECTZ80C/U160].

Conflict of interest

All authors declare no competing interests.

References

1. Netzband R, Pager CT. Epitranscriptomic marks: Emerging modulators of RNA virus gene expression. WIREs RNA. 2020;11(3):e1576.

2. McIntyre W, Netzband R, Bonenfant G, Biegel JM, Miller C, Fuchs G, et al. Positive-sense RNA viruses reveal the complexity and dynamics of the cellular and viral epitranscriptomes during infection. Nucleic Acids Res. 2018 Jun 20;46(11):5776–91.

3. Kennedy EM, Courtney DG, Tsai K, Cullen BR. Viral Epitranscriptomics. J Virol. 91(9):e02263-16.

4. Decroly E, Ferron F, Lescar J, Canard B. Conventional and unconventional mechanisms for capping viral mRNA. Nat Rev Microbiol. 2012 Jan;10(1):51–65.

5. Decroly E, Canard B. Biochemical principles and inhibitors to interfere with viral capping pathways. Curr Opin Virol. 2017 Jun 1;24:87–96.

6. Devarkar SC, Wang C, Miller MT, Ramanathan A, Jiang F, Khan AG, et al. Structural basis for m7G recognition and 2′-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. Proc Natl Acad Sci U S A. 2016 Jan 19;113(3):596–601.

7. Schuberth-Wagner C, Ludwig J, Bruder AK, Herzner A-M, Zillinger T, Goldeck M, et al. A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2’O-Methylated Self RNA. Immunity. 2015 Jul 21;43(1):41–51.

8. Daffis S, Szretter KJ, Schriewer J, Li J, Youn S, Errett J, et al. 2′-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature. 2010 Nov;468(7322):452–6.

9. Abbas YM, Pichlmair A, Górna MW, Superti-Furga G, Nagar B. Structural basis for viral 5′-PPP-RNA recognition by human IFIT proteins. Nature. 2013 Feb;494(7435):60–4.

10. Kumar P, Sweeney TR, Skabkin MA, Skabkina OV, Hellen CUT, Pestova TV. Inhibition of translation by IFIT family members is determined by their ability to interact selectively with the 5′-terminal regions of cap0-, cap1- and 5′ppp- mRNAs. Nucleic Acids Res. 2014 Mar 1;42(5):3228–45.

11. Johnson B, VanBlargan LA, Xu W, White JP, Shan C, Shi P-Y, et al. Human IFIT3 Modulates IFIT1 RNA Binding Specificity and Protein Stability. Immunity. 2018 Mar 20;48(3):487-499.e5.

12. Shuman S, Hurwitz J. Mechanism of mRNA capping by vaccinia virus guanylyltransferase: characterization of an enzyme--guanylate intermediate. Proc Natl Acad Sci U S A. 1981 Jan;78(1):187–91.

13. Egloff M-P, Benarroch D, Selisko B, Romette J-L, Canard B. An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. EMBO J. 2002 Jun 3;21(11):2757–68.

14. Decroly E, Debarnot C, Ferron F, Bouvet M, Coutard B, Imbert I, et al. Crystal Structure and Functional Analysis of the SARS-Coronavirus RNA Cap 2′-O-Methyltransferase nsp10/nsp16 Complex. PLOS Pathog. 2011 May 26;7(5):e1002059.

15. Paesen GC, Collet A, Sallamand C, Debart F, Vasseur J-J, Canard B, et al. X-ray structure and activities of an essential Mononegavirales L-protein domain. Nat Commun. 2015 Nov 9;6(1):8749.

16. Valle C, Martin B, Ferron F, Roig-Zamboni V, Desmyter A, Debart F, et al. First insights into the structural features of Ebola virus methyltransferase activities. Nucleic Acids Res. 2021 Feb 22;49(3):1737–48.

17. Barral K, Sallamand C, Petzold C, Coutard B, Collet A, Thillier Y, et al. Development of specific dengue virus 2'-O- and N7-methyltransferase assays for antiviral drug screening. Antiviral Res. 2013 Sep;99(3):292–300.

18. Coutard B, Barral K, Lichière J, Selisko B, Martin B, Aouadi W, et al. Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. J Virol [Internet]. 2016 Dec 28 [cited 2022 Jan 10]; Available from: https://journals.asm.org/doi/abs/10.1128/JVI.02202-16

19. Dong H, Chang DC, Hua MHC, Lim SP, Chionh YH, Hia F, et al. 2′-O Methylation of Internal Adenosine by Flavivirus NS5 Methyltransferase. PLOS Pathog. 2012 Apr 5;8(4):e1002642.

20. Martin B, Coutard B, Guez T, Paesen GC, Canard B, Debart F, et al. The methyltransferase

domain of the Sudan ebolavirus L protein specifically targets internal adenosines of RNA substrates, in addition to the cap structure. Nucleic Acids Res. 2018 Sep 6;46(15):7902–12.

21. Ayadi L, Motorin Y, Marchand V. Quantification of 2'-O-Me Residues in RNA Using Next-Generation Sequencing (Illumina RiboMethSeq Protocol). Methods Mol Biol Clifton NJ. 2018;1649:29–48.

22. Yang SL, DeFalco L, Anderson DE, Zhang Y, Aw JGA, Lim SY, et al. Comprehensive mapping of SARS-CoV-2 interactions in vivo reveals functional virus-host interactions. Nat Commun. 2021 Aug 25;12(1):5113.

23. Ringeard M, Marchand V, Decroly E, Motorin Y, Bennasser Y. FTSJ3 is an RNA 2'-O-methyltransferase recruited by HIV to avoid innate immune sensing. Nature. 2019 Jan;565(7740):500–4.

24. Wu B, Peisley A, Richards C, Yao H, Zeng X, Lin C, et al. Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5. Cell. 2013 Jan 17;152(1–2):276–89.

25. Silverman RH. Viral Encounters with 2′,5′-Oligoadenylate Synthetase and RNase L during the Interferon Antiviral Response. J Virol. 2007 Dec 1;81(23):12720–9.

26. Espert L, Degols G, Gongora C, Blondel D, Williams BR, Silverman RH, et al. ISG20, a new interferon-induced RNase specific for single-stranded RNA, defines an alternative antiviral pathway against RNA genomic viruses. J Biol Chem. 2003 May 2;278(18):16151–8.

27. Weiss CM, Trobaugh DW, Sun C, Lucas TM, Diamond MS, Ryman KD, et al. The Interferon-Induced Exonuclease ISG20 Exerts Antiviral Activity through Upregulation of Type I Interferon Response Proteins. mSphere. 2018 Sep 19;3(5):e00209-18.

28. Jiang D, Weidner JM, Qing M, Pan X-B, Guo H, Xu C, et al. Identification of Five Interferon-Induced Cellular Proteins That Inhibit West Nile Virus and Dengue Virus Infections. J Virol. 2010 Aug 15;84(16):8332–41.

29. Xu H, Lei Y, Zhong S, Peng F-Y, Zhou Z, Li K, et al. [Antiviral activities of ISG20 against hepatitis C virus]. Zhonghua Gan Zang Bing Za Zhi Zhonghua Ganzangbing Zazhi Chin J

Hepatol. 2013 Jan;21(1):33–7.

30. Espert L, Degols G, Lin Y-L, Vincent T, Benkirane M, Mechti N. Interferon-induced exonuclease ISG20 exhibits an antiviral activity against human immunodeficiency virus type 1. J Gen Virol. 2005 Aug;86(Pt 8):2221–9.

31. Espert L, Eldin P, Gongora C, Bayard B, Harper F, Chelbi-Alix MK, et al. The exonuclease ISG20 mainly localizes in the nucleolus and the Cajal (Coiled) bodies and is associated with nuclear SMN protein-containing complexes. J Cell Biochem. 2006 Aug 1;98(5):1320–33.

32. Horio T, Murai M, Inoue T, Hamasaki T, Tanaka T, Ohgi T. Crystal structure of human ISG20, an interferon-induced antiviral ribonuclease. FEBS Lett. 2004 Nov 5;577(1–2):111–6.

33. Nguyen LH, Espert L, Mechti N, Wilson DM. The Human Interferon- and Estrogen-Regulated ISG20/HEM45 Gene Product Degrades Single-Stranded RNA and DNA in Vitro. Biochemistry. 2001 Jun 1;40(24):7174–9.

34. Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, et al. Reversible methylation of m6Am in the 5' cap controls mRNA stability. Nature. 2017 Jan 19;541(7637):371–5.

35. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004 Oct;25(13):1605–12.

36. Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, et al. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. Protein Sci Publ Protein Soc. 2021 Jan;30(1):70–82.

37. He J, Choe S, Walker R, Di Marzio P, Morgan DO, Landau NR. Human immunodeficiency virus type 1 viral protein R (Vpr) arrests cells in the G2 phase of the cell cycle by inhibiting p34cdc2 activity. J Virol. 1995 Nov;69(11):6705–11.

38. Deymier S, Louvat C, Fiorini F, Cimarelli A. ISG20: an enigmatic antiviral RNase targeting multiple viruses. FEBS Open Bio. 2022 Feb 16;

39. Wu N, Nguyen X-N, Wang L, Appourchaux R, Zhang C, Panthu B, et al. The interferon stimulated gene 20 protein (ISG20) is an innate defense antiviral factor that discriminates self

versus non-self translation. PLoS Pathog. 2019 Oct;15(10):e1008093.

40. Imam H, Kim G-W, Mir SA, Khan M, Siddiqui A. Interferon-stimulated gene 20 (ISG20) selectively degrades N6-methyladenosine modified Hepatitis B Virus transcripts. PLoS Pathog. 2020 Feb;16(2):e1008338.

41. Deffrasnes C, Marsh GA, Foo CH, Rootes CL, Gould CM, Grusovin J, et al. Genome-wide siRNA Screening at Biosafety Level 4 Reveals a Crucial Role for Fibrillarin in Henipavirus Infection. PLOS Pathog. 2016 Mar 24;12(3):e1005478.

42. Wickenhagen A, Sugrue E, Lytras S, Kuchi S, Noerenberg M, Turnbull ML, et al. A prenylated dsRNA sensor protects against severe COVID-19. Science. 374(6567):eabj3624.

43. Hammoudeh SM, Hammoudeh AM, Bhamidimarri PM, Al Safar H, Mahboub B, Künstner A, et al. Systems Immunology Analysis Reveals the Contribution of Pulmonary and Extrapulmonary Tissues to the Immunopathogenesis of Severe COVID-19 Patients. Front Immunol. 2021;12:2370.

44. Flynn RA, Belk JA, Qi Y, Yasumoto Y, Wei J, Alfajaro MM, et al. Discovery and functional interrogation of SARS-CoV-2 RNA-host protein interactions. Cell. 2021 Apr 29;184(9):2394-2411.e16.

45. Martin SA, Paoletti E, Moss B. Purification of mRNA guanylyltransferase and mRNA (guanine-7-) methyltransferase from vaccinia virions. J Biol Chem. 1975 Dec 25;250(24):9322–9.

46. Stepinski J, Waddell C, Stolarski R, Darzynkiewicz E, Rhoads RE. Synthesis and properties of mRNAs containing the novel “anti-reverse” cap analogs 7-methyl(3’-O-methyl)GpppG and 7-methyl (3’-deoxy)GpppG. RNA N Y N. 2001 Oct;7(10):1486–95.

47. Jaïs PH, Decroly E, Jacquet E, Le Boulch M, Jaïs A, Jean-Jean O, et al. C3P3-G1: first generation of a eukaryotic artificial cytoplasmic expression system. Nucleic Acids Res. 2019 Mar 18;47(5):2681–98.

48. Ross J, Sullivan TD. Half-lives of beta and gamma globin messenger RNAs and of protein synthetic capacity in cultured human reticulocytes. Blood. 1985 Nov;66(5):1149–54.

49. Gallie DR. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. Genes Dev. 1991 Nov;5(11):2108–16.

50. Karikó K, Muramatsu H, Welsh FA, Ludwig J, Kato H, Akira S, et al. Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. Mol Ther J Am Soc Gene Ther. 2008 Nov;16(11):1833–40.

51. Karikó K, Buckstein M, Ni H, Weissman D. Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. Immunity. 2005 Aug;23(2):165–75.

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*	

Table S2: *Substrates used for ISG20 characterization.*

Name	**Sequence from 5' to 3'**	**Modifications**
A_{27}	AAAAAAAAAAAAAAAAAAAAAAAAAAA	No
U_{27}	UUUUUUUUUUUUUUUUUUUUUUUUUUU	No
C_{27}	CCCCCCCCCCCCCCCCCCCCCCCCCCC	No
$A_{20}A_mA_6$	AAAAAAAAAAAAAAAAAAA**A_m**AAAAAA	2'O-methylation
$A_{20}A^{N6m}A_6$	AAAAAAAAAAAAAAAAAAA**A^{N6m}**AAAAAA	N6-methylation
RNA_1	AGACGAUGCGGAAAACUCUAACAAGAU	No
RNA_2	UGACGGCCCGGAAAACCGGGCC	No
DNA_1	AAAAAAAAAAAAAAAAAAAAAAAAAAA	No

RNA_3	CUAUCCCCAUGUGAUUUUAAUAGCUUCUUAGG AGAAUGAC	No
RNA_{3RC}	GUCAUUCUCCUAAGAAGCUA	No
DNA_2	CTATCCCCATGTGATTTTAATAGCTTCTTAGGAGA ATGAC	No
RNA_4-A_{21}	AGAUCGUGAACUUACAAACUAUACAAA	No
RNA_4-A_{21m}	AGAUCGUGAACUUACAAACU**A$_m$**UACAAA	2'O-methylation
RNA_5-C_{21}	AGAUCGUGAACUUACAAACUCUACAAA	No
RNA_5-C_{21m}	AGAUCGUGAACUUACAAACU**C$_m$**UACAAA	2'O-methylation

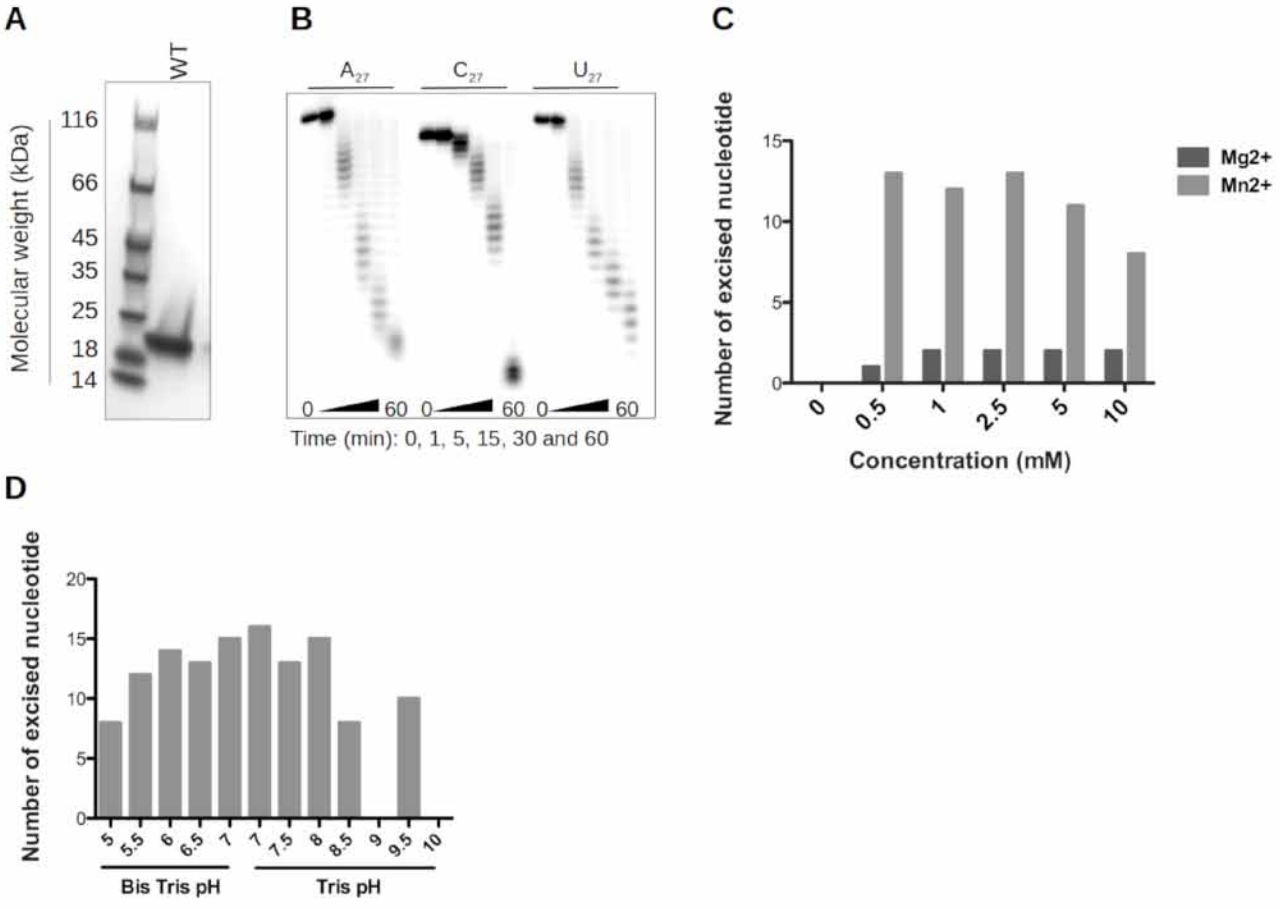

Figure S1: ISG20 purification and optimization of the exonuclease assay. *Recombinant human wild-type (WT) ISG20 was expressed in E. coli and purified on a NTa column. A) Purified proteins were separated on 15% SDS-PAGE gels followed by*

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).

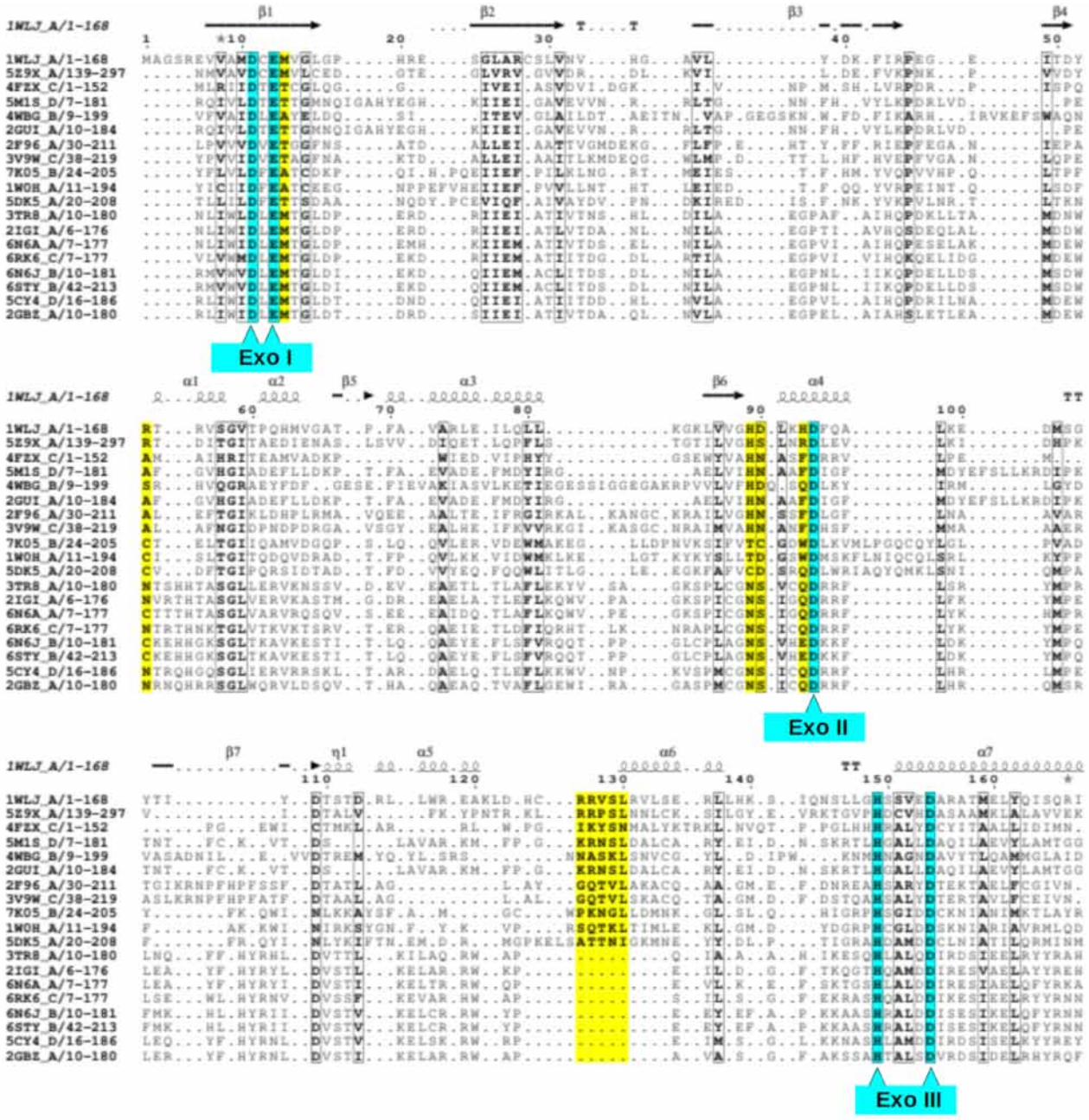

Figure S2: ISG20 sequence alignment. *Sequence and structure homologies were assessed with HHpred (https://toolkit.tuebingen.mpg.de/tools/hhpred), hits were recovered and curated using TCoffee alignment. The alignment was processed with ESPript3 server. Sequences are annotated according to their PDB accession number. 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.*

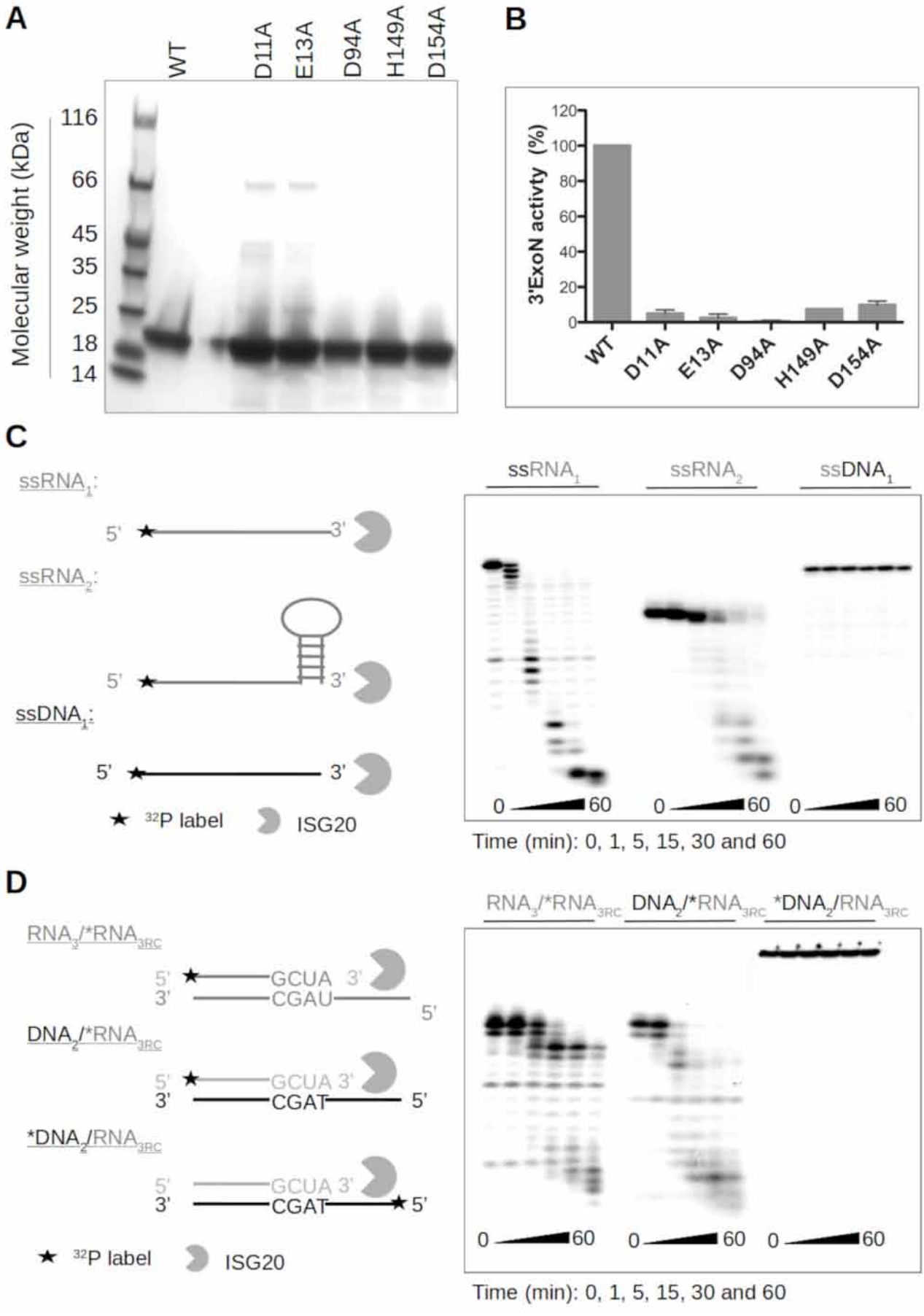
A
WT
D11A
E13A
D94A
H149A
D154A
Molecular weight (kDa)
116
66
45
35
25
18
14
B
3'ExoN activty (%)
120
100
80
60
40
20
0
WT
D11A
E13A
D94A
H149A
D154A
C
ssRNA1:
5'
3'
ssRNA2:
5'
3'
ssDNA1:
5'
3'
32P label
ISG20
ssRNA1
ssRNA2
ssDNA1
0
60
0
60
0
60
Time (min): 0, 1, 5, 15, 30 and 60
D
RNA3/*RNA3RC
5'
GCUA
3'
3'
CGAU
5'
DNA2/*RNA3RC
5'
GCUA
3'
3'
CGAT
5'
*DNA2/RNA3RC
5'
GCUA
3'
3'
CGAT
5'
32P label
ISG20
RNA3/*RNA3RC
DNA2/*RNA3RC
*DNA2/RNA3RC
0
60
0
60
0
60
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. 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 represents the radiolabeled position) and their degradation profiles on 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}).

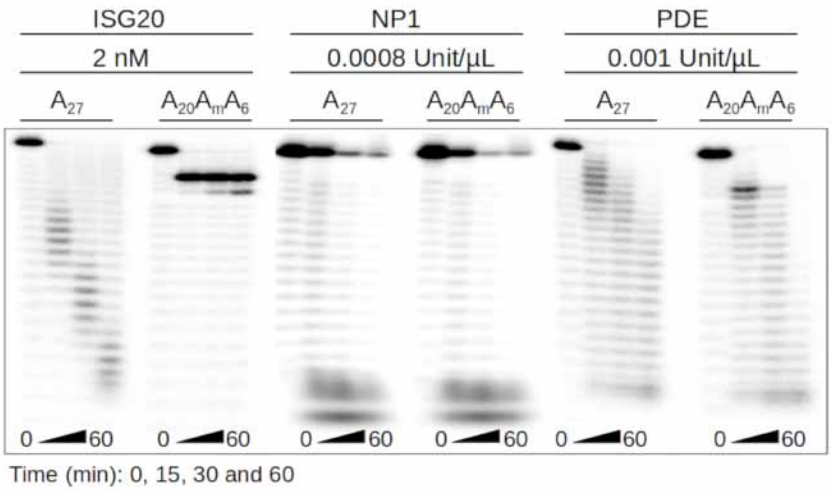

Figure S4: Comparative hydrolysis of 2'O-methylated RNA by different 3' exonucleases. *The two 5'end radiolabeled RNA substrates (A_{27} and $A_{20}A_mA_6$) were incubated with ISG20, NP1, or PDE for different times, and the reaction products were analyzed by PAGE and autoradiography.*

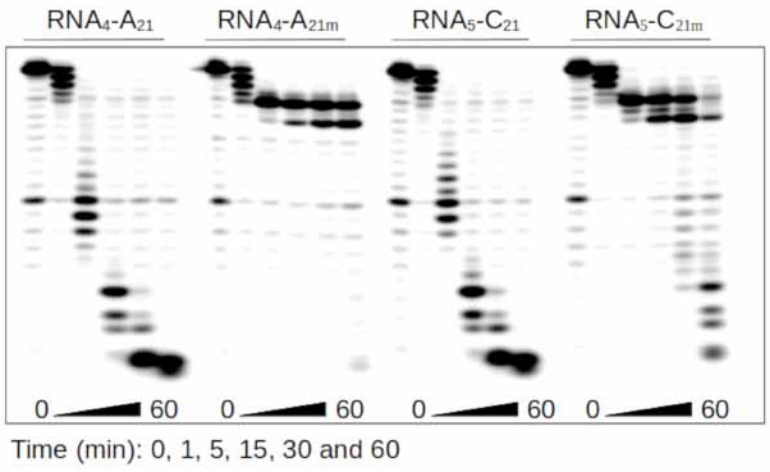

Figure S5: Effect of heteropolymeric RNA 2'O-methylation on ISG20 exonuclease activity. *Evaluation of the ISG20 exonuclease activity on methylated and non-methylated RNA_4-A_{21} and RNA_4-C_{21} monitored by PAGE and autoradiography.*

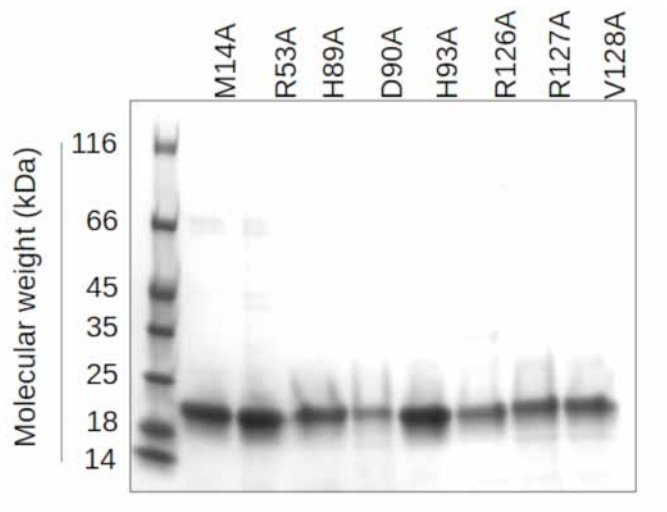

Figure S6: SDS-page of ISG20 RBD mutants. *Mutants of recombinant ISG20 RBD were engineered, expressed in E. coli, and purified on NTa columns. Purified proteins were analyzed on 15% SDS-PAGE gels followed by Coomassie Blue staining.*

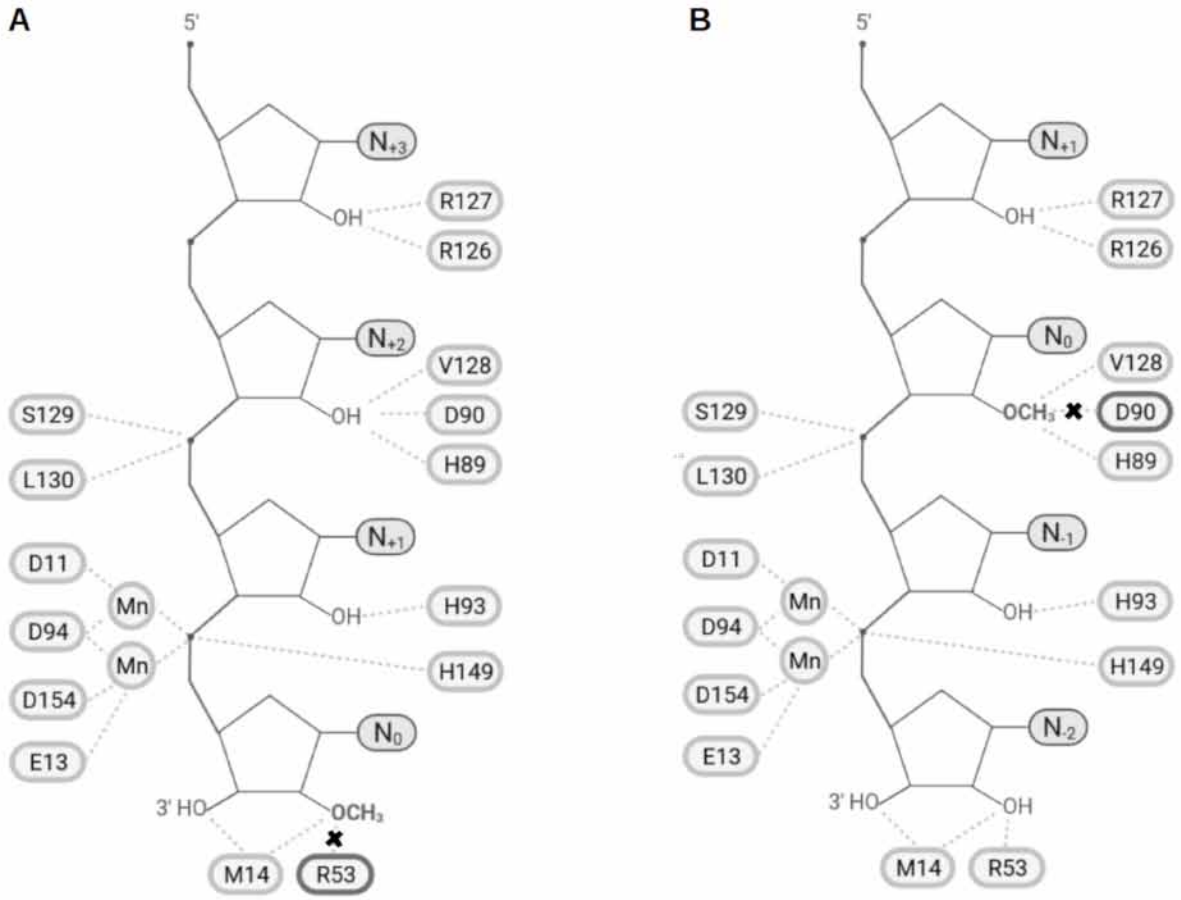

Figure S7: Representation of the N_0 and N_{-2} stops induced by RNA 2'O-methylation on internal residues. *Cartoon model of ISG20 with an RNA carrying 2'O-methylation marks performed on Biorender. A) Depiction of the N_0 stop resulting from the destruction of the interaction (black cross) between the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

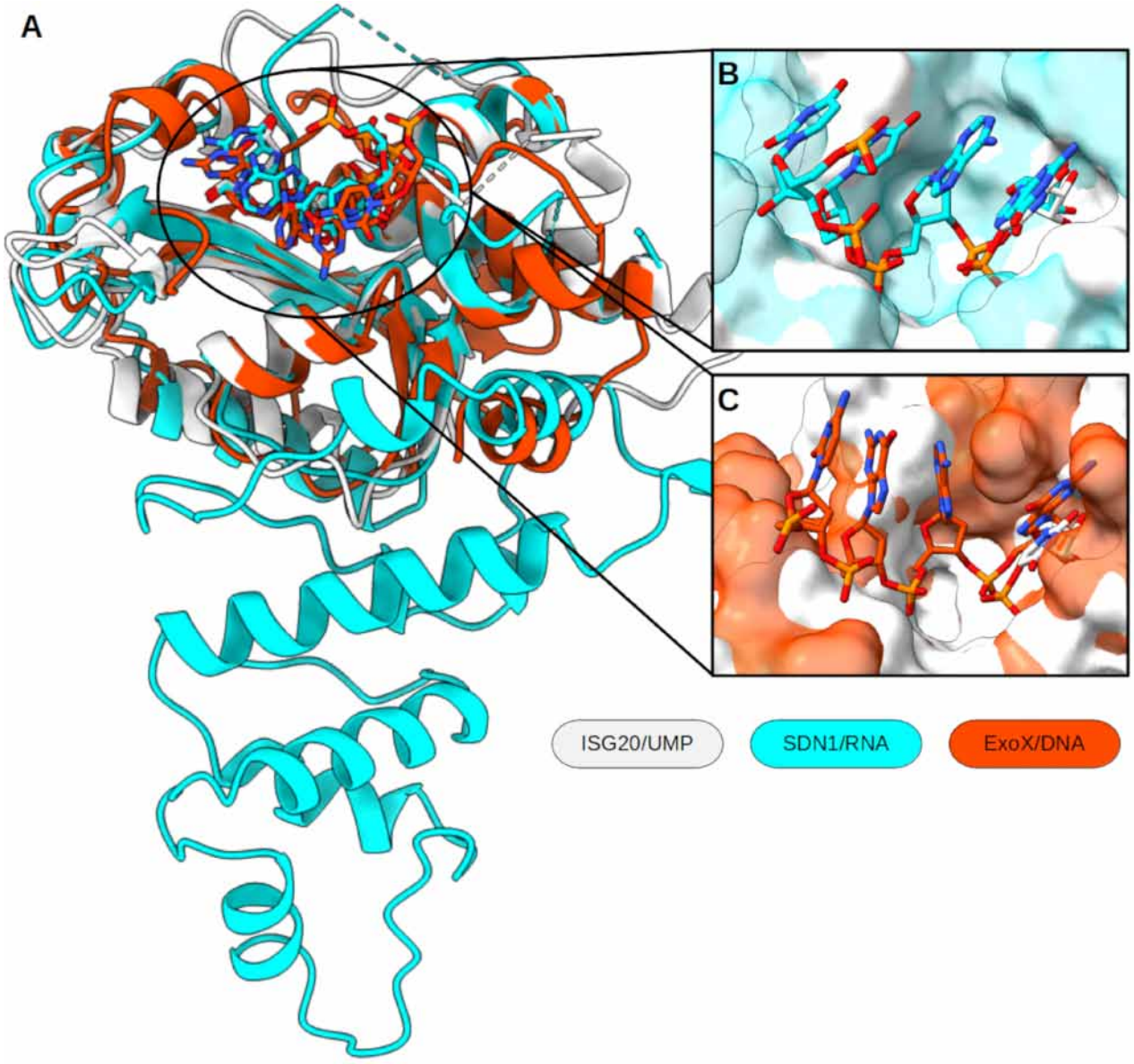

Figure S8: Structural alignment of ISG20 with its exonuclease homologues. *Superimposition of ISG20 (PDB: 1WLJ, in off white), SDN1 in complex with RNA (PDB: 5Z9X, sea green), and ExoX in complex with DNA (PDB: 4FZX, orange) represented in ribbons (A), zoom on the catalytic cavity containing the RNA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

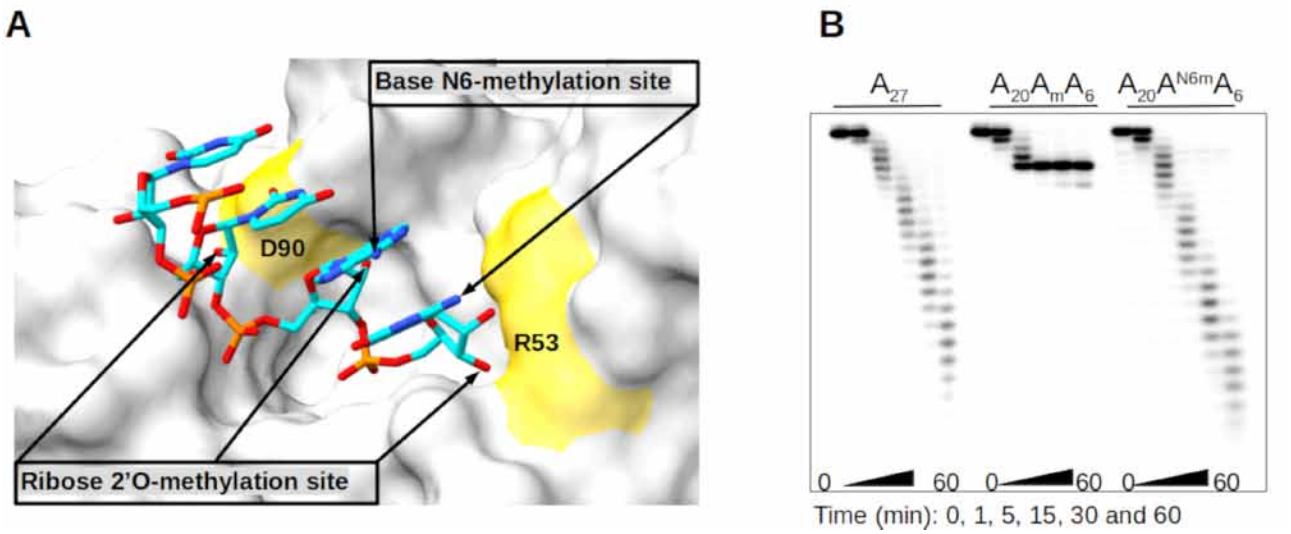

Figure S9: ISG20 exonuclease activity is impaired by RNA 2'O-methylation, but not by N6 methylation. *A) Model of ISG20 in interaction with an RNA substrate built based on the superimposition of ISG20 (PDB: 1WLJ) and SDN1 in complex with RNA (PDB: 5Z9X). Surface representation of ISG20 (off-white) containing 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.*

Chapter 3

Structure-function analysis of the nsp14 N7-guanine methyltransferase reveals an essential role in betacoronavirus replication

Natacha S. Ogando[1#], Priscila El Kazzi[2#], Jessika C. Zevenhoven-Dobbe[1], Brenda W. Bontes[1], Clara C. Posthuma1, Volker Thiel[3,4], Bruno Canard[2], François Ferron[2,3#], Etienne Decroly[2£#] & Eric J. Snijder[1*#]

[1] Department of Medical Microbiology, Leiden University Medical Center, Leiden, The Netherlands, 2333ZA
[2] Architecture et Fonction des Macromolécules Biologiques, Centre National de la Recherche Scientifique, Aix-Marseille Université, Marseille, France, 13288
[3] European Virus Bioinformatics Center, Jena, Germany, 07743.
[4] Institute of Virology and Immunology (IVI), Bern, Switzerland,3350.
[5] Department of Infectious Diseases and Pathobiology, Vetsuisse Faculty, University of Bern, Bern, Switzerland, 3012.

# authors contributed equally
* and £ corresponding authors

PNAS2021 Vol. 118 No. 49 e2108709118

Significance

The ongoing severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) pandemic emphasizes the urgent need to develop efficient broad-spectrum anti-CoV drugs. The structure–function characterization of conserved CoV replicative enzymes is key to identifying the most suitable drug targets. Using a multidisciplinary comparative approach and different betacoronaviruses, we characterized the key conserved residues of the nsp14 (N7-guanine)–methyltransferase, a poorly defined subunit of the CoV messenger RNA–synthesizing machinery. Our study highlights the unique structural features of this enzyme and establishes its essential role in betacoronavirus replication, while identifying two residues that are critical for the replication of the four betacoronaviruses tested, including SARS-CoV-2.

Abstract

As coronaviruses (CoVs) replicate in the host cell cytoplasm, they rely on their own capping machinery to ensure the efficient translation of their messenger RNAs (mRNAs), protect them from degradation by cellular 5′ exoribonucleases (ExoNs), and escape innate immune sensing. The CoV nonstructural protein 14 (nsp14) is a bifunctional replicase subunit harboring an N-terminal 3′-to-5′ ExoN domain and a C-terminal (N7-guanine)–methyltransferase (N7-MTase) domain that is presumably involved in viral mRNA capping. Here, we aimed to integrate structural, biochemical, and virological data to assess the importance of conserved N7-MTase residues for nsp14's enzymatic activities and virus viability. We revisited the crystal structure of severe acute respiratory syndrome (SARS)–CoV nsp14 to perform an in silico comparative analysis between betacoronaviruses. We identified several residues likely involved in the formation of the N7-MTase catalytic pocket, which presents a fold distinct from the Rossmann fold observed in most known MTases. Next, for SARS-CoV and Middle East respiratory syndrome CoV, site-directed mutagenesis of selected residues was used to assess their importance for in vitro enzymatic activity. Most of the engineered mutations abolished N7-MTase activity, while not affecting nsp14-ExoN activity. Upon reverse engineering of these mutations into different betacoronavirus genomes, we identified two substitutions (R310A and F426A in SARS-CoV nsp14) abrogating virus viability and one mutation (H424A) yielding a crippled phenotype across all viruses tested. Our results identify the N7-MTase as a critical enzyme for betacoronavirus replication and define key residues of its catalytic pocket that can be targeted to design inhibitors with a potential pan-coronaviral activity spectrum.

Introduction:

At their 5′ end, all eukaryotic messenger RNAs (mRNAs) carry an N7‑methylguanosine cap that ensures their translation by mediating mRNA recognition during the formation of the ribosomal preinitiation complex. The cotranscriptional capping of cellular pre‑mRNAs occurs in the nucleus and is also critical for pre-mRNA splicing and nuclear export (reviewed in refs. 1–3). The mRNA cap consists of an N7-methylated 5′ guanosine moiety that is linked to the first nucleotide of the transcript by a 5′-5′ triphosphate bridge (4). Its synthesis requires (presumably) the consecutive involvement of triphosphatase, guanylyltransferase (GTAse), and N7-guanine methyltransferase (MTase) activities to produce a cap-0 structure. The first nucleotides of mammalian mRNAs are then methylated on the 2′OH position to yield a cap-1 structure that identifies the transcript as "self" and prevents the activation of innate immune sensors (reviewed in refs. 2 and 5). Furthermore, the cap structure promotes mRNA stability by providing protection from cellular 5′ exoribonucleases (ExoNs).

Viruses rely on host ribosomes for their gene expression and have adopted different strategies to ensure translation of their own mRNAs. These include using the canonical, nuclear capping pathway, so-called "cap-snatching" mechanisms, and the replacement of the cap by a ribosome-recruiting RNA structure (reviewed in refs. 2, 6, and 7). Various cytosolically replicating virus families have evolved their own capping machinery. The latter applies to the coronavirus (CoV) family, which includes the severe acute respiratory syndrome CoV 2 (SARS-CoV-2), the causative agent of COVID-19 (8, 9), and a range of other CoVs infecting human or animal hosts (10, 11). This century alone, the CoV family has given rise to three major zoonotic introductions: SARS-CoV-2, the Middle East respiratory syndrome CoV (MERS-CoV) discovered in 2012, and SARS-CoV emerging in Southeast Asia in 2002. All three belong to the genus Betacoronavirus, which is abundantly represented among CoVs circulating in bat species (12–15). Despite their demonstrated potential to cross species barriers, prophylactic and therapeutic solutions for CoV infections to prevent or rapidly contain the current COVID-19 pandemic were not available.

The positive-sense CoV genome is unusually large (~30 kb), and its 5′ proximal two-thirds encodes for two replicase polyproteins that are posttranslationally cleaved into 16 nonstructural proteins (nsp) (16, 17). The CoV replicative enzymes, including the nsp12 RNA-dependent RNA polymerase, assemble into a protein complex that is embedded within virus-induced replication

organelles (18–20) and directs the synthesis and capping of newly made viral genomes as well as subgenomic mRNAs that serve to express additional CoV genes. Capping is thought to involve the successive action of multiple CoV enzymes: 1) the nsp13 RNA triphosphatase removing the γ-phosphate from the nascent 5′-triphosphorylated RNA (21, 22); 2) an RNA GTAse producing a GpppN cap by transferring guanosine monophosphate (GMP) to the RNA's dephosphorylated 5′ end, a role recently attributed to the nsp12 nucleotidyltransferase domain but remaining to be confirmed (23–25); 3) the nsp14 (N7-guanine)–MTase (N7-MTase) methylating the N7 position of the cap while using S-adenosyl methionine (SAM) as methyl donor; and 4) the nsp16 ribose 2′-O-MTase converting the cap-0 into a cap-1 structure [7mGpppN2′Om (26, 27)] by performing additional methylation with the assistance of nsp10 as cofactor (26, 28, 29).

Over the past 15 y, the CoV-capping machinery has mainly been analyzed in vitro, in particular for SARS-CoV, but its characterization in the context of the viral replication cycle has remained limited to a handful of studies. This applies in particular to the CoV N7-MTase domain, expressed as part of the ~60-kDa nsp14, a bifunctional replicase subunit also containing an N-terminal 3′-to-5′ ExoN domain implicated in promoting the fidelity of CoV replication (30, 31). Following the discovery of an N7-MTase activity associated with nsp14's C-terminal domain (27), the protein was found to methylate non-methylated cap analogs or guanosine triphosphate (GTP) substrates in the presence of SAM in biochemical assays (26, 32, 33). While the association of nsp10 with nsp14 enhances its ExoN activity, the in vitro N7-MTase activity does not depend on nsp10 cofactor (26, 34). Biochemical and structural characterization of the N7-MTase and ExoN domains demonstrated that the two domains are functionally distinct (35–38). Nevertheless, truncations and alanine substitutions in the ExoN domain can severely affect SAM binding and N7-MTase activity (27, 33). The notion that the two enzymatic domains are structurally intertwined was also supported by the SARS-CoV nsp14 crystal structure (35, 36), which was found to be composed of 1) a flexible N-terminal subdomain forming the nsp10 binding site (amino acids [aa] 1 to 58); 2) the 3′-to-5′ ExoN domain (aa 1 to 291); 3) a flexible hinge region consisting of a loop that connects the N- and C-terminal domains and three strands protruding from the C-terminal domain (aa 285 to 300 and aa 407 to 430); and 4) the C-terminal N7-MTase domain (aa 292 to 527) (refs. 35, 36; Figure 1*A*).

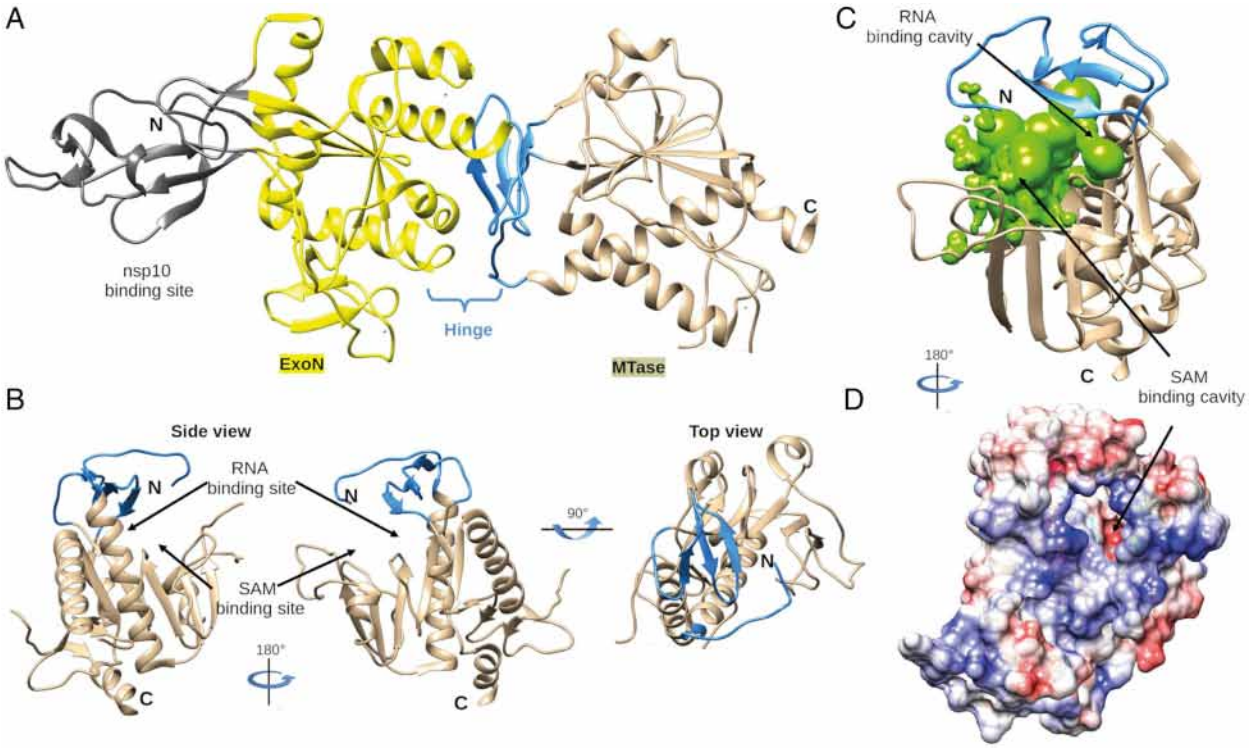

Figure 1: *Global architecture of CoV nsp14. (A) Architecture of SARS-CoV nsp14 (PDB: 5NFY) showing the nsp10-binding site (gray), N-terminal ExoN domain (yellow), hinge subdomain (blue), and C-terminal N7-MTase domain (brown). (B) Side and top view of the hinge region and N7-MTase domain. The three strands of the hinge (blue) protrude from the N7-MTase domain (brown). (C) Analysis of the volume of the N7-MTase active site, with the cavity highlighted in green and the hinge subdomain in blue. (D) Electrostatic surface representation of the CoV nsp14 hinge region and N7-MTase domain. The surface electrostatic potential is calculated by Adaptive Poisson–Boltzmann Solver from −10 (red) to +10 (blue) kT/e.*

Interestingly, the structural analysis of the SARS-CoV-nsp14 N7-MTase revealed a non-Rossmann fold (36), distinguishing this enzyme from commonly known cellular and viral MTases (39, 40). Despite the biochemical characterization of the CoV N7-MTase, the assessment of its importance for virus replication has remained limited to studies with a few point mutations introduced into nsp14 of murine hepatitis virus, a model betacoronavirus (41–44). These studies highlighted two motifs important for CoV replication: 1) the presumed SAM-binding motif I (DxGxPxG/A, with x being any aa; Figure 2*C*, motif III), first discovered by superimposition of an SARS-CoV nsp14 N7-MTase structure model with the crystal structures of cellular N7-MTases (27) and 2) nsp14 residues 420 to 428 (Figure 2*C*, part of motif VI) that, based on the SARS-CoV crystal structure, seem to form a constricted pocket holding the cap's GTP moiety (35). Comparative analysis of N7-MTase domains revealed that a number of residues

crucial for substrate and ligand binding are conserved among homologous enzymes in more distant CoV (Figure 2*A*) and other nidoviruses (45–47).

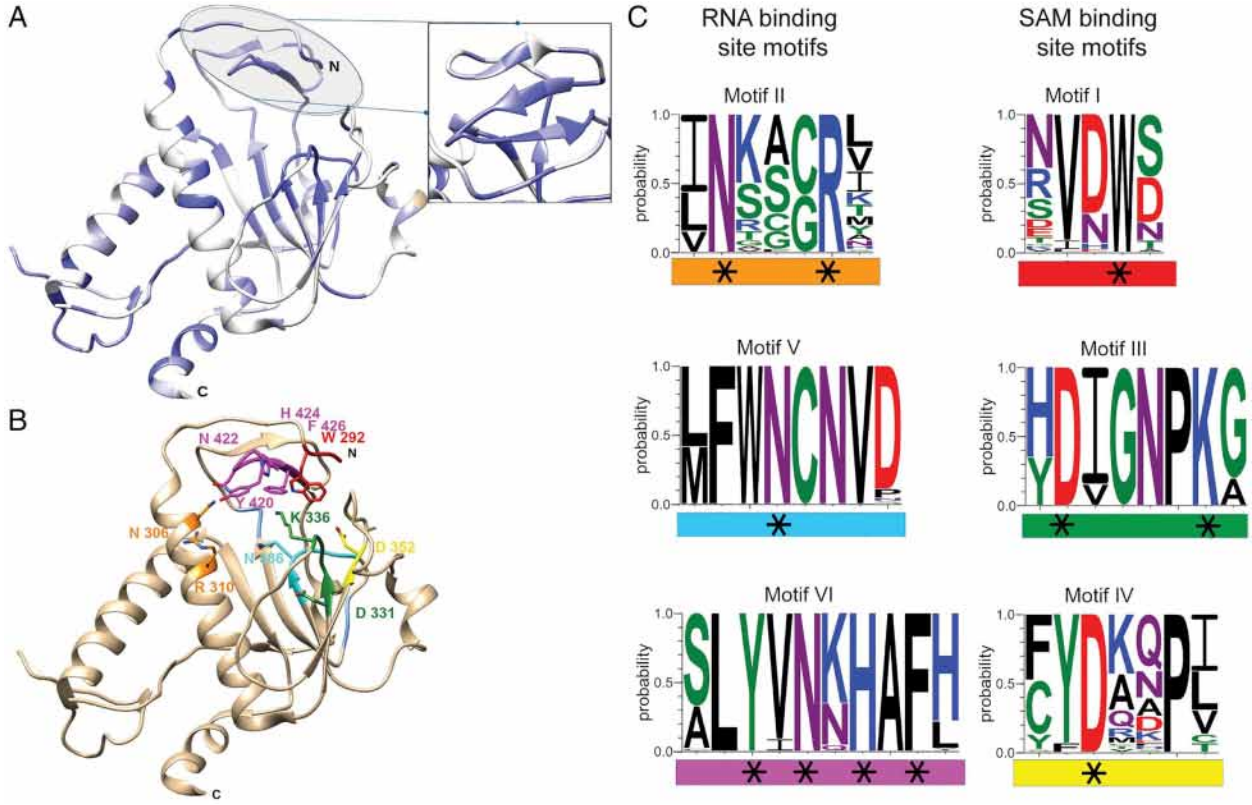

Figure 2: *CoV-wide nsp14 N7-MTase conservation and structural analysis. (A) CoV nsp14 amino acid sequence conservation plotted on the structure (PDB: 5NFY) of the SARS-CoV hinge region and N7-MTase domain (dark blue to white shading representing 100% to less than 50% sequence identity). A list of sequences used for this comparison is presented in SI Appendix, Table S1. (B) Close-up of identified, conserved motifs and residues in the N7-MTase catalytic pocket. (C) WebLogo representation of six conserved motifs (I to VI) identified in the N7-MTase catalytic pocket. Each motif is highlighted with a specific color (matching that in B) and categorized as a proposed SAM- or RNA-binding motif. Black stars highlight charged or aromatic residues most likely involved in ligand binding or catalytic activity.*

Because of its conservation and unique structural features, the CoV N7-MTase constitutes an attractive target for antiviral drug development (48–50) to combat SARS-CoV-2 or future emerging CoV threats. Only a few compounds have been reported to inhibit nsp14 N7-MTase activity in vitro (26, 32, 48–50). Evaluation of their antiviral activity revealed limited inhibition of CoV replication in cell culture, suggesting poor bioavailability and/or specificity (48, 51). Structural, biochemical, and virological studies of CoV N7-MTase structure and function have not been integrated thus far. Here, we set out to define the catalytic pocket, characterize its

involvement in enzymatic activity, and use these observations to probe the enzyme's importance for CoV replication. Using four different betacoronaviruses (SARS-CoV, MERS-CoV, mouse hepatitis virus [MHV], and SARS-CoV-2), we identified conserved features and residues supporting N7-MTase activity and viral replication, thus providing a solid framework for future efforts to design broad-spectrum inhibitors of this critical CoV enzyme.

Results

Identification of Key Residues for RNA and SAM Binding by the CoV N7-MTase.

The previously resolved SARS-CoV nsp14 structure (35, 36) revealed how the ExoN and N7-MTase domains are structurally interconnected, with possible functional implications (Figure 1). Thus far, a structure of nsp14 in complex with 5′-capped RNA is lacking. Due to some structural peculiarities, it was unclear which conserved residues may be mechanistically involved in N7-methylation and how important these may be for overall CoV replication. Therefore, we first revisited the core structure of the SARS-CoV N7-MTase to guide a subsequent biochemical and virological comparison across multiple betacoronaviruses.

In the SARS-CoV nsp14 structure (35), the ExoN core presents a fold characteristic of the DED/EDh family of exonucleases (31, 52, 53). However, the N7-MTase domain does not exhibit the canonical "Rossmann fold" that is common among RNA virus MTases, RNA cap-0 MTases at large, and all five classes of SAM-dependent MTases (54, 55). A hinge region that is highly conserved across CoVs is present at the interface of nsp14's ExoN and N7-MTase domains (Figure 1*A*) and constitutes a unique structural feature of this bifunctional CoV protein. It not only connects the two domains but also forms an extension that protrudes from the surface of the N7-MTase domain (Figure 1*B*). Although the overall structure suggests ExoN and N7-MTase to be separate domains, the successful expression and purification of truncated forms of the N7-MTase domain, with or without the hinge subdomain, have not been reported (27, 56). This might be related to the hydrophobic nature of the hinge, which is likely important for protein stability and folding. Several studies reported that the replacement of ExoN catalytic residues does not impair the N7-MTase activity, suggesting that the functional interplay between the two domains is limited (26, 27, 33, 37, 38, 48). Whereas the hinge region allows lateral and rotational movement of the two nsp14 domains, one side of the hinge also constitutes the "ceiling" of the N7-MTase active site (Figure 1 *B* and *C*).

The structures of SARS-CoV nsp14 in complex with SAM and GpppA [Protein Data Bank (PDB): 5C8S and 5C8T; (35)] have defined the enzyme's cap-binding pocket. However, the crystal packing profoundly constrained the structural characterization of the N7-MTase domain, and the overall low resolution left uncertainties regarding the positioning of the RNA ligand. Therefore, we performed a thorough structural analysis of the enzyme's cavity, supported by CoV-wide nsp14 sequence comparisons, in order to define conserved N7-MTase residues that may be involved in enzymatic activity (Figure S2). Several aspects were taken into consideration while delimiting the SAM and RNA-binding sites: the general geometry of the cavity, its electrostatic properties, and the conservation of specific amino acid residues. We used Surfnet software (57) to define the volume corresponding to the ligand-binding cavity (Figure 1*C*). This volume is shaped as a dual bulb, with the larger pocket accommodating the capped RNA and the smaller one forming the SAM-binding site. An electrostatic surface analysis shows positive charges lining the wall of the putative RNA-binding cavity (Figure 1*D* and *SI Appendix*, Figure S1), which would be consistent with its function. Likewise, positive charges that might accommodate the carbocyclic part of the methyl donor were identified in the SAM-binding pocket (Figure 1*D*). Additionally, conserved hydrophobic residues (Figure 2*C*, motif I) were mapped to a deep hydrophobic cavity, supposedly accommodating the SAM base by a stacking interaction with F426 (SARS-CoV numbering). Finally, the integration of the structural models with CoV-wide N7-MTase sequence comparisons (*SI Appendix*, Table S1 and Figure S2) allowed the identification of conserved potential key residues within each cavity (Figure 2*A*, blue regions). Based on their conservation and positioning, six conserved motifs (I to VI) were defined, each containing a series of specific, charged, or aromatic residues that have their side chain pointing toward the cavity (Figure 2 *B* and *C*). Their features suggested that they can facilitate the methyl transfer from SAM onto the cap's guanine residue at the 5′ end of the RNA substrate by stabilizing and/or correctly positioning the cap structure. The following potential key residues were identified (amino acid numbers matching those in SARS-CoV nsp14): motif I, W292; motif II, N306 and R310; motif III, D331 and K336; motif IV, D352; motif V, N386; and motif VI, Y420, N422, H424, and F426 (Figure 2 *B* and *C*). To assess the possible impact of their replacement on nsp14 folding, we analyzed the predicted impact of single-site substitutions with alanine on the thermostability of SARS-CoV nsp14 (see *SI Appendix*, Table S2 and Figure 5*A*). Except for R310, all replacements yielded positive ΔΔG values, suggesting that these mutations may affect MTase stability by altering either its fold or the cavity for SAM or RNA binding (see *SI Appendix*, Table S2 and Figure 5*A*). Noticeably, mutations in motifs I and VI,

which are spatially close as part of the hinge and most likely involved in the binding of capped RNA, resulted in the largest ΔΔG gains. Similar observations were made when the impact of substitutions with other amino acids was evaluated for other betacoronaviruses (*SI Appendix*, Table S3).

Identification of Residues Crucial for In Vitro N7-MTase Activity.

To experimentally verify the outcome of our structural analysis (Figs. 1 and 2), we probed the functional importance of selected residues through targeted mutagenesis and in vitro N7-MTase assays. Based on their conservation, charge, position, and potential role for RNA or SAM binding in the catalytic pocket (Figure 2 *B* and *C*), 11 and 9 N7-MTase residues were replaced with alanine in recombinant SARS-CoV and MERS-CoV nsp14, respectively. N-terminally H-tagged proteins were expressed in *Escherichia coli* and purified using immobilized metal affinity chromatography followed by size-exclusion chromatography (Figure 3 *A* and *B*).

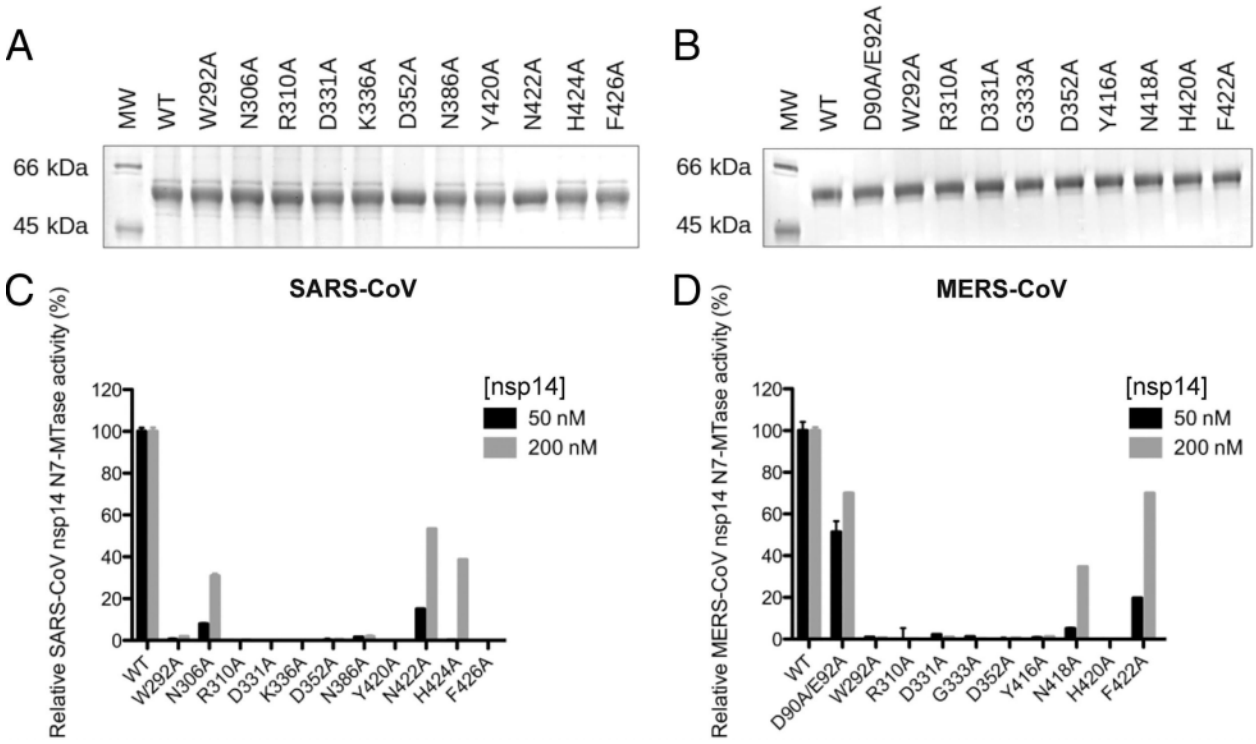

Figure 3: *Expression and in vitro N7-MTase activity of SARS-CoV and MERS-CoV nsp14 mutants. Recombinant SARS-CoV (A) and MERS-CoV (B) wild-type (WT) and mutant nsp14 proteins were expressed in E. coli and purified. Proteins were loaded (2 and 1 μg for SARS-CoV and MERS-CoV, respectively) and analyzed using 10% SDS–polyacrylamide electrophoresis gels stained with Coomassie blue. The in vitro N7-MTase activity of SARS-CoV*

(C) and MERS-CoV (D) nsp14 mutants was determined using an assay with a GpppACCCC synthetic RNA substrate and radiolabeled SAM as methyl donor. Nsp14 concentrations of 50 and 200 nM were used, as indicated. N7-MTase activities were compared to those of the respective wild-type nsp14 controls. For MERS-CoV, ExoN knockout mutant D90A/E92A was included as a control. MW, molecular weight marker.

We evaluated the N7-MTase activity of nsp14 mutants in an assay using a GpppACCCC-capped RNA substrate and radiolabeled [3H]SAM. The transfer of the [3H]methyl group onto the RNA substrate was quantified using filter-binding assays (Figure 3 *C* and *D*), as described previously (26, 34), and compared to the enzymatic activity of wild-type SARS-CoV or MERS-CoV nsp14. With the exception of N306A (30% residual activity), N422A (53% residual activity), and H424A (40% remaining), all SARS-CoV mutations tested almost completely abrogated nsp14 N7-MTase activity (Figure 3*C*). In the case of MERS-CoV nsp14, only mutants N418A and F422A retained partial N7-MTase activity, 34% and 70%, respectively, while again all other mutations rendered the enzymatic activity barely detectable (Figure 3*D*). In terms of residual activity, differences were observed for some pairs of equivalent SARS-CoV and MERS-CoV mutants (e.g., the H and F in motif VI), but overall, the results were fully in line with the outcome of our structural analysis. Thus, our data confirmed and extended a previous study (35) and showed that N7-MTase activity is affected by mutations that either may inhibit SAM binding (W292A, D331A, G333A, K336A, and D352A in SARS-CoV) or likely interfere with RNA chain stabilization (N306A, R310A, Y420A, N422A, and F426A) in the catalytic pocket.

Revisiting the Interplay between the N7-MTase and ExoN Domains of nsp14.

Despite the notion that the ExoN and N7-MTase domains of CoV nsp14 may be functionally independent (27, 33, 35, 36), they are structurally interconnected by the hinge region (Figure 1). Therefore, we evaluated the impact of all of our N7-MTase mutations on ExoN functionality, using an in vitro assay with 5′-radiolabeled RNA substrate H4 (34), a 22-nucleotide RNA of which the largest part folds into a hairpin structure. Its degradation was monitored using denaturing polyacrylamide gel electrophoresis and autoradiography (Figure 4). Nsp10 was added as a cofactor that importantly stimulates nsp14 ExoN activity (34, 35, 37), as again confirmed in the "nsp14 only" control assay (Figure 4). As expected, in time course experiments, we observed the progressive 3′-to-5′ degradation of the RNA substrate by the wild-type nsp10-nsp14 pair of both SARS-CoV (Figure 4*A*) and MERS-CoV (Figure 4*B*). In the same assay, most of our N7-MTase mutations barely affected ExoN activity (Figure 4 *A* and *B*), also supporting the

notion that these mutant proteins had folded correctly. In contrast, the ExoN activity of SARS-CoV mutants R310A, Y420A, and H424A and MERS-CoV mutant W292A was strongly or partially affected, as indicated by the reduced amount of hydrolysis products at the bottom of the gel. Meanwhile, incorporation of the H420A mutation completely abrogated MERS-CoV nsp14 ExoN activity. Four of the five mutations (Y420A and H424A in SARS-CoV, and W292A and H420A in MERS-CoV) that affected ExoN activity mapped to the hinge region (Figure 2, motif I and VI). Based on the structural analysis, we assume that these mutations affect either the overall nsp14 folding or—more likely—constrain the flexibility of the hinge subdomain with negative consequences for ExoN functionality (35, 36). Conversely, a MERS-CoV ExoN knockout mutant (D90A/E92A), which was included as a control, was found to modestly impact N7-MTase activity (Figure 3*D*). Taken together, our data suggest that although the N7-MTase sequence is well conserved among betacoronaviruses [(35, 37); *SI Appendix*, Figure S2], the differences observed between SARS-CoV and MERS-CoV must be caused by a certain level of structural variability or differences in recombinant protein stability.

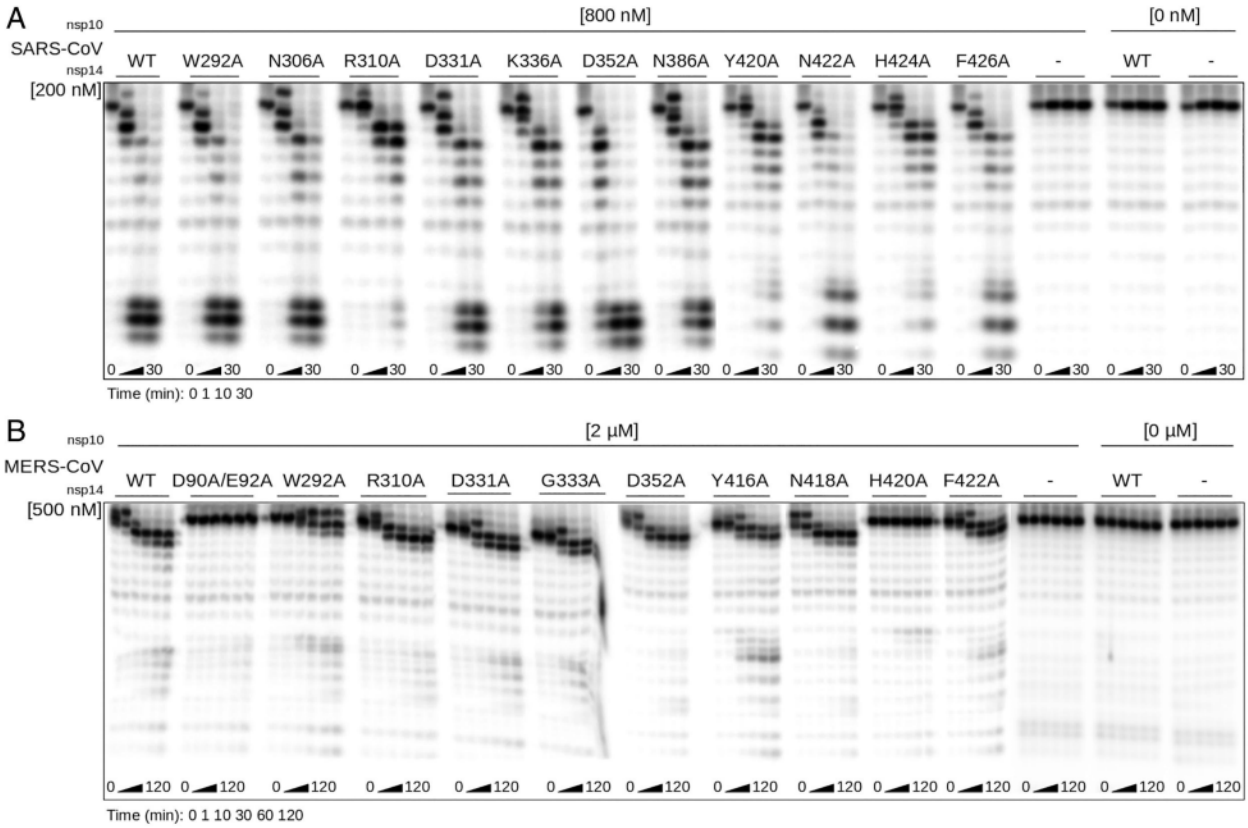

Figure 4: *In vitro ExoN activity of SARS-CoV and MERS-CoV N7-MTase mutants. The in vitro ExoN activity of SARS-CoV (A) and MERS-CoV (B) mutant nsp14 proteins (Figure 3) was determined by monitoring the degradation of a 5′-radiolabeled RNA substrate (see Materials and Methods). An nsp14 concentration of 200 or 500 nM was*

used (as indicated) and a fourfold molar excess of the corresponding nsp10 was added. A time course assay was performed using time points 0, 1, 10, and 30 min for SARS-CoV, and 0, 1, 10, 30, 60, and 120 min for MERS-CoV nsp14. Reaction products were analyzed by denaturing gel electrophoresis and autoradiography. WT, wild type control.

The nsp14 N7-MTase Is Critical for SARS-CoV Viability.

As summarized in the Introduction and *Results*, most prior biochemical and structural studies of the CoV N7-MTase were performed using SARS-CoV nsp14, whereas mutagenesis, in the context of virus replication (using reverse genetics), was restricted to MHV studies in which, for different reasons, the conserved D and G residues in motif III and the Y residue in motif VI were targeted (41, 43, 58). To establish a connection between the biochemical and virological data on the N7-MTase, we first introduced 12 single N7-MTase mutations into the SARS-CoV genome, using a bacterial artificial chromosome (BAC)–based reverse genetics system. Each mutant was engineered in duplicate and launched by in vitro–transcribing full-length RNA that was electroporated into baby hamster kidney (BHK-21) cells. To propagate viral progeny, if released, transfected BHK-21 cells were mixed with Vero E6 cells and incubated up to 6 d. Each mutant was launched at least four times, using RNA from two independent clones in two independent experiments, and mutant phenotypes are summarized in Figure 5*A*.

A

Coronavirus N7-MTase motif	I	II		III			IV	V	VI			
Conserved residue replaced with Alanine	W	N	R	D	G	K	D	N	Y	N	H	F
nsp14 aa # SARS-CoV and SARS-CoV-2	**292**	**306**	**310**	**331**	**333**	**336**	**352**	**386**	**420**	**422**	**424**	**426**
nsp14 aa # MERS-CoV	292		310	331	333		352		416	418	420	422
nsp14 aa # MHV	291		309	330	332				414		418	420
Predicted impact of Ala substitution on **SARS-CoV nsp14 fold** (ΔΔG in kcal/mol)†	1.8	0.72	0	0.46		0.38	1.12	0.99	3.35	0.92	2.21	1.57
Predicted to interact with	SAM	RNA		SAM			SAM	RNA	RNA			
Mutant *in vitro* N7-MTase activity SARS-CoV*	--	+/-	--	--		--	--	--	--	+	+/-	--
Mutant *in vitro* N7-MTase activity MERS-CoV*	--		--	--	--		--		--	+/-	--	+
Mutant *in vitro* ExoN activity SARS-CoV**	+	+	+/-	+		+	+	+	+/-	+	+/-	+
Mutant *in vitro* ExoN activity MERS-CoV**	+/-		+	+	+		+		+	+	--	+
Virus mutant viability SARS-CoV§	+/-	++	--	--	--	++/-	+/-	--	--	++	+/-	--
Virus mutant viability MERS-CoV§	--		--	+/-	+/-		++/-		+/-	++	+/-	--
Virus mutant viability MHV§	--		--	++	--				--		++/-	--
Virus mutant viability SARS-CoV-2§			--								++/-	--

B

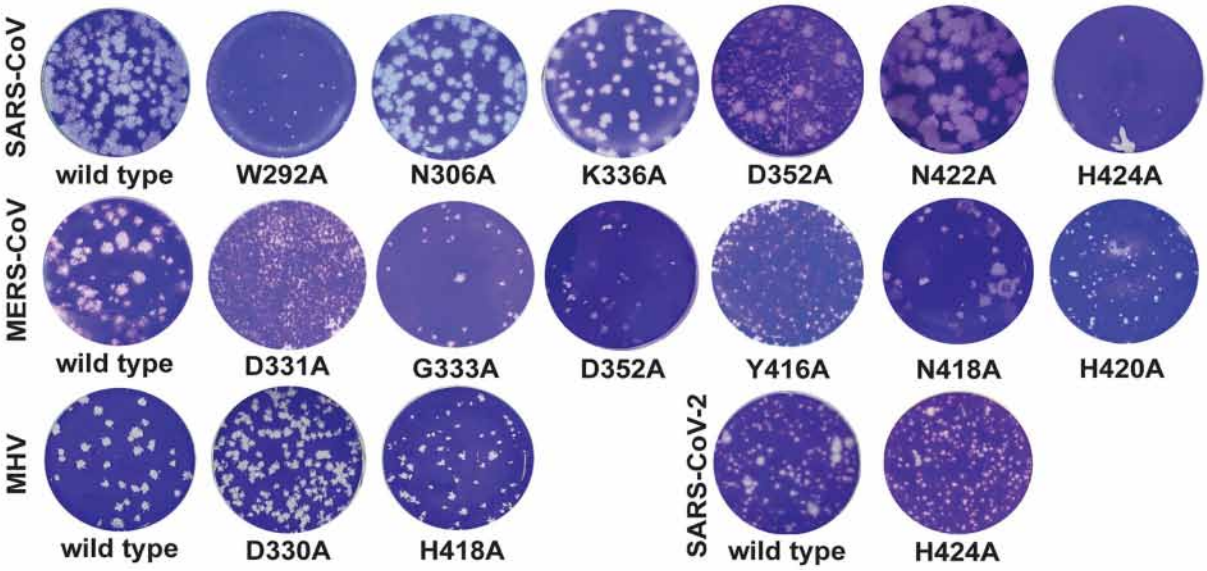

Figure 5: *Virological characterization of betacoronavirus N7-MTase mutants. (A) Summary of results obtained from in silico, biochemical, and virological studies of CoV mutants. † represents ΔΔG values from SI Appendix, Table S2. * represents N7-MTase activity of each mutant compared to the wild-type control enzyme and scored +, +/-, or -- when >50%, between 10 and 50%, or <10%, respectively. Values used here correspond to results obtained using the high-enzyme concentration (Figure 3 B and D, 200 nM). ** represents ExoN activity of each mutant evaluated relative to the wild-type control enzyme and scored +, +/-, or -- for equal, reduced, and abolished ExoN activity, respectively. § represents mutant virus phenotypes, as deduced from plaque assays, scored as: --, nonviable; +/-, severely crippled; ++/-, mildly crippled; and ++, similar to the wild-type control. Empty cells indicate mutants that were not generated. (B) Plaque phenotype of the progeny of viable N7-MTase mutants. Plaque assays were performed using supernatants harvested from transfected cells at 3- (MERS-CoV in HuH7 cells and SARS-CoV and SARS-CoV-2 in Vero E6 cells) or 4-d (MHV in 17Cl1) post transfection.*

In line with the biochemical data, the nonviable phenotype of 6 of the 12 SARS-CoV mutants (Figure 5*B*) provided clear support for the importance of key residues in N7-MTase motifs II

(R310), III (D331 and G333), V (N386), and VI (Y420 and F426). As anticipated, mutations in the canonical, SAM-binding motif III (DxGxPxG/A) completely abrogated SARS-CoV replication (Figure 5*A*), apparently confirming the critical role of D331, which was postulated to be a key residue for methylation upon the discovery of the CoV N7-MTase (27). On the other hand, D331A was the only non-viable SARS-CoV mutant for which reversion to wild-type was occasionally observed, suggesting that a very low level of viral RNA synthesis remained possible despite this mutation (see also *Discussion*).

Remarkably, SARS-CoV mutations N306A, K336A, and N422A in motifs II, III, and VI, respectively, were found to yield viruses with plaque phenotypes and progeny titers similar to those of the wild-type control (Figure 5), despite the major impact of these mutations on in vitro N7-MTase activity (Figure 3*C*). Likewise, the viable but severely crippled (small plaque) virus phenotypes of motif I mutant W292A and motif VI mutant H424A were surprising (Figure 5*B*), although for the latter the biochemical assays did reveal some activity when performed with an increased enzyme concentration [Figure 3*C*; (35)]. Interestingly, mutant D352A yielded a mixed-size plaque phenotype, suggesting rapid (pseudo)reversion in a minor fraction of this mutant's progeny (Figure 5*B*). For all six viable mutants, the presence of the original mutation in the viral progeny was confirmed by sequence analysis of the full-length, nsp14-coding region of the viral genome. No other mutations were detected in this region of the genome. For non-viable mutants, transfected cells were incubated and monitored for 6 d and absence of viral activity was also confirmed by immunofluorescence microscopy with antibodies specific for double-stranded RNA and SARS-CoV nsp4.

In general, our data demonstrated the importance of the N7-MTase domain for SARS-CoV viability and confirmed the importance of the motifs and key residues identified using structural biology and biochemical approaches (summary presented in Figure 5*A*). Nevertheless, for several mutants, the data from different types of assays did not readily align, which prompted us to expand the reverse genetics efforts to other betacoronaviruses.

Phenotypic Differences between Betacoronaviruses N7-MTase Mutants Suggest Complex Structure–Function Relationships.

Even when targeting highly conserved viral functions, the introduction of equivalent mutations in closely related viruses can sometimes yield remarkably different mutant phenotypes. A recent example is the inactivation of the nsp14 ExoN, which is tolerated by MHV and SARS-CoV but not by MERS-CoV and SARS-CoV-2, the latter virus having an nsp14 sequence that is 95%

identical to that of SARS-CoV (37). To expand our understanding of the impact of N7-MTase mutagenesis, we engineered, launched, and analyzed a set of MERS-CoV and MHV mutants using technical procedures similar to those described for SARS-CoV (*Materials and Methods*). In this case, the production of viable progeny was facilitated by coculturing transfected BHK-21 cells with host cells appropriate for the amplification of MHV (17clone1 cells) or MERS-CoV (Huh7 cells). Again, each mutant was launched at least four times (from duplicate full-length copy DNA [cDNA] clones) and the results are summarized in Figure 5.

The mutations tested for MERS-CoV and MHV had a large, predicted impact in our folding free-energy analysis (*SI Appendix*, Tables S2 and S3) and/or yielded a nonviable or crippled phenotype in our SARS-CoV study (Figure 5*A*). We evaluated whether these residues were equally critical for the replication of other betacoronaviruses. For clarity, from this point forward, we will refer to the conserved key residues of each motif instead of using nsp14 amino acid numbers, which are slightly different when comparing SARS-CoV, MERS-CoV, and MHV (Figure 5*A*).

In contrast to the SARS-CoV result, the replacement of the W in SAM-binding site motif I was lethal for both MERS-CoV and MHV. Strikingly, the mutagenesis of the D and G in motif III (SAM-binding site) yielded the opposite outcome: Both were not tolerated in SARS-CoV but resulted in crippled but viable or even wild-type–like phenotypes for MERS-CoV and MHV, respectively (Figure 5*B*). These results again indicated that CoV N7-MTase active site mutants can be (partially) viable, even in the absence of detectable in vitro enzymatic activity (Figure 3*D*). Similar to our observations for SARS-CoV, the replacement of the D in motif IV and the N in motif VI had moderate or no impact, respectively, on the production of MERS-CoV progeny (Figure 5). Replacement of the conserved H in motif VI (RNA-binding site) consistently crippled replication across SARS-CoV, MERS-CoV, and MHV (Figure 5*B*), while replacement of the conserved Y in the same motif was partially tolerated by MERS-CoV but not by SARS-CoV and MHV.

Our betacoronavirus comparison identified only two N7-MTase mutations that consistently abrogated the replication of all three viruses tested: the R-to-A in motif II and the F-to-A in motif VI, which both map to the putative RNA-binding site. This was surprising in the case of MERS-CoV, given the fact that this mutation (F422A in MERS-CoV) allowed substantial N7-MTase activity in the in vitro assay (Figure 3*D*). When SARS-CoV-2 emerged during the course of this study, the three mutations that produced a similar phenotype across SARS-CoV, MERS-CoV, and MHV (R310A, H424A, and F426A, using SARS-CoV numbering) were also

engineered for this newly discovered CoV. Again, the R310A and F426A replacements were found to fully abrogate virus replication, while H424A yielded a crippled phenotype in SARS-CoV-2 (Figure 5).

Discussion

Most viral MTases belong to the Rossmann fold family (55, 59), a ubiquitous, higher-order structure among dinucleotide-binding enzymes (55, 60). The CoV nsp14 N7-MTase was the first identified example of a non-Rossmann fold viral MTase (35, 36, 45), and the only one thus far for which some structural and functional information had been gathered. While some viral N7-MTase crystal structures have been resolved (35, 36, 61–63), their biochemical properties and signature sequences critical for RNA binding or enzymatic activity remain poorly defined compared to, for example, the 2′-O-MTases, an example of which is found in CoV nsp16 (reviewed in ref. 6). Likewise, the biological role and relevance of the CoV N7-MTase have not been explored in much detail. In recent studies and reviews, often related to SARS-CoV-2, the enzyme is widely assumed to secure the translation of CoV subgenomic mRNAs and genome, which obviously is a critical step for any positive-stranded RNA virus. However, direct biochemical evidence showing that CoV mRNAs indeed carry an N7-methylated cap at their 5′ end is still lacking. The presence of such a cap on CoV RNAs was first postulated following RNase T1 and T2 digestion studies with 32P-labeled MHV RNA 40 y ago (64). Additional support came from immunoprecipitation experiments using a cap-specific monoclonal antibody (recognizing both the rare nucleoside 2,2,7-trimethylguanosine and 7-methylguanosine cap structures) that brought down the mRNAs of equine torovirus (65), a distant CoV relative for which—perhaps strikingly—an N7-MTase domain still remains to be identified (45). The presence of enzymes required for capping in CoVs and many of their relatives (6, 17, 45, 47, 66), and the in vitro activity profile of recombinant CoV nsp14 (26, 27, 32, 33, 37, 38) lend additional credibility to CoV capping and cap methylation but do not exclude the possibility that the CoV N7-MTase may target other substrates as well.

To enhance our overall understanding of nsp14 N7-MTase structure and function, also in the light of its emergence as an important drug target in the battle against SARS-CoV-2 (50, 67–69), we now revisited the SARS-CoV nsp14 X-ray structure to define the most likely residues involved in N7-MTase substrate binding and catalysis. Instead of a βαβ architecture (a seven-stranded β-sheet surrounded by six α-helices) and the canonical MTase motifs, the CoV N7-MTase incorporates 12 β-strands and five α-helices that form a five-stranded β-sheet core

(36, 45). The overall nsp14 structure reveals two domains interconnected by a hinge that may confer the flexibility needed to orchestrate the different functions of the protein during CoV replication (36). Furthermore, the protein binds to nsp10, a critical cofactor for nsp14's ExoN activity (34, 70). The conversion of a 5′-terminal GMP cap (GpppN) into a cap-0 structure (7mGpppN) involves multiple steps: stabilization of the RNA chain, SAM binding, methyl transfer to the N7 position of the cap, release of the methylated RNA substrate, and S-adenosyl-homocysteine (SAH) release. Our structural analysis identified several residues with their side chains pointing toward the catalytic pocket, which could be classified as likely RNA- or SAM-binding motifs (Figure 2 *B* and *C*). Taking into account the amino acid sequence conservation between MHV, SARS-CoV, SARS-CoV-2, and MERS-CoV (Figure 2*A* and alignment in *SI Appendix*, Figure S2) and the structures available to date (35, 36, 71, 72), we surmised these CoV N7-MTases to have an overall similar fold and structural organization. The impact of alanine substitutions of selected key residues in these motifs was then evaluated both in vitro, using SARS-CoV or MERS-CoV recombinant nsp14, and in the context of the viral replication cycle, by engineering the corresponding virus mutants in different betacoronaviruses.

Although the biochemical and virological data presented in this study clearly provide support for the predictions derived from our structural analysis, the overall interpretation of the dataset undeniably is much more complex than anticipated (Figure 5*A*). Replacement of conserved SARS-CoV and MERS-CoV N7-MTase residues largely or completely abrogated enzymatic activity in vitro (Figure 3 *C* and *D*), supporting their identification as key residues for the enzyme's functionality when the protein is expressed alone (N7-MTase activity) or when tested in complex with nsp10 (ExoN activity). However, for several SARS-CoV and MERS-CoV mutations, the data on enzymatic activity in vitro and virus mutant viability appeared to be at odds with each other (Figure 5*A*). One possible interpretation is that (very) low levels of N7-MTase activity may still suffice to support viral replication in cell culture models. Alternatively, the in vitro N7-MTase assays may have suffered from technical complications, such as suboptimal or incorrect (mutant) N7-MTase domain folding. This could be different for nsp14 expressed in the context of the virus-infected cell and in the presence of its natural interaction partners, in particular other members of the viral replication and transcription complex. It is conceivable that the impact of nsp14 mutations on the fold and/or critical protein–protein or protein–RNA interactions of the N7-MTase domain could fluctuate between different assay systems. This might explain a stronger (e.g., MERS-CoV mutant F422A) or less dramatic effect in the virus-infected cell compared to what is observed in enzymatic assays

(Figure 5*A*). Mutations mapping to motif VI (hinge region) yielded inconsistent results in comparison to prior in vitro studies (26, 27, 32–35), which might be attributed (in part) to different in vitro assay conditions. Such technical explanations, however, do not apply when introducing equivalent substitutions in different betacoronaviruses and evaluating them in the context of the viral replication cycle. Also, here apparent inconsistencies were observed in terms of the variable impact of certain mutations on the overall replication of virus mutants. The results obtained with mutations in motif III (the presumed SAM-binding motif DxGxPxG/A) were a striking example: The viral phenotype for the D-to-A mutant (D331A in SARS-CoV and MERS-CoV and D330A in MHV) ranged from nonviable for SARS-CoV, via severely crippled for MERS-CoV, to wild-type–like for MHV (Figure 5). SARS-CoV residue D331 was first identified as important for N7-MTase activity by the superimposition of nsp14 with cellular N7-MTase structures (27). However, a previous MHV study (43) had already documented that the replacement of the corresponding residue D330 did not affect MHV replication and pointed to G332 as a more important residue in motif III, which was confirmed in this study (Figure 5). These results are consistent with the SARS-CoV nsp14 crystal structure showing that residue G333 in the DxG motif (G332 in MHV) is in direct contact with the SAM methyl donor (35), although apparently its replacement is not sufficient to render all betacoronaviruses nonviable. These results stress the importance to achieve a series of high-resolution structures of these different proteins in order to determine the subtle mechanistic differences.

The only other N7-MTase position probed by reverse genetics so far was the conserved tyrosine in motif VI (Figure 2*C*; Y414 in MHV). This residue attracted attention by the intriguing, serendipitous finding that its replacement with histidine did not affect the replication of MHV strain A59 in cell culture but strongly reduced replication and virulence in mice (41). Also, a Y414A substitution was tolerated in MHV-A59 (44, 58), but in our study, Y414A prevented the recovery of infectious progeny for MHV strain JHM, which exhibits less robust RNA synthesis and overall replication than MHV-A59. The results for the corresponding SARS-CoV (nonviable) and MERS-CoV (crippled) mutants were also variable, adding to the complexity of the overall picture.

A substantial set of N7-MTase mutations was monitored for “side effects” at the level of in vitro ExoN activity (Figure 4), although for SARS-CoV and MHV these would unlikely explain a lack of viability as ExoN knockout mutants, for both these viruses are only mildly crippled (42, 58, 73). Strikingly, for MERS-CoV, which does not tolerate ExoN inactivation (37), two of the N7-MTase mutations (G333A in motif III and H420A in motif VI) abolished detectable ExoN

activity in vitro (Figure 4*B*) but still allowed a certain level of virus replication (small-plaque phenotype), an observation that clearly warrants further investigation. In more general terms, the ExoN biochemical assay (Figure 4) suggested that the functional separation between the two enzyme domains may be less strict than previously concluded, as also recently hypothesized following an in silico and biochemical analysis using SARS-CoV-2 nsp14 ExoN domain (72, 73). Alternatively, structural variation may explain the discrepancies observed. The impact of SARS-CoV N7-MTase motif VI mutations on ExoN activity was major, highlighting the peculiar structural organization of nsp14, in which part of the N7-MTase substrate-binding cavity maps to the hinge that connects the N7-MTase and ExoN domains (Figure 1). For other N7-MTase motifs probed, the functional separation from ExoN was confirmed, as also deduced from previous studies (27, 33, 35, 38).

In our reverse genetics studies with four betacoronaviruses, a consistent phenotype was observed only for N7-MTase mutants carrying replacements of the conserved R in motif II (nonviable) and the conserved H and F in motif VI (crippled and nonviable, respectively). SARS-CoV residue R310 was previously reported to play a role in SAM binding (33), whereas F426 was proposed to entrench and stabilize the guanosine's purine moiety in the proximity of SAM (35). Our analysis (Figure 2) redefined both residues as part of putative RNA-binding site motifs II and VI, respectively, and they were found to be essential for in vitro N7-MTase activity in SARS-CoV. Our results highlight the importance of the nsp14 N7-MTase for CoV replication, but the variable impact of the replacement of several conserved residues suggests a substantial degree of conformational or functional flexibility in the enzyme's active site. Other factors, such as interactions of nsp14 with other replicase subunits, may also contribute to the observed phenotypic differences between equivalent N7-MTase mutants of different betacoronaviruses. Likewise, the translation of in vitro N7-MTase activity to virus viability is not straightforward and suggests complex structure–function relationships for the structurally unique CoV N7-MTase. Given both its essential role in CoV replication and its emerging status as a target for antiviral drug development efforts, it will be important to further expand the integrated biochemical and virological analysis to support the rational design of broad-spectrum inhibitors of the CoV N7-MTase.

Materials and Methods

Bioinformatics Analysis.

A total of 47 CoV nsp14 sequences were retrieved (a complete list is provided in *SI Appendix*, Table S1) and aligned using MAFFT. Delineation of motif I to VI was done manually using Seaview and WebLogo (74, 75). Structure analysis [PDB: 5NFY; (36)], volume estimation, cavity determination, and sequence conservation were plotted onto the structure using UCSF Chimera (76). Electrostatic surface calculations were done using APBS (77). Predicting the structural impact of mutations was done using the PoPMuSiC server (dezyme.com/en/software) (78). This program introduces single-site mutations into a protein's structure and estimates the change in ΔΔGs values of such mutations. In the next step, all possible single-site mutations (4,731 mutations) were sorted by their ΔΔGs, but only those in the conserved motifs in the vicinity of the catalytic pocket were used for further studies. PopMuSic predictions were cross-validated with SNAP2 to assess the impact of single–amino acid substitutions on protein function (79).

Recombinant Protein Expression and Purification.

Recombinant SARS- and MERS-CoV nsp10 and nsp14 were expressed in *E. coli* and purified, as described previously (26), MERS-CoV-nsp14 (37, 49) and MERS-nsp10 (29, 80). Vectors for mutant nsp14 expression were generated by QuikChange site-directed mutagenesis using Accuzyme DNA polymerase (Bioline) and verified by sequence analysis. For each recombinant protein used, two batches were produced and tested in enzymatic assays.

In Vitro nsp14 N7-MTase Activity Assay.

Reaction mixtures contained 50 or 200 nM of SARS-CoV or MERS-CoV recombinant nsp14, 7 nM GpppACCCC synthetic RNA substrate, 40 mM Tris·HCl (pH 8.0), 10 mM DTT, 5 mM MgCl2, 1.9 μM SAM, and 0.1 μM 3H-SAM (PerkinElmer). After a 30-min incubation at 30 °C, the assay was stopped by the addition of a 10-fold volume of ice-cold 100 μM SAH (Thermo Fisher Scientific). Samples were spotted on DEAE filter mats (PerkinElmer) and washed twice with 10 mM ammonium formate (Sigma-Aldrich) (pH 8.0), twice with MilliQ water, and once with absolute ethanol (Sigma-Aldrich) (26), and MTase activity was quantified using a Wallac scintillation counter. To determine relative enzyme activities, the incorporation measurements for

mutant proteins were normalized to values obtained with wild-type nsp14. Samples were measured in triplicate in each experiment.

In Vitro nsp14 ExoN Assay.

Synthetic RNA substrate H4 (34) was radiolabeled at its 5′ end using T4 polynucleotide kinase (Epicentre) and [γ-32P]ATP (PerkinElmer) and used as substrate in ExoN activity assays. To this end, recombinant SARS-CoV or MERS-CoV nsp14 and nsp10 were mixed in a 1:4 concentration ratio of nsp14:nsp10, as indicated in Figure 4. The proteins were added to 500 nM radiolabeled substrate in reaction buffer (40 mM Tris·HCl [pH 7.5], 5 mM MgCl2, and 1 mM DTT). The protein mix was left for 10 min at room temperature to allow the formation of the complex. Assays were performed at 37 °C and stopped by addition of a 3× volume of loading buffer containing 96% formamide and 10 mM EDTA. Samples were analyzed on 7 M urea containing 14% (weight/volume) polyacrylamide gels (acrylamide/bisacrylamide ratio, 19:1) buffered with 0.5× Tris-taurine-EDTA and run at high voltage (1,600 V). Results were visualized by phosphorimaging using a Typhoon-9410 variable-mode scanner (GE Healthcare).

Cell Culture.

BHK-21 cells (ATCC CCL10), Vero E6 (ATCC; CCL-81), HuH7 cells, and mouse 17 Cl1 cells were grown as described previously (19, 37, 81, 82). In order to amplify viral progeny and titrate recombinant CoVs by plaque assay, Vero E6 cells were used for SARS-CoV and SARS-CoV-2, HuH7 cells for MERS-CoV, and 17Cl1 cells for MHV. Cells were cultured in Eagle's minimal essential medium (EMEM; Lonza) with 8% fetal calf serum (FCS; Bodinco) supplemented with 100 IU/mL penicillin, 100 µg/mL streptomycin (Sigma), and 2 mM L-Glutamine (PAA Laboratories). After infection, complete EMEM medium containing 2% FCS was used.

Viruses and Reverse Genetics.

Mutations in the nsp14-coding region were engineered by two-step en passant recombineering in *E. coli* (83) using a BAC vector with a full-length cDNA copy of a betacoronavirus genome. Virus isolates used were MERS-CoV strain EMC/2012 (84, 85), SARS-CoV Frankfurt-1 (86), MHV-JHM-IA (87), and SARS-CoV-2 BetaCoV/Wuhan/IVDC-HB-01/2019 (88). When designing mutations, additional, translationally silent marker mutations were introduced near the site of mutagenesis, in order to analyze possible reversion and rule out potential contaminations with parental virus. For each mutant, two independent BAC clones were obtained, verified by

sequencing of the full-length nsp14-coding region, and used for in vitro transcription (mMessage-mMachine T7 Kit; Ambion) and virus launching. Transfections with full-length RNA transcripts were performed as described before (37). Briefly, 5 μg RNA was electroporated into BHK-21 cells using an Amaxa nucleofector 2b (program A-031) and Nucleofection T solution kit (Lonza). Transfected BHK-21 cells were mixed in a 1:1 ratio with cells susceptible to CoV infection: Vero E6 cells (for SARS-CoV and SARS-CoV-2), HuH7 cells for MERS-CoV, or 17Cl1 cells (for MHV). Cell culture supernatants were collected when full cytopathic effect was observed, or at 6 d, post-transfection and progeny virus titers were determined by plaque assay (89). Viral replication was also monitored by immunofluorescence microscopy using antibodies recognizing double-stranded RNA (90) and nonstructural or structural CoV proteins (37, 82, 91). To confirm the presence of the original mutations in viral progeny, supernatant from transfected cells was used to infect fresh cells, after which intracellular RNA was isolated with TriPure isolation reagent (Roche Applied Science). Next, the nsp14-coding region was amplified using standard RT-PCR methods, and the purified amplicon was sequenced by Sanger sequencing. All work with live (recombinant) class-3 CoVs was done in a biosafety level 3 laboratory at Leiden University Medical Center.

References

1. A. Ghosh, C. D. Lima, Enzymology of RNA cap synthesis. Wiley Interdiscip. Rev. RNA 1, 152–172 (2010).
2. A. Ramanathan, G. B. Robb, S.-H. Chan, mRNA capping: Biological functions and applications. Nucleic Acids Res. 44, 7511–7526 (2016).
3. F. Ferron, E. Decroly, B. Selisko, B. Canard, The viral RNA capping machinery as a target for antiviral drugs. Antiviral Res. 96, 21–31 (2012).
4. A. J. Shatkin, Capping of eucaryotic mRNAs. Cell 9, 645–653 (1976).
5. E. Kindler, V. Thiel, To sense or not to sense viral RNA--Essentials of coronavirus innate immune evasion. Curr. Opin. Microbiol. 20, 69–75 (2014).
6. E. Decroly, F. Ferron, J. Lescar, B. Canard, Conventional and unconventional mechanisms for capping viral mRNA. Nat. Rev. Microbiol. 10, 51–65 (2011).
7. E. V. Koonin, B. Moss, Viruses know more than one way to don a cap. Proc. Natl. Acad. Sci. U.S.A. 107, 3283–3284 (2010).
8. P. Zhou et al., A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579, 270–273 (2020).
9. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, The species severe acute respiratory syndrome-related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. Nat. Microbiol. 5, 536–544 (2020).
10. L. van der Hoek, Human coronaviruses: What do they cause? Antivir. Ther. 12, 651–658 (2007).
11. Y. Wang, M. Grunewald, S. Perlman, Coronaviruses: An updated overview of their replication and pathogenesis. Methods Mol. Biol. 2203, 1–29 (2020).
12. X.-Y. Ge et al., Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. Nature 503, 535–538 (2013).
13. V. D. Menachery et al., A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. Nat. Med. 21, 1508–1513 (2015).
14. B. Hu et al., Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS Pathog. 13, e1006698 (2017).
15. J. Cui, F. Li, Z.-L. Shi, Origin and evolution of pathogenic coronaviruses. Nat. Rev. Microbiol. 17, 181–192 (2019).
16. E. J. Snijder et al., Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J. Mol. Biol. 331, 991–1004 (2003).
17. E. J. Snijder, E. Decroly, J. Ziebuhr, The nonstructural proteins directing coronavirus RNA synthesis and processing. Adv. Virus Res. 96, 59–126 (2016).

18. K. Knoops et al., SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. PLoS Biol. 6, e226 (2008).
19. E. J. Snijder et al., A unifying structural and functional model of the coronavirus replication organelle: Tracking down RNA synthesis. PLoS Biol. 18, e3000715 (2020).
20. S. Klein et al., SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. Nat. Commun. 11, 5885 (2020).
21. A. Seybert, A. Hegyi, S. G. Siddell, J. Ziebuhr, The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5′-to-3′ polarity. RNA 6, 1056–1068 (2000).
22. K. A. Ivanov, J. Ziebuhr, Human coronavirus 229E nonstructural protein 13: Characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5′-triphosphatase activities. J. Virol. 78, 7833–7838 (2004).
23. K. C. Lehmann et al., Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. Nucleic Acids Res. 43, 8416–8434 (2015).
24. L. Yan et al., Cryo-EM structure of an extended SARS-CoV-2 replication and transcription complex reveals an intermediate state in cap synthesis. Cell 184, 184–193.e10 (2021).
25. A. Shannon et al., Protein-primed RNA synthesis in SARS-CoVs and structural basis for inhibition by AT-527. bioRxiv [Preprint] (2021). https://www.biorxiv.org/content/10.1101/2021.03.23.436564v1 (Accessed 22 November 2021).
26. M. Bouvet et al., In vitro reconstitution of SARS-coronavirus mRNA cap methylation. PLoS Pathog. 6, e1000863 (2010).
27. Y. Chen et al., Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. Proc. Natl. Acad. Sci. U.S.A. 106, 3484–3489 (2009).
28. E. Decroly et al., Coronavirus nonstructural protein 16 is a cap-0 binding enzyme possessing (nucleoside-2′O)-methyltransferase activity. J. Virol. 82, 8071–8084 (2008).
29. Y. Chen et al., Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2′-O-methylation by nsp16/nsp10 protein complex. PLoS Pathog. 7, e1002294 (2011).
30. M. R. Denison, R. L. Graham, E. F. Donaldson, L. D. Eckerle, R. S. Baric, Coronaviruses: An RNA proofreading machine regulates replication fidelity and diversity. RNA Biol. 8, 270–279 (2011).
31. N. S. Ogando et al., The curious case of the nidovirus exoribonuclease: Its role in RNA synthesis and replication fidelity. Front. Microbiol. 10, 1813 (2019).
32. X. Jin et al., Characterization of the guanine-N7 methyltransferase activity of coronavirus nsp14 on nucleotide GTP. Virus Res. 176, 45–52 (2013).

33. Y. Chen et al., Structure-function analysis of severe acute respiratory syndrome coronavirus RNA cap guanine-N7-methyltransferase. J. Virol. 87, 6296–6305 (2013).
34. M. Bouvet et al., RNA 3′-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. Proc. Natl. Acad. Sci. U.S.A. 109, 9372–9377 (2012).
35. Y. Ma et al., Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. Proc. Natl. Acad. Sci. U.S.A. 112, 9436–9441 (2015).
36. F. Ferron et al., Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. Proc. Natl. Acad. Sci. U.S.A. 115, E162–E171 (2018).
37. N. S. Ogando et al., The enzymatic activity of the nsp14 exoribonuclease is critical for replication of MERS-CoV and SARS-CoV-2. J. Virol. 94, e01246-20(2020).
38. M. Saramago et al., New targets for drug design: Importance of nsp14/nsp10 complex formation for the 3′-5′ exoribonucleolytic activity on SARS-CoV-2. FEBS J. 288, 5130–5147 (2021).
39. L. Xie, P. E. Bourne, Detecting evolutionary relationships across existing fold space, using sequence order-independent profile-profile alignments. Proc. Natl. Acad. Sci. U.S.A. 105, 5441–5446 (2008).
40. R. Gana, S. Rao, H. Huang, C. Wu, S. Vasudevan, Structural and functional studies of S-adenosyl-L-methionine binding proteins: A ligand-centric approach. BMC Struct. Biol. 13, 6 (2013).
41. S. M. Sperry et al., Single-amino-acid substitutions in open reading frame (ORF) 1b-nsp14 and ORF 2a proteins of the coronavirus mouse hepatitis virus are attenuating in mice. J. Virol. 79, 3391–3400 (2005).
42. L. D. Eckerle, X. Lu, S. M. Sperry, L. Choi, M. R. Denison, High fidelity of murine hepatitis virus replication is decreased in nsp14 exoribonuclease mutants. J. Virol. 81, 12135–12144 (2007).
43. J. B. Case, A. W. Ashbrook, T. S. Dermody, M. R. Denison, Mutagenesis of S-adenosyl-l-methionine-binding residues in coronavirus nsp14 N7-methyltransferase demonstrates differing requirements for genome translation and resistance to innate immunity. J. Virol. 90, 7248–7256 (2016).
44. Z. Zhang et al., Live attenuated coronavirus vaccines deficient in N7-Methyltransferase activity induce both humoral and cellular immune responses in mice. Emerg. Microbes Infect. 10, 1626–1637 (2021).
45. F. Ferron, H. J. Debat, A. Shannon, E. Decroly, B. Canard, A N7-guanine RNA cap methyltransferase signature-sequence as a genetic marker of large genome, non-mammalian Tobaniviridae. NAR Genom. Bioinform. 2, lqz022 (2020).

46. P. T. Nga et al., Discovery of the first insect nidovirus, a missing evolutionary link in the emergence of the largest RNA virus genomes. PLoS Pathog. 7, e1002215 (2011).
47. A. Saberi, A. A. Gulyaeva, J. L. Brubacher, P. A. Newmark, A. E. Gorbalenya, A planarian nidovirus expands the limits of RNA genome size. PLoS Pathog. 14, e1007314 (2018).
48. Y. Sun et al., Yeast-based assays for the high-throughput screening of inhibitors of coronavirus RNA cap guanine-N7-methyltransferase. Antiviral Res. 104, 156–164 (2014).
49. W. Aouadi et al., Toward the identification of viral cap-methyltransferase inhibitors by fluorescence screening assay. Antiviral Res. 144, 330–339 (2017).
50. R. Ahmed-Belkacem et al., Synthesis of adenine dinucleosides SAM analogs as specific inhibitors of SARS-CoV nsp14 RNA cap guanine-N7-methyltransferase. Eur. J. Med. Chem. 201, 112557 (2020).
51. R. He et al., Potent and selective inhibition of SARS coronavirus replication by aurintricarboxylic acid. Biochem. Biophys. Res. Commun. 320, 1199–1203 (2004). Correction in: Biochem. Biophys. Res. Commun. 324, 1152–1153 (2004).
52. M. H. Barnes, P. Spacciapoli, D. H. Li, N. C. Brown, The 3′-5′ exonuclease site of DNA polymerase III from gram-positive bacteria: Definition of a novel motif structure. Gene 165, 45–50 (1995).
53. Y. Zuo, M. P. Deutscher, Exoribonuclease superfamilies: Structural analysis and phylogenetic distribution. Nucleic Acids Res. 29, 1017–1026 (2001).
54. M. Byszewska, M. Śmietański, E. Purta, J. M. Bujnicki, RNA methyltransferases involved in 5′ cap biosynthesis. RNA Biol. 11, 1597–1607 (2014).
55. B. P. S. Chouhan, S. Maimaiti, M. Gade, P. Laurino, Rossmann-fold methyltransferases: Taking a “β-Turn” around their cofactor, S-adenosylmethionine. Biochemistry 58, 166–170 (2019).
56. N. H. Moeller et al., Structure and dynamics of SARS-CoV-2 proofreading exoribonuclease ExoN. bioRxiv [Preprint] (2021). https://www.biorxiv.org/content/10.1101/2021.04.02.438274v1 (Accessed 22 November 2021).
57. R. A. Laskowski, SURFNET: A program for visualizing molecular surfaces, cavities, and intermolecular interactions. J. Mol. Graph 13, 323–330, 307–328 (1995).
58. L. D. Eckerle, S. M. Brockway, S. M. Sperry, X. Lu, M. R. Denison, Effects of mutagenesis of murine hepatitis virus nsp1 and nsp14 on replication in culture. Adv. Exp. Med. Biol. 581, 55–60 (2006).
59. K. E. Medvedev, L. N. Kinch, N. V. Grishin, Functional and evolutionary analysis of viral proteins containing a Rossmann-like fold. Protein Sci. 27, 1450–1463 (2018).
60. S. T. Rao, M. G. Rossmann, Comparison of super-secondary structures in proteins. J. Mol. Biol. 76, 241–256 (1973).

61. G. Sutton, J. M. Grimes, D. I. Stuart, P. Roy, Bluetongue virus VP4 is an RNA-capping assembly line. Nat. Struct. Mol. Biol. 14, 449–451 (2007).
62. Y. Tao, D. L. Farsetta, M. L. Nibert, S. C. Harrison, RNA synthesis in a cage--Structural studies of reovirus polymerase lambda3. Cell 111, 733–745 (2002).
63. M.-P. Egloff, D. Benarroch, B. Selisko, J.-L. Romette, B. Canard, An RNA cap (nucleoside-2′-O-)-methyltransferase in the flavivirus RNA polymerase NS5: Crystal structure and functional characterization. EMBO J. 21, 2757–2768 (2002).
64. M. M. Lai, S. A. Stohlman, Comparative analysis of RNA genomes of mouse hepatitis viruses. J. Virol. 38, 661–670 (1981).
65. A. L. W. van Vliet, S. L. Smits, P. J. M. Rottier, R. J. de Groot, Discontinuous and non-discontinuous subgenomic RNA transcription in a nidovirus. EMBO J. 21, 6571–6580 (2002).
66. A. Seybert et al., A complex zinc finger controls the enzymatic activities of nidovirus helicases. J. Virol. 79, 696–704 (2005).
67. L.-A. Pearson et al., Development of a high-throughput screening assay to identify inhibitors of the SARS-CoV-2 Guanine-N7-methyltransferase using rapidfire mass spectrometry. SLAS Discov. 26, 749–756 (2021).
68. K. Devkota et al., Probing the SAM binding site of SARS-CoV-2 Nsp14 in vitro using SAM competitive inhibitors guides developing selective bisubstrate inhibitors. SLAS Discov. 26, 1200–1211 (2021).
69. S. Basu et al., Identifying SARS-CoV-2 antiviral compounds by screening for small molecule inhibitors of Nsp14 RNA cap methyltransferase. Biochem. J. 478, 2481–2497 (2021).
70. M. Bouvet et al., Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes. J. Biol. Chem. 289, 25783–25796 (2014).
71. L. Yan et al., Coupling of N7-methyltransferase and 3′-5′ exoribonuclease with SARS-CoV-2 polymerase reveals mechanisms for capping and proofreading. Cell 184, 3474–3485.e11 (2021).
72. C. Liu et al., Structural basis of mismatch recognition by a SARS-CoV-2 proofreading enzyme. Science 373, 1142–1146 (2021).
73. E. C. Smith, H. Blanc, M. C. Surdel, M. Vignuzzi, M. R. Denison, Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: Evidence for proofreading and potential therapeutics. PLoS Pathog. 9, e1003565 (2013). Correction in: PLoS Pathog. 10, e1004342 (2014).
74. M. Gouy, S. Guindon, O. Gascuel, SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27, 221–224 (2010).
75. G. E. Crooks, G. Hon, J.-M. Chandonia, S. E. Brenner, WebLogo: A sequence logo generator. Genome Res. 14, 1188–1190 (2004).

76. E. F. Pettersen et al., UCSF Chimera--A visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612 (2004).
77. E. Jurrus et al., Improvements to the APBS biomolecular solvation software suite. Protein Sci. 27, 112–128 (2018).
78. Y. Dehouck, J. M. Kwasigroch, D. Gilis, M. Rooman, PoPMuSiC 2.1: A web server for the estimation of protein stability changes upon mutation and sequence optimality. BMC Bioinformatics 12, 151 (2011).
79. M. Hecht, Y. Bromberg, B. Rost, Better prediction of functional effects for sequence variants. BMC Genomics 16 (suppl. 8), S1 (2015).
80. Y. Wang et al., Coronavirus nsp10/nsp16 methyltransferase can be targeted by nsp10-derived peptide in vitro and in vivo to reduce replication and pathogenesis. J. Virol. 89, 8416–8427 (2015).
81. D. D. Nedialkova, A. E. Gorbalenya, E. J. Snijder, Arterivirus Nsp1 modulates the accumulation of minus-strand templates to control the relative abundance of viral mRNAs. PLoS Pathog. 6, e1000772 (2010).
82. A. H. de Wilde et al., MERS-coronavirus replication induces severe in vitro cytopathology and is strongly inhibited by cyclosporin A or interferon-α treatment. J. Gen. Virol. 94, 1749–1760 (2013).
83. B. K. Tischer, G. A. Smith, N. Osterrieder, En passant mutagenesis: A two step markerless red recombination system. Methods Mol. Biol. 634, 421–430 (2010).
84. F. Almazán et al., Engineering a replication-competent, propagation-defective Middle East respiratory syndrome coronavirus as a vaccine candidate. MBio 4, e00650-13 (2013).
85. H. H. Rabouw et al., Middle East respiratory coronavirus accessory protein 4a inhibits PKR-mediated antiviral stress responses. PLoS Pathog. 12, e1005982 (2016).
86. S. Pfefferle et al., Reverse genetic characterization of the natural genomic deletion in SARS-Coronavirus strain Frankfurt-1 open reading frame 7b reveals an attenuating function of the 7b protein in-vitro and in-vivo. Virol. J. 6, 131 (2009).
87. A. R. Fehr et al., The nsp3 macrodomain promotes virulence in mice with coronavirus-induced encephalitis. J. Virol. 89, 1523–1536 (2015).
88. T. Thi Nhu Thao et al., Rapid reconstruction of SARS-CoV-2 using a synthetic genomics platform. Nature 582, 561–565 (2020).
89. S. H. E. van den Worm et al., Reverse genetics of SARS-related coronavirus using vaccinia virus-based recombination. PLoS One 7, e32857 (2012).
90. F. Weber, V. Wagner, S. B. Rasmussen, R. Hartmann, S. R. Paludan, Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. J. Virol. 80, 5059–5064 (2006).

91. M. J. van Hemert et al., SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity in vitro. PLoS Pathog. 4, e1000054 (2008).

Supplementary Information

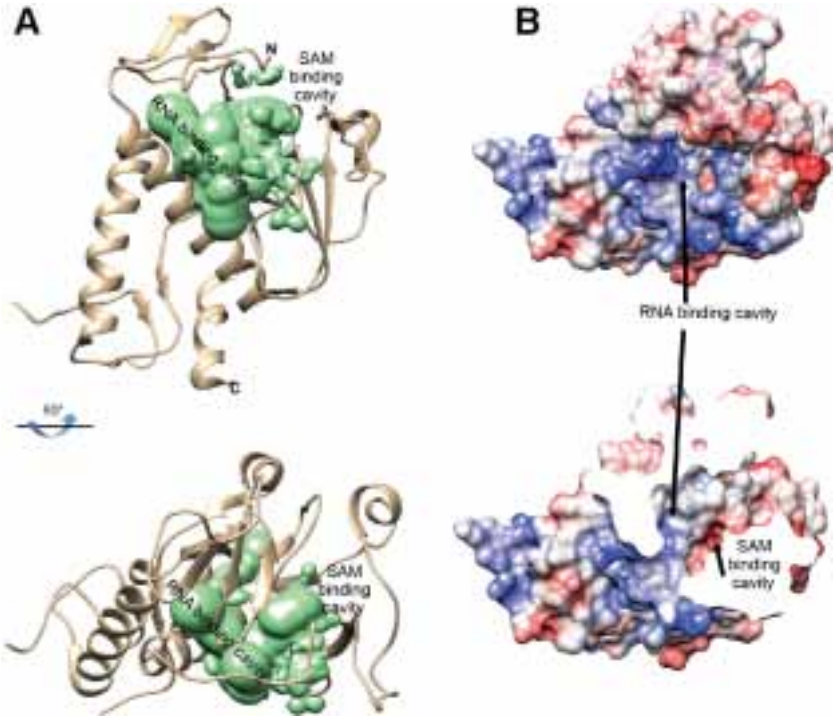

Figure: S1. *Structural analysis of the catalytic site and hinge region of the SARS-CoV nsp14 N7- MTase domain. A) Determination of the volume of the enzyme's catalytic site (with the volume depicting the mold of the cavity shown in green). B) Electrostatic surface representation with the surface electrostatic potential calculated by APBS from -10 (red) to +10 (blue) kT/e.*

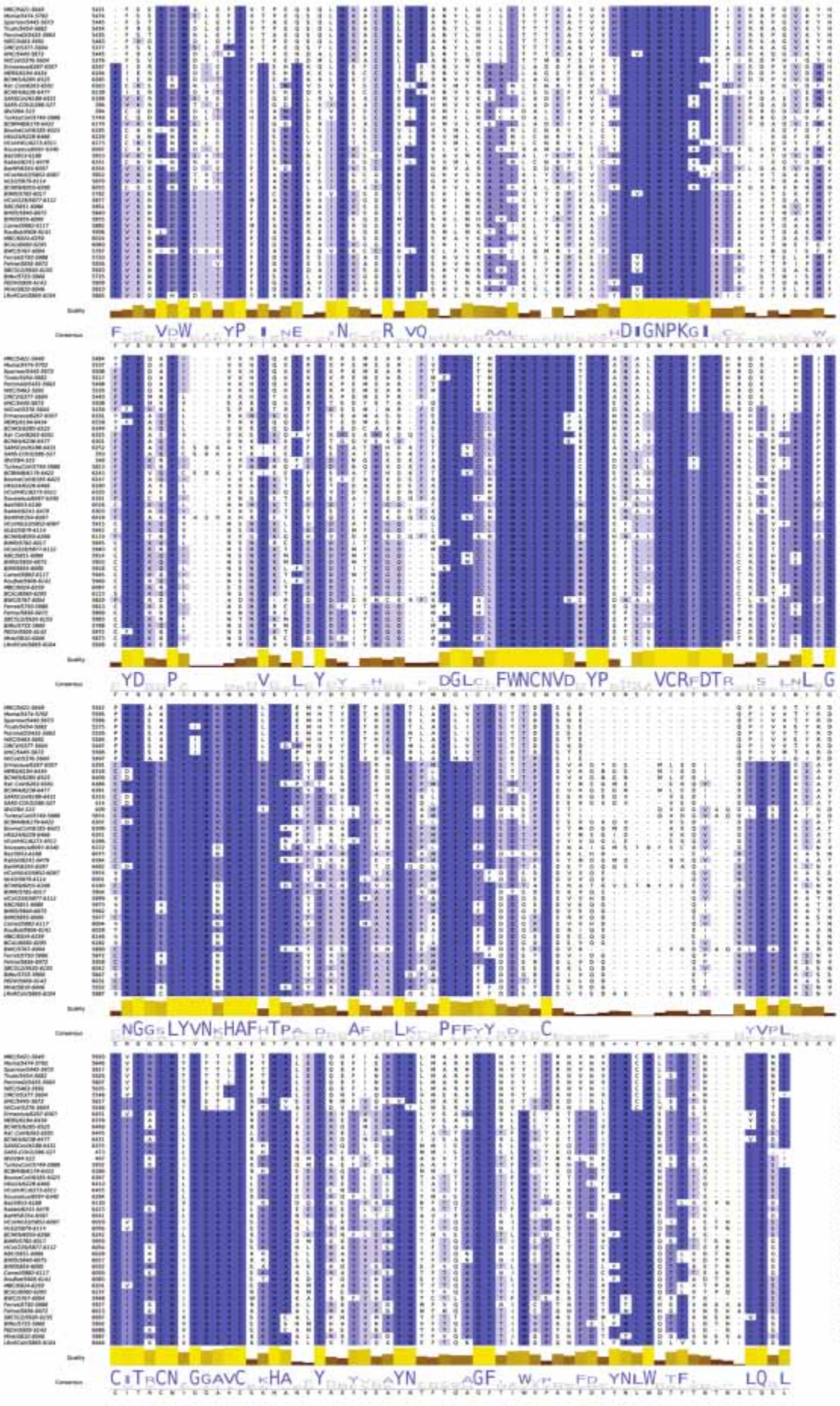

Figure: S2. *Alignment of the nsp14 N7-MTase sequences from 47 selected coronaviruses (listed in Table S1). Highly conserved residues are boxed in dark blue (above 70% conservation), while partially conserved residues are displayed in lighter shades of blue. A consensus logo sequence is presented at the bottom of the alignment. The N7-MTase motifs I to VI are highlighted in red, orange, green, yellow, blue, and pink, using the same color scheme as in Figure: 2.*

Supplementary table 1- *List of CoV genomes and accession numbers used to extract nsp14 sequences for structural studies*

Virus	**Acession number**
Alphacoronavirus BtMs-AlphaCoV/GS 2013	A0A0U1WHG4
Avian infectious bronchitis virus (IBV)	P0C6Y2
Bat coronavirus 1A	YP_001718603.1
Bat coronavirus BM48-31	E0XIZ2
Bat coronavirus CDPHE15/USA/200 6	YP_008439224.1
Bat coronavirus HKU4	P0C6W3
Bat coronavirus HKU5	P0C6W4
Bat coronavirus HKU9	P0C6W5
Bat Hp- betacoronavirus	A0A088DIE1
Beluga whale coronavirus SW1	YP_001876435.1
Betacoronavirus Erinaceus	U5KNA9
Betacoronavirus HKU24	A0A0A7UXR0
Bovine coronavirus	P0C6W8
BtMr-AlphaCoV/ SAX2011	A0A0U1UZC3
BtNv-AlphaCoV/SC2013	YP_009201729.1
BtRf-AlphaCoV/HuB2013	YP_009199789.1
Camel alphacoronavirus	ALA50136.1
Common moorhen coronavirus HKU21	H9BR34
Feline infectious peritonitis virus	AGZ84515.1
Ferret coronavirus	YP_009256195.1
Human coronavirus 229E (HCoV-229E)	P0C6X1
Human coronavirus HKU1 (HCoV-HKU1)	P0C6X2

Human coronavirus NL63 (HCoV-NL63)	P0C6X5
Lucheng Rn rat coronavirus	YP_009336483.1
Magpie-robin coronavirus HKU18	H9BR07
Middle East respiratory syndrome-related coronavirus	K9N7C7
Miniopterus bat coronavirus HKU8	YP_001718610.1
Mink coronavirus strain WD1127	YP_009019180
Munia coronavirus HKU13-3514	YP_002308505.1
Murine coronavirus (strain A59) (MHV- A59)	P0C6X9
Night heron coronavirus HKU19	H9BR16
Porcine deltacoronavirus	A0A140ESF0
Porcine epidemic diarrhea virus	NP_839967
Rabbit coronavirus HKU14	H9AA60
Rat coronavirus Parker	YP_009924380.1
Rhinolophus bat coronavirus HKU2	A8JNZ0
Rousettus bat coronavirus	A0A1B3Q5W8
Rousettus bat coronavirus HKU10	AFU92103
Scotophilus bat coronavirus 512	YP_001351683
Severe acute respiratory syndrome coronavirus (SARS-CoV)	P0C6X7
Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2)	P0DTD1
Sparrow coronavirus HKU17	H9BQZ9
Swine acute diarrhea syndrome related coronavirus BtRf2	AVM80482.1
Thrush coronavirus HKU12	B6VDX7
Turkey coronavirus	YP_001941187
White-eye coronavirus HKU16	YP_005352837.1
Wigeon coronavirus HKU20	H9BR24

Table S2. *Projected impact on folding free energy by alanine substitutions of the identified core residues on the SARS-CoV N7-MTase structure, as calculated by PoPMusic (1).*

N7-MTase motif	Mutation	Secondary structure*	Solvent accessibility (%)	ΔΔG (kcal/mol)
	wild-type	-	-	0
I	**W292A**	T	27.44	1.8
II	**N306A**	H	39.77	0.72
	R310A	H	58.37	0
III	**D331A**	H	4.20	0.46
	K336A	H	65.79	0.38
IV	**D352A**	C	14.21	1.12
V	**N386A**	C	32.59	0.99
VI	**Y420A**	S	86.78	3.35
	N422A	S	10.04	0.92
	H424A	S	43.12	2.21
	F426A	C	49.59	1.57

*classified as: T, turn; H, α-helix; C, random coil; and S, β-sheet.

***Supplementary Table 3A**: Predicted functional impact of individual mutations of key conserved residues in motifs I-III of SARS-CoV, MERS-CoV, and MHV nsp14, predicted with SNAP2 (2).*

	SARS-CoV	MERS-CoV	MHV	SARS-CoV	MERS-CoV	MHV	SARS-CoV	MERS-CoV	MHV	SARS-CoV	MERS-CoV	MHV
aa	**W292**	**W292**	**W291**	**N306**	**N306**	**N305**	**R310**	**R310**	**R309**	**D331**	**D331**	**D330**
A	72	66	78	51	26 *	54	66	30 *	63	75	67	78
R	82	81	88	75	32 *	82				91	80	90
N	80	63	86				71	41	71	77	78	83
D	88	72	91	67	58	78	85	74	84			
C	59	62	69	48	22 *	58	65	18 *	63	73	65	76
Q	76	74	84	48	35 *	65	60	32 *	61	81	79	84
E	82	72	87	73	63	82	81	61	78	70	73	74
G	82	82	87	53	41	64	81	31 *	79	79	84	88
H	78	62	86	50	36 *	63	59	29 *	60	87	84	89
I	72	69	79	71	35 *	77	73	35 *	71	88	85	89
L	73	75	82	72	35 *	79	73	21 *	70	90	81	91
K	85	84	90	63	59	80	41	20 *	45	91	91	92
M	66	71	78	63	34 *	73	66	29 *	65	87	84	89
F	45	55	63	74	52	80	81	52	78	89	80	90
P	83	89	94	77	64	83	86	70	85	91	91	93
S	78	77	78	41	22 *	58	70	32 *	68	74	74	80
T	78	80	86	49	24 *	64	69	32 *	66	79	77	83
W				81	70	86	87	66	85	92	88	93
Y	29 *	50	52	70	54	77	79	53	77	90	83	91

V	71	60	78	67	26 *	76	74	38 *	72	87	83	88

*All data are predicted above 70 % expected accuracy except for * above 53 % - Positive value indicates a destabilising effect - Negative value indicates a neutral effect.*

***Supplementary Table 3B**: Predicted functional impact of individual mutations of key conserved residues in motifs III-VI of SARS-CoV, MERS-CoV, and MHV nsp14, predicted with SNAP2 (2).*

	SARS-CoV	MERS-CoV	MHV	SARS-CoV	MERS-CoV	MHV	SARS-CoV	MERS-CoV	MHV	SARS-CoV	MERS-CoV	MHV
aa	**K336**	**K336**	**K335**	**D352**	**D352**	**D349**	**N386**	**N382**	**N380**	**Y420**	**Y416**	**Y414**
A	31 *	20 *	39 *	34 *	63	73	60	45	56	71	66	74
R	-17 *	-20 *	-1 *	45	85	87	78	73	77	89	85	91
N	28 *	26 *	44	22 *	73	69				86	71	88
D	69	63	76				62	56	63	90	73	92
C	22 *	7 *	-16 *	27 *	57	69	60	15	57	61	54	64
Q	16 *	6 *	29 *	43	74	73	47	52	64	82	77	86
E	52	45	60	26 *	67	53	75	70	74	87	82	89
G	52	41	60	49	77	78	51	55	49	87	81	88
H	7 *	-6 *	19 *	21 *	74	80	63	51	63	59	70	78
I	-15 *	-70	-40 *	59	78	86	77	53	76	69	62	70
L	22 *	8 *	33 *	61	80	88	80	49	78	74	70	74
K				48	87	88	76	71	76	90	86	91
M	10 *	-39 *	18 *	55	78	84	74	52	73	78	72	79
F	52	41	55	23 *	71	87	81	67	80	42	47	40
P	62	30 *	69	72	89	89	71	76	82	88	91	94
S	22 *	13 *	36 *	29 *	65	65	44	31	42	85	69	86
T	26 *	13 *	39 *	34 *	69	69	51	30 *	33 *	86	79	88
W	67	56	67	66	81	90	87	78 *	86	68	71	71
Y	48	35 *	51	32 *	76	86	79	38 *	78			

V	22 *	8 *	32 *	53	76	83	75	41	75	68	49	70

*All data are predicted above 70 % expected accuracy except for * above 53 % - Positive value indicates a destabilising effect - Negative value indicates a neutral effect*.

***Supplementary Table 3C**: Predicted functional impact of individual mutations of key conserved residues in motif VI of SARS-CoV, MERS-CoV, and MHV nsp14, predicted with SNAP2 (2).*

	SARS-CoV	MERS-CoV	MHV	SARS-CoV	MERS-CoV	MHV	SARS-CoV	MERS-CoV	MHV
aa	**N422**	**N418**	**N416**	**H424**	**H420**	**H418**	**F426**	**F422**	**F420**
A	66	33 *	69	63	62	64	56	59	67
R	81	72	84	74	26 *	35 *	83	81	85
N				63	70	65	82	79	83
D	64	65	79	84	86 *	85	89	88	90
C	61	41	66	64	20	64	48 *	39 *	50
Q	64	4 *	70	65	69	67	78	72	80
E	80	71	83	79	81	80	84	72	85
G	64	20 *	64	75	66	76	81	79	83
H	60	49	67				59	76	79
I	79	71	83	76	69	76	58 *	10 *	60
L	81	72	85	76	41	75	63	50	59
K	79	69	81	81	81	82	85	82	87
M	76	63	80	72	65	72	60	49	47
F	69	74	85	64	75	76			
P	84	76	87	90	90	90	89	87	91
S	52	-3 *	59	70	71	72	79	75	80
T	59	42	65	64	72	76	78	60	80
W	88	81	90	81	82	82	69	69	67
Y	79	71	83	65	70	67	49	49	49
V	78	65	81	71	54	72	62 *	37 *	65

*All data are predicted above 70 % expected accuracy except for * above 53 % - Positive value indicates a destabilising effect - Negative value indicates a neutral effect.*

SI References

1. Y. Dehouck, J. M. Kwasigroch, D. Gilis, M. Rooman, PoPMuSiC 2.1: a web server for the estimation of protein stability changes upon mutation and sequence optimality. *BMC Bioinformatics* **12**, 151 (2011).

2. M. Hecht, Y. Bromberg, B. Rost, Better prediction of functional effects for sequence variants. *BMC Genomics* **16 Suppl 8**, S1 (2015).

Chapter 4

General Discussion

Upon viral infection, the vRNA can be detected as non-self by TLRs, RLRs, NLRs, and other RNA sensors. Pathogen sensing induces a cascade of events leading to the INF-I expression and response. However, viruses have evolved different coping strategies to escape innate immunity recognition and INF-I response. One of these strategies is the epitranscriptomic modifications of the RNA, till now more than 100s of these are identified. Among which are RNA methylations and 5'-end-capping (process described in chapter 1). The latter consists of guanosine residue bridged to the first nucleotide (N_1) of most vRNAs via a 5'-5' triphosphate bond ($GpppN_1$)[66]. This structure is methylated at the N7 position of the guanosine moiety by a specific N7-MTase, to generate the cap-0 ($^{N7m}GpppN_1$) structure. The N_1 of the latter is often 2'O-methylated for some viruses by a 2'O-MTase to generate a cap-1 structure ($^{N7m}GpppN_m$)[66]. The role of the cap structure has been well established. The N7-methylation of the cap allows its recognition by the ribosomal EIF4e protein and is thus essential for mRNA translation into protein, it also protects vRNA from decapping enzymes and decay by 5' to 3' RNases. Adding a 2'O-methyl group to the cap-0 structure marks the viral RNA as self and prevents the detection of RIG-I and/or MDA5 and limits the antiviral activity of IFIT1/3 proteins[49,51,51]. Given the synergistic "proviral" role of N7- and 2'O-cap methylations, it is not surprising that many viruses have acquired the cap-1 structure.

In the viral world, several molecular pathways have been selected to produce cap-1 structures. For example, in flaviviruses, filoviruses or coronaviruses, the viral genome encodes a cap-dependent 2'O-methyltransferase, whereas, in orthomyxoviruses or bunyaviruses, the viral machinery steals cap-1 structures from cellular mRNA (process described in chapter 1). Numerous viral MTases involved in cap methylation have been biochemically and structurally characterized. Structural studies revealed that most of these SAM-dependent enzymes share a similar Rossmann fold organization[109,110] except for the SARS-CoV nsp14 N7-MTase domain[119]. The nsp14 is composed of an N-terminal 3'ExoN domain and a C-terminal N7-MTase domain, which are functionally independent and crucial for viral survival. The outstanding non-Rossmann fold organization is a unique feature shared only among CoV nsp14s[115,119]. While the crystal structure of some CoV N7-MTase has been resolved[114,244], their biochemical properties and signature sequences critical for RNA binding or enzymatic activity remain poorly defined and this is due to their unusual organization. Having more information on the key motifs or residues constituting the CoVs N7-MTase domain could play a fundamental role in the design of broad antivirals targeting major CoV pathogens.

Interestingly, it has been brought to our attention that in addition to their involvement in

the RNA capping pathway, the 2'O-MTase of WNV, DEN, ZIKV, and Ebola virus has been shown in an *in vitro* assay to induce the 2'O-methylation of adenosine residues within the RNA genome[33,95,129]. With the advent of mass spectrometry and sequencing technologies, these internal epitranscriptomic marks were confirmed on the genome of numerous RNA viruses including DENV, ZIKV, SARS-CoV-2, hepatitis C, poliovirus, HIV-1, and MLV. The discovery of internal 2'O-methylations of the genome of these last four viruses is quite surprising since they do not code for any known 2'O-MTase, thus suggesting an intervention of host MTases. Indeed, HIV-1 has recently been shown to subvert the host MTase FTSJ3 to ensure 17 confident internal Nm on its own RNA genome[30]. Although the viral RNA 2'O-methylation landscape is still poorly characterized, the current knowledge we have about its biological function is quite intricating and mysterious as it displays both pro and antiviral activity. Hence, the antiviral effect of these epitranscriptomic marks can be emphasized by the inhibitory effect it exerts on the activity of some viral polymerases namely the MLV reverse transcriptase and DENV RdRp[33,130]. Adding to this, it is likely that the vRNA 2'O-methylation downregulates mRNA translation as is observed with the cellular mRNA[133]. Regarding the proviral role of vRNA 2'O-methylation, it has been shown for HIV-1 that it shields the viral RNA from being recognized by the innate immune sensor MDA5 which in turn limits IFN expression[30]. In this study, we were interested in the role of viral RNA methylation, in particular, the 2'O- and N7-methylation in viral replication.

In chapter 2, we showed that RNA 2'O-methylation may interplay with the antiviral action of an interferon-induced restriction factor, namely ISG20. Through an *in vitro* study, we demonstrated that ISG20 ExoN activity is impaired by RNA 2'O-methylation. Interestingly, by evaluating the activity of ISG20 on an RNA containing a 2'O-methylated residue 7 nucleotide upstream of the 3'end ($A_{20}N_mA_6$), the inhibition of substrate decay was marked by a pause at N_{-2} and N_0 of the methylated nucleotide. This last observation suggested that ISG20 inhibition is due to steric hindrance mechanisms rather than direct inhibition of the catalytic activity of ISG20. To understand the molecular basis of ISG20 inhibition by RNA 2'O-methylation, we first made use of the currently available structure of ISG20 with a UMP[229] to address the stop observed at N_0. We, therefore, modeled a methyl group on the 2'O of UMP ribose and saw that the modification led to a potential steric clash with the M14 and R53 residues of ISG20. Through a mutagenesis study, we confirmed this observation, however, we found that the residue R53 was the only culprit causing the stop at N_0. With the structural data available from the ISG20, we could not go further in the understanding of the stop observed at the N_{-2} position. Therefore, another approach was used, we modeled the RNA from the structure of SDN1 (ISG20 homolog) in the ISG20

catalytic pocket which implemented our knowledge of how it interacts with its substrate. The model that we thus built, allowed us to hypothesize that the cause of the pause observed in N_{-2} could result from a steric clash between the 2'O-methylated residue of the RNA and residues H89, D90, or V128. Among the identified residues, we demonstrated that D90 had some difficulties in accommodating the internal RNA 2'O-methylated residues, which promotes destabilization and limits the advancement of the substrate in the ISG20 catalytic pocket, therefore, causing the stop at N_{-2}. Combining these observations, we concluded that RNA 2'O-methylation limits ISG20 ExoN activity through a mechanism of steric hindrance. Moreover, the observed inhibition seems to RNA 2'O-methylation as the ^{N6m}A modification does not seem to affect ISG20 ExoN activity nor antiviral activity[241], which suggests that viruses that induce internal, using their own[33,129] or the host MTase[30], could limit the antiviral activity of ISG20. To validate our *in vitro* data, we extrapolated them on an infectious model, in particular HIV-1, whose RNA 2'O-methylation mechanism is well characterized[30]. We first demonstrated in an enzymatic assay that hypomethylated RNAs produced and extracted from FTSJ3-deficient HEK293 cells were more sensitive to ISG20-mediated degradation than those from cells expressing the FTSJ3, thus highlighting the shielding effect of RNA 2'O-methylation. This was further confirmed by showing that ISG20 lacking the key residues in 2'O-methyl sensing (ISG20 double mutant R53A/D90A) degrades both hypo- and hypermethylated HIV-1 RNA with the same efficacy. These fruitful biochemistry results prompted us to move on to cellular trials, where we showed that ectopically expressed ISG20 drastically inhibits hypomethylated VSV-G pseudotyped HIV-1 replication in infected cells as a consequence of viral RNA decay. Collectively, these results indicate the proviral role of HIV-1 RNA internal 2'O-methylation in counteracting ISG20 antiviral activity. As this IFN-induced restriction factor (ISG20) was shown to specifically recognize non-self vRNA by a yet unknown molecular basis and subsequently induces an appropriate antiviral response[6]. Through this study, we propose that RNA 2'O-methylation is a self-signature that allows ISG20 to discriminate foreign RNA. Finally, our results allowed us to elucidate a new consequence of viral RNA 2'O-methylation originally identified as the subversion of "self" sensing to impair IFN-I expression. Our results also highlight the dual proviral role of 2'O-methylation in the HIV-1 life cycle. It is questionable whether this also applies to viruses that have been proven to harbor internal 2'O-methylation and those who depend on host 2'O-MTase for survival[32,34,35].

In the second part of this study, we took the time to define and characterize the key motifs of CoV nsp14 N7-MTase to improve our understanding of its biological role among different

CoVs (chapter 3). When this study was carried out, the only available structural data CoVs nsp14 was that of the SARS-CoV[118,244]. These structures have defined the enzyme's SAM and cap-binding pockets. However, crystal packing constraints and the overall low resolution have left some uncertainties in the structural characterization of the N7-MTase domain, particularly at the level of the RNA ligand positioning. Therefore, to define CoVs N7-MTase SAM and RNA-binding motifs, a thorough structural analysis of the SARS-CoV N7-MTase cavity was assessed, supported by CoV-wide nsp14 sequence comparisons. Our study allows us to identify the main residues potentially conserved in each cavity, and classify them according to their characteristic features. Thus, the binding domain of SAM consists of motif I (W292), motif III (D331 and K336), and motif IV (D352), while motif II (N306 and R310) motif V (N386), and motif VI (Y420, N422, H424, and F426) are of the RNA binding site. Interestingly, it was previously reported the R310 residue of SARS-CoV nsp14 plays a role in SAM-binding[118], while F426 was proposed to root and stabilize the guanosine of the cap near the SAM moiety[114]. However, our structural analysis redefined both residues as part of the putative RNA binding site motifs II and VI, respectively. We then characterized the effect of alanine substitutions of these key residues both *in vitro*, by using purified recombinant SARS-CoV or MERS-CoV nsp14 proteins, and in an infectious context, by modifying the corresponding viral mutants in different beta-coronaviruses (this part was performed in E. Snijder lab in Leiden). The substitutions inserted in the conserved SARS-CoV and MERS-CoV N7-MTase residues drastically abrogated enzymatic activity *in vitro*, supporting their identification as key residues for the enzyme's functionality. Strikingly, the *in vitro* enzymatic data of several SARS-CoV and MERS-CoV nsp14 mutants, and the viability of these viral mutants appeared contradictory. One possible explanation is that low levels of N7-MTase activity might be sufficient for viral replication in our cell culture models. Alternatively, our *in vitro* tool could have suffered from technical complications, such as a folding defect of the N7-MTase mutants. This problem could be overcome in the context of the cell infected by the virus, in particular in the presence of its natural interaction partners. It is therefore conceivable that the impact of nsp14 mutations on the fold and/or critical protein-protein or protein-RNA interactions of the N7-MTase domain can vary between different assay systems. In contrast, consistent results were observed between *in vitro* and in infected cells for the mutation of the conserved R in motif II (nonviable) and conserved H and F in motif VI (paralyzed and non-viable, respectively) across all tested CoV. Thus indicating the importance of these residues for N7-MTase activity. Overall our results highlight the N7-MTase as an essential enzyme for betacoronavirus replication and define key

residues of its catalytic pocket that can be targeted for the rational design of inhibitors with a potential pan-coronaviral activity spectrum. Moreover, given its role in CoV replication and its emerging status as a target for antiviral drug design, further integrated biochemical and virological analysis are needed.

Finally, this thesis project highlights the key role of the epitranscriptomic marks particularly that of N7- and 2'O-methylation. These marks have important regulatory functions that are still poorly understood because they regulate both essential metabolic properties for viral replication such as the ability of viruses to produce RNA "translatable into proteins" and key mechanisms in the processes host-pathogens interactions. The regulation of these epitranscriptomic modifications is therefore very important and complex because they have both pro and antiviral roles.

References

1. Wierzbicki, A. T. The role of long non-coding RNA in transcriptional gene silencing. *Curr. Opin. Plant Biol.* **15**, 517–522 (2012).
2. Zhao, J.-H., Hua, C.-L., Fang, Y.-Y. & Guo, H.-S. The dual edge of RNA silencing suppressors in the virus-host interactions. *Curr. Opin. Virol.* 17, 39–44 (2016).
3. Banerjee, S. *et al.* RNA Interference: A Novel Source of Resistance to Combat Plant Parasitic Nematodes. *Front. Plant Sci.* 8, 834 (2017).
4. Netzband, R. & Pager, C. T. Epitranscriptomic marks: Emerging modulators of RNA virus gene expression. *WIREs RNA* 11, e1576 (2020).
5. Li, N. & Rana, T. M. Regulation of antiviral innate immunity by chemical modification of viral RNA. *WIREs RNA* n/a, e1720.
6. Wu, N. *et al.* The interferon stimulated gene 20 protein (ISG20) is an innate defense antiviral factor that discriminates self versus non-self translation. *PLoS Pathog.* 15, e1008093 (2019).
7. Soares, A. R., Kikkert, M., Kellner-Kaiser, S. & Ribeiro, D. Editorial: Viruses and Epitranscriptomes: Regulation of Infection and Antiviral Response. *Front. Cell Dev. Biol.* 10, (2022).
8. Shatkin, A. J. Capping of eucaryotic mRNAs. *Cell* 9, 645–653 (1976).
9. Nallagatla, S. R., Toroney, R. & Bevilacqua, P. C. A brilliant disguise for self RNA: 5'-end and internal modifications of primary transcripts suppress elements of innate immunity. *RNA Biol.* 5, 140–144 (2008).
10. Rehwinkel, J. *et al.* RIG-I detects viral genomic RNA during negative-strand RNA virus infection. *Cell* 140, 397–408 (2010).
11. Schibler, U. & Perry, R. P. The 5'-termini of heterogeneous nuclear RNA: a comparison among molecules of different sizes and ages. *Nucleic Acids Res.* 4, 4133–4149 (1977).
12. Filipowicz, W. *et al.* A protein binding the methylated 5'-terminal sequence, m7GpppN, of eukaryotic messenger RNA. *Proc. Natl. Acad. Sci. U. S. A.* 73, 1559–1563 (1976).
13. George, C. X., John, L. & Samuel, C. E. An RNA editor, adenosine deaminase acting on double-stranded RNA (ADAR1). *J. Interferon Cytokine Res. Off. J. Int. Soc. Interferon Cytokine Res.* 34, 437–446 (2014).
14. Li, X., Ma, S. & Yi, C. Pseudouridine: the fifth RNA nucleotide with renewed interests. *Curr. Opin. Chem. Biol.* 33, 108–116 (2016).
15. Slotkin, W. & Nishikura, K. Adenosine-to-inosine RNA editing and human disease. *Genome Med.* 5, 105 (2013).
16. Spenkuch, F., Motorin, Y. & Helm, M. Pseudouridine: still mysterious, but never a fake (uridine)! *RNA Biol.* 11, 1540–1554 (2014).

17. Shi, H., Wei, J. & He, C. Where, When, and How: Context-Dependent Functions of RNA Methylation Writers, Readers, and Erasers. *Mol. Cell* 74, 640–650 (2019).
18. Roundtree, I. A., Evans, M. E., Pan, T. & He, C. Dynamic RNA Modifications in Gene Expression Regulation. *Cell* 169, 1187–1200 (2017).
19. Kennedy, E. M., Courtney, D. G., Tsai, K. & Cullen, B. R. Viral Epitranscriptomics. *J. Virol.* 91, e02263-16 (2017).
20. Gonzales-van Horn, S. R. & Sarnow, P. Making the Mark: The Role of Adenosine Modifications in the Life Cycle of RNA Viruses. *Cell Host Microbe* 21, 661–669 (2017).
21. Gokhale, N. S. & Horner, S. M. RNA modifications go viral. *PLoS Pathog.* 13, e1006188 (2017).
22. Meyer, K. D. & Jaffrey, S. R. Rethinking m6A Readers, Writers, and Erasers. *Annu. Rev. Cell Dev. Biol.* 33, 319–342 (2017).
23. Jia, G. *et al.* N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* 7, 885–887 (2011).
24. Zheng, G. *et al.* ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* 49, 18–29 (2013).
25. Trixl, L. & Lusser, A. The dynamic RNA modification 5-methylcytosine and its emerging role as an epitranscriptomic mark. *Wiley Interdiscip. Rev. RNA* 10, e1510 (2019).
26. Yang, X. *et al.* 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m5C reader. *Cell Res.* 27, 606–625 (2017).
27. Bujnicki, J. M., Feder, M., Ayres, C. L. & Redman, K. L. Sequence-structure-function studies of tRNA:m5C methyltransferase Trm4p and its relationship to DNA:m5C and RNA:m5U methyltransferases. *Nucleic Acids Res.* 32, 2453–2463 (2004).
28. Motorin, Y., Lyko, F. & Helm, M. 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. *Nucleic Acids Res.* 38, 1415–1430 (2010).
29. McIntyre, W. *et al.* Positive-sense RNA viruses reveal the complexity and dynamics of the cellular and viral epitranscriptomes during infection. *Nucleic Acids Res.* 46, 5776–5791 (2018).
30. Ringeard, M., Marchand, V., Decroly, E., Motorin, Y. & Bennasser, Y. FTSJ3 is an RNA 2'-O-methyltransferase recruited by HIV to avoid innate immune sensing. *Nature* 565, 500–504 (2019).
31. Courtney, D. G. *et al.* Extensive Epitranscriptomic Methylation of A and C Residues on Murine Leukemia Virus Transcripts Enhances Viral Gene Expression. *mBio* 10, e01209-19 (2019).
32. Yang, S. L. *et al.* Comprehensive mapping of SARS-CoV-2 interactions in vivo reveals functional virus-host interactions. *Nat. Commun.* 12, 5113 (2021).

33. Dong, H. *et al.* 2′-O Methylation of Internal Adenosine by Flavivirus NS5 Methyltransferase. *PLOS Pathog.* 8, e1002642 (2012).
34. Deffrasnes, C. *et al.* Genome-wide siRNA Screening at Biosafety Level 4 Reveals a Crucial Role for Fibrillarin in Henipavirus Infection. *PLOS Pathog.* 12, e1005478 (2016).
35. Li, P. *et al.* RNA 2'-O-Methyltransferase Fibrillarin Facilitates Virus Entry Into Macrophages Through Inhibiting Type I Interferon Response. *Front. Immunol.* 13, 793582 (2022).
36. Bentley, D. L. Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr. Opin. Cell Biol.* 17, 251–256 (2005).
37. Gu, M., Rajashankar, K. R. & Lima, C. D. Structure of the Saccharomyces cerevisiae Cet1-Ceg1 mRNA Capping Apparatus. *Structure* 18, 216–227 (2010).
38. Chu, C. *et al.* Structure of the guanylyltransferase domain of human mRNA capping enzyme. *Proc. Natl. Acad. Sci. U. S. A.* 108, 10104–10108 (2011).
39. Cramer, P. *et al.* Coordination between transcription and pre-mRNA processing. *FEBS Lett.* 498, 179–182 (2001).
40. Cowling, V. H. Regulation of mRNA cap methylation. *Biochem. J.* 425, 295–302 (2009).
41. Beelman, C. A. & Parker, R. Degradation of mRNA in eukaryotes. *Cell* 81, 179–183 (1995).
42. Houseley, J. & Tollervey, D. The Many Pathways of RNA Degradation. *Cell* 136, 763–776 (2009).
43. Perry, K. L., Watkins, K. P. & Agabian, N. Trypanosome mRNAs have unusual 'cap 4' structures acquired by addition of a spliced leader. *Proc. Natl. Acad. Sci. U. S. A.* 84, 8190–8194 (1987).
44. Byszewska, M., Śmietański, M., Purta, E. & Bujnicki, J. M. RNA methyltransferases involved in 5' cap biosynthesis. *RNA Biol.* 11, 1597–1607 (2014).
45. Banerjee, A. K. 5'-terminal cap structure in eucaryotic messenger ribonucleic acids. *Microbiol. Rev.* 44, 175–205 (1980).
46. Perry, R. P. & Kelley, D. E. Kinetics of formation of 5' terminal caps in mRNA. *Cell* 8, 433–442 (1976).
47. Smietanski, M. *et al.* Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nat. Commun.* 5, 3004 (2014).
48. Werner, M. *et al.* 2'-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. *Nucleic Acids Res.* 39, 4756–4768 (2011).
49. Daffis, S. *et al.* 2′-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* 468, 452–456 (2010).

50. Züst, R. *et al.* Ribose 2′-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat. Immunol.* 12, 137–143 (2011).
51. Schuberth-Wagner, C. *et al.* A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2'O-Methylated Self RNA. *Immunity* 43, 41–51 (2015).
52. Jiao, X., Chang, J. H., Kilic, T., Tong, L. & Kiledjian, M. A mammalian pre-mRNA 5' end capping quality control mechanism and an unexpected link of capping to pre-mRNA processing. *Mol. Cell* 50, 104–115 (2013).
53. Jiao, X. *et al.* Identification of a quality-control mechanism for mRNA 5'-end capping. *Nature* 467, 608–611 (2010).
54. Xiang, S. *et al.* Structure and function of the 5'-->3' exoribonuclease Rat1 and its activating partner Rai1. *Nature* 458, 784–788 (2009).
55. Chang, J. H. *et al.* Dxo1 is a new type of eukaryotic enzyme with both decapping and 5'-3' exoribonuclease activity. *Nat. Struct. Mol. Biol.* 19, 1011–1017 (2012).
56. Brengues, M., Teixeira, D. & Parker, R. Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science* 310, 486–489 (2005).
57. van Dijk, E. *et al.* Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. *EMBO J.* 21, 6915–6924 (2002).
58. Balagopal, V. & Parker, R. Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs. *Curr. Opin. Cell Biol.* 21, 403–408 (2009).
59. Kiss, D. L. *et al.* Cap homeostasis is independent of poly(A) tail length. *Nucleic Acids Res.* 44, 304–314 (2016).
60. Otsuka, Y., Kedersha, N. L. & Schoenberg, D. R. Identification of a cytoplasmic complex that adds a cap onto 5'-monophosphate RNA. *Mol. Cell. Biol.* 29, 2155–2167 (2009).
61. Trotman, J. B. & Schoenberg, D. R. A recap of RNA recapping. *WIREs RNA* 10, e1504 (2019).
62. Trotman, J. B., Giltmier, A. J., Mukherjee, C. & Schoenberg, D. R. RNA guanine-7 methyltransferase catalyzes the methylation of cytoplasmically recapped RNAs. *Nucleic Acids Res.* 45, 10726–10739 (2017).
63. Ignatochkina, A. V., Takagi, Y., Liu, Y., Nagata, K. & Ho, C. K. The messenger RNA decapping and recapping pathway in Trypanosoma. *Proc. Natl. Acad. Sci. U. S. A.* 112, 6967–6972 (2015).
64. Mukherjee, C. *et al.* Identification of cytoplasmic capping targets reveals a role for cap homeostasis in translation and mRNA stability. *Cell Rep.* 2, 674–684 (2012).
65. Thoma, C. *et al.* Generation of stable mRNA fragments and translation of N-truncated proteins induced by antisense oligodeoxynucleotides. *Mol. Cell* 8, 865–872 (2001).

66. Decroly, E., Ferron, F., Lescar, J. & Canard, B. Conventional and unconventional mechanisms for capping viral mRNA. *Nat. Rev. Microbiol.* 10, 51–65 (2012).
67. Yedavalli, V. S. R. K. & Jeang, K.-T. Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs. *Proc. Natl. Acad. Sci.* 107, 14787–14792 (2010).
68. Mauer, J. *et al.* Reversible methylation of m6Am in the 5' cap controls mRNA stability. *Nature* 541, 371–375 (2017).
69. Wittmann, S., Chatel, H., Fortin, M. G. & Laliberté, J. F. Interaction of the viral protein genome linked of turnip mosaic potyvirus with the translational eukaryotic initiation factor (iso) 4E of Arabidopsis thaliana using the yeast two-hybrid system. *Virology* 234, 84–92 (1997).
70. Goodfellow, I. *et al.* Calicivirus translation initiation requires an interaction between VPg and eIF 4 E. *EMBO Rep.* 6, 968–972 (2005).
71. Balvay, L., Rifo, R. S., Ricci, E. P., Decimo, D. & Ohlmann, T. Structural and functional diversity of viral IRESes. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* 1789, 542–557 (2009).
72. Guidotti, L. G. & Chisari, F. V. Noncytolytic control of viral infections by the innate and adaptive immune response. *Annu. Rev. Immunol.* 19, 65–91 (2001).
73. Malmgaard, L. Induction and Regulation of IFNs During Viral Infections. *J. Interferon Cytokine Res.* 24, 439–454 (2004).
74. Wang, J. *et al.* Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA. *Nucleic Acids Res.* 47, e130 (2019).
75. Wiedermannová, J., Julius, C. & Yuzenkova, Y. The expanding field of non-canonical RNA capping: new enzymes and mechanisms. *R. Soc. Open Sci.* 8, 201979.
76. Jiao, X. *et al.* 5′ End Nicotinamide Adenine Dinucleotide Cap in Human Cells Promotes RNA Decay through DXO-Mediated deNADding. *Cell* 168, 1015-1027.e10 (2017).
77. Plotch, S. J., Bouloy, M. & Krug, R. M. Transfer of 5'-terminal cap of globin mRNA to influenza viral complementary RNA during transcription in vitro. *Proc. Natl. Acad. Sci. U. S. A.* 76, 1618–1622 (1979).
78. Caton, A. J. & Robertson, J. S. Structure of the host-derived sequences present at the 5' ends of influenza virus mRNA. *Nucleic Acids Res.* 8, 2591–2603 (1980).
79. Guilligay, D. *et al.* The structural basis for cap binding by influenza virus polymerase subunit PB2. *Nat. Struct. Mol. Biol.* 15, 500–506 (2008).
80. Dias, A. *et al.* The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature* 458, 914–918 (2009).
81. Yuan, P. *et al.* Crystal structure of an avian influenza polymerase PAN reveals an endonuclease active site. *Nature* 458, 909–913 (2009).

82. Wang, J. *et al.* Development and validation of a hypoxia-related prognostic signature for breast cancer. *Oncol. Lett.* 20, 1906–1914 (2020).
83. De la Peña, M., Kyrieleis, O. J. P. & Cusack, S. Structural insights into the mechanism and evolution of the vaccinia virus mRNA cap N7 methyl-transferase. *EMBO J.* 26, 4913–4925 (2007).
84. Mao, X. & Shuman, S. Intrinsic RNA (guanine-7) methyltransferase activity of the vaccinia virus capping enzyme D1 subunit is stimulated by the D12 subunit. Identification of amino acid residues in the D1 protein required for subunit association and methyl group transfer. *J. Biol. Chem.* 269, 24472–24479 (1994).
85. Schnierle, B. S., Gershon, P. D. & Moss, B. Cap-specific mRNA (nucleoside-O2'-)-methyltransferase and poly(A) polymerase stimulatory activities of vaccinia virus are mediated by a single protein. *Proc. Natl. Acad. Sci. U. S. A.* 89, 2897–2901 (1992).
86. Ahola, T., Laakkonen, P., Vihinen, H. & Kääriäinen, L. Critical residues of Semliki Forest virus RNA capping enzyme involved in methyltransferase and guanylyltransferase-like activities. *J. Virol.* 71, 392–397 (1997).
87. Vasiljeva, L., Merits, A., Auvinen, P. & Kääriäinen, L. Identification of a Novel Function of the AlphavirusCapping Apparatus: RNA 5′-TRIPHOSPHATASE ACTIVITY OF Nsp2 *. *J. Biol. Chem.* 275, 17281–17287 (2000).
88. Devarkar, S. C. *et al.* Structural basis for m7G recognition and 2′-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proc. Natl. Acad. Sci. U. S. A.* 113, 596–601 (2016).
89. Ogino, T. & Banerjee, A. K. The HR motif in the RNA-dependent RNA polymerase L protein of Chandipura virus is required for unconventional mRNA-capping activity. *J. Gen. Virol.* 91, 1311–1314 (2010).
90. Abraham, G., Rhodes, D. P. & Banerjee, A. K. Novel initiation of RNA synthesis in vitro by vesicular stomatitis virus. *Nature* 255, 37–40 (1975).
91. Gupta, K. C. & Roy, P. Alternate capping mechanisms for transcription of spring viremia of carp virus: evidence for independent mRNA initiation. *J. Virol.* 33, 292–303 (1980).
92. Barik, S. The structure of the 5' terminal cap of the respiratory syncytial virus mRNA. *J. Gen. Virol.* 74 (Pt 3), 485–490 (1993).
93. Rahmeh, A. A., Li, J., Kranzusch, P. J. & Whelan, S. P. J. Ribose 2'-O methylation of the vesicular stomatitis virus mRNA cap precedes and facilitates subsequent guanine-N-7 methylation by the large polymerase protein. *J. Virol.* 83, 11043–11050 (2009).
94. Sutto-Ortiz, P. *et al.* The methyltransferase domain of the Respiratory Syncytial Virus L protein catalyzes cap N7 and 2'-O-methylation. *PLoS Pathog.* 17, e1009562 (2021).

95. Martin, B. *et al.* The methyltransferase domain of the Sudan ebolavirus L protein specifically targets internal adenosines of RNA substrates, in addition to the cap structure. *Nucleic Acids Res.* 46, 7902–7912 (2018).

96. Ge, X.-Y. *et al.* Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 503, 535–538 (2013).

97. Menachery, V. D. *et al.* A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat. Med.* 21, 1508–1513 (2015).

98. Hu, B. *et al.* Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. *PLoS Pathog.* 13, e1006698 (2017).

99. Cui, J., Li, F. & Shi, Z.-L. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* 17, 181–192 (2019).

100. Bouvet, M. *et al.* In Vitro Reconstitution of SARS-Coronavirus mRNA Cap Methylation. *PLOS Pathog.* 6, e1000863 (2010).

101. Decroly, E. *et al.* Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex. *PLoS Pathog.* 7, e1002059 (2011).

102. Chen, Y. *et al.* Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *Proc. Natl. Acad. Sci. U. S. A.* 106, 3484–3489 (2009).

103. Seybert, A., Hegyi, A., Siddell, S. G. & Ziebuhr, J. The human coronavirus 229E superfamily 1 helicase has RNA and DNA duplex-unwinding activities with 5'-to-3' polarity. *RNA N. Y. N* 6, 1056–1068 (2000).

104. Ivanov, K. A. & Ziebuhr, J. Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities. *J. Virol.* 78, 7833–7838 (2004).

105. Lehmann, K. C. *et al.* Discovery of an essential nucleotidylating activity associated with a newly delineated conserved domain in the RNA polymerase-containing protein of all nidoviruses. *Nucleic Acids Res.* 43, 8416–8434 (2015).

106. Yan, L. *et al.* Cryo-EM Structure of an Extended SARS-CoV-2 Replication and Transcription Complex Reveals an Intermediate State in Cap Synthesis. *Cell* 184, 184-193.e10 (2021).

107. Shannon, A. *et al.* Protein-primed RNA synthesis in SARS-CoVs and structural basis for inhibition by AT-527. 2021.03.23.436564 (2021) doi:10.1101/2021.03.23.436564.

108. The mechanism of RNA capping by SARS-CoV-2. https://www.researchsquare.com (2022) doi:10.21203/rs.3.rs-1336910/v1.

109. Chouhan, B. P. S., Maimaiti, S., Gade, M. & Laurino, P. Rossmann-Fold Methyltransferases: Taking a 'β-Turn' around Their Cofactor, S-Adenosylmethionine. *Biochemistry* 58, 166–170 (2019).

110. Medvedev, K. E., Kinch, L. N. & Grishin, N. V. Functional and evolutionary analysis of viral proteins containing a Rossmann-like fold. *Protein Sci. Publ. Protein Soc.* 27, 1450–1463 (2018).

111. Rao, S. T. & Rossmann, M. G. Comparison of super-secondary structures in proteins. *J. Mol. Biol.* 76, 241–256 (1973).

112. Denison, M. R., Graham, R. L., Donaldson, E. F., Eckerle, L. D. & Baric, R. S. Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol.* 8, 270–279 (2011).

113. Ogando, N. S. *et al.* The Curious Case of the Nidovirus Exoribonuclease: Its Role in RNA Synthesis and Replication Fidelity. *Front. Microbiol.* 10, 1813 (2019).

114. Ma, Y. *et al.* Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. *Proc. Natl. Acad. Sci. U. S. A.* 112, 9436–9441 (2015).

115. Ferron, F. *et al.* Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proc. Natl. Acad. Sci.* 115, E162–E171 (2018).

116. Ogando, N. S. *et al.* The Enzymatic Activity of the nsp14 Exoribonuclease Is Critical for Replication of MERS-CoV and SARS-CoV-2. *J. Virol.* 94, e01246-20 (2020).

117. Saramago, M. *et al.* New targets for drug design: importance of nsp14/nsp10 complex formation for the 3'-5' exoribonucleolytic activity on SARS-CoV-2. *FEBS J.* 288, 5130–5147 (2021).

118. Chen, Y. *et al.* Structure-function analysis of severe acute respiratory syndrome coronavirus RNA cap guanine-N7-methyltransferase. *J. Virol.* 87, 6296–6305 (2013).

119. Ferron, F., Debat, H. J., Shannon, A., Decroly, E. & Canard, B. A N7-guanine RNA cap methyltransferase signature-sequence as a genetic marker of large genome, non-mammalian Tobaniviridae. *NAR Genomics Bioinforma.* 2, lqz022 (2020).

120. Case, J. B., Ashbrook, A. W., Dermody, T. S. & Denison, M. R. Mutagenesis of S-Adenosyl-l-Methionine-Binding Residues in Coronavirus nsp14 N7-Methyltransferase Demonstrates Differing Requirements for Genome Translation and Resistance to Innate Immunity. *J. Virol.* 90, 7248–7256 (2016).

121. Bouvet, M. *et al.* RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *Proc. Natl. Acad. Sci. U. S. A.* 109, 9372–9377 (2012).

122. Darzacq, X. *et al.* Cajal body-specific small nuclear RNAs: a novel class of 2'-O-methylation and pseudouridylation guide RNAs. *EMBO J.* 21, 2746–2756 (2002).

123. Rebane, A., Roomere, H. & Metspalu, A. Locations of several novel 2'-O-methylated nucleotides in human 28S rRNA. *BMC Mol. Biol.* 3, 1 (2002).

124. Somme, J. *et al.* Characterization of two homologous 2'-O-methyltransferases showing different specificities for their tRNA substrates. *RNA N. Y. N* 20, 1257–1271 (2014).

125. Dai, Q. *et al.* Nm-seq maps 2′-O-methylation sites in human mRNA with base precision. *Nat. Methods* 14, 695–698 (2017).

126. Zhao, Y., Mo, B. & Chen, X. Mechanisms that impact microRNA stability in plants. *RNA Biol.* 9, 1218–1223 (2012).

127. Horwich, M. D. *et al.* The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr. Biol. CB* 17, 1265–1272 (2007).

128. Marchand, V., Blanloeil-Oillo, F., Helm, M. & Motorin, Y. Illumina-based RiboMethSeq approach for mapping of 2'-O-Me residues in RNA. *Nucleic Acids Res.* 44, e135 (2016).

129. Coutard, B. *et al.* Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. *J. Virol.* (2016) doi:10.1128/JVI.02202-16.

130. Dong, Z.-W. *et al.* RTL-P: a sensitive approach for detecting sites of 2′-O-methylation in RNA molecules. *Nucleic Acids Res.* 40, e157 (2012).

131. Hoernes, T. P. *et al.* Nucleotide modifications within bacterial messenger RNAs regulate their translation and are able to rewire the genetic code. *Nucleic Acids Res.* 44, 852–862 (2016).

132. Hoernes, T. P. *et al.* Eukaryotic Translation Elongation is Modulated by Single Natural Nucleotide Derivatives in the Coding Sequences of mRNAs. *Genes* 10, E84 (2019).

133. Choi, J. *et al.* 2'-O-methylation in mRNA disrupts tRNA decoding during translation elongation. *Nat. Struct. Mol. Biol.* 25, 208–216 (2018).

134. Mayer-Barber, K. D. *et al.* Host-directed therapy of tuberculosis based on interleukin-1 and type I interferon crosstalk. *Nature* 511, 99–103 (2014).

135. Mayer-Barber, K. D. & Yan, B. Clash of the Cytokine Titans: counter-regulation of interleukin-1 and type I interferon-mediated inflammatory responses. *Cell. Mol. Immunol.* 14, 22–35 (2017).

136. O'Neill, L. A. J., Golenbock, D. & Bowie, A. G. The history of Toll-like receptors — redefining innate immunity. *Nat. Rev. Immunol.* 13, 453–460 (2013).

137. Kawai, T. & Akira, S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat. Immunol.* 11, 373–384 (2010).

138. Akira, S., Uematsu, S. & Takeuchi, O. Pathogen Recognition and Innate Immunity. *Cell* 124, 783–801 (2006).

139. Kaisho, T. & Akira, S. Toll-like receptor function and signaling. *J. Allergy Clin. Immunol.* 117, 979–987; quiz 988 (2006).

140. Beutler, B. A. TLRs and innate immunity. *Blood* 113, 1399–1407 (2009).

141. Choe, J., Kelker, M. S. & Wilson, I. A. Crystal Structure of Human Toll-Like Receptor 3 (TLR3) Ectodomain. *Science* 309, 581–585 (2005).

142. Tanji, H., Ohto, U., Shibata, T., Miyake, K. & Shimizu, T. Structural Reorganization of the Toll-Like Receptor 8 Dimer Induced by Agonistic Ligands. *Science* 339, 1426–1429 (2013).

143. Zhang, Z. *et al.* Structural Analysis Reveals that Toll-like Receptor 7 Is a Dual Receptor for Guanosine and Single-Stranded RNA. *Immunity* 45, 737–748 (2016).

144. Bergstrøm, B. *et al.* TLR8 Senses Staphylococcus aureus RNA in Human Primary Monocytes and Macrophages and Induces IFN-β Production via a TAK1–IKKβ–IRF5 Signaling Pathway. *J. Immunol.* 195, 1100–1111 (2015).

145. Takaoka, A. *et al.* Integral role of IRF-5 in the gene induction programme activated by Toll-like receptors. *Nature* 434, 243–249 (2005).

146. Alexopoulou, L., Holt, A. C., Medzhitov, R. & Flavell, R. A. Recognition of double-stranded RNA and activation of NF-κB by Toll-like receptor 3. *Nature* 413, 732–738 (2001).

147. Daffis, S., Samuel, M. A., Suthar, M. S., Gale, M. & Diamond, M. S. Toll-Like Receptor 3 Has a Protective Role against West Nile Virus Infection. *J. Virol.* 82, 10349–10358 (2008).

148. Wang, T. *et al.* Toll-like receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis. *Nat. Med.* 10, 1366–1373 (2004).

149. Abe, Y. *et al.* The Toll-Like Receptor 3-Mediated Antiviral Response Is Important for Protection against Poliovirus Infection in Poliovirus Receptor Transgenic Mice. *J. Virol.* 86, 185–194 (2012).

150. Tatematsu, M., Nishikawa, F., Seya, T. & Matsumoto, M. Toll-like receptor 3 recognizes incomplete stem structures in single-stranded viral RNA. *Nat. Commun.* 4, 1833 (2013).

151. Diebold, S. S., Kaisho, T., Hemmi, H., Akira, S. & Reis e Sousa, C. Innate Antiviral Responses by Means of TLR7-Mediated Recognition of Single-Stranded RNA. *Science* 303, 1529–1531 (2004).

152. Heil, F. *et al.* Species-Specific Recognition of Single-Stranded RNA via Toll-like Receptor 7 and 8. *Science* 303, 1526–1529 (2004).

153. Tanji, H. *et al.* Toll-like receptor 8 senses degradation products of single-stranded RNA. *Nat. Struct. Mol. Biol.* 22, 109–115 (2015).

154. Zhang, Z. *et al.* Structural Analyses of Toll-like Receptor 7 Reveal Detailed RNA Sequence Specificity and Recognition Mechanism of Agonistic Ligands. *Cell Rep.* 25, 3371-3381.e5 (2018).

155. Lee, H. K., Lund, J. M., Ramanathan, B., Mizushima, N. & Iwasaki, A. Autophagy-Dependent Viral Recognition by Plasmacytoid Dendritic Cells. *Science* 315, 1398–1401 (2007).

156. Lund, J. M. *et al.* Recognition of single-stranded RNA viruses by Toll-like receptor 7. *Proc. Natl. Acad. Sci.* 101, 5598–5603 (2004).

157. Li, Y. *et al.* Extraordinary GU-rich single-strand RNA identified from SARS coronavirus contributes an excessive innate immune response. *Microbes Infect.* 15, 88–95 (2013).

158. Campbell, G. R., To, R. K., Hanna, J. & Spector, S. A. SARS-CoV-2, SARS-CoV-1, and HIV-1 derived ssRNA sequences activate the NLRP3 inflammasome in human macrophages through a non-classical pathway. *iScience* 24, 102295 (2021).

159. Robbins, M. *et al.* 2'-O-methyl-modified RNAs act as TLR7 antagonists. *Mol. Ther. J. Am. Soc. Gene Ther.* 15, 1663–1669 (2007).

160. van der Sluis, R. M. *et al.* TLR2 and TLR7 mediate distinct immunopathological and antiviral plasmacytoid dendritic cell responses to SARS-CoV-2 infection. *EMBO J.* e109622 (2022) doi:10.15252/embj.2021109622.

161. Englmeier, L. & Subburayalu, J. What's happening where when SARS-CoV-2 infects: are TLR7 and MAFB sufficient to explain patient vulnerability? *Immun. Ageing* 19, 6 (2022).

162. Andrejeva, J. *et al.* The V proteins of paramyxoviruses bind the IFN-inducible RNA helicase, mda-5, and inhibit its activation of the IFN-β promoter. *Proc. Natl. Acad. Sci.* 101, 17264–17269 (2004).

163. Cui, Y. *et al.* The Stat3/5 Locus Encodes Novel Endoplasmic Reticulum and Helicase-like Proteins That Are Preferentially Expressed in Normal and Neoplastic Mammary Tissue. *Genomics* 78, 129–134 (2001).

164. Yoneyama, M. *et al.* The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat. Immunol.* 5, 730–737 (2004).

165. Liu, G. *et al.* Nuclear-resident RIG-I senses viral replication inducing antiviral immunity. *Nat. Commun.* 9, 3199 (2018).

166. Saito, T. *et al.* Regulation of innate antiviral defenses through a shared repressor domain in RIG-I and LGP2. *Proc. Natl. Acad. Sci. U. S. A.* 104, 582–587 (2007).

167. Bruns, A. M. & Horvath, C. M. LGP2 synergy with MDA5 in RLR-mediated RNA recognition and antiviral signaling. *Cytokine* 74, 198–206 (2015).

168. Kolakofsky, D., Kowalinski, E. & Cusack, S. A structure-based model of RIG-I activation. *RNA* 18, 2118–2127 (2012).

169. Chow, K. T., Gale, M. & Loo, Y.-M. RIG-I and Other RNA Sensors in Antiviral Immunity. *Annu. Rev. Immunol.* 36, 667–694 (2018).

170. Horner, S. M., Liu, H. M., Park, H. S., Briley, J. & Gale, M. Mitochondrial-associated endoplasmic reticulum membranes (MAM) form innate immune synapses and are targeted by hepatitis C virus. *Proc. Natl. Acad. Sci.* 108, 14590–14595 (2011).

171. Tan, X., Sun, L., Chen, J. & Chen, Z. J. Detection of Microbial Infections Through Innate Immune Sensing of Nucleic Acids. *Annu. Rev. Microbiol.* 72, 447–478 (2018).

172. Schlee, M. & Hartmann, G. Discriminating self from non-self in nucleic acid sensing. *Nat. Rev. Immunol.* 16, 566–580 (2016).

173. Rehwinkel, J. & Gack, M. U. RIG-I-like receptors: their regulation and roles in RNA sensing. *Nat. Rev. Immunol.* 20, 537–551 (2020).

174. Kato, H. *et al.* Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid–inducible gene-I and melanoma differentiation–associated gene 5. *J. Exp. Med.* 205, 1601–1610 (2008).

175. Kato, H. *et al.* Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* 441, 101–105 (2006).

176. Loo, Y.-M. *et al.* Distinct RIG-I and MDA5 Signaling by RNA Viruses in Innate Immunity. *J. Virol.* 82, 335–345 (2008).

177. Errett, J. S., Suthar, M. S., McMillan, A., Diamond, M. S. & Gale, M. The essential, nonredundant roles of RIG-I and MDA5 in detecting and controlling West Nile virus infection. *J. Virol.* 87, 11416–11425 (2013).

178. Liu, G., Park, H.-S., Pyo, H.-M., Liu, Q. & Zhou, Y. Influenza A Virus Panhandle Structure Is Directly Involved in RIG-I Activation and Interferon Induction. *J. Virol.* 89, 6067–6079 (2015).

179. Schlee, M. *et al.* Recognition of 5′ Triphosphate by RIG-I Helicase Requires Short Blunt Double-Stranded RNA as Contained in Panhandle of Negative-Strand Virus. *Immunity* 31, 25–34 (2009).

180. Saito, T., Owen, D. M., Jiang, F., Marcotrigiano, J. & Gale Jr., M. Innate immunity induced by composition-dependent RIG-I recognition of hepatitis C virus RNA. *Nature* 454, 523–527 (2008).

181. Jiang, M. *et al.* Self-Recognition of an Inducible Host lncRNA by RIG-I Feedback Restricts Innate Immune Response. *Cell* 173, 906-919.e13 (2018).

182. Zhang, Y. *et al.* RIG-I Detects Kaposi's Sarcoma-Associated Herpesvirus Transcripts in a RNA Polymerase III-Independent Manner. *mBio* 9, e00823-18 (2018).

183. Xu, J. *et al.* Identification of a Natural Viral RNA Motif That Optimizes Sensing of Viral RNA by RIG-I. *mBio* 6, e01265-15 (2015).

184. Runge, S. *et al.* In Vivo Ligands of MDA5 and RIG-I in Measles Virus-Infected Cells. *PLOS Pathog.* 10, e1004081 (2014).

185. Chen, Y. G. *et al.* Sensing Self and Foreign Circular RNAs by Intron Identity. *Mol. Cell* 67, 228-238.e5 (2017).

186. Malathi, K., Dong, B., Gale, M. & Silverman, R. H. Small self-RNA generated by RNase L amplifies antiviral innate immunity. *Nature* 448, 816–819 (2007).

187. Wu, B. *et al.* Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5. *Cell* 152, 276–289 (2013).

188. Pichlmair, A. *et al.* Activation of MDA5 Requires Higher-Order RNA Structures Generated during Virus Infection. *J. Virol.* 83, 10761–10769 (2009).

189. Wies, E. *et al.* Dephosphorylation of the RNA Sensors RIG-I and MDA5 by the Phosphatase PP1 Is Essential for Innate Immune Signaling. *Immunity* 38, 437–449 (2013).

190. Jiang, X. *et al.* Ubiquitin-Induced Oligomerization of the RNA Sensors RIG-I and MDA5 Activates Antiviral Innate Immune Response. *Immunity* 36, 959–973 (2012).

191. Gack, M. U. *et al.* TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature* 446, 916–920 (2007).

192. Oshiumi, H., Matsumoto, M., Hatakeyama, S. & Seya, T. Riplet/RNF135, a RING Finger Protein, Ubiquitinates RIG-I to Promote Interferon-β Induction during the Early Phase of Viral Infection *. *J. Biol. Chem.* 284, 807–817 (2009).

193. Oshiumi, H., Miyashita, M., Matsumoto, M. & Seya, T. A Distinct Role of Riplet-Mediated K63-Linked Polyubiquitination of the RIG-I Repressor Domain in Human Antiviral Innate Immune Responses. *PLOS Pathog.* 9, e1003533 (2013).

194. Lang, X. *et al.* TRIM65-catalized ubiquitination is essential for MDA5-mediated antiviral innate immunity. *J. Exp. Med.* 214, 459–473 (2016).

195. Murali, A. *et al.* Structure and Function of LGP2, a DEX(D/H) Helicase That Regulates the Innate Immunity Response *. *J. Biol. Chem.* 283, 15825–15833 (2008).

196. Komuro, A. & Horvath, C. M. RNA- and Virus-Independent Inhibition of Antiviral Signaling by RNA Helicase LGP2. *J. Virol.* 80, 12332–12342 (2006).

197. Satoh, T. *et al.* LGP2 is a positive regulator of RIG-I– and MDA5-mediated antiviral responses. *Proc. Natl. Acad. Sci.* 107, 1512–1517 (2010).

198. Venkataraman, T. *et al.* Loss of DExD/H Box RNA Helicase LGP2 Manifests Disparate Antiviral Responses. *J. Immunol.* 178, 6444–6455 (2007).

199. Uchikawa, E. *et al.* Structural Analysis of dsRNA Binding to Anti-viral Pattern Recognition Receptors LGP2 and MDA5. *Mol. Cell* 62, 586–602 (2016).

200. Wen, H., Miao, E. A. & Ting, J. P.-Y. Mechanisms of NOD-like receptor-associated inflammasome activation. *Immunity* 39, 432–441 (2013).

201. Franchi, L., Warner, N., Viani, K. & Nuñez, G. Function of Nod-like Receptors in Microbial Recognition and Host Defense. *Immunol. Rev.* 227, 106–128 (2009).

202. Kanneganti, T.-D., Lamkanfi, M. & Núñez, G. Intracellular NOD-like receptors in host defense and disease. *Immunity* 27, 549–559 (2007).

203. Sabbah, A. *et al.* Activation of innate immune antiviral responses by Nod2. *Nat. Immunol.* 10, 1073–1080 (2009).

204. Wang, P. *et al.* Nlrp6 regulates intestinal antiviral innate immunity. *Science* 350, 826–830 (2015).

205. Wang, Y. *et al.* Mitochondria-localised ZNFX1 functions as a dsRNA sensor to initiate antiviral responses through MAVS. *Nat. Cell Biol.* 21, 1346–1356 (2019).

206. Zhang, Z., Yuan, B., Lu, N., Facchinetti, V. & Liu, Y.-J. DHX9 pairs with IPS-1 to sense double-stranded RNA in myeloid dendritic cells. *J. Immunol. Baltim. Md 1950* 187, 4501–4508 (2011).

207. Mitoma, H. *et al.* The DHX33 RNA helicase senses cytosolic RNA and activates the NLRP3 inflammasome. *Immunity* 39, 123–135 (2013).

208. Lu, H. *et al.* DHX15 Senses Double-Stranded RNA in Myeloid Dendritic Cells. *J. Immunol.* 193, 1364–1372 (2014).

209. Zhang, Z. *et al.* DDX1, DDX21, and DHX36 helicases form a complex with the adaptor molecule TRIF to sense dsRNA in dendritic cells. *Immunity* 34, 866–878 (2011).

210. Gringhuis, S. I. *et al.* HIV-1 blocks the signaling adaptor MAVS to evade antiviral host defense after sensing of abortive HIV-1 RNA by the host helicase DDX3. *Nat. Immunol.* 18, 225–235 (2017).

211. Cao, L. *et al.* The Nuclear Matrix Protein SAFA Surveils Viral RNA and Facilitates Immunity by Activating Antiviral Enhancers and Super-enhancers. *Cell Host Microbe* 26, 369-384.e8 (2019).

212. Toczydlowska-Socha, D. *et al.* Human RNA cap1 methyltransferase CMTr1 cooperates with RNA helicase DHX15 to modify RNAs with highly structured 5' termini. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 373, 20180161 (2018).

213. Inesta-Vaquera, F. *et al.* DHX15 regulates CMTR1-dependent gene expression and cell proliferation. *Life Sci. Alliance* 1, e201800092 (2018).

214. Isaacs, A. & Lindenmann, J. Virus interference. I. The interferon. *Proc. R. Soc. Lond. B Biol. Sci.* 147, 258–267 (1957).

215. Roberts, R. M., Liu, L., Guo, Q., Leaman, D. & Bixby, J. The Evolution of the Type I Interferons1. *J. Interferon Cytokine Res.* 18, 805–816 (1998).

216. Pestka, S., Krause, C. D. & Walter, M. R. Interferons, interferon-like cytokines, and their receptors. *Immunol. Rev.* 202, 8–32 (2004).

217. Bonjardim, C. A., Ferreira, P. C. P. & Kroon, E. G. Interferons: signaling, antiviral and viral evasion. *Immunol. Lett.* 122, 1–11 (2009).

218. Deal, E. M., Lahl, K., Narváez, C. F., Butcher, E. C. & Greenberg, H. B. Plasmacytoid dendritic cells promote rotavirus-induced human and murine B cell responses. *J. Clin. Invest.* 123, 2464–2474 (2013).

219. Li, X. L., Blackford, J. A. & Hassel, B. A. RNase L mediates the antiviral effect of interferon through a selective reduction in viral RNA during encephalomyocarditis virus infection. *J. Virol.* 72, 2752–2759 (1998).

220. Gongora, C. *et al.* Molecular cloning of a new interferon-induced PML nuclear body-associated protein. *J. Biol. Chem.* 272, 19457–19463 (1997).

221. Gongora, C., Degols, G., Espert, L., Hua, T. D. & Mechti, N. A unique ISRE, in the TATA-less human Isg20 promoter, confers IRF-1-mediated responsiveness to both interferon type I and type II. *Nucleic Acids Res.* 28, 2333–2341 (2000).

222. Imaizumi, T. *et al.* Expression of interferon-stimulated gene 20 in vascular endothelial cells. *Microbiol. Immunol.* 52, 30–35 (2008).

223. Pentecost, B. T. Expression and estrogen regulation of the HEM45 MRNA in human tumor lines and in the rat uterus. *J. Steroid Biochem. Mol. Biol.* 64, 25–33 (1998).

224. Yu, J. *et al.* Identification and validation of a novel glycolysis-related gene signature for predicting the prognosis in ovarian cancer. *Cancer Cell Int.* 21, 353 (2021).

225. Miyashita, H., Fukumoto, M., Kuwahara, Y., Takahashi, T. & Fukumoto, M. ISG20 is overexpressed in clinically relevant radioresistant oral cancer cells. *Int. J. Clin. Exp. Pathol.* 13, 1633–1639 (2020).

226. Lin, S.-L. *et al.* Stimulation of Interferon-Stimulated Gene 20 by Thyroid Hormone Enhances Angiogenesis in Liver Cancer. *Neoplasia N. Y. N* 20, 57–68 (2018).

227. Moser, M. J., Holley, W. R., Chatterjee, A. & Mian, I. S. The Proofreading Domain of Escherichia coli DNA Polymerase I and Other DNA and/Or RNA Exonuclease Domains. *Nucleic Acids Res.* 25, 5110–5118 (1997).

228. Zuo, Y. & Deutscher, M. P. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res.* 29, 1017–1026 (2001).

229. Horio, T. *et al.* Crystal structure of human ISG20, an interferon-induced antiviral ribonuclease. *FEBS Lett.* 577, 111–116 (2004).

230. Chen, J. *et al.* Structural and biochemical insights into small RNA 3′ end trimming by Arabidopsis SDN1. *Nat. Commun.* 9, 3585 (2018).

231. Nguyen, L. H., Espert, L., Mechti, N. & Wilson, D. M. The Human Interferon- and Estrogen-Regulated ISG20/HEM45 Gene Product Degrades Single-Stranded RNA and DNA in Vitro. *Biochemistry* 40, 7174–7179 (2001).

232. Steitz, T. A. & Steitz, J. A. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl. Acad. Sci. U. S. A.* 90, 6498–6502 (1993).

233. Liu, Y. *et al.* Interferon-inducible ribonuclease ISG20 inhibits hepatitis B virus replication through directly binding to the epsilon stem-loop structure of viral RNA. *PLOS Pathog.* 13, e1006296 (2017).

234. Stadler, D. *et al.* Interferon-induced degradation of the persistent hepatitis B virus cccDNA form depends on ISG20. *EMBO Rep.* 22, e49568 (2021).

235. Jiang, D. *et al.* Identification of Five Interferon-Induced Cellular Proteins That Inhibit West Nile Virus and Dengue Virus Infections. *J. Virol.* 84, 8332–8341 (2010).

236. Xu, H. *et al.* [Antiviral activities of ISG20 against hepatitis C virus]. *Zhonghua Gan Zang Bing Za Zhi Zhonghua Ganzangbing Zazhi Chin. J. Hepatol.* 21, 33–37 (2013).

237. Espert, L. *et al.* Interferon-induced exonuclease ISG20 exhibits an antiviral activity against human immunodeficiency virus type 1. *J. Gen. Virol.* 86, 2221–2229 (2005).

238. Weiss, C. M. *et al.* The Interferon-Induced Exonuclease ISG20 Exerts Antiviral Activity through Upregulation of Type I Interferon Response Proteins. *mSphere* 3, e00209-18 (2018).

239. Lee, J.-H. *et al.* Identification of a novel ionizing radiation-induced nuclease, AEN, and its functional characterization in apoptosis. *Biochem. Biophys. Res. Commun.* 337, 39–47 (2005).

240. Couté, Y. *et al.* ISG20L2, a novel vertebrate nucleolar exoribonuclease involved in ribosome biogenesis. *Mol. Cell. Proteomics MCP* 7, 546–559 (2008).

241. Imam, H., Kim, G.-W., Mir, S. A., Khan, M. & Siddiqui, A. Interferon-stimulated gene 20 (ISG20) selectively degrades N6-methyladenosine modified Hepatitis B Virus transcripts. *PLoS Pathog.* 16, e1008338 (2020).

242. Daniels, P. W., Hama Soor, T., Levicky, Q., Hettema, E. H. & Mitchell, P. Contribution of domain structure to the function of the yeast DEDD family exoribonuclease and RNase T functional homolog, Rex1. *RNA* 28, 493–507 (2022).

243. Ramachandran, V. & Chen, X. Degradation of microRNAs by a Family of Exoribonucleases in Arabidopsis. *Science* 321, 1490–1492 (2008).

244. Ferron, F. *et al.* Structural and molecular basis of mismatch correction and ribavirin excision from coronavirus RNA. *Proc. Natl. Acad. Sci. U. S. A.* 115, E162–E171 (2018).