

# AIX-MARSEILLE UNIVERSITÉ

ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTÉ  
(ED62)

Thèse présentée pour obtenir le grade universitaire de  
**Docteur en science**

Spécialité : Maladies infectieuses

Présentée et défendue publiquement le 18 Mars 2022 par

**Oney ORTEGA GRANDA**

---

*Positive strand RNA virus nonstructural proteins (nsps):  
involvement in viral escape and development of antiviral  
tools.*

---

Réalisée dans l'équipe "Réplicases Virales : Structure Mécanisme et Drug-Design",  
Laboratoire "Architecture et Fonction des Macromolécules Biologiques (AFMB)"  
Sous la direction du **Dr. Nadia RABAH**

**Membres du jury de la Thèse :**

Dr. Hélène DUTARTRE (Rapporteur)  
Dr. Vincent PARISSI (Rapporteur)  
Dr. Laurance BRIANT (Examineur)  
Dr. Françoise DEBART (Examineur)  
Dr. Hélène MALET (Examineur)  
Dr. Jean-Claude GUILLEMOT (Examineur)  
Dr. Nadia RABAH (Directrice de thèse)  
Dr. Bruno CANARD (Co-Directeur de thèse)

## ACKNOWLEDGMENTS

First & foremost, I want to thank all the jury members for agreeing to be part of the evaluation of this thesis.

I would like to thank la Fondation Méditerranée Infection for giving me a PhD scholarship in the research field of infectious diseases to work on a project proposed by Nadia Rabah. I also want to thank Yves Bourne for accepting me to do my PhD at the AFMB lab.

With great pleasure, I express my gratitude to my supervisor, Docteur Nadia Rabah for giving me the opportunity to do a PhD with her. I am thankful for her excellent mentorship, guidance, constant encouragement and support. In particular, she made sure that I learn all the necessary skills to survive in France, sorry, in Marseille “car ici c’est pas la capital”. Nadia has incredible energy and vision. I realized during my research work that she is an open-minded advisor, always open to new ideas. Nadia is a person who always has a smile to offer you and when you think things are not going well, she always finds the positive part so that the enthusiasm does not wane. I clearly understand why the others (I won't take the names hahah) are so jealous that I have Nadia as a supervisor, because she is the best. I am thankful to her to have made a substantial impact on my career.

I am extremely thankful to the big boss Bruno Canard for giving me the warm welcome in his group, for his valuable support, help and advice, for being the great scientist, cool boss and person that he is.

A special thanks to François Ferron who always has the answers for teaching me everything on bioinformatics and analysis of protein structures, and providing constant advice for my hardware or software problems during these years.

I want to express here a big thank you to Barbara Selisko for many things. First, for giving me the opportunity to meet her in Havana in August 2015, which opened the doors to the world for me. I thank her for all the support I received from her since I arrived to this lab, for having been my “security valve” at the beginning with whom I could speak Spanish and for helping me whenever I needed it. I just hope that I became “le doctorant indépendant & autonome” que tu voulais, au moins un peu hahahah.

I would also like to express my gratitude to Karine Alvarez for her generous help, advice, scientific discussions and the time spent with the HPLC.

I can't find the right words to say thanks to my "love-hate" pal in the lab, la Bhawnii, but at least I want to say that I am so happy and grateful to life for having this opportunity to share all this time with you in the lab and also outside. For all the discussions, fights, reconciliations but above all, always after yelling and protesting, for your availability to help me and follow me. Thank you Bhawna SAMA for all that you have endured with me and what is left to endure, my little sister, gros bisous.

Merci to Priscila El Kazzi and Adrien Delpal for all tea and coffees and to all the nice and interesting discussions.

I am thankful to all my colleagues in the lab for all the good vibes Ashleigh Shannon, Alice Decombe, Camille Falcou, Léa Lo Bello, Pierre Gauffre and le plus marseillais et supporteur de la meilleure équipe de foot d'Europe Axel Privat Meunier (allez l'OM forever).

Je tiens également à remercier Véronique Fattorini (notre Véro) for all the help in the lab. Also thank you for always reminding me the things that I forgot with a smile, that never changed.

I also want to say thanks to other lab members - Coralie Valle for all the help, for the nice time that we shared in the office and for your efforts to "améliorer mon beau français". To Etienne Decroly, Bruno Coutard, Cécilia Eydoux, Maria Mate, Mikaël Feracci, Jean-Claude Guillemot for all the advise and the knowledge shared at some point with me.

I cannot forget to say thank you to Sergio Hernandez and Priscila Sutto for all their help in the lab since the beginning and all the beers and parties shared.

Je voudrais aussi remercier à Chantal, Fabienne at Patricia pour faire leur excellent travail qui rend le travail quotidien plus facile pour tout le monde.

I am also thankful to all the past lab members whom I met who made my life happy in this lab - Nhung, Van, Afroditi and Candice.

Also I want to say thanks to my Marseille gang Christiana and Linda.

Last, but not least, I thank my family :

Sandra Madariaga because you are a part of my family for all these years. I could also copy the acknowledgement from my Bachelor's thesis or my Master's and put it here because you have always been there to support me, to listen to me and to advise me without changing at all during these passing years (maybe just a little !)

Sergio (Chencho) I could write a thesis just dedicated to you, to express how happy I feel for having met you, por todo lo vivido y no vivido. For being there in the moments that I needed you the most, for all your emotional support and affection. You are one of the best gifts that I take from this adventure of my thesis.

Finally, to the most important people in my life : my parents, quiero dar el más grande de los agradecimientos a mis padres por todo el sacrificio que han hecho a lo largo de la vida para que este y los sueños anteriores se hayan hecho realidad. En especial a mi papá, que aunque la vida no quiso que viera el final de la historia estoy seguro que orgulloso esta donde quiere que se encuentre porque siempre lo estuvo. Agradecido estoy y estaré de tenerlos como mis padres. También quiero agradecer a las otras cuatro personas mas importantes en mi vida: mi hermana, mi sobrino, mi cuñado y ahora a la pequeña de la casa la Ana Paula por todas las alegrías, apoyo y la gran familia que son.



## Abbreviations

2-5A : 2-5 oligoadenylates  
2'PDE : 2'phosphodiesterase  
3CLpro : Chymotrypsin-like enzyme  
4-PBAK : 4-(Piperidin-4-yl) butanoic acid hydrochloride  
ACE2 : Angiotensin-converting enzyme 2  
AdoHcy / SAH : S-adenosylhomocysteine  
AdoMet / SAM: S-adenosylmethionine  
ADP : Adenosine diphosphate  
ALC1- Amplified in liver cancer 1  
Alpha-CoV : Alphacoronavirus  
APN : Aminopeptidase N  
ARDTs : ADP-ribosyl transferases  
ARH3 : ADP-ribosyl hydrolase 3  
AUD : Alphavirus-unique domain  
Beta-CoV : Betacoronavirus  
BFV : Barmah Forest Virus  
Bme : Brucella melitensis  
bMeOH : Betamercaptoethanol  
BSA : Bovine serum albumin  
CAPS : 3-(Cyclohexylamino)-1-propanesulfonic acid  
CAPSO : 3-(Cyclohexylamino)-2-hydroxy-1-propanesulfonic acid  
cDNA : complementary DNA  
CHEZ : 2-(Cyclohexylamino) ethanesulfonic acid  
CHIKV : Chikungunya virus  
CIAP : Calf Intestinal Alkaline Phosphatase  
COVID-19 : Coronavirus disease 2019  
CoVs : Coronaviruses  
CSEs : Conserved sequence elements  
Ct : Threshold cycle  
CV : Column volume  
DAPI :4',6-diamidino-2-phenylindole  
Delta-CoV : Deltacoronavirus  
DMVs: Double-membrane vesicles  
DENV : Dengue virus  
DNA : Deoxyribonucleic acid  
DNase : deoxyribonuclease  
DPP4 : Dipeptidyl-peptidase 4  
DPUP : Domain Preceding Ubl2 and PL2 pro  
ds : Double stranded  
E : Envelope proteins  
*E. coli* : *Escherichia coli*  
ECL :enhanced chemiluminescence  
EEEV : Eastern equine encephalitis virus  
eIF-4E : Eukaryotic translation initiation factor 4E  
eIF2 $\alpha$  : eukaryotic initiation factor 2 $\alpha$   
ELISA :enzyme-linked immunosorbent assay

EMA : European Medicines Agency  
 EMA : European Medicines Agency  
 EMCV: Encephalomyocarditis virus  
 EPPS : 2-(Cyclohexylamino) ethanesulfonic acid  
 ER : Endoplasmic reticulum  
 ERGIC : ER-Golgi intermediate compartment  
 FBS : *Fetal bovine serum*  
 FC : Fold change  
 G3BP : Ras-GTPase-Activating Protein SH3-Domain-Binding Protein  
 Gamma-CoV : Gammacoronavirus  
 GTase : Guanylyltransferase  
 H2BJ : Histone H2B type 1-J  
 H3K27 : Histone 3, lysine 27  
 HAC : Hepatitis A virus  
 HCoV-229E : Human coronavirus 229E  
 HCoV-HKU1 : Human coronavirus HKU1  
 HCoV-NL63 : Human coronavirus NL63  
 HCoV-OC43 : Human coronavirus OC43  
 HCV : Hepatitis C virus  
 HEL : Helicase  
 HEV : Hepatitis E virus  
 HIV : Human immunodeficiency virus  
 hPARP : Human Poly ADP-ribose polymerase  
 HPLC : High Performance Liquid Chromatography  
 HRV : Human Rotavirus  
 HVD : *proline-rich hinge/hypervariable region*  
 IFITs : Innate immune restriction factors  
 IFN : Interferon  
 IFNAR : Interferon- $\alpha/\beta$  receptor  
 IgG : Immunoglobulin G  
 IKKs : Inhibitory kappa B kinases  
 IL : Interleukin  
 IMAC : Immobilized metal affinity chromatography  
 IMPDH : Inosine-5'-monophosphate dehydrogenase  
 IRES : Internal ribosome entry site  
 IRF : Interferon regulatory transcription factor  
 ISGF3 : Interferon stimulated gene factor 3  
 ISGs : Interferon- stimulated genes  
 ISRE : Interferon-sensitive response element  
 ITC : Isothermal titration calorimetry  
 JAK : Janus kinases  
 K<sub>d</sub> : Dissociation constant ,  
 kDa : Kilo Dalton  
 LB: Lysogeny broth  
 M : Membrane proteins  
<sup>m7</sup>GTP : 7-methyl-GTP  
 MALDI-TOF : Matrix Assisted Laser Desorption Ionization - Time of Flight  
 MAR : Mono-ADP-ribose

MARYlation: Mono-ADP-ribosylation  
 MAVS : Mitochondrial anti-viral proteins  
 MAYV : Mayaro virus  
 MDA5 : Melanoma differentiation-associated protein 5  
 MERS-CoV : Middle east respiratory syndrome coronavirus  
 MES : 2-(N-morpholino) ethanesulfonic acid  
 mGpppN : N7 position of guanosine  
 mGpppNm : 2'O position of the ribose of the first nucleotide  
 MHV : Murine hepatitis virus  
 MHV-68 : Murine gamma herpesvirus 68  
 MOPS : 3-(N-morpholino) propanesulfonic acid  
 MOPSO : 3-Morpholino-2-hydroxypropanesulfonic acid  
 Mpro : Main protease  
 mRNA : Messenger RNA  
 MTase : Methyltransferase  
 MyD88 : Myeloid differentiation primary response 88  
 N : Nucleocapsid protein  
 N : Stoichiometry  
 NAB : Nucleic-acid binding  
 NAD<sup>+</sup> : Nicotinamide adenine dinucleotide  
 NEMO : NF-kappa-B essential modulator  
 NF-κB or NF-kappa-B : nuclear factor kappa-light-chain-enhancer of activated B cells  
 NMR : Nuclear magnetic resonance  
 nsPs : Non-structural proteins  
 NTP : Nucleoside triphosphate  
 NTPase : Nucleoside-triphosphatase  
 Nts : Nucleotides  
 OAS : Oligoadenylate synthetase  
 OD : Optical Density  
 ONNV : O'nyong-nyong virus  
 ORF : Open reading frame  
 PAMPS : Pathogen-associated molecular pattern  
 PAR : Poly-ADP-ribose  
 PARG : PAR glycohydrolase  
 PARPs : Poly ADP-ribose polymerases  
 PARylation: Poly-ADP-ribosylation  
 PBS : Phosphate-buffered saline  
 PCP : Papain-like cysteine protease  
 PCP: Putative protease region  
 PKR : Protein kinase R  
 PL1pro : papain-like protease domains  
 PL2<sup>pro</sup> : Papain-like protease 2 domain  
 PMSF : Phenylmethylsulfonyl fluoride  
 Pp : Polyprotein  
 PRRs : Pattern recognition receptors  
 PTM : Post-translational modification  
 PVDF : Poly(vinylidene fluoride)  
 RBD : Receptor-binding domain

RdRp : RNA dependent RNA polymerase  
 RFS : Ribosomal frameshift signal  
 RIG-I : Retinoic acid inducible gene – I  
 RLRs : RIG-like receptors  
 RNA : Ribonucleic acid  
 RNase L : Ribonuclease L  
 RPM : Revolutions Per Minute  
 RRV : Ross River virus  
 Rso : Ralstonia solanacearum  
 RT : Room temperature  
 RT-PCR : Reverse transcriptase PCR  
 RTC : Replication-transcription complex  
 RTPase : RNA triphosphatase  
 RVF : Rift valley fever virus  
 S : Glycoprotein spike  
 SARS-CoV : Severe acute respiratory syndrome coronavirus  
 SARS-CoV-2 : Severe acute respiratory syndrome coronavirus -2  
 SDS PAGE : Sodium dodecyl sulphate–polyacrylamide gel electrophoresis  
 SFV : Semliki Forest virus  
 sgRNA : subgenomic RNA  
 SINV : Sindbis virus  
 STAT 1 or 2 : Signal transducer and activator of transcription 1 or 2  
 SUD : SARS-unique domain  
 TAPS : [(2-Hydroxy-1,1-bis(hydroxymethyl)ethyl)amino]-1-propanesulfonic acid  
 TBK-1 : Tank binding kinase 1  
 TBS-T : Tris buffer saline with Tween  
 TEAB : Triethylammonium bicarbonate  
 Th1/2 : Effector T helper 1/2  
 TLC : Thin-layer chromatography  
 TLRs : Toll-like receptors  
 TM : Transmembrane region  
 TRAF : Tumor necrosis factor (TNF) receptor-associated factor  
 TRIF : TIR-domain-containing adapter-inducing interferon- $\beta$   
 TSA : Thermal shift assay  
 TYK2 : Tyrosine kinase 2  
 Ubl1 : Ubiquitin-like domain 1  
 UTR : Untranslated region  
 VEEV : Venezuelan equine encephalitis virus  
 VLPs : Virus-like particles  
 VPg : Viral protein genome-linked  
 vRNA : Viral RNA  
 VSV : Vesicular stomatitis virus  
 WEEV : Western equine encephalitis virus  
 WHO : World Health Organization  
 XD : *Macro* domain  
 ZBD : Zinc binding domain  
 ZIKV : Zika virus  
 ZRV : Endogenous zebrafish retrovirus

$\beta$ SM : Beta-CoVs-specific marker

$\Delta H$  : Enthalpy

$\Delta S$  : Entropy

$\Delta T_m$  : Melting temperature

## Table of figures

**General Introduction**

<b>Figure 1. Antiviral innate immune response.....</b>	<b>16</b>
<b>Figure 2. Diversity of capstructures .....</b>	<b>18</b>
<b>Figure 3. Viral activation and escape of the OAS/RNase L pathway.....</b>	<b>20</b>
<b>Figure 4. The Protein ADP-Ribosylation process.....</b>	<b>22</b>
<b>Figure 5. Different roles of PARPs in IFN and ISGs production.....</b>	<b>24</b>

**Hepatitis E virus: Escaping the OAS/RNase L system**

<b>Figure 6. The global distribution of the four human HEV genotypes. Data adapted from WHO.....</b>	<b>28</b>
<b>Figure 7. HEV genomic organization.....</b>	<b>31</b>
<b>Figure 8. Hepatitis E virus life cycle.....</b>	<b>32</b>

**Article 1. Hepatitis E Virus ORF1 harbors a 2'-5'-Phosphodiesterase (PDE) activity, interfering with Type I Interferon Induction.**

<b>Figure 1. A. Position of H-x- (S / T) -x patterns in ORF1. B. Multiple alignment of the 2H motifs of proteins.....</b>	<b>52</b>
<b>Figure 2. Cellular localisation of ORF1 proteins harboring PDE motifs observed by Immunofluorescence.....</b>	<b>54</b>
<b>Figure 3. Effect of ORF1 HEV domains containing PDE motifs on <math>\beta</math>-IFN induction following poly I:C treatment.....</b>	<b>56</b>
<b>Figure 4. Purification and activity assessment of HEV ORF1 HEL and various XD length fragments.....</b>	<b>58</b>
<b>Figure 5. HPLC hydrolysis profiles of 2-5A by recombinant HEV ORF1 proteins.....</b>	<b>59</b>

## **Alphavirus: Capping process as a specific antiviral target for drug development**

<b>Figure 9. Geographical distribution of the main human pathological alphaviruses.....</b>	<b>68</b>
<b>Figure 10. Enzootic and epizootic/epidemic transmission cycles of Alphavirus.....</b>	<b>70</b>
<b>Figure 11. Genome organization of alphavirus.....</b>	<b>71</b>
<b>Figure 12. Alphavirus replication cycle.....</b>	<b>73</b>
<b>Figure 13. Alphavirus nsP1 capping reactions.....</b>	<b>75</b>
<b>Figure 14. Domain organization of alphavirus nsPs: nsP1, nsP2, nsP3 and nsP4.....</b>	<b>76</b>

## **Macro domain key player in coronavirus infections**

<b>Figure 15. Structure of SARS-CoV and SARS-CoV-2 <i>Macro</i> domain in complex with ADP-ribose molecule.....</b>	<b>104</b>
<b>Figure 16. Coronavirus phylogenetic tree.....</b>	<b>108</b>
<b>Figure 17. Coronavirus genomic organization.....</b>	<b>109</b>
<b>Figure 18. Replication cycle of coronavirus.....</b>	<b>111</b>

## **Article 4. *Macro* domain residue F156: a hallmark of SARS-CoV-2 de-MARylation specificity**

<b>Figure 1: Sequence alignment and structure superposition of SARS-CoV with SARS-CoV-2 <i>Macro</i> domain proteins.....</b>	<b>140</b>
<b>Figure 2: Expression and purification of recombinant SARS-CoV and SARS-CoV-2 <i>Macro</i> domains.....</b>	<b>141</b>
<b>Figure 3: ADP-ribose binding to SARS-CoV and SARS-CoV-2 recombinant <i>Macro</i> domains.....</b>	<b>142</b>
<b>Figure 4: de-MARylation of automodified PARP3 by CoV <i>Macro</i> domains.....</b>	<b>143</b>

**Article 5. MOPS derivatives as a scaffold for Macro domain targeted inhibition**

<b>Figure1: SARS-CoV <i>Macro</i> domain in complex with MOPS .....</b>	<b>153</b>
<b>Figure 2. ELISA for ADP-ribose inhibitors high scale pre-screen.....</b>	<b>155</b>
<b>Figure 3. MOPS analogs screen of ADP-ribose binding inhibition towards <i>Macro</i> domains of SARS-CoV and SARS-CoV-2.....</b>	<b>156</b>
<b>Figure 4. MOPSO and CAPSO inhibit PAR-hPARP1 binding to <i>Macro</i> domains of SARS-CoV and SARS-CoV-2.....</b>	<b>158</b>
<b>Figure 5. Effect of CAPSO on the thermal stability of SARS-CoV and SARS-CoV-2 <i>Macro</i> domains.....</b>	<b>160</b>
<b>Figure 6. Molecular docking and amino acids interaction in SARS-CoV <i>Macro</i> domain.....</b>	<b>162</b>



**Abstract**

Positive stranded RNA viruses comprise many (re)-emerging human pathogens that pose a public health problem. No effective antiviral drugs are available to treat a number of these viruses. In the course of the infection, the presence of viral RNA replication intermediates in the cytoplasm triggers the antiviral response by activating different pathways. Hence, viral RNA is recognized by cytoplasmic receptors such as: (i) protein kinase R (PKR), which inhibits protein translation and mediates signaling to enhance IFN production and activate Interferon stimulated genes (ISGs). (ii) Oligoadenylate synthetase (OAS), leading to the synthesis of short 2-5 oligonucleotides (2-5A). The latter activate the latent cellular RNase L, which degrades viral and cellular RNAs, restricting viral infection. (iii) Retinoic acid inducible gene – I (RIG-I) and (Melanoma differentiation-associated protein 5) MDA5 sensors, promoting the induction of ISGs and ultimately infected cells apoptosis. However, viruses have developed multiple strategies to evade the innate immune response and increase their replication capacity. Non-structural proteins play a key role in this interference by interacting specifically with host factors. Here we investigate the role of non-structural proteins from alphavirus-like super family and coronavirus in viral escape. In alphavirus, we studied the capping enzyme nsp1, carrying methyl- and guanylyltransferase activities and involved in translational shut down escape. We characterized the role of residues involved in ribavirin resistance and brought some light on the RNA substrate requirements for guanylyltransferase reaction. For Hepatitis E virus (HEV), member of the alpha-like super family, our data show that the domains containing Hx(S/T)x motifs, a hallmark of phosphodiesterase (PDE) activity, can modify/degrade 2-5A. Also, these domains inhibit IFN production in mammalian cells. These findings suggest that HEV could escape innate immune response through the inhibition of OAS/RNase L system. In parallel, we characterized SARS-CoV-2 *Macro* domain. *Macro* domains can bind and hydrolyze ADP-ribose derivatives, important for IFN production and ISGs induction, making it an attractive anti-viral target. Through mutagenesis study we highlighted the importance of ADP-ribose orientation within the binding cleft for hydrolysis activity. Moreover, based on *Macro* domain 3D structure we designed and tested potential inhibitors, identifying

key molecular determinants for drug optimization studies. These studies can be extrapolated to other RNA viruses and can lead more broadly to the development of antiviral tools against other human pathogens, presenting a real public health problem.

### **Resumé**

Les virus à ARN simple brin de polarité positive comprennent de nombreux agents pathogènes humains (ré)émergents et représentent de réelles menaces de santé publique. A date, aucun médicament antiviral n'est disponible. Au cours de l'infection, la présence d'intermédiaires de réplication de l'ARN viral dans le cytoplasme déclenche la réponse antivirale en activant différentes voies. Par conséquent, l'ARN viral est reconnu par des récepteurs cytoplasmiques tels que : (i) la protéine kinase R (PKR), qui inhibe la traduction des protéines et médie la signalisation pour améliorer la production d'IFN et activer les gènes stimulés par l'interféron (ISG). (ii) Oligoadenylate synthetase (OAS), conduisant la synthèse d'oligonucléotides courts 2'-5' (2-5A). Ces derniers activent la RNase L cellulaire, latente, qui dégrade les ARN viraux et cellulaires, limitant l'infection virale. (iii) les senseurs RIG-I et MDA5, favorisant l'induction des ISGs et conduisant à l'apoptose des cellules infectées. Cependant, les virus ont développé de multiples stratégies pour échapper à la réponse immunitaire innée et augmenter leur capacité de réplication. Les protéines non structurales (nsP) jouent un rôle clé dans cette interférence en interagissant spécifiquement avec les facteurs de l'hôte. Ici, nous étudions le rôle des nsps de la super famille de alphavirus-like et du coronavirus dans l'évasion virale. Chez les alphavirus, nous caractérisons l'enzyme de coiffage nsp1, porteuse des activités méthyl et guanylyltransférase, et impliquée dans l'échappement de l'arrêt de la traduction. De ce fait, nous avons caractérisé le rôle des résidus impliqués dans la résistance à la ribavirine et avons mis en lumière les spécificités de substrats ARN pour la réaction de guanylyl transfer. Pour le virus de l'hépatite E (VHE), membre de la super famille de alphavirus-like, nos données montrent que les domaines contenant des motifs Hx(S/T)x, une caractéristique de l'activité phosphodiesterase (PDE), peuvent modifier/dégrader les 2-5A. De plus, ces domaines inhibent la production d'IFN dans les cellules de mammifères. Ces résultats suggèrent que le VHE pourrait échapper à la réponse immunitaire innée par l'inhibition du système OAS/RNase L.

En parallèle, nous avons caractérisé le domaine *Macro* SARS-CoV-2. Les *Macro* domaines peuvent lier et hydrolyser les dérivés de l'ADP-ribose, importants pour la production d'IFN et l'induction d'ISG, ce qui en fait une cible antivirale attrayante. Grâce à une étude de mutagenèse, nous avons mis en évidence l'importance de l'orientation ADP-ribose dans la clé de liaison pour l'activité d'hydrolyse. De plus, sur la base de la structure 3D du domaine *Macro*, nous avons conçu et testé des inhibiteurs potentiels, identifiant les déterminants moléculaires clés pour les études d'optimisation de ces inhibiteurs. Finalement, ces études peuvent être extrapolées à d'autres virus à ARN et peuvent conduire plus largement au développement d'outils antiviraux contre d'autres agents pathogènes humains.

## INDEX

Abbreviations.....	1
Table of figures.....	6
Abstract/Resumé.....	9
1 General Introduction.....	14
1.1 Positive sense RNA viruses.....	14
1.2 Host response to viral infection.....	15
1.3 Viral escape from the innate immune response.....	17
1.3.1 The addition of the cap structure in viruses.....	17
1.3.1.1 Role of the cap structure.....	19
1.3.2 Evading the OAS/RNase L system.....	19
1.3.3 ADP-rybosylation and poly (ADP-ribose) polymerases.....	21
1.3.3.1 ADP-ribosylation key player in immune response.....	22
1.4 Thesis objectives.....	25
2 Hepatitis E virus: Escaping the OAS/RNase L system.....	26
2.1 History of hepatitis E virus.....	26
2.2 Taxonomic classification and geographical distribution of HEV.....	26
2.3 Epidemiology of HEV.....	28
2.4 Clinical manifestation HEV infections.....	29
2.5 Genome organization of HEV.....	30
2.5.1 Life cycle of HEV.....	31
2.5.2 HEV-ORF1.....	33
2.5.2.1 Methyltransferase.....	33
2.5.2.2 Y domain.....	34
2.5.2.3 Putative protease region : PCP.....	34
2.5.2.4 Hyper variable region (HVD).....	35
2.5.2.5 <i>Macro</i> domain.....	36
2.5.2.6 The helicase /NTPase domain.....	36
2.5.2.7 RNA dependent RNA polymerase (RdRp).....	37
2.5.3 HEV-ORF2.....	38
2.5.4 HEV-ORF3.....	38
2.5.5 HEV-ORF4.....	39
2.6 Treatment and Vaccine.....	39
2.7 Objectives.....	40
2.8 Article 1.....	43
3 Alphavirus: Capping process as a specific antiviral target for drug development.....	66
3.1 Alphavirus History.....	66
3.2 Classification and geographical distribution.....	66
3.3 Clinical manifestation and treatments.....	68
3.4 Genome organization.....	70
3.4.1 life cycle.....	71
3.4.2 Virus replication.....	72
3.4.3 nsPs: Viral replication complex.....	74
3.4.3.1 nsP1.....	74
3.4.3.2 nsP2.....	75
3.4.3.3 nsP3.....	76

---

3.4.3.4 nsP4.....	77
3.4.4 Structural proteins involved in assembly and viral release.....	77
3.5 Objectives.....	78
3.6 Article 2.....	80
3.7 Article 3.....	93
4 Macro domain key player in coronavirus infections.....	103
4.1 Macro Domains.....	103
4.1.1 Viral <i>Macro</i> Domains: prominent antiviral targets.....	105
4.2 Coronavirus: Classification and epidemiology.....	106
4.2.1 Genome organization and replication cycle.....	108
4.2.1.1 Replication cycle.....	109
4.2.1.2 nsp3.....	112
4.3 Coronavirus: treatments and vaccines.....	113
4.4 Objectives.....	115
4.5 Article 4.....	118
4.6 Article 5.....	144
5 General Conclusions and perspectives.....	169
6 References.....	174

# 1 General Introduction

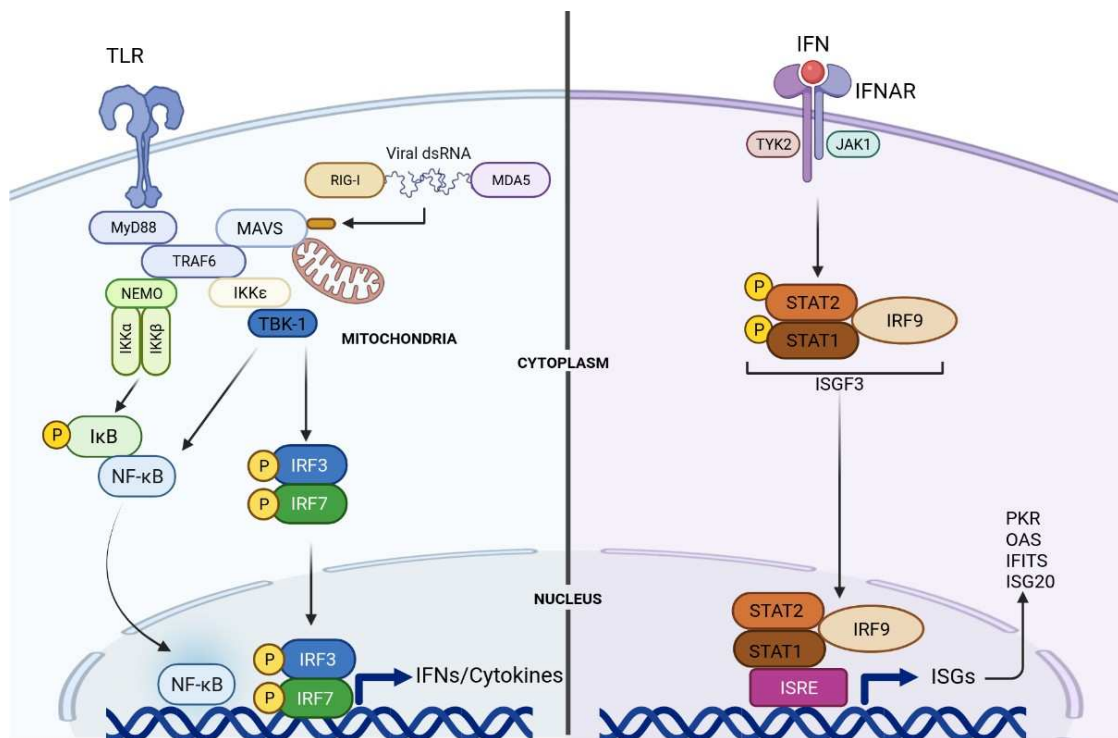
## 1.1 Positive sense RNA viruses

Viruses are small infectious agents capable of multiplication inside living host cells. Viruses are present in all kind of environments in the nature. Therefore, they have the capability of targeting organisms from all kingdoms of life. On the basis of genetic material, viruses can be classified as having DNA or RNA genomes, either single or double stranded. Positive strand (+) RNA viruses form the largest group in this classification. In these viruses the genetic material, i.e RNA, is in the same sense as that of the host cell mRNA and could be directly translated to viral proteins by using host cell machinery (1). (+) RNA viruses shares similarities, in particular at the level of the replication strategy, which can be generalized as follows: (i) The viral genome has a cap at the 5' end as well as a polyadenylation tail at the 3' end. (ii) Once released into the cytoplasm, the cellular machinery will translate the first two thirds of the genome into non-structural proteins (nsPs), which will form the viral replication complex associated to membranes. These membrane associated replication complexes are the sites where all viral replication factors are concentrated and protected from host cell defense mechanisms (2). (iii) The viral RNA dependent RNA polymerase (RdRp) ensures the transcription of the genomic RNA into RNA of negative polarity (-) RNA, which will serve as a template for the production of new genomic RNAs. This RNA (-) will also allow the synthesis of subgenomic RNA (sgRNA), leading to the production of structural proteins. The latter forming the capsid of new virions. Several human pathogens such as severe acute respiratory syndrome coronavirus (SARS-CoV), SARS-CoV-2, middle east respiratory syndrome coronavirus (MERS-CoV), Poliovirus, Hepatitis A, C and E viruses (HAV, HCV, and HEV), Rubella virus and Alphavirus have (+) RNA genomes.

## 1.2 Host response to viral infection

Infection of cells by (+) RNA viruses triggers the antiviral response of the cell following the recognition of pathogen-associated molecular pattern (PAMPS) by pattern recognition receptors (PRRs). Two types of PRRs are activated by (+) RNA viruses: (i) transmembrane Toll-like receptors (TLRs), present in endosomes and on the cell surface (ii) receptors related to the retinoic acid inducible gene I-like receptor family (RIG-I-like), such as RIG-I or MDA5, containing a RNA interaction domain, allowing them to locate viral RNA in the cytoplasm. Activation of these PRRs induces a reaction cascade involving the recruitment of, among other actors, mitochondrial anti-viral proteins (MAVS) and the regulatory factor of Interferon 3 (IRF 3). This interaction allows the recruitment of NF-kappa-B essential modulator (NEMO) that activates the IKKs complex. The latter leads to the activation of the transcription factor NF-kB. The transcription factor NF-kB is translocated to the nucleus, stimulating genes implicated in immune and inflammatory responses (3). On the other hand, the activation of TLR on the cell surface and in the cytoplasm allows the recruitment of effectors such as Myeloid differentiation primary response 88 (MyD88) and Tumor necrosis factor (TNF) receptor-associated factor (TRAF), thereby activating Tank binding kinase 1 (TBK-1). TBK1 phosphorylates IRF3 and IRF7, allowing their dimerization and translocation to the nucleus. Activation of these factors leads to the induction of type I interferon (IFN). Type I IFN, namely IFN $\alpha$  and IFN $\beta$ , play an important role in the antiviral response (4). Type I IFNs will act in a paracrine and an autocrine ways, by binding to IFN receptor (IFNAR). The latter, once activated, allows the recruitment of tyrosine kinases, namely Janus kinases (JAK1) and Tyrosine kinase 2 (TYK2), leading to the phosphorylation of Signal transducer and activator of transcription factors 1 and 2 (STAT1 and STAT2). These dimerize and associate with a third factor, interferon Regulatory Factor 9 (IRF9) forming the interferon stimulated gene factor 3 complex (ISGF3). The ISGF3 complex is translocated into the nucleus, where it binds to sequences of Interferon-sensitive response element (ISRE), stimulating the transcription of the interferon-

stimulated genes (ISGs), among which 2-5-oligoadenylate synthetase (OAS), protein kinase R (PKR) and innate immune restriction factors (IFITs) (Figure 1) (5). IFITs are an IRF-inducible family of genes with antiviral activity against several RNA viruses. The accumulation of double-stranded replication intermediates (dsRNA), generated during viral replication, will act as a signal to activate PKR and OAS. PKR inhibits protein translation, by phosphorylating the initiation factor eIF2 $\alpha$ , and stimulating signaling mediated by NF- $\kappa$ B to activate IFN-stimulated genes. OAS converts ATP to specific oligoadenylates (2-5A), linked by a 2'-5' phosphoester bond. These 2-5A will activate a cytoplasmic ribonuclease L (RNase L) (6). Most viruses induce an IFN and ISG response, although the kinetics and magnitude of this induction might vary. However, it is unclear how early in infection these defenses. The timing of the IFN and ISGs response during initial viral replication are difficult to determine in some virus.



**Figure 1. Antiviral innate immune response.** The activation of innate immunity receptors such as TLRs and "RIG-like receptors (RLRs) leads to the activation of a



signaling cascade, allowing the production of IFNs and cytokines that participate in the antiviral defense. See text for details. The image was created in *BioRender.com*.

### **1.3 Viral escape from the innate immune response**

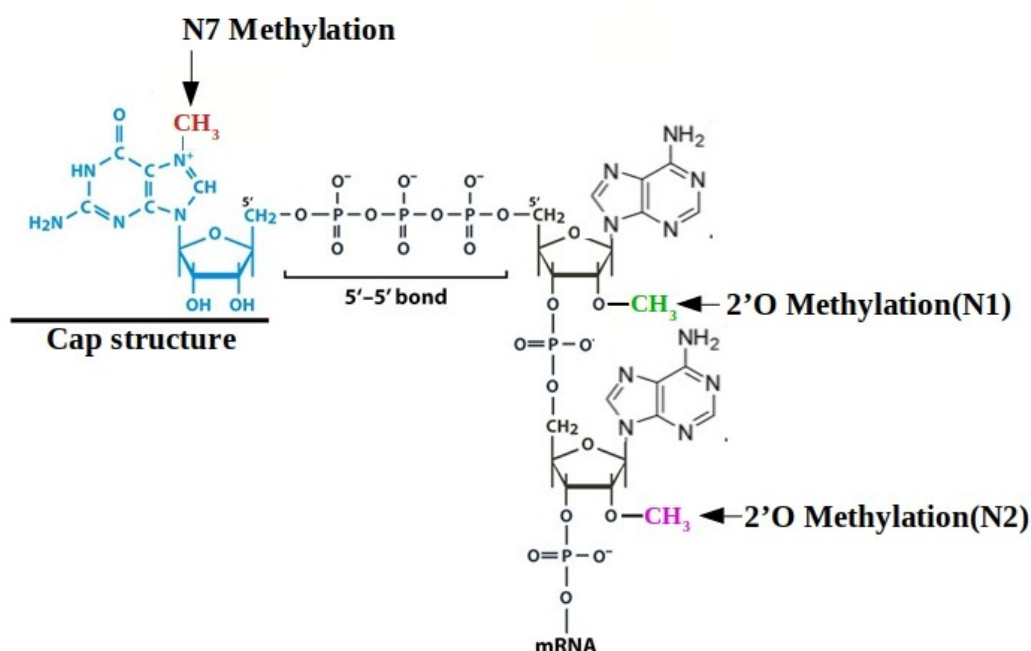
The progression of the viral infection depends on the ability of the virus to counter the effect of PRRs activation and IFNs production. Viruses have developed various mechanisms in order to avoid the innate immune response. This evasion mechanisms involve multiple viral actors (molecular determinants), acting on many cellular targets. Among viral escape mechanisms: (i) The addition of a cap structure at the 5' end of the viral RNA. (ii) The inhibition of OAS/RNase L pathway. (iii) The interference with IFN and ISGs induction through post-translational modifications, such as ADP-ribosylation. Viral nsPs will play a crucial role in the evasion process, by interacting specifically with host factors.

#### **1.3.1 The addition of the cap structure in viruses.**

New viral transcripts undergo post-transcriptional changes and modifications. The cap is a chemical structure made up of a guanosine attached to the first nucleotide at the 5' end of the mRNA by a 5'-5' triphosphate bond. This structure can be methylated at several positions: (i) the N7 position of guanosine (mGpppN) forming a cap-0 and (ii) the 2'O position of the ribose of the first nucleotide (mGpppNm) or the second nucleotide (mGpppNmNm) forming a cap-1 or cap-2 structures, respectively (figure 2).

There are three different strategies that allow the virus to acquire a cap structure at the 5' end of its mRNA (10). First, for retroviruses and most DNA viruses, whose mRNAs are synthesized by RNA polymerase II, the synthesis of the cap structure is carried out by cellular enzymes. A second strategy characteristic of the segmented negative strand

RNA virus, from the *Bunyavirales* order and *Orthomyxoviridae* family, consists of stealing a cap present on the cellular mRNA through a "cap-snatching" mechanism. The last mechanism consists of encoding its own enzymatic machinery necessary for the formation of the cap structure. This strategy is used in particular by viruses that have a cytoplasmic replicative cycle, such as alphavirus, coronaviruses and flaviviruses. However, not all viruses will synthesize a cap structure in the 5' end of their mRNAs. Hence, an alternative strategy, called "cap-independent", ensuring the translation of viral proteins and the protection of their transcripts do exist in some viral families. For example, viruses of the *Calciviridae* family mask the 5' end of their RNAs by covalently binding to a VPg protein (viral protein genome-linked), which interacts directly with the eIF4E protein and thus initiates the translation of viral proteins (8). Other viruses, such as hepaciviruses and pestiviruses, present at the 5' end of their transcripts a specific fold called IRES (internal ribosome entry site) that allows the translation of viral proteins to be initiated without passing through the translation factor eIF-4E (12).



**Figure 2. Diversity of cap structures.** The cap structure is formed by a guanosine (blue nucleoside) methylated at the N7 position (red) linked a tri-phosphate bond to

the 5' end of the mRNA. First and second can also be methylated but at the 2' position (green and pink). The image was created in *BioRender.com*.

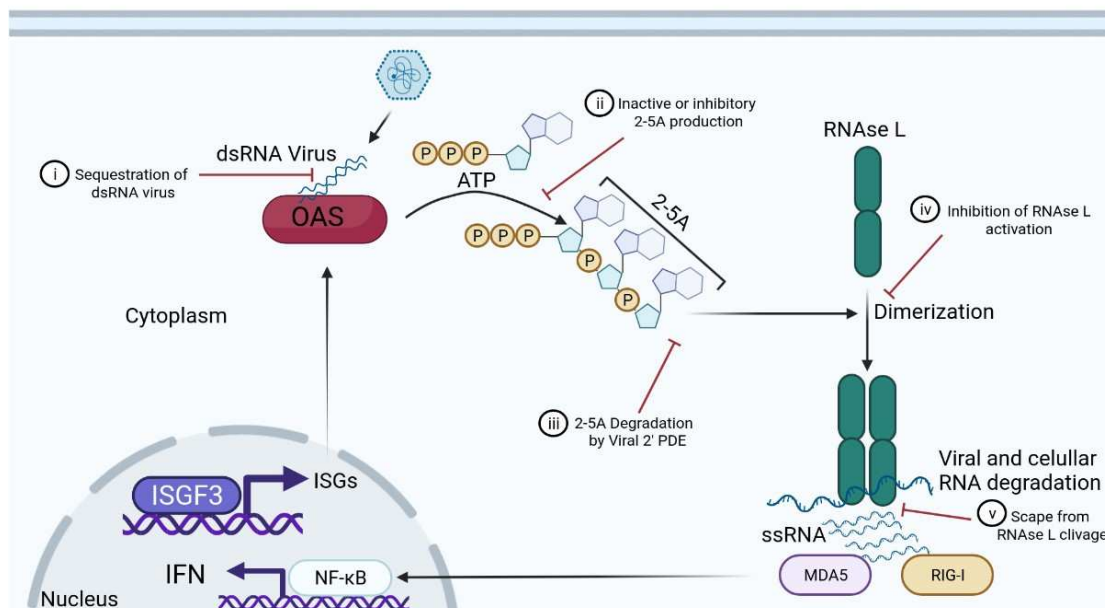
### **1.3.1.1 Role of the cap structure**

Cap structure plays an important role in many biological processes. It is necessary for mRNA maturation, alternative splicing, nuclear export, translation and stability. In the cytoplasm, cap structures are implicated in the recruitment of ribosomes and translation factors notably eIF-4E, allowing efficient translation of proteins (10). Viruses using the cellular ribosomal machinery for the translation of their proteins, will recruit eIF-4E through the interaction of their cap structure at the 5' of mRNA. The cap structure stabilizes vRNAs by protecting them from degradation by cellular 5' end exonucleases. Moreover, the addition of the cap structure to the 5' end of the viral RNAs limits their detection by innate immunity sensors. The methylation at the 2'O position of the first nucleotide represents a marker of "self". Viral RNAs lacking this modification will be detected by the innate immunity sensors, leading to the production of IFN (15). ISG proteins produced as a result of the activation of the IFN pathway, bind and sequester vRNAs lacking 2'O methylation, limiting their translation (11). Some viral 2'O MTases are capable of methylating RNA sequence (12). The role of these methylations is not yet well understood. It has recently been shown, for human immunodeficiency virus (HIV), that methylations at internal positions recruit a 2'O cellular methyltransferase enzyme (MTase) (FtsJ3), allowing the methylation of the viral RNA at 17 specific positions. These specific internal methylations participate in the evasion process, limiting viral detection by the MDA5 receptor and the induction of the IFN pathway (13).

### **1.3.2 Evading the OAS/RNase L system**

One of the key pathways in the activation of type I INFs is the interferon dependent Oligoadenylate synthetase / endoribonuclease L (OAS/RNase L) pathway (figure 3). In response to viral infection, IFN secretion induces the expression of OAS enzymes.

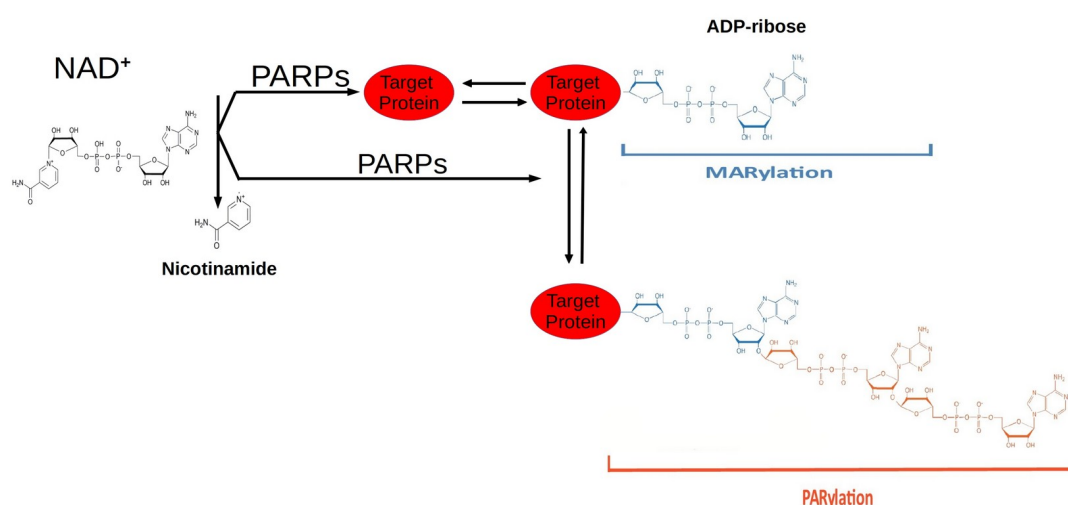
Three OAS family members are produced as inactive enzyme and become active by binding to dsRNAs (formed during viral replication), generating specific 2-5 oligoadenylates (2-5A) from ATP. The 2-5A will bind to the monomeric cellular ribonuclease RNase L, leading to its dimerization and activation. The activated RNase L cleaves cellular and viral RNAs. The cleavage of cellular RNA allows the generation of small RNAs that will activate RIG-I and MDA5 and promote ISGs expression and apoptosis in the infected cell (14). Viral proteins, such as VP35 from Ebola virus, NS1 from influenza A virus, E3L in vaccinia virus and the transactivator protein Tat of HIV, interfere with the activation of the OAS/RNase L system. This interference can occur through sequestration of dsRNA intermediates or direct inhibition/interference with RNase L or OAS (figure 3) (15). Moreover, RNase L degradation escape can occur by the variation of the genomic sequence, reduction of possible cleavage sites or formation of secondary structures resistant to hydrolysis. This is the case for hepatitis C virus and poliovirus, respectively (16). In addition, recent studies have revealed that the non-structural proteins VP3 (Rotavirus), ns2 (murine hepatitis virus) and ns4b (MERS-coronavirus) inhibit the OAS/RNase L pathway by directly degrading 2-5 A involving a 2'phosphodiesterase (2'PDE) activity (17,18).



**Figure 3. Viral activation and escape of the OAS/RNase L pathway.** RNA intermediates resulting from viral replication activate OAS enzymes. OAS convert ATP into 2-5A. 2-5A trigger the activation of RNase L. It degrades viral and cellular RNAs. The generated RNA fragments can activate MDA5 leading to interferon production creating a positive feed-back to the antiviral defense. OAS/RNase L pathway escape: (i) Sequestration of viral dsRNA intermediates. (ii) OAS hijacking by the synthesis of inactive or inhibitory 2–5A. (iii) 2-5A degradation through a PDE activity. (iv) Direct inhibition of RNase L (v) Escape from RNase L cleavage through genome adaptation. The image was created in *BioRender.com*.

### 1.3.3 ADP-rybosylation and poly (ADP-ribose) polymerases

ADP-ribosylation is a post-translational modification reaction corresponding to the transfer of single mono ADP-ribose (MAR) residue (mono-ADP-ribosylation or MARYlation) or multiple poly ADP-ribose (PAR) residues (poly-ADP-ribosylation or PARylation) from NAD<sup>+</sup> to a protein (figure 4). The ADP-ribosyltransferase diphtheria toxin-like family are the enzymes responsible for the catalysis of ADP-ribosylation referred to as poly (ADP-ribose) polymerases (PARPs). This post-translational modification regulates many key pathological and biological processes. Among these are DNA repair, gene transcription, cell differentiation and signal transduction (19). In humans, 17 PARPs are described. Some of them such as PARP3, PARP4, PARP6, PARP10, PARP14 and PARP15 promote the MARYlation, while PARP1, PARP2, PARP5A and PARP5B catalyze PARylation. PARPs have distinct subcellular localizations, indicating their distinctive targets and functions (20). Hence, PARP1 is restricted to the nucleus, while PARP6, PARP8, PARP12, PARP13, PARP15 and PARP16 are mostly located in the cytoplasm. The rest of PARPs are present in both.



**Figure 4. The Protein ADP-Ribosylation process.** The Poly (ADP-ribose) polymerases (PARPs) catalyze the MADylation or PARylation on specific amino acid residues of target proteins using NAD<sup>+</sup> as substrate. These post-translational modifications are reversed by enzymes that remove covalently linked ADP-ribose from PARylation and MADylation proteins. The image was created in *BioRender.com*.

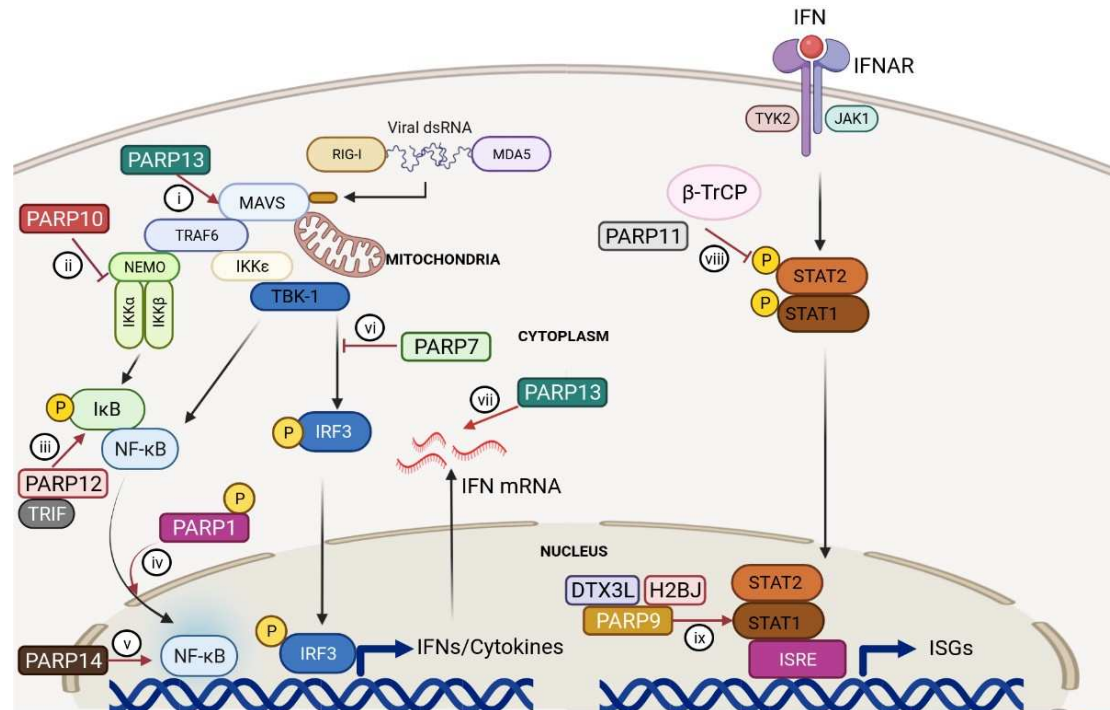
### 1.3.3.1 ADP-ribosylation key player in immune response

In addition to the above described functions, PARPs are able to restrict or promote virus replication and innate immune response. Thus, PARPs modulate IFNs and pro-inflammatory cytokines induction through different mechanisms (figure 5). PARP13 promotes RIG-I dimerization and oligomerization upon innate immunity activation. PARP13 restricts the replication of several RNA viruses by targeting viral or host RNA for degradation (21). In the case of PARP1, it can PARylate NF- $\kappa$ B (nuclear factor kappa-light-chain-enhancer of activated B cells) and promote its activity in the transcription of various cytokines such as IFN- $\alpha/\beta$  (see above), to counteract infection. Further, NF- $\kappa$ B-dependent gene expression is increased by PARP12 after binding to TRIF (TIR-domain-containing adapter-inducing interferon- $\beta$ ). Also,

PARP1 plays a role in the regulation of T cell differentiation into effector T helper 1 (Th1) and Th2; and regulatory T cells (22). Additionally, PARP1 modulates maturation and function of dendritic cells by increasing Interleukin (IL) IL-10 and IL-12 production (21). On the other hand, PARP9 and PARP14 co-regulate pro-inflammatory response in human macrophages (10). PARP9 enhances the expression of ISGs by interacting with ubiquitinated histone H2B type 1-J (H2BJ). PARP14 recruits Pol II and boosts the acetylation of the H3K27 (histone 3, lysine 27), which is associated with active transcription, to promote IFN expression. Moreover, PARP14 interacts and ADP-ribosylates STAT1, reducing the phosphorylation of this pro-inflammatory mediator (22). PARPs appear also to have pro-viral effects by inhibiting different steps in IFNs production and IFN-I signaling (figure 5). As an example, PARP7 seems to ADP-ribosylate TBK-1 to inhibit IRF3 phosphorylation. In the case of PARP10, it prevents the activation of IKKs  $\alpha/\beta$  by interacting with the ADP-ribosylated NEMO. On the other side, PARP11 ADP-ribosylates the E3 ubiquitin ligase  $\beta$ -TrCP, which targets IFNAR for proteasome-dependent degradation after ubiquitinating it (21).

Many reports suggested the contribution of PARPs to the immune response, but the specific molecular mechanisms behind this effect remain not completely understood. However, it is clear that PARPs have strong antiviral role during viral infection. Thus, it was demonstrated that PARP12 is implicated in the inhibition of the replication cycle of Vesicular stomatitis virus (VSV), Murine gamma herpesvirus 68 (MHV-68), Venezuelan equine encephalitis virus (VEEV), Sindbis virus (SINV), encephalomyocarditis virus (EMCV), Rift Valley fever virus (RVF), Chikungunya virus (CHIKV) and Zika virus (ZIKV) (18–21). However, ADP-ribosylation activity turns over rapidly in the cell by some enzymes that remove covalently linked ADP-ribose from proteins. These enzymes include the PAR glycohydrolase (PARG), ADP-ribosyl hydrolase 3 (ARH3), TARG/C6orf130, NUDIX family of hydrolases and MacroD1 and MacroD2 (28). Several of these enzymes contain a *Macro* domain fold that interact and remove ADP-ribose derivatives. Also, conserved *Macro* domains are

present in several virus families as coronaviruses (CoVs), Alphavirus and Hepatitis E virus (HEV).



**Figure 5. Different roles of PARPs in IFN and ISGs production.** PARPs modulate IFN and proinflammatory cytokines induction by different mechanisms. (i) PARP13 promotes RIG-I oligomerization. (ii) PARP10 prevents the activation of IKKs. (iii) PARP12 enhances NFκB-dependent gene expression. (iv) PARP1 can poly-ADP-ribosylate NFκB and promote its activity. (v) PARP14 enhances IFN production. (vi) PARP7 inhibits iTBK-1 from phosphorylating IRF3. (vii) PARP13 target IFN mRNA to degradation. (viii) PARP11 counters the IFNAR by ADP-ribosylating the E3 ubiquitin ligase β-TrCP. (ix) PARP9 and DTX3L enhance ISGs expression. See main text for details. The image was adapted from *Fehr et al, 2020* (21) and created in *BioRender.com*.



## 1.4 Thesis objectives

### General aim:

This thesis is based on the characterization of nsPs from (+) RNA viruses. Specifically, the identification and the characterization of enzymatic activities and molecular motifs of Alpha-like virus and coronavirus, implicated in the inactivation of the innate immune response. Alpha-like virus and coronavirus are used here as a model to elucidate the function of nsPs in viral replication and escape. HEV serves as a model for the study of viral escape conducted through the inhibition of the OAS/RNaseL pathway. In VEEV we studied the mechanism of addition of the cap structure, unique in these viruses and constituting an attractive therapeutic target. Furthermore, coronavirus is the chosen model to study the role of the *Macro* domain in viral escape via ADP ribosylation.

### Thesis goals:

To achieve the above listed aims, it is necessary to reach the following goals:

- (i) Identify and characterize enzymatic activities and molecular motifs of HEV, implicated in the degradation 2-5A oligoadenylates, activators of OAS/RNase L pathway.
- (ii) Understand drug resistance associated to capping enzymes alteration and explore the specificity of nsP1-mediated capping reaction, important for viral RNA translation.
- (iii) Conduct a biochemical characterization of SARS-CoV-2 *Macro* domain, involved in immuno-modulation.

## **2 Hepatitis E virus: Escaping the OAS/RNase L system**

### **2.1 History of hepatitis E virus**

The recognition of viral hepatitis as an infectious disease dates from the 8th century AC. By the middle of the 20th century, it was clear that acute hepatitis consisted of two separate diseases: infectious hepatitis and serum hepatitis, acquired through enteric and parenteral routes, respectively. They were tentatively named hepatitis A and hepatitis B. In the 1970s, the agents responsible for these diseases were discovered and named hepatitis A virus (HAV) and hepatitis B virus (HBV), respectively. It was not until the early 1980s, when the development of sensitive serological tests for these agents led to the realization that a large proportion of cases with post-transfusion hepatitis were not related to any of these agents. Such cases were labelled provisionally as caused by a non-A, non-B post-transfusion hepatitis agent (29). Thus, begins the search for a new causative agent of viral hepatitis, now called hepatitis C virus (HCV). The first recorded outbreak of epidemic jaundice probably caused by hepatitis E (HEV) occurred on the island of Martinique in 1858. The onset of the disease in pregnant women and adults characterized the hepatitis E outbreaks and confirm that this outbreak was caused by HEV and not by HAV or HBV (30).

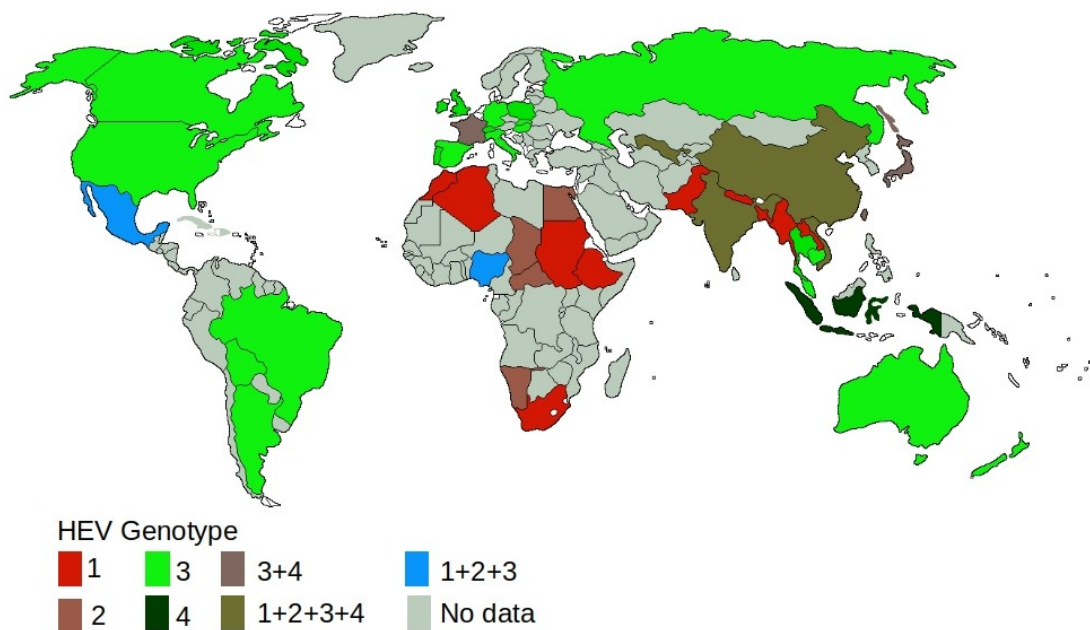
### **2.2 Taxonomic classification and geographical distribution of HEV**

In 2014, a new proposal was published for the classification of the *Hepeviridae* family (31). This new classification divides the *Hepeviridae* family into two genera: *Orthohepevirus* (all HEV isolates from birds and mammals) and *Piscihepevirus* (HEV from trout). Within the genus *Orthohepevirus*, four different species (A, B, C,

D) are designated to include isolates from different hosts. In the *Orthohepevirus A* group, four genotypes of HEV (HEV-1 to 4) were initially described. The HEV-1 is the most conserved among the HEV-1 to 4 that infect humans (32). HEV-1 and 2 are restricted to humans without known animal reservoirs. In the case of HEV-3 and 4, they are zoonotic with an expanded host range (31).

Strains of HEV with unique viral nucleotides sequences isolates, in Japan from wild boar, were designated as HEV-5 and 6. A HEV-7 isolated from the camel is classified in *Orthohepevirus A* group (31). Hence, HEV-1-4 and 7 are able to infect humans (32). HEV-like virus called avian HEV are classified under the *Orthohepevirus B* group. Despite sequence identity, avian HEV shares common epitopes of the capsid protein with mammalian HEV (33). *Orthohepevirus* species C and D include HEV strains isolated from ferrets, rats and bats, respectively. Despite the heterogeneity of HEV strains, there is only one serotype and the classification of HEV strains is in transition due to the different criteria used (2).

The geographic distributions of the four genotypes that infect humans are different (figure 6). In the case of genotype 1 and 2, strains were mainly isolated from Africa and Asia, including the HEV Sar55. Although HEV-1 circulation is also endemic in some region of Latin America as Chile, Uruguay, Venezuela and Cuba (35). HEV-2 included one strain from Mexico and some African variants (36,37). HEV-3 is the most frequent in South America and is found mainly in industrialized countries (38). The HEV-3 have been documented in humans and animals in Venezuela and Uruguay but only in humans in Argentina, Brazil and Colombia (39,40). HEV-3 is present in some non-endemic countries in Europe and United states, probably broth by traveling to endemic regions (35). HEV-4 was initially restricted to China, but recently strains were also isolated from France, Spain, Italy, Vietnam, Indonesia, Japan and India (38,41).



**Figure 6. The global distribution of the four human HEV genotypes.** Data adapted from WHO (185). The image was created in *BioRender.com*.

## 2.3 Epidemiology of HEV

HEV infection is one of the most common causes of acute hepatitis and has a wide distribution throughout the world. WHO estimates that 2.3 billion people have already been infected with HEV, and 70,000 deaths are attributed to this virus annually (42). Two distinct epidemiological patterns have been observed in different regions of the world. These patterns appear to be correlated with the distribution of viral genotypes, the routes of transmission, the source of infection and the prevalence of the disease. Genotypes and targeted hosts determine the clinical and epidemiological characteristics of HEV infection. In subtropical areas, hepatitis E occurs as outbreaks; and sporadic cases are transmitted by the fecal-oral route due to the ingestion of contaminated water containing HEV-1 or 2 (43). Generally, outbreaks of HEV-1 and 2 have been described in areas with limited access to water and inadequate sanitation services (49). The prevalence of antibodies against HEV (anti-HEV IgG) in Asia and

Africa is between 3% and 27% (44). In addition, apart from the fecal-oral and vertical transmission there are studies about the parenteral transmission of HEV-1 and 2 (45). Furthermore, HEV can be transmitted between humans through infected blood especially the HEV-3 in Europe (46). Transfusion-associated infections have also been reported in Japan (HEV-3 and 4) and China (HEV-1 and 4). Regions with adequate sanitation and a well-controlled water supply are considered low-endemic areas for hepatitis E infection. In this region such as the Americas, East Asia and Europe the prevalence of anti-HEV is 7% to 10% (47). Autonomous transmission cases of HEV infection appear to be associated with occasional zoonotic transmission in these regions by the HEV-3 and 4, the most common being from pigs to humans (48). The prevalence rates of anti-HEV in Europe could be explained by the consumption of pork, as the seroprevalence in France and Germany is 17% and 35%, respectively (49). Overall morbidity rates are highest among adolescents and young adults (10-40 years), but lowest in children and the elderly. High morbidity and severity have been observed among pregnant women and patients with pre-existing chronic liver disease, due to fulminant hepatitis. Laboratory diagnostic methods for HEV infection include viral particle analysis (electron immune-microscopy and immunofluorescence), virus-specific antibodies (IgG, IgM, or IgA), HEV nucleic acid and viral antigen analysis (50). Only one vaccine against hepatitis E virus infection has been developed and is licensed in China, but it is not yet available elsewhere (51).

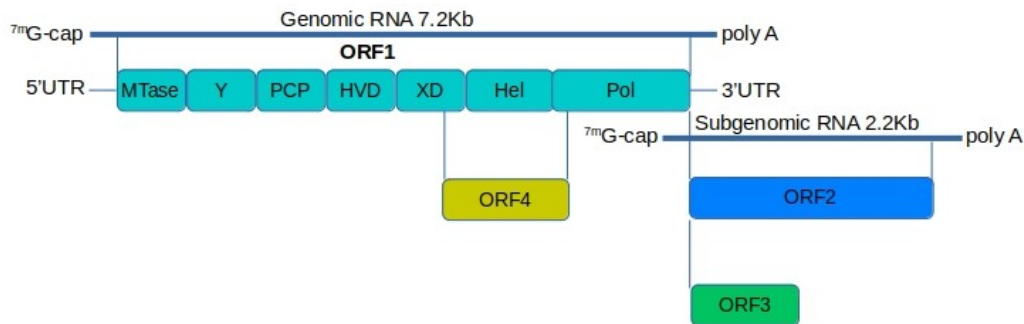
## **2.4 Clinical manifestation HEV infections.**

HEV infections are asymptomatic in almost all cases. Generally, it is a self-limited disease lasting a few weeks in most patients. The classic presentation of hepatitis E is the acute icteric hepatitis with a prevalence of 5% to 30% in infected patients. The incubation period for hepatitis E generally ranges from 14 to 60 days. Three phases were described during the infection process: latent, acute, and recovery (52). The latent phase of hepatitis E infection generally lasts 15 to 75 days. The symptoms associated with HEV infection include: anorexia, nausea, vomiting, fatigue, malaise,

arthralgia, myalgia, headache, photophobia, pharyngitis and cough. They often precede the onset of jaundice, appearing one to two weeks later. During the entire jaundice phase, the liver is enlarged, painful and with discomfort in the right upper quadrant (53). Despite the symptoms, the disease normally resolves spontaneously. The mortality rates during an outbreak are between 0.5 to 4% (54). HEV-infected patients are usually middle-aged or elderly men (> 55 years) in developed countries. In Europe, 5% to 33% of patients infected with HEV-3 or 4 developed symptoms and signs such as jaundice (55). In developing countries, young male adults (15-30 years) are mainly infected with the HEV-1 and 2. Liver disease has a poor prognosis in both developing and developed countries. Pregnant women are the most susceptible group among the infected HEV patients, specially during second and third trimesters of pregnancy. Complications such as hemorrhage or eclampsia and fulminant liver failure appear in this period reaching mortality rate around 25% (56). HEV-1 is associated with spontaneous abortion, preterm birth, late fetal death, and perinatal mortality during pregnancy (45).

## **2.5 Genome organization of HEV.**

HEV is an ~7.2 kb, (+) RNA virus. The mRNA is protected by a cap at the 5' end and at the 3' end contains a polyadenylated tail. HEV exists as non-enveloped or lipid-derived membrane envelope coated state. It contains three partially overlapping Open reading frames (ORFs) (figure 7). ORF1 encodes nonstructural proteins (nsPs): methyltransferase (MTase), Y domain (Y), papain-like cysteine protease (PCP), proline-rich hinge/hypervariable region (HVD), *Macro* domain (XD), helicase (HEL/NTPase) and RNA dependent RNA polymerase (RdRp) (57). ORF2 encodes viral capsid proteins and ORF3 encodes a small protein that shares several structural features with class I viroporins, involved in virion morphogenesis and egress (58). An ORF4, integrated in ORF1, was reported to be used under specific conditions, only in HEV-1 strains (59).

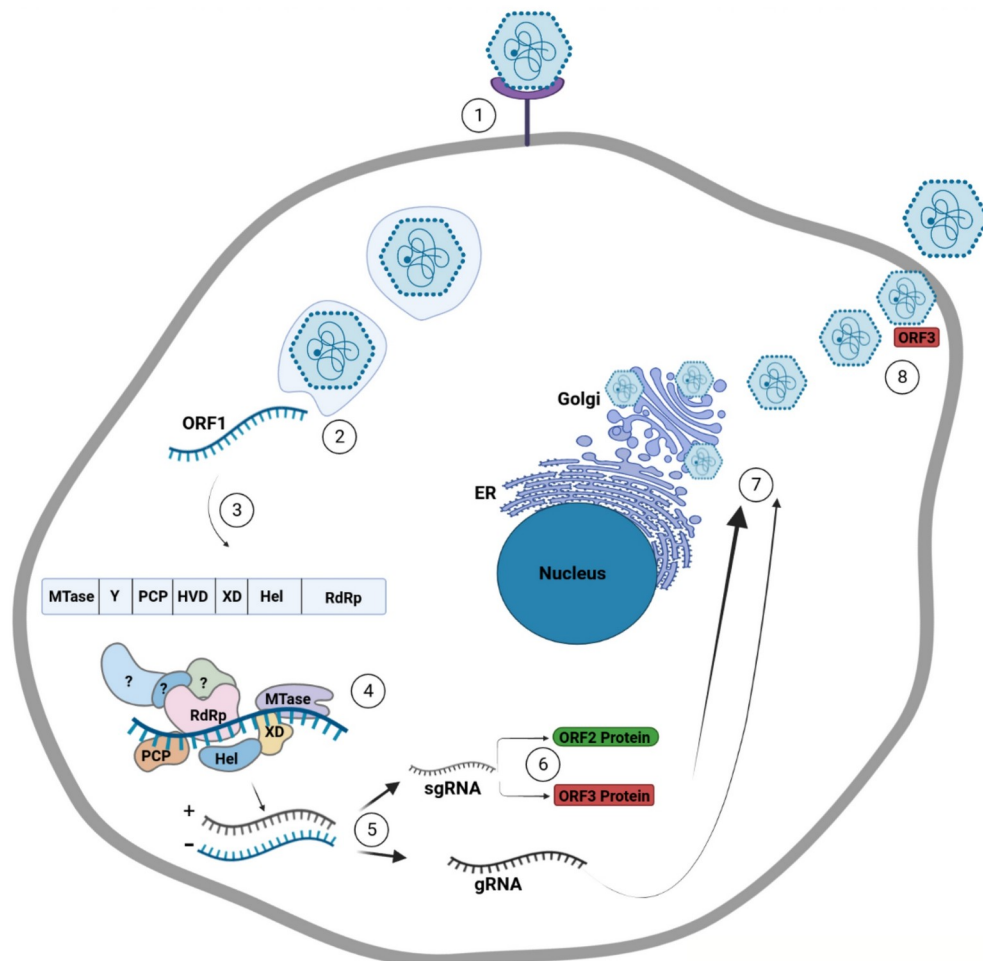


**Figure 7. HEV genomic organization.** The RNA genome consists of a 5' untranslated region (UTR), Four ORFs and a 3' UTR. The sgRNA contains ORF2 and ORF3. The RNA is capped at the 5' end and polyadenylated at the 3' end. The image was created in *BioRender.com*.

### 2.5.1 Life cycle of HEV

The life cycle of HEV is not fully elucidated, it is based on the replication cycle of alphavirus-like virus. The enveloped HEV enters the cell through clathrin-dependent and dynamin-dependent receptor-mediated endocytosis (figure 8). Upon entry, the envelope of HEV undergoes lysosome-mediated lipid degradation. The virus is uncoated in a very poorly understood manner, releasing the viral RNA. It has been reported that ORF1 polyprotein localizes to endoplasmic reticulum (ER) membranes, which probably is the site of HEV replication. The ORF1 polyprotein containing the RdRp is translated from the positive sense (+) strand. RdRp then transcribes the full-length (–) RNA. The latter serves as a template for transcribing more full-length viral RNAs and sgRNA. The interaction of ORF3 product with the endosomal sorting complexes, probably promotes the budding of progeny virions into multivesicular bodies. Then, the multivesicular bodies fuse with the plasma membrane to release virions from the cell. The lipid envelope of HEV is thought to be derived from the

trans-Golgi network. HEV enveloped viral particles are associated to ORF3 protein (60).



**Figure 8. Hepatitis E virus life cycle.** (1) Viral attachment to heparan sulfate proteoglycans (2) endocytosis and release (+) RNA genome into the cytosol. (3) Translation of the ORF1 poly-protein. (4) HEV Replication complex is constituted by ORF1 proteins/domains and probably host factors associated to ER membranes to produce various RNAs intermediates (5) Synthesis of sgRNAs and genomic RNA. (6) Synthesize of ORF2 and ORF3 protein products. (7) Packaging and virion assembly. (8) Release of the virus into the bloodstream and the bile. The image was created in *BioRender.com*.



## 2.5.2 HEV-ORF1

The genome starts at the 5' end with an untranslated region (UTR) of 25 nucleotides (nt). ORF1 is the largest ORF within the HEV genome with 5082 nucleotides. ORF1 encodes a 1693 amino acids polyprotein, corresponding to nsPs (50). Seven putative domains are identified encoded within the ORF1 (mentioned above). Whether ORF1 product function as one polyprotein or needs to be processed onto smaller units by a viral or a cellular protease is still a matter of debate. Interactome studies, using two yeast hybrids system, support polyprotein proteolysis and showed its importance for replication (62). Using vaccinia virus expression system, it was demonstrated that ORF1 polyprotein is cleaved, generating 107 kDa and 78 kDa fragments in HepG2 cells (63). While the identical construct of ORF1 expressed in bacteria or cell free system showed no processing (64). Other study showed that after transfection of infectious HEV RNA into HepG2 cells, small fragments of 35, 36, and 38 kDa were obtained using MTase, RdRp and HEL/NTPase antibodies (10).

### 2.5.2.1 Methyltransferase

MTase domain is a 11kDa domain, harboring MTase and guanylyltransferase (GTase) activities and responsible for genome capping. This reaction takes place in three stages: first, GTP is methylated in the presence of the methyl donor, S-adenosylmethionine (AdoMet) to generate 7-methyl-GTP ( $m^7$ GTP) and S-adenosylhomocysteine (AdoHcy). Second, the  $m^7$ GTP is complexed to the protein by a phosphoamide bond ( $m^7$ GMP-MTase). Third,  $m^7$ GMP is transferred to RNA by guanylyltransferase activity, forming a cap-0-like structure ( $m^7$ GpppNpRNA) (66). HEV MTase follows the mechanism employed by the alphavirus-like supergroup (67). The transfer of  $m^7$ GMP to HEV RNA remains to be demonstrated. Note that the shorter constructs, MTase-Y or MTase-PCP, are not active emphasizing the importance of protein-protein interactions within the polyprotein of the ORF1 products (68). RNA capping has been shown critical for the HEV infectivity in non-human primates and for replication in cultured hepatoma cells (69). MTase protein

tightly interacts with membranes and is essential for establishing the replication complex (7). Mutations within the MTase domain of HEV-1 were identified in patients with acute liver failure (ALF) and increased viral load (70).

### **2.5.2.2 Y domain**

Y domain of HEV is poorly characterized. Structural comparison of conserved region of Y domain in alphavirus-like family, plant virus, animal and humans suggested that this domain is an extension of C-terminus MTase domain (71). Bioinformatics studies of alpha-like viruses sequences have linked this domain to the “Iceberg region” of Alphaviruses, present downstream of the MTase domain. This “iceberg region” contributes to the assembly of viral replication components by the presence of putative membrane binding association sites, as shown for other viruses (70). Mutagenesis studies on SINV and Semliki Forest viruses (SFV), respectively, revealed that this region contains membrane anchor peptides (residues 245-264 of SFV) as well as palmitoylation sites (Cys 418-420 for SFV), essential for methyl transferase activity, replication and viral infectivity (72). In (+) RNA viruses palmitoylation of a conserved cysteine of MTase-Y region is necessary for the cytoplasmic membrane-binding and viral replication (73). Substitution of the conserved cysteine abolishes RNA replication of HEV in HepG2C3A human hepatoma cells (74). The extensive mutation of the Y domain residues revealed the importance of the putative helix (residues 410-416 in HEV-3) and the presence of potential palmitoylation sites (Cys 336,337) but also of a secondary structural loop of viral RNA, formed by nucleotides 778 and 992, in viral infectivity and replication (73).

### **2.5.2.3 Putative protease region : PCP**

HEV PCP and its functions are still under debate. The PCP domain is very small and spans over less than 200 residues. The sequence similarity among the papain-like viral

endopeptidases is very low. However, the presence of a Cys / His catalytic dyad is very conserved, suggesting a cysteine protease activity. PCP region was identified as a putative protease from the sequence alignment with Rubella virus. In the absence of conclusive biochemical data using purified proteins, the question of the proteolysis of the ORF1 by the PCP remains unclear. In addition, a deubiquitinase activity carried out by PCP domain was identified, acting as a putative IFN antagonist (75). Deubiquitination activity seems to allow the deISGylation of the gene stimulated by interferons 15 (ISG15), reinforcing viral escape (76). This deubiquitination activity would be provided in conjunction with the MTase domain. The latter would contribute by the putative zinc finger motif.

#### **2.5.2.4 Hyper variable region (HVD)**

HVD is located in ORF1 directly downstream of the PCP domain. HVD has sequence heterogeneities and contains hypervariable motifs diverging up to 71% across HEV genotypes. It also differs in length among HEV strains, earning it the name “hypervariable region” (71). Nevertheless, HVDs from all HEV genotypes share the characteristic of being rich in proline. HVD region encodes approximately 68 to 86 amino acids. The most important structural role attributed to HVD is that it forms a flexible hinge between the ORF1 regions (78). The sequence divergence among HVD zoonotic HEV-3, HEV-4 and human HEV-1 suggests the possible impact/role of HVD in host tropism (77). The exchange of HVD sequences between genotypes affects the efficiency of replication. Studies in hepatoma cells have reported that large deletions of HVD severely affect HEV replication (79). Insertion, deletion or recombination of HVD has been associated with pathogenesis in a chronic hepatitis E patient. Analyses of HEV isolates from chronically infected patients show that these insertions are from the viral or human origins (80). Although small HVD sequence deletion seem not to have impact on virion infectivity, they could affect the replication by interacting with host or other viral proteins. The particular role of the HVD in HEV replication have to be elucidated.

#### 2.5.2.5 **Macro domain**

The *Macro* domain or X domain (XD) is separated from the PCP domain by HVD. It has a fold typical of the XD of histones, such as MacroH2A, involved, among other things, in the regulation of gene expression, DNA repair, microtubule formation, X chromosome inactivation and apoptosis (81). HEV XD and viral XD, in general, consist of a mixture of  $\alpha$ -helices and  $\beta$ -strands that delimit a pocket containing the active site of ADP-ribose, capable of binding ADP ribose derivatives with different affinities (82). Functional studies have shown that the HEV XD possesses a de-ribosylation activity on mono and poly-ADP-ribosylated protein substrates conferred by the C-terminal end of the HEL domain (83). Mutations in the putative catalytic site of the HEV XD severely reduce or completely abrogate viral replication (84). Moreover, in HEK293T cells, the phosphorylation of the transcription factor IRF3 for IFN induction is inhibited by overexpression of XD (85). In other studies, the transfection of the HEV XD drastically decreased IFN production and inhibits ferritin secretion. The latter also is involved in the antiviral response (86). Still, the molecular mechanisms involved in viral escape remain unknown to this day.

#### 2.5.2.6 **The helicase /NTPase domain**

The HEV HEL/NTPase (HEL) sequences are mapped between XD and RdRp domains. HEL domain exhibits the four motifs typical of helicases superfamily I (78). HEL contains highly conserved “Walker A” motif GVPGSGK located at position 975-982, responsible for binding to purine nucleoside triphosphate (NTP). The “Walker B” motif, characteristic of NTPase-RNA or DNA dependent activity, is located at position 1092-1032. This activity was confirmed by mutagenesis and binding studies using different nucleotide substrates (87). Mutations in Walker HEV motifs abolish RNA replication in hepatoma cells (80). The HEL domain harbours a 5'-3' helicase activity and a RNA triphosphatase activity. The latter removes a phosphate moiety from the 5' end of nascent RNA. This process would be required for cap transfer

( $m^7$ GMP) to RNA, described above. The HEL structure of the tomato mosaic virus, member of Alpha-like super family, has been resolved and reveals an organization into two domains: (i) a C-terminal domain with two RecA motifs consisting of 4-6  $\beta$  sheets sandwiched by  $\alpha$  helix, involved in helicase activity (ii) an N-terminal domain containing a flexible loop connected to an  $\alpha$  helix and 6  $\beta$  sheets, probably involved in interactions with the host factors. Mutations of residues within the HEL domain strongly impacts replication as well as viral infectivity suggesting that it may play a role in anti-viral immunity (88). Furthermore, in hepatitis E cases of fulminant hepatic failure mutation in the HEV HEL domain are frequently detected (89).

#### **2.5.2.7 RNA dependent RNA polymerase (RdRp)**

HEV RdRp is located downstream of the helicase gene. It spans nucleotides 1249–1671. Immunofluorescence experiments indicate that RdRp is localized at the membrane of the endoplasmic reticulum (ER) upstream of the secretion pathway, suggesting the involvement of ER membrane in HEV replication (63). Deletion of the hydrophobic residues (1585-1601) in RdRp abrogates the localization at ER (59). The catalytic triad of HEV RdRp is constituted by a highly conserved motif (GDD), essential for RNA replication (90). The HEV RdRp binds to the 3' UTR of the sense strand to produce the antisense intermediate strand and to the 5' UTR of the antisense strand to produce the full-length sense RNA genome. RNA binding assays have shown the importance of the secondary structural loops of the cis-regulatory element in the 3' UTR. Activity tests confirmed the importance of 3' UTR as well as the poly-A tail as a matrix for polymerization (91). The production / purification of the HEV RdRp under denaturing conditions underlined the importance of the regulatory elements at the 5' end of sgRNA with a hierarchy of affinities towards 3' UTR, 5' UTR then the sgRNA, respectively (92). HEV RdRp has been shown to form complex with others ORF1 proteins as MTase, PCP and HEL. Also, it interacts with immune host factors, as IFN induced protein IFIT1 (93). Some mutations in the RdRp sequence have been related to pathogenesis and adverse clinical outcomes. A point

mutation in RdRp (G1634R) of HEV-3 detected in immunocompromised patients, with chronic infections was correlated with higher viral load and initial resistance to ribavirin (94). In addition, this mutation promotes the replication of HEV in cell system (95). In pregnant women infected with the HEV-1, the higher presence of HEV ARN in samples were related to mutations in the RdRp (C1483W and N1530T) (85). Until now, Ribavirin is the only available drug for treatment of chronic HEV infection, targeting RdRp activity (95).

### **2.5.3 HEV-ORF2**

ORF2 has a length of 1983 nt, starting at 37 nt after ORF1 and ending at 65 nt before the poly-A tail. ORF2 encode viral capsid protein. Capsid protein translocates into the ER by a signal peptide sequence located in the N-terminus. The C-terminus part of the capsid protein is implicated in encapsidation and viral stabilization (95). The cellular localization of the capsid protein was investigated. It appears to be widely distributed in subcellular organelles such as ER, Golgi and even the nucleus (96). This protein presents three conserved potential N-glycosylation Asn-X-Ser/Thr sites, the relevance of the N- glycosylation of the ORF2 protein is still not elucidated. The capsid protein is the major component and the most immunogenic HEV protein (133). HEV capsid protein represent a potential epitope for neutralizing antibodies (97). Truncated ORF2 proteins generated by baculovirus or bacterial expression systems have been tested in clinical trials for vaccine, but without success (98).

### **2.5.4 HEV-ORF3**

ORF3 is the smallest HEV ORFs overlapping ORF2 by approximately 300 nt. ORF3 encodes a small transmembrane multifunctional protein of 13kDa that resembles class I viroporins (VP13). It forms a multimeric complex localized at ER Membranes. The function of VP13 has not yet been elucidated. However, studies have suggested that VP13 is required for viral particle release during HEV infection, although the function of ORF3 in cells culture has not appeared essential for HEV RNA replication or viral

assembly (99). Immuno-mapping VP13 demonstrated that the last 32 aa of VP13 represent an immunodominant region (100).

### 2.5.5 HEV-ORF4

Recently, a new ORF4 (nt 2835-3308) was identified in HEV-1. The translation of ORF4 is regulated by an IRES-like sequences. Transiently expressing ORF4 produces a 20 kDa protein (83). ORF4 product form a protein complex with multiple viral proteins, including XD, HEL and RdRp. The complex stimulates RdRp, in Huh7 cells, in response to ER stress (13). In another work, it was demonstrated that the presence of ORF4 is not required for active *in vivo* and *in vitro* replication of HEV in rats (102). ORF4 product with loss of the ubiquitination site was observed in patients with acute hepatitis or fulminant hepatic failure. This observation suggests a negative contribution of proteasome-resistant ORF4 in patient outcome (103).

## 2.6 Treatment and Vaccine

Various stages of the HEV cell cycle can be potential targets for antiviral drug development. Acute infection generally does not require treatment. Chronic hepatitis E can lead to spontaneous resolution in some cases, but it can also lead to rapid progression into cirrhosis and death. Therefore, it is important to consider treatment in these patients. Ribavirin is currently the only treatment option for many patients and the most useful antiviral available for HEV infection (104). Studies have shown that ribavirin alone has led to a sustained antiviral response in at least 2/3 of cases with chronic infection. However, treatment failure associated to ribavirin use was also reported (63).

Regarding vaccination, there are several vaccine candidates based on ORF2 capsid protein, which generate anti-HEV antibodies and provide cross-genotype protection against infection (105). However, most HEV vaccine programs based on the

recombinant HEV capsid were discontinued at the pre-clinical stages due to mixed data or incomplete protection rate. HEV 239 vaccine was generated from a truncated version of pORF2 aa 368–607 expressed in an *Escherichia coli* (*E. coli*) system, which self-assemble into virus-like particles (VLPs) *in vitro*. The vaccine is commercialized, since 2012, in China only, under the name Hecolin (Xiamen Innovax Biotech, Xiamen, China (106). The lack of robust HEV cell culture system constrains the development of attenuated or inactive vaccine. Moreover, other region of HEV have been explored as vaccine candidates. Recombinant ORF3 provides only a partial protection and the generated antibodies were not able to neutralize the virus in cell models (107). The antibodies generated against ORF1 proteins could not recognize HEV particles during infection (108). Thus, the ORF2 protein is the most attractive candidate for the moment.

## 2.7 Objectives

Hepatitis E, caused by the HEV, is among the emerging neglected diseases. At present, very little data is available on the molecular mechanisms that allow the replication and the propagation of HEV. In addition, the role of the different proteins encoded by the genome has been putatively assigned, by sequence similarity with the group of viruses related to alphaviruses. The main obstacle to the characterization of functional domains is the lack of a suitable animal model, a robust virus propagation system, but also the difficulty in producing functionally active proteins. HEV infection leads to acute hepatitis, which is eventually brought under control by the immune system in the majority of people. However, in so-called vulnerable individuals, including pregnant women, adolescents, immunosuppressed or transplant patients, HEV infection can become chronic with fatal consequences, underscoring the importance of the immune response in this infection. During infection, the nsPs, encoded by ORF1, are the first to be expressed and appear to play a crucial role in regulating the host's immune response, leading the replication cascade can progress.



The aim of this study is to identify and characterize viral molecular motifs and enzymatic activities harboured by HEV ORF1 products and implicated in immune evasion.

The development of this project will rely on:

- (i) The evaluation of the effect of ORF1 products on the modulation of the innate immune response in a mammalian system.
- (ii) The biochemical characterization of target ORF1 domains, explaining the effect observed in cells.
- (iii) Mutagenesis analysis of the selected targets for a structure/ function assignment.
- (iv) The assessment of the effect of these mutations on type I IFN susceptibility and ISGs activation in cells.

Finally, the study conducted during this thesis will contribute to the understanding of HEV evasion mechanisms and will pinpoint viral targets important for the development of specific antiviral drugs.

# **Results**

## 2.8 Article 1

### ***Hepatitis E Virus ORF1 harbors a 2'-5'-Phosphodiesterase (PDE) activity, interfering with Type I Interferon Induction***

**Oney Ortega Granda<sup>a</sup>**, Sandra Madariaga Zarza<sup>b</sup>, Benjamin Morin<sup>a\*</sup>, Jean-Louis Mege<sup>b</sup>, Laurence Briant<sup>c</sup>, Françoise Debart<sup>d</sup>, Bruno Canard<sup>a</sup>, Karine Alvarez<sup>a</sup> and Nadia Rabah<sup>a,e,#</sup>

<sup>a</sup>Aix Marseille Université, CNRS, AFMB UMR 7257, Marseille, France

<sup>b</sup>Institut hospitalo-universitaire Méditerranée infection, Marseille, France

<sup>c</sup>IRIM, CNRS UMR9004-University Montpellier, Montpellier

<sup>d</sup>IBMM, CNRS, University of Montpellier, ENSCM, Montpellier, France

<sup>e</sup>Université de Toulon, 83130 La Garde, France

\*Present address: Agenesis Inc., Lexington, MA 02421, USA.

# Address correspondence to Nadia Rabah, [nadia.rabah@univ-tln.fr](mailto:nadia.rabah@univ-tln.fr)

***In preparation***

### ***Abstract***

Hepatitis E (HEV) is an important etiological agent causing acute and chronic hepatitis. HEV belongs to the alphavirus-like super family and possesses a single stranded RNA genome with three open reading frames (ORF1 to 3). ORF1 encodes the viral nonstructural polyprotein, essential for RNA replication and crucial for the down regulation of the innate immune response, through inhibition of interferon (IFN) production. The molecular mechanisms underlying this evasion are still unknown. IFN inducible OAS/RNase L system is one of the first antiviral responses. It relies on the synthesis of short 2'-5' oligonucleotides (2-5A) by oligoadenylate synthetases (OAS). 2-5A activate the latent cellular RNase L, which then degrades viral and cellular RNAs, restricting viral infection. Viruses evolved different strategies to escape the OAS/RNase L system, one of which is the degradation of 2-5A by harbouring a 2'-5'-phosphodiesterase (PDE) activity. PDE enzymes bear two conserved catalytic Hx(S/T)x motifs (where x is preferentially a hydrophobic amino acid). In the present study, we conducted sequence analysis of HEV ORF1 polyprotein, revealing the presence of Hx(S/T)x conserved motifs in three HEV ORF1 domains, namely *Macro*, helicase (HEL) and RNA dependent polymerase (RdRp) domains. We then determine the implication of these motifs on IFN induction. Recombinant proteins corresponding to the domains of interest were produced in *E. coli* to assess the presence of 2-5A degradation activity. Our results show that 2-5A oligotrimers can be modified only in the presence of the *Macro* domain, extended with the hyper variable region, combined to HEL domain. Finally, this study will contribute to the understanding of HEV evasion mechanisms, which could ultimately be extrapolated to other RNA viruses with the aim of developing antiviral tools.

### **Introduction**

HEV (Hepatitis E Virus) is now recognized as a global health problem in both developing and industrialized regions, including South and East Asia, East Africa, Mexico, Western Europe and USA (1). The World Health Organization (WHO) estimates that annually 20 million people are infected with HEV, encompassing 3.3 million symptomatic cases and up to 70,000 deaths per year (2). In developing countries, morbidity concerns mainly adolescents, young adults and pregnant women, in the latter group liver failure rate can reach 25%. In developed countries, clinical manifestations are associated with the development of chronic hepatitis (potentially fatal) in immuno-compromised patients (3). HEV is classified in the *Hepeviridae* family, genus *Orthohepevirus* and species *Orthohepevirus A*. HEV is actually subdivided into eight genotypes (HEV-1 to HEV-8). HEV-1 and HEV-2 are limited to humans and primates and circulates mostly in developing countries, where the virus is transmitted by fecal contaminated drinking water. While HEV-3 and HEV-4 follow a zoonotic transmission and might be contracted by blood transfusions or organ transplants, from HEV infected donors (4,5). Genotypes HEV-5, HEV-6 and HEV-7 have been isolated from wild boar and camelids (6). Their zoonotic potential is still to be defined.

The HEV virion is small, non-enveloped and contains a single-stranded RNA of a positive polarity 6.6 to 7.3 kb in length. Viral RNA is capped (7 methylguanylate) at its 5' end and harbours a poly A tail at the 3' end. In total, four open reading frames (ORF1-4) were described (7). ORF1 encodes a ~187kDa polyprotein containing seven potential non-structural proteins (nsPs) domains including: methyltransferase (Met), Y-domain (Y), papain-related cysteine peptidase (PCP), proline-rich hypervariable (HVD), a *Macro* (XD), helicase (Hel) as well as a RNA dependent RNA polymerase (RdRp) domains. The last third of HEV genome gives rise to a capped and bicistronic subgenomic RNA (sgRNA), containing the overlapping ORF2 and 3. ORF2 encodes the capsid structural proteins. ORF3 leads to the synthesis of a phosphoprotein, involved in virion release and viral escape. ORF4, present only in HEV-1, generates a product stimulating the activity of RdRp (6,8). The functions of ORF1 HEV nsPs were initially deduced, by comparison and similarity to the Alpha-like supergroup (9). Some functions have been confirmed by cloning, expression and partial purification of individual domains (10,11). However, it remains to be elucidated whether the activity is carried by the entire polyprotein or proteolytically generated individual domains (12,13).

In the course of RNA virus of positive polarity ((+) RNA) infection, as HEV, double-stranded RNA (dsRNA) replication intermediates accumulate inside the cell and act as signals to activate numerous innate immune responses, among which the 2-5-oligoadenylate (2-5A) synthetase (OAS)/RNase L pathway. The enzymatic function of OAS proteins is to synthesize, from ATP, 5'-triphosphorylated 2-5A. The latter are linked by 2'-5' phosphodiester bonds, instead of a 3'-5' link found usually in RNA and DNA. These 2-5A bind with high affinity to the inactive, monomeric form of the endoribonuclease L (RNaseL), causing its dimerization and activation (14). Once active, RNase L cleaves viral and ultimately cellular RNAs, triggering apoptosis of the infected cell (15). The small RNAs generated by RNase L action are sensed by the retinoic acid-inducible gene I (RIG-I)-like receptors, amplifying thereby IFN production, and OAS expression (16,17). A 3D characterization of RNase L, in the presence of different 2-5A analogues, allowed the characterization of RNase L/ 2-5A interaction prerequisites for binding and activation (18). These studies highlighted that RNase L can only be activated by 2-5A trimers or tetramers which are mono, di or tri-phosphorylated at their 5' end, the maximum activity being achieved with the trimer triphosphate (2-5) p<sub>3</sub>A<sub>3</sub> (19).

Numerous studies have highlighted the importance OAS/RNase L pathway inhibition in viral escape (20,21). Viruses can antagonize this antiviral machinery by several mechanisms, including dsRNA sequestration direct RNase L inhibition. Viral proteins antagonising the OAS/RNase L pathway include: (i) Vaccinia E3L protein, the latter decreases the available level of dsRNA (22); (ii) VP3 protein of group A rotaviruses (RVA), ns2 group 2 betacoronaviruses, and mouse hepatitis virus (MHV), hydrolysing 2-5A by an encoded phosphodiesterase (PDE) activity (23,24). Viral PDEs belong to the type 2H family possessing two conserved motifs H-x- (S / T) -x, where x is a preferably hydrophobic amino acid. The histidine residue of the catalytic units would directly contribute to the catalysis, while the serine or the threonine residue would stabilize the substrate in the active site pocket (25). PDEs generally adopt a  $\beta$ - $\alpha$ - $\beta$ - $\alpha$ - $\beta$ - $\beta$  fold, with the two histidine residues on antiparallel sheets. These viral PDE activities are thought to derive from ancestral RNA ligases (LigT-like) family found in bacteria and archaea, and involved in the metabolism and maturation of RNAs (26).

Herein, we conducted bioinformatics analyses of the HEV ORF1 polypeptide. This led to identification of three H-x- (S / T) -x motifs, which might play a role in the modulation of the innate immune system through the inhibition of the OAS / RNase L pathway. We then investigated the impact of domains harboring these motifs on IFN induction and 2-5A degradation.

### ***Material and methods***

#### **Cloning**

The full-length DNA of XD, HVD-XD and HEL gene encoding the corresponding proteins were amplified from Sar 55 (HEV-1) replicon by PCR with the corresponding forward and reverse primers. Amplified DNA fragments were cloned by enzymatic restriction into pet28b or pEf6 vectors (Invitrogen) for bacterial and mammalian expression, respectively.

#### **Cell Culture and transfection**

Human hepatoma cells (Huh7.5) were cultured in Dulbecco's Modified Eagle's Medium supplemented with 100 µg/ml penicillin-streptomycin (GIBCO) and 10% Fetal Bovine Serum (FBS), at 37°C, in humidified atmosphere containing 5% CO<sub>2</sub>. The day prior transfection 4x10<sup>5</sup> cells were seeded into 10 cm<sup>2</sup> plates. 70-80% confluent plates were transfected with 4µg of recombinant pEf6 plasmids encoding the protein sequences of XD, HVD-XD, HEL, and RdRp domains. The empty vector serves as mock control. After 48h of incubation, medium was changed and the cells were further transfected with 5µg/ml Poly(I:C). Cells were harvested 6 hours post-transfection. Transfection reagent Lipofectamine™2000 (Invitrogen) was used for both transfections according to the manufacturer's protocol.

#### **Immunofluorescence**

Cells were fixed in 3% paraformaldehyde for 20 minutes at room temperature (RT) and washed two times with PBS. The fixed cells were permeabilized with PBS-Tween20 (0.1%) for 10 minutes, washed two times then incubated with the blocking buffer (PBS-Tween20 (0.1%) - BSA (3%)) for 30min at RT. Monoclonal anti-FLAG M2 antibody (Sigma - F3165) diluted 1:500 was added to the cells for 1h at 37°C. Following two washes, cells were incubated 30 min with 4',6-diamidino-2 phenylindole (1:500) (DAPI, Invitrogen), to label the nucleus, and Phalloidin (1:50) (Alexa Fluor 488, Invitrogen) for actin. The expressed

recombinant protein were incubate with monoclonal anti-FLAG M2 antibody (1:500), followed by the fluorescent conjugated goat anti-mouse IgG (1:500) (Alexa Fluor 633, Invitrogen). Fluorescence signals were analyzed using LSM 8000 Airyscan confocal microscope (Zeiss LSM 800) under oil immersion objective (63 × 1.5).

### **RT-PCR (Reverse transcriptase PCR)**

RNA extraction was carried out with RNeasy MiniKit (Qiagen). Purified RNA was quantified by nanodrop spectrophotometer (Nanodrop technologies). 80ng of total RNA was reverse transcribed using the MMLV-RT kit according to the manufacturer's protocol (Invitrogen). Expression of the  $\beta$ -IFN gene was evaluated by real-time qPCR using the Applied Biosystems 7900HT. Specific primers was used for  $\beta$ -IFN gene (Forward primer : 5'-GTTCCCTTAGGATTTCCACTCTGACTATGGTCC-3' ; Reverse primer : 5'-GAACTTTGACATCCCTGAGGAGATTAAGCAGC-3') and  $\beta$ -actin gene (Forward primer 5'-GGAAATCGTGCGTGACATTA-3'; Reverse primer 5'-AGGAGGAAGGCTGGAAGAG-3'), as interested and housekeeping genes respectively. Threshold cycles ( $C_t$ ) were obtained using CFX Touch Real-Time PCR detection system (Bio-Rad). Results were normalized using  $\beta$ -actin gene expression and are expressed as Fold Change (FC),  $FC = 2^{-\Delta\Delta C_t}$ , where  $\Delta\Delta C_t = (C_{t\beta\text{-IFN}} - C_{t\text{actin}})_{\text{stimulated}} - (C_{t\beta\text{-IFN}} - C_{t\text{actin}})_{\text{unstimulated}}$ . Statistical analysis was performed using GraphPad Prism software. In vitro data were analyzed using Ordinary one-way ANOVA with Dunnett's multiple comparisons test. \* $P < 0.05$  \*\* $P < 0.01$ ; \*\*\*\* $P < 0.0001$ .

### **Recombinant protein expression**

Recombinant plasmids were transformed into the competent *E. coli* (C41 (DE3) plys) strain and cultured in Lysogeny Broth (LB) containing 100  $\mu\text{g/mL}$  of Kanamycin at 220 rpm at 37°C during 4 hours. Logarithmic phase (OD600 of 0.6) cultures were induced with 0.5mM of IPTG. Bacterial pellets were re-suspended in lysis buffer (20 mM Tris-HCl, pH7, 5 mM betamercaptoethanol (bMeOH), 150 mM NaCl, 5% glycerol, 20 $\mu\text{g/ml}$  DNase, 0.25 mg/ml lysozyme and 1 mM PMSF). Cells were then subjected to ultrasonication (40%, 1 min, 15sec on /15sec off). Immobilized metal affinity chromatography (IMAC) purification was performed using 1ml Ni-NTA column (Qiagen) in a batch mode. The filtrated supernatant was loaded onto the column pre-equilibrated with binding buffer (20 mM Tris-HCl, pH7, 5 mM bMeOH, 150 mM NaCl, 5% glycerol and 10mM imidazole). After a first column wash with



10 volumes of wash buffer (binding buffer containing 60 mM imidazole), a second washing was performed with high salt concentration 1M NaCl in the same buffer. Recombinant proteins were eluted with 250 mM imidazole.

### **ADP-ribose binding**

ADP-ribose binding to XD and XD-HVD of HEV was assessed using dot blot assay. Briefly, serial dilutions (from 250 to 3.9 pmol) of *Macro* domain containing proteins were spotted onto a nitrocellulose membrane (Schleicher & Schuell). Following a saturation step in the dot blot buffer (10 mM Tris-HCL pH 7.5, 150 mM NaCl<sub>2</sub> and 0.5% Tween-20), supplemented with 4% skimmed milk, during 1h at RT, the membrane was incubated with auto-ADP-ribosylated hPARP-1 for 1h at RT. The primary antibody anti-PAR binding reagent (Sigma) diluted 1:1500 in dot blot buffer with 1 % skimmed milk was added for 2h at RT. Then, the secondary antibody, anti-rabbit IgG (Dako), diluted up to 1:2000 was incubated with the membrane for 1 h, at RT. After each step, three washes with dot blot buffer were performed. The signals were revealed using ECL reagent (Cat. # 170–5061, Bio-Rad). The results were visualized using Amersham™ ImageQuant™ 800 Immager system. BSA was used as a negative control of ADP-ribose binding.

### **NTPase assay of HEL**

ATPase activity of purified HEV HEL was performed using a thin-layer chromatography (TLC) assay. The reaction was carried out in 15µL final volume of 500 mM MOPS-Na buffer (pH 7.0, 50 mM MgCl<sub>2</sub>, 50 mM DTT, 1% Tween-20) and 0.4µ Ci of [ $\gamma$ -<sup>32</sup>P] ATP or 0.4µCi of [ $\alpha$ -<sup>32</sup>P] ATP. The HEV HEL and DENV NS3, used as control, were used at 3µM at 22°C for 1hour. After an inactivation step at 95°C for 5 minutes, hydrolysis products were subjected to TLC polyethyleneimine-cellulose plate (Merck) with 0,45M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> as a mobile phase. Signals visualized using an Amersham™ Typhoon™ Biomolecular Imager (Cytiva).

### **2-5 PDE activity assay**

2-5 PDE activity was assessed by mixing 3 µM of different combination of recombinant XD, XD-HVD, HEL, RdRp domains, with 20uM of purified 2-5A trimer in the reaction buffer (20 mM Hepes, pH 7.5, 10 mM MgCl<sub>2</sub>, 1 mM DTT). After 16 hours at 22°C, as described in Zhao et al. (24), the reaction was stopped by heating at 95 °C for 5 min. Reaction products were applied to Microcon 10K centrifugal filter column and centrifuged at 3500g for 30 min at

4°C. The filtrate was diluted in triethylammonium bicarbonate (TEAB) 0,05M, pH 8.0, then analyzed on an HPLC column (Novapak C18, 4 µm, 3.9 × 150 mm), using a linear gradient of 10 to 60% (TEAB) 0.05M/acetonitrile, at a flow rate of 1 mL/min in 60 min. Elution profiles of reaction products were recorded at 260 nm. Recombinant non-structural protein 3 of dengue virus 2 (DENV NS3) (27) and commercial CIAP (Calf Intestinal Alkaline Phosphatase, 20 U/µL, Invitrogen) were used for the generation of mono, di and tri phosphate 2-5A of different lengths, serving as controls.

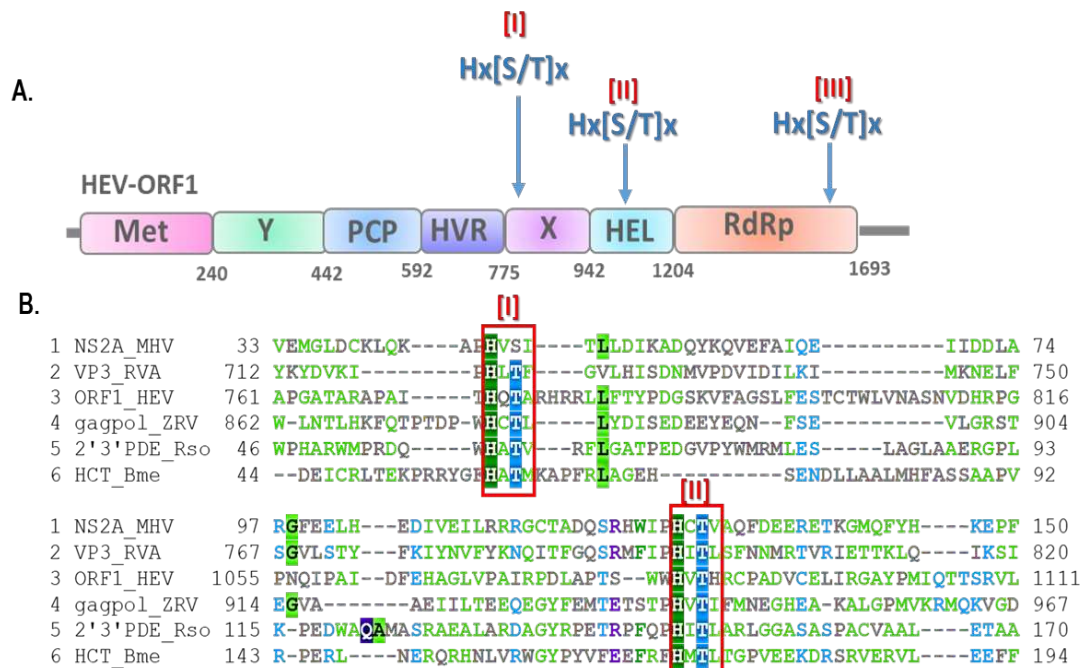
### ***Results and Discussion***

#### **HEV ORF1 proteins contains H-x- (S / T) -x motifs, a hallmark of PDE enzymes:**

The innate immune response plays an important role in HEV infection, with type I IFN activation ( $\alpha$  and  $\beta$  IFN) as a key component. Viral replication is directly correlated with a decrease in IFN production, showing that viruses evolved effective escape strategies to resist to a sustained INF response in infected cells (28). Numerous studies have highlighted the role of HEV ORF1, ORF2 and ORF3 in modulating immune response (29–31). Hence, transfected full ORF2 protein in HEK293T cells inhibited Retinoic acid inducible gene – I (RIG-I) dependent interferon response and antagonized Toll-like receptor (TLR) pathways (32). Immunoprecipitation studies have shown that the phosphoprotein, encoded by HEV ORF3, interacts with STAT1 altering IFN $\alpha$ -dependent STAT1 phosphorylation and hence activating ISGs (33). At present, the replication cycle of HEV has not been established and the proposed model is based on the replication cycle of alpha-like viruses. This model implies that ORF1 is the first ORF to be translated. This is followed by genome replication and the translation of ORF2 and 3 from sgRNA (34). The timeline, described herein, imposes a crucial role for ORF1 products in viral replication and immune response evasion. This assumption was further supported, in the case of HEV, by data showing that the overexpression of ORF1 domains decreases IFN response (30). Although, the mechanism underneath this action is still unknown, it suggests the presence of molecular determinants in HEV ORF1 products that inhibit host's innate immunity. Since, one of the IFN response escape pathway is the inhibition of the OAS/RNase L system by 2-5A degradation (35), we decided to search for the presence of potential PDE in ORF1 polyprotein.

ScanProsite Proteins and Proteomes Software tool was used to scan HEV ORF1 polyprotein and compare it to the Prosite database. The analysis revealed the presence of three Hx- (S / T) -x motifs, characteristic of PDE activity. The first motif is present at the junction of the

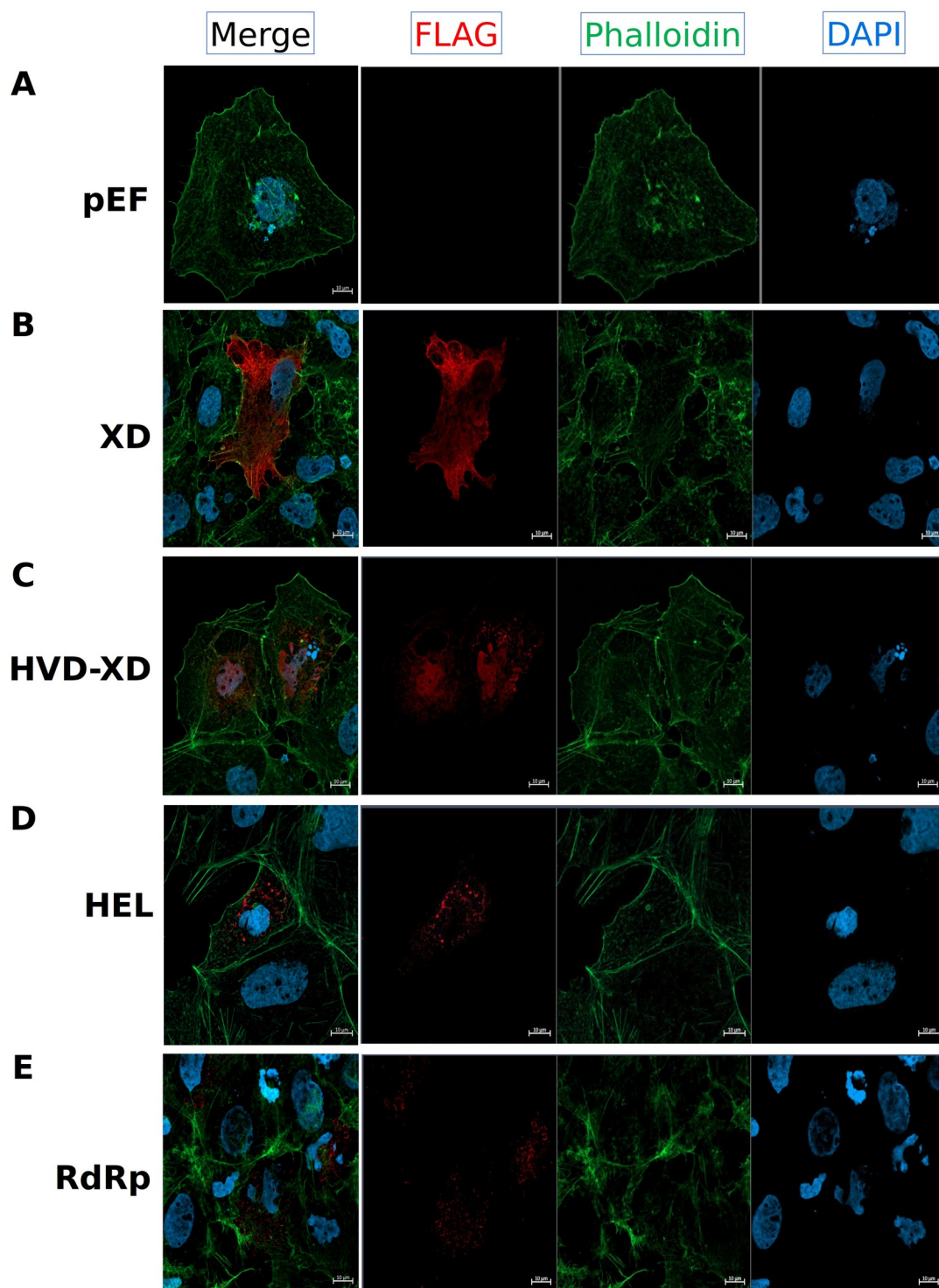
HVD and the XD domain (I), the second is located in the HEL domain (II) and the third at the C-terminal end of the RdRp (III) (Figure 1A). These motifs are conserved in different HEV-1 isolates but also present in the HEV-2 to HEV-4 genotypes, suggesting that they may have a functional role in viral replication. The comparison of the sequences delimiting the potential H-x- (S / T) -x motifs, show that these align with the catalytic ligand-binding cleft of 2H phosphoesterases (2'PDE) of different groups related to ARN ligases (LigT-like), present in bacteria and archaea (26) (Figure 1B). Type 2'PDE hydrolyze the (2-5A) phosphoester bond using two conserved H-x- (S / T) -x motifs. At the evolutionary level, the H-x-(S/ T)-x motifs would derive from an ancestral duplication of a single motif which was accompanied by a sequence divergence between the motifs allowing species-specific adaptation. In the At5g40190 family of *Arabidopsis* and the AGR\_C\_4233 protein of *Agrobacterium*, for example, histidine is present in only one of two motifs, required for PDE activity (26). Other organisms, such as retroviruses, exhibit very long stretches of several hundred amino acids between the two motifs (36). The implication of these motifs in functions related to RNA metabolism suggests that they may present a binding site to RNA and / or nucleotides (ATP, 2-5A, NADP + and ADP derivatives ribose). These interactions could also play an important role in the escape from innate immunity.



**Figure 1. A. Position of H-x- (S / T) -x patterns in ORF1. B. Multiple alignment of the 2H motifs of proteins.** Proteins from NS2A (MHV, murine hepatitis virus), VP3 (HRV, human Rotavirus), gagRdRp (ZRV, Endogenous zebrafish retrovirus), ORF1 of HEV (HVD-XD domain for the motif [I] and HEL domain for motif [II], 2'3' RNA cyclic PDE (Rso, Ralstonia solanacearum), HPC (Bme, Brucella melitensis). Alignment was performed using the algorithm muscle from the Seaview program.

**HEV ORF1 proteins have different subcellular localisation in Huh7.5 cells**

Subcellular localization of HEV replication complex have been poorly defined to date. Previous studies in heterologous expression system revealed that ORF1 protein products are membrane associated and partially colocalize within ER and ER-Golgi intermediate compartments (37). Also, the ORF1 protein products colocalize with exosomes and multivesicular bodies, suggesting exosomal membranes are required for the establishment of the viral replication complex. Further, these finding and the partial colocalization of ORF1 products with ORF2 and ORF3 proteins, suggest a close connection between HEV replication and assembly sites (11). In order to address the behavior of ORF1 domains, containing PDE motifs, in Huh 7.5 cells, the expression of the domains of interest was assessed by immunofluorescence assay (Figure 2). Hence, the expression of XD, HVD-XD, HEL and RdRp domains with a FLAG tag reveal differences in localisation profiles. The FLAG tag is expressed only in fusion to a target protein (Figure 2, line A) and can be detected as a red signal (Figure 2, lines B-E). The XD, when expressed alone, is diffuse in the cytoplasm with an accentuated perinuclear staining (Figure 2, lines B). However, when XD is preceded by the HVD (HVD-XD), the cytoplasmic staining is not diffuse any more but restricted to dot-like structures in the perinuclear region. Moreover, our results show that HVD-XD is able to translocate into the nucleus (Figure 2, lines C). This observation suggests that XD has targets in the nucleus important for viral replication or immune response suppression. HEL and RdRp individual domains were detected in dot-like structures surrounding the nucleus, in a manner similar to XD alone (Figure 2, lines D and E). Previously, alphavirus nsPs involved in immune response evasion were reported to be localized in the nucleus. Hence, Semliki Forest virus (SFV) nsP2 localize to the nucleus and blocks cellular RNA export to the cytoplasm (38,39). Also, nsP3 of SFV can be detected in cytopathic vacuoles at the perinuclear area, in infected cells (40). NsP2 and nsP3 from alphavirus have been involved in the interference with cellular transcription and translation. Accordingly, nsP2 and nsP3 can antagonize cellular antiviral responses triggered by alphavirus infection. In addition, nsP3 blocks the formation of host cellular stress granules involved in innate antiviral (41,42). Although additional controls using organelle markers are needed to ascertain the cellular localization of PDE motif containing domains, the results presented here open new expectations on the role of the studied domains on the HEV replication.



**Figure 2. Cellular localisation of ORF1 proteins harbouring PDE motifs observed by Immunofluorescence.** Pictures were generated using confocal microscopy (scale bar, 10  $\mu$ m). ORF1 proteins (XD, HVD-XD, HEL, RdRp) in fusion to the FLAG tag are represented

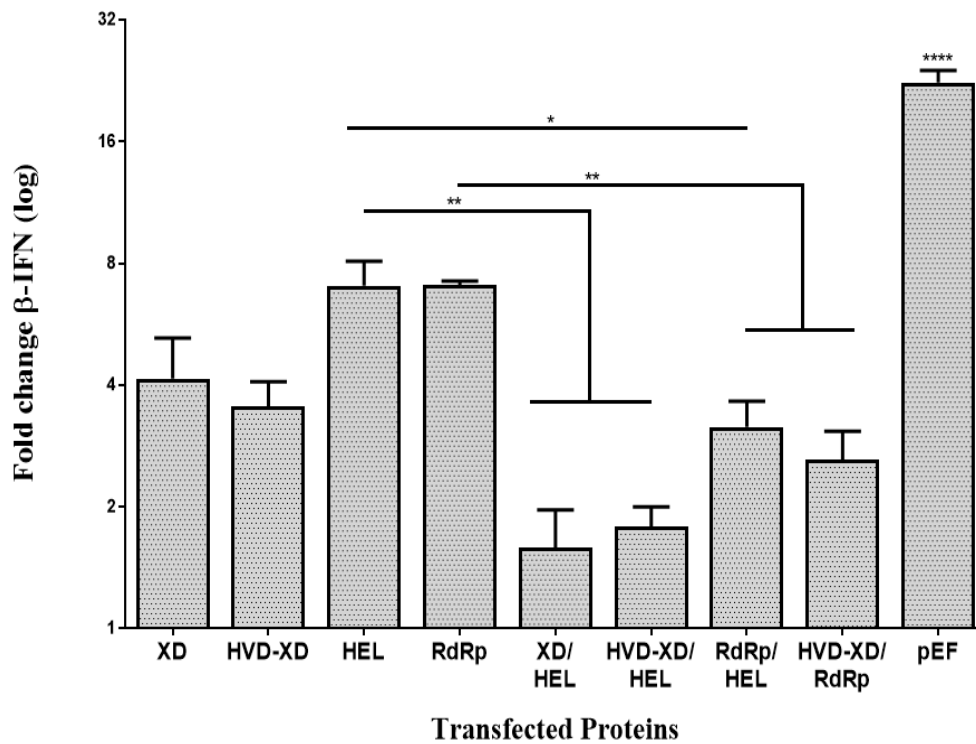


in red. The nucleus is blue following DAPI staining. Actin filaments are stained in green following phalloidin treatment. pEF represent the Mock control. Cells were doubly labeled with monoclonal anti-FLAG M2 antibody, DAPI and phalloidin; followed by Alexa Fluor 633 or Alexa Fluor 488-conjugated secondary antibodies for FLAG tag and phalloidin, respectively. Co-localizations appear in the Merge panels.

### **HEV ORF1 proteins inhibit IFN expression in Huh 7.5 cells**

The investigation of the effect of PDE containing motifs ORF1 protein domains on the modulation of the innate immune response were carried out in Huh7.5 mammalian cells, endogenously expressing immuno-detectable amounts of RNase L (43). The cells were transfected with the constructs used for immunofluorescence study either individually or by combining different domains, to ensure the presence of two PDE motifs required for catalysis. Activation of OAS, RNase L and type I IFNs was triggered by stimulating the transfected cells with Poly I:C. The effect of our ORF1-PDE domains of interest on the inhibition of  $\beta$ -IFN induction was then evaluated by quantification of the  $\beta$ -IFN mRNA using real time PCR (RT-qPCR), after mRNA extraction from transfected cells. Results are represented as relative fold change of  $\beta$ -IFN genes expression (Figure 3). One can notice that the expression of individual PDE containing HEV domains reduces 65 to 83 %  $\beta$ -IFN induction, when comparing to the mock control. This effect is prominent in *Macro* domain containing fragments (XD and XD-HVD). Our results confirm previous published data stating that XD-HVD reduces  $\beta$ -IFN induction. However, in contrast to our data, they did not notice a reduction in  $\beta$ -IFN induction in the case of HEL and RdRp (30). This might suggest that  $\beta$ -IFN repression mechanisms could be tissue specific. Interestingly, co-transfection of different protein domains had a synergic effect on  $\beta$ -IFN repression (Figure 3). All the tested combinations show increased  $\beta$ -IFN antagonism compared to the corresponding individual counterpart. Hence,  $\beta$ -IFN induction is decreased by more than 90% in XD/HEL and XD-HVD/HEL combinations. This result emphasizes the role of *Macro* domain in viral escape. The latter was shown to be achieved by different mechanisms, including: (i) the interference with PARPS antiviral activity, targeting viral replication, transcription and translation (44); (ii) de-MARYlation of Ras-GTPase-Activating Protein SH3-Domain-Binding Protein (G3BP), involved in stress granules formations (45); (iii)  $\beta$ -IFN system antagonism (46). In HEV, this antagonistic effect might be reached through the interference with the RIG-I-like receptors (RLRs) pathway, by the inhibition of RIG-I and MDA5 induction; or the abolition of

interferon regulatory factor 3 (IRF-3) phosphorylation, a prerequisite for  $\beta$ -IFN expression (30). The synergic effect on  $\beta$ -IFN repression observed in our study suggest that other molecular events, contributing to viral escape, are still to be discovered. Regarding the fact that our ORF1 proteins of interest harbour PDE catalytic motifs, we hypothesized that a potential 2-5A degradation mechanism could promote  $\beta$ -IFN repression by the inhibition of the OAS/RNase L system.

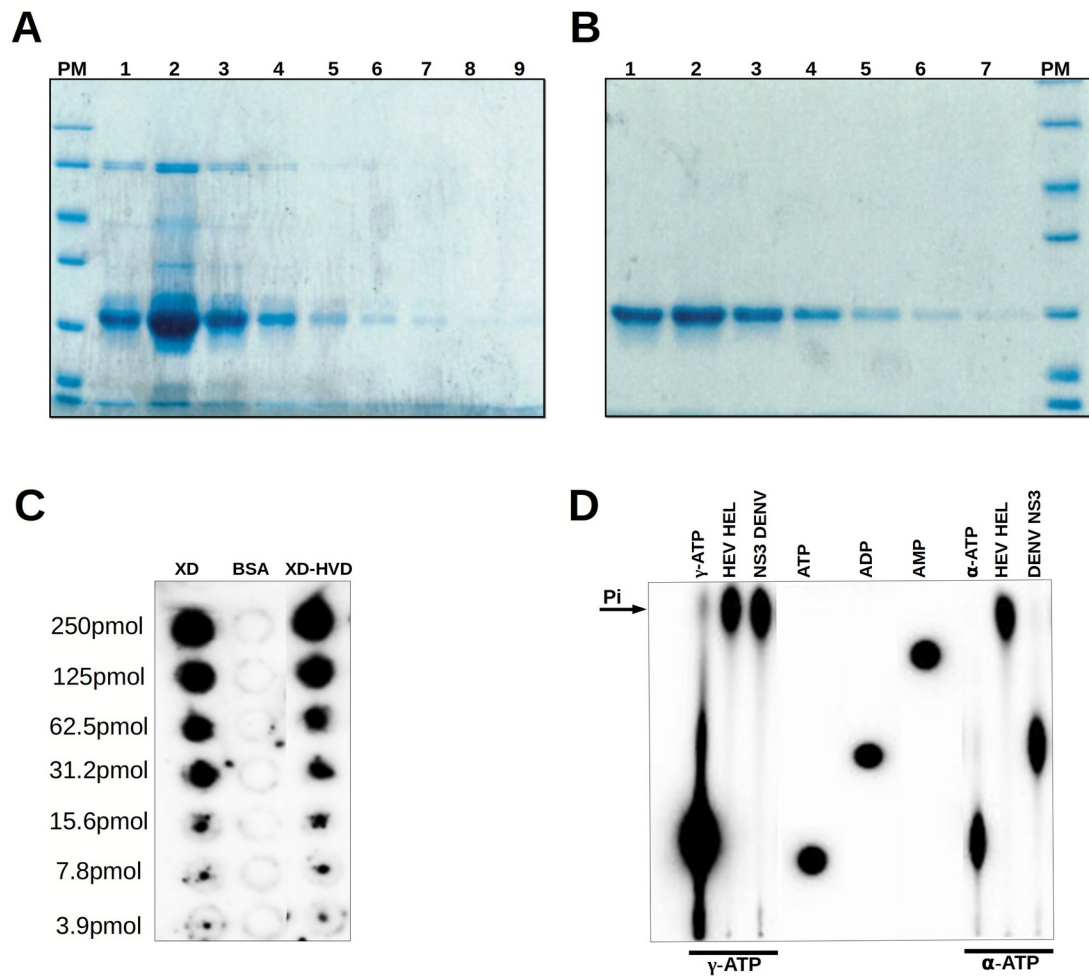


**Figure 3: Effect of ORF1 HEV domains containing PDE motifs on  $\beta$ -IFN induction following poly I:C treatment.** Expression level are represented as the mean fold change ( $FC = 2^{-\Delta\Delta Ct}$ , where  $\Delta\Delta Ct = (C_{t_{target}} - C_{t_{actin}})_{stimulated} - (C_{t_{target}} - C_{t_{actin}})_{unstimulated}$ )  $\pm$  standard deviation of mean (SEM). Ordinary one-way ANOVA with Dunnett's multiple comparisons test was used to compare different conditions. Significant differences are denoted by asterisks \* $P < 0.05$  \*\* $P < 0.01$ ; \*\*\*\* $P < 0.0001$ . The graph represents averages of three independent experiments. The error bars represent the difference of SEM.



### **Recombinant HVD-XD and HEL degrade 2-5A**

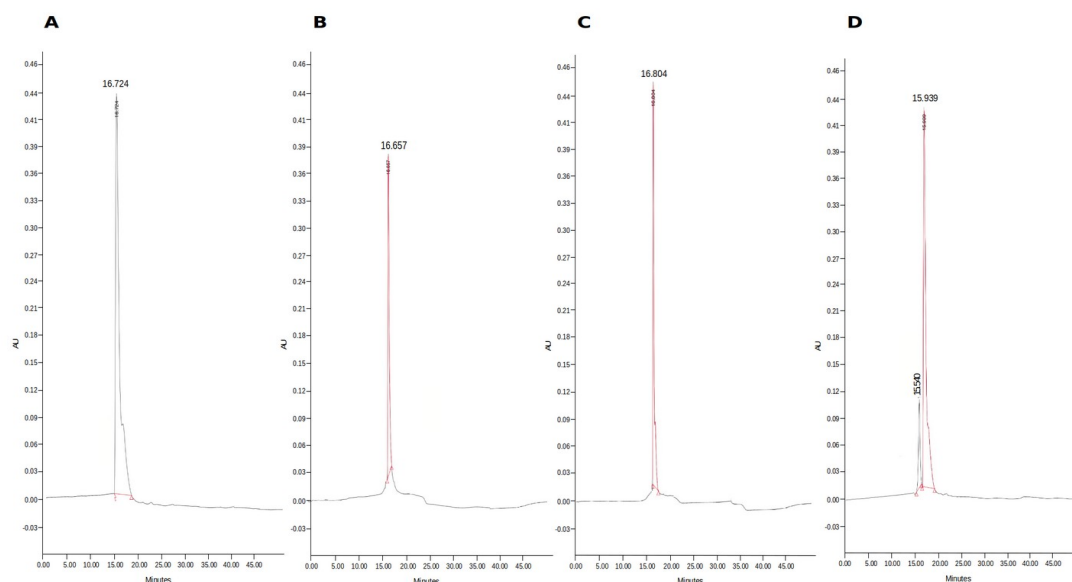
In order to test our hypothesis, *i.e.* assess the presence of a PDE activity, degrading 2-5A, we carried the expression and the purification of ORF1-PDE proteins in bacterial system. Unfortunately, we were not able to get a soluble RdRp or a N-terminal truncated RdRp domains, inferring the need of continuing the optimisation of expression and purification conditions. Recombinant XD, HVD-XD and HEL proteins were purified in high yield (2mg/L) using two purification steps (Figure 4A and B). Purified recombinant XD, HVD-XD and HEL proteins were subjected to biochemical characterization, in order to verify the integrity of their functional domains (Figure 4 C and D). Hence, XD and HVD-XD are able to bind ADP-ribose from the auto-poly ADP-ribosylated hPARP-1 in a concentration depending manner (Figure 4 C). The binding was achieved without the addition of HEL domain, as advocated by Li and Coutard 2016 (47). This discrepancy might be explained by the difference in constructs design, XD was shortened at the N-terminus by 17 amino acids in the previously published study (47). In the case of HEL domain, ATPase assay was performed (Figure 4 C). Results show clearly that recombinant HEV HEL possesses a NTPase activity. It is able to dephosphorylate completely ATP, releasing the  $\gamma$ -phosphate from  $\gamma$ -<sup>32</sup>P-ATP as well as the  $\alpha$ -phosphate from  $\alpha$ -<sup>32</sup>P-ATP. Dengue NS3, used as a reaction control, is just able to remove the  $\gamma$ -phosphate, converting ATP to ADP (Figure 4 D).



**Figure 4: Purification and activity assessment of HEV ORF1 HEL and various XD length fragments.** (A and B) SDS-PAGE-Coomassie blue staining of purified fractions, after gel filtration column, for XD and HEL, (C) XD and HVD-XD ADP-ribose binding assay (D) NTPase activity assessment for HEL domain.

The putative 2-5 PDE activity of the expressed domains was evaluated by 2-5 PDE assay *in vitro*, as described previously (24). For that end, recombinant proteins XD, HVD-XD and HEL were incubated in the presence of purified 2-5A trimer ((2-5) p<sub>3</sub>A<sub>3</sub>). After the incubation period the reaction products were subjected to HPLC analysis to seek for any 2-5A modification. The results are summarized in Figure 5. Elution peaks were compared to controls established using 2-5A treated enzymatically in order to obtain tri/di and dephosphorylated 2-5A trimer as well as dimers. These desphosphorylated controls were

obtained by mono or total dephosphorylation using DENV NS3 and various phosphatase enzymes. A well characterized PDE12-2'5' enzyme was added as a positive control (data not shown) (48).



**Figure 5. HPLC hydrolysis profiles of 2-5A by recombinant HEV ORF1 proteins. (A)** 2-5A ((2-5) p<sub>3</sub>A<sub>3</sub>) untreated negative control). **(B)** HVD-XD. **(C)** Hel. **(D)** HVD-XD and HEL.

The elution profile of the untreated 2-5A trimer is represented in Figure 5A. Similar retention times (16.657 and 16.804 min) were obtained when 2-5A trimers were incubated in the presence of XD-HVD (B) and (C) HEL recombinant proteins, respectively, indicating an absence of 2-5A degradation when the domains are used individually. Comparable retention time was observed when recombinant XD was used (data not shown). Interestingly, when XD-HVD and HEL were combined, two elution peaks were recorded (Figure 5D). A prominent peak eluted at 15.939 min and a small peak at 15.540 min. When comparing to controls, these two peaks might correspond to:

- (i) Tri and di phosphorylated 2-5A dimer, with retention times of 15.821 and 15.120, respectively. Both tri and di phosphorylated 2-5A dimers are non-RNaseL activating (49,50). These results support the presence of a 2'5'-PDE activity, which when linked to the NTPase activity of HEL leads to the production of a dephosphorylated dimer.

- (ii) The prominent peak could also be assimilated to a di phosphate form of 2-5A trimer (15.809).

Hence, at this point, two hypotheses can be drowned. First, HEL activity removes one phosphate from the tri-phosphorylated 2-5A trimer only in the presence of XD (prominent peak), probably needed for the correct recognition and positioning of the 2-5A oligoadenylate. The product resulting from the reaction is further not completely degraded into the 2-5A dimer by the putative PDE activity (small peak). Second, the putative 2'PDE activity converts the 2-5A trimer into the 2-5A dimer (prominent peak). The latter is then subjected to a partial HEL dephosphorylation. The first step being a prerequisite, allowing to give access to the HEL active site, which might be not optimized for 2-5A oligonucleotides. This assumption is supported by previous studies showing that HEV HEL has a NTPase activity, which is able to remove the  $\alpha$ -phosphate from the ATP but only the  $\gamma$ - $\beta$ -phosphate from the RNA (51,52).

PDE enzymes degrade 2-5A trimer into 2-5A dimer releasing 5'-AMP residues one by one until the 5'-terminal ATP moiety is released (23,24). However, in our case no AMP, nor ATP are detected. This might be due to HEL NTPase activity, which can dephosphorylate released AMP and/or ATP and ultimately adenosine. The latter is known to possess an immunomodulatory role in favour of viral pathogenesis (53,54). However, further analysis are required to improve the identification and characterization of the 2-5A trimer degradation products when treated with HVD-XD and HEL domains.

In conclusion, the results obtained here show that HEV ORF1 contain three PDE motifs (H-x-(S / T) -x). When transfected into Huh7.5 cells, the domains harbouring these motifs antagonize type I IFN response. IFN repression might be due to the inhibition of the OAS / RNase L pathway, through the partial degradation of 2-5A. Further studies are needed to confirm our results and complete the characterization of HEV ORF1 containing PDE domains. Understanding viral activities involved in the inactivation of cellular defense might allow the identification of targets that can be exploited for the generation of antiviral drugs against HEV.

## References

1. Salines M, Andraud M, Rose N. From the epidemiology of hepatitis E virus (HEV) within the swine reservoir to public health risk mitigation strategies: a comprehensive review. *Vet Res.* 2017;48:31.
2. Lin S, Zhang Y-J. Advances in Hepatitis E Virus Biology and Pathogenesis. *Viruses.* 2021 Feb 9;13(2):267.
3. Aslan AT, Balaban HY. Hepatitis E virus: Epidemiology, diagnosis, clinical manifestations, and treatment. *World J Gastroenterol.* 2020 Oct 7;26(37):5543–60.
4. Janahi EM, Parkar SFD, Mustafa S, Eisa ZM. Implications of Hepatitis E Virus in Blood Transfusions, Hemodialysis, and Solid Organ Transplants. *Medicina (Kaunas).* 2020 Apr 25;56(5):206.
5. Kamar N, Izopet J, Pavio N, Aggarwal R, Labrique A, Wedemeyer H, et al. Hepatitis E virus infection. *Nat Rev Dis Primers.* 2017 Nov 16;3:17086.
6. Sridhar S, Teng JLL, Chiu T-H, Lau SKP, Woo PCY. Hepatitis E Virus Genotypes and Evolution: Emergence of Camel Hepatitis E Variants. *Int J Mol Sci.* 2017 Apr 20;18(4):E869.
7. Oechslin N, Moradpour D, Gouttenoire J. On the Host Side of the Hepatitis E Virus Life Cycle. *Cells.* 2020 May 22;9(5):E1294.
8. LeDesma R, Nimgaonkar I, Ploss A. Hepatitis E Virus Replication. *Viruses.* 2019 Aug 6;11(8):E719.
9. Koonin EV, Gorbalenya AE, Purdy MA, Rozanov MN, Reyes GR, Bradley DW. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of an additional group of positive-strand RNA plant and animal viruses. *Proc Natl Acad Sci U S A.* 1992 Sep 1;89(17):8259–63.
10. Ansari IH, Nanda SK, Durgapal H, Agrawal S, Mohanty SK, Gupta D, et al. Cloning, sequencing, and expression of the hepatitis E virus (HEV) nonstructural open reading frame 1 (ORF1). *J Med Virol.* 2000 Mar;60(3):275–83.
11. Szkolnicka D, Pollán A, Da Silva N, Oechslin N, Gouttenoire J, Moradpour D. Recombinant Hepatitis E Viruses Harboring Tags in the ORF1 Protein. *Journal of Virology.* 93(19):e00459-19.
12. Sehgal D, Thomas S, Chakraborty M, Jameel S. Expression and processing of the Hepatitis E virus ORF1 nonstructural polyprotein. *Virol J.* 2006 Dec;3(1):38.

13. Habersetzer J, Debbah M, Fogeron M-L, Böckmann A, Bressanelli S, Fieulaine S. In vitro translation of virally-encoded replication polyproteins to recapitulate polyprotein maturation processes. *Protein Expression and Purification*. 2020 Nov;175:105694.
14. Dong B, Silverman RH. 2-5A-dependent RNase molecules dimerize during activation by 2-5A. *J Biol Chem*. 1995 Feb 24;270(8):4133–7.
15. Keel AY, Jha BK, Kieft JS. Structural architecture of an RNA that competitively inhibits RNase L. *RNA*. 2012 Jan;18(1):88–99.
16. Rice AD, Turner PC, Embury JE, Moldawer LL, Baker HV, Moyer RW. Roles of Vaccinia Virus Genes E3L and K3L and Host Genes PKR and RNase L during Intratracheal Infection of C57BL/6 Mice. *Journal of Virology*. 2011 Jan 1;85(1):550–67.
17. Khapersky DA, McCormick C. Timing Is Everything: Coordinated Control of Host Shutoff by Influenza A Virus NS1 and PA-X Proteins. *J Virol*. 2015 Apr 15;89(13):6528–31.
18. Tanaka N, Nakanishi M, Kusakabe Y, Goto Y, Kitade Y, Nakamura KT. Structural basis for recognition of 2',5'-linked oligoadenylates by human ribonuclease L. *EMBO J*. 2004 Oct 13;23(20):3929–38.
19. Tanaka N, Nakanishi M, Kusakabe Y, Goto Y, Kitade Y, Nakamura KT. Structural basis for recognition of 2',5'-linked oligoadenylates by human ribonuclease L. *EMBO J*. 2004 Oct 13;23(20):3929–38.
20. Li Y, Banerjee S, Goldstein SA, Dong B, Gaughan C, Rath S, et al. Ribonuclease L mediates the cell-lethal phenotype of double-stranded RNA editing enzyme ADAR1 deficiency in a human cell line. Nilsen TW, editor. *eLife*. 2017 Mar 31;6:e25687.
21. Thornbrough JM, Jha BK, Yount B, Goldstein SA, Li Y, Elliott R, et al. Middle East Respiratory Syndrome Coronavirus NS4b Protein Inhibits Host RNase L Activation. *mBio*. 2016 Mar 29;7(2):e00258.
22. Xiang Y, Condit RC, Vijaysri S, Jacobs B, Williams BRG, Silverman RH. Blockade of Interferon Induction and Action by the E3L Double-Stranded RNA Binding Proteins of Vaccinia Virus. *J Virol*. 2002 May;76(10):5251–9.
23. Zhang R, Jha BK, Ogden KM, Dong B, Zhao L, Elliott R, et al. Homologous 2',5'-phosphodiesterases from disparate RNA viruses antagonize antiviral innate immunity. *Proc Natl Acad Sci U S A*. 2013 Aug 6;110(32):13114–9.

24. Zhao L, Jha BK, Wu A, Elliott R, Ziebuhr J, Gorbalenya AE, et al. Antagonism of the interferon-induced OAS-RNase L pathway by murine coronavirus ns2 protein is required for virus replication and liver pathology. *Cell Host Microbe*. 2012 Jun 14;11(6):607–16.
25. Hofmann A, Tarasov S, Grella M, Ruvinov S, Nasr F, Filipowicz W, et al. Biophysical characterization of cyclic nucleotide phosphodiesterases. *Biochem Biophys Res Commun*. 2002 Mar 8;291(4):875–83.
26. Mazumder R, Iyer LM, Vasudevan S, Aravind L. Detection of novel members, structure–function analysis and evolutionary classification of the 2H phosphoesterase superfamily. *Nucleic Acids Res*. 2002 Dec 1;30(23):5229–43.
27. Benarroch D, Selisko B, Locatelli GA, Maga G, Romette J-L, Canard B. The RNA helicase, nucleotide 5'-triphosphatase, and RNA 5'-triphosphatase activities of Dengue virus protein NS3 are Mg<sup>2+</sup>-dependent and require a functional Walker B motif in the helicase catalytic core. *Virology*. 2004 Oct 25;328(2):208–18.
28. Knegendorf L, Drave SA, Thi VLD, Debing Y, Brown RJP, Vondran FWR, et al. Hepatitis E virus replication and interferon responses in human placental cells. *Hepatology Communications*. 2018;2(2):173–87.
29. Qi Y, Zhang F, Zhang L, Harrison TJ, Huang W, Zhao C, et al. Hepatitis E Virus Produced from Cell Culture Has a Lipid Envelope. *PLOS ONE*. 2015 Jul 10;10(7):e0132503.
30. Nan Y, Yu Y, Ma Z, Khattar SK, Fredericksen B, Zhang Y-J. Hepatitis E Virus Inhibits Type I Interferon Induction by ORF1 Products. *Journal of Virology*. 2014 Oct 15;88(20):11924–32.
31. Kang S, Myoung J. Host Innate Immunity against Hepatitis E Virus and Viral Evasion Mechanisms. *J Microbiol Biotechnol*. 2017 Oct 28;27(10):1727–35.
32. Hingane S, Joshi N, Surjit M, Ranjith-Kumar CT. Hepatitis E Virus ORF2 Inhibits RIG-I Mediated Interferon Response. *Frontiers in Microbiology*. 2020;11:656.
33. Dong C, Zafrullah M, Mixson-Hayden T, Dai X, Liang J, Meng J, et al. Suppression of interferon- $\alpha$  signaling by hepatitis E virus. *Hepatology*. 2012 May;55(5):1324–32.
34. Nimgaonkar I, Ding Q, Schwartz RE, Ploss A. Hepatitis E virus: advances and challenges. *Nat Rev Gastroenterol Hepatol*. 2018 Feb;15(2):96–110.
35. Drappier M, Michiels T. Inhibition of the OAS/RNase L pathway by viruses. *Current Opinion in Virology*. 2015 Dec;15:19–26.

36. Myllykoski M, Kursula P. Structural aspects of nucleotide ligand binding by a bacterial 2H phosphoesterase. Permyakov EA, editor. PLoS ONE. 2017 Jan 31;12(1):e0170355.
37. Perttilä J, Spuul P, Ahola T. Early secretory pathway localization and lack of processing for hepatitis E virus replication protein pORF1. Journal of General Virology. 2013 Apr 1;94(4):807–16.
38. Rikkinen M, Peränen J, Kääriäinen L. Nuclear targeting of Semliki Forest virus nsP2. Arch Virol Suppl. 1994;9:369–77.
39. Breakwell L, Dosenovic P, Karlsson Hedestam GB, D’Amato M, Liljeström P, Fazakerley J, et al. Semliki Forest virus nonstructural protein 2 is involved in suppression of the type I interferon response. J Virol. 2007 Aug;81(16):8677–84.
40. Kujala P, Ikäheimonen A, Ehsani N, Vihinen H, Auvinen P, Kääriäinen L. Biogenesis of the Semliki Forest virus RNA replication complex. J Virol. 2001 Apr;75(8):3873–84.
41. Sanz MA, García-Moreno M, Carrasco L. Inhibition of host protein synthesis by Sindbis virus: correlation with viral RNA replication and release of nuclear proteins to the cytoplasm. Cellular Microbiology. 2015;17(4):520–41.
42. Fros JJ, Domeradzka NE, Baggen J, Geertsema C, Flipse J, Vlak JM, et al. Chikungunya virus nsP3 blocks stress granule assembly by recruitment of G3BP into cytoplasmic foci. J Virol. 2012 Oct;86(19):10873–9.
43. Kwon Y-C, Kang J-I, Hwang SB, Ahn B-Y. The ribonuclease L-dependent antiviral roles of human 2’,5’-oligoadenylate synthetase family members against hepatitis C virus. FEBS Lett. 2013 Jan 16;587(2):156–64.
44. Malgras M, Garcia M, Jousselin C, Bodet C, Lévêque N. The Antiviral Activities of Poly-ADP-Ribose Polymerases. Viruses. 2021 Mar 30;13(4):582.
45. Jayabalan AK, Adivarahan S, Koppula A, Abraham R, Batish M, Zenklusen D, et al. Stress granule formation, disassembly, and composition are regulated by alphavirus ADP-ribosylhydrolase activity. Proc Natl Acad Sci USA. 2021 Feb 9;118(6):e2021719118.
46. Alhammad YMO, Kashipathy MM, Roy A, Gagné J-P, McDonald P, Gao P, et al. The SARS-CoV-2 Conserved Macrodomain Is a Mono-ADP-Ribosylhydrolase. J Virol. 2021 Jan 13;95(3):e01969-20.
47. Li C, Debing Y, Jankevicius G, Neyts J, Ahel I, Coutard B, et al. Viral Macro Domains Reverse Protein ADP-Ribosylation. Perlman S, editor. J Virol. 2016 Oct 1;90(19):8478.



48. Poulsen JB, Kjær KH, Justesen J, Martensen PM. Enzyme assays for synthesis and degradation of 2-5As and other 2'-5' oligonucleotides. *BMC Biochem.* 2015 Dec;16(1):15.
49. Huang H, Zeqiraj E, Dong B, Jha BK, Duffy NM, Orlicky S, et al. Dimeric Structure of Pseudokinase RNase L Bound to 2-5A Reveals a Basis for Interferon-Induced Antiviral Activity. *Molecular Cell.* 2014 Jan;53(2):221–34.
50. Tanaka N, Nakanishi M, Kusakabe Y, Goto Y, Kitade Y, Nakamura KT. Structural basis for recognition of 2',5'-linked oligoadenylates by human ribonuclease L. *EMBO J.* 2004 Oct 13;23(20):3929–38.
51. Karpe YA, Lole KS. NTPase and 5' to 3' RNA Duplex-Unwinding Activities of the Hepatitis E Virus Helicase Domain. *J Virol.* 2010 Apr;84(7):3595–602.
52. RNA 5'-Triphosphatase Activity of the Hepatitis E Virus Helicase Domain [Internet]. [cited 2021 Sep 16]. Available from: <https://journals.asm.org/doi/epub/10.1128/JVI.00492-10>
53. Passos DF, Bernardes VM, da Silva JLG, Schetinger MRC, Leal DBR. Adenosine signaling and adenosine deaminase regulation of immune responses: impact on the immunopathogenesis of HIV infection. *Purinergic Signalling.* 2018 Dec 1;14(4):309–20.
54. Lin Y-H, Tai C-C, Brož V, Tang C-K, Chen P, Wu CP, et al. Adenosine Receptor Modulates Permissiveness of Baculovirus (Budded Virus) Infection via Regulation of Energy Metabolism in *Bombyx mori*. *Frontiers in Immunology.* 2020;11:763.

## **3 Alphavirus: Capping process as a specific antiviral target for drug development**

### **3.1 Alphavirus History**

Discovery of alphaviruses dates back to 1930, when the first member, *i.e* Western equine encephalitis virus (WEEV), was isolated in the United States. However, it was not until the WEEV epidemic in 1941, incurring the loss of around 3 million equines and affected 4000 humans, that the family was brought into the focus of researchers (109). Although various outbreaks must have taken place previously. Following the discovery of WEEV, other alphavirus members were discovered, including Eastern equine encephalitis virus (EEEV) in 1933 in eastern United States (168); VEEV in 1935, after outbreaks in Venezuela, Columbia and Trinidad (169); SFV in 1942 in Uganda; SINV was isolated in Cairo (Egypt) in 1952 (170) and at the same time CHIKV in la Reunion island (171), O'nyong-nyong virus (ONNV) emerged as a disease causing entity in 1959 in Uganda (110). This unusually distinct virus, is more closely related to CHIKV than to other alphavirus members. Since 1959, this virus has caused at least three outbreaks in Africa affecting more than 2 million individuals. Another member of this genus Mayaro virus (MAYV) was first isolated in 1954 and is enzootic in South American regions (110). Alphaviruses have recently been identified to infect marine organisms like rainbow trout and Atlantic salmon pointing towards the wide range of their possible host organisms (111).

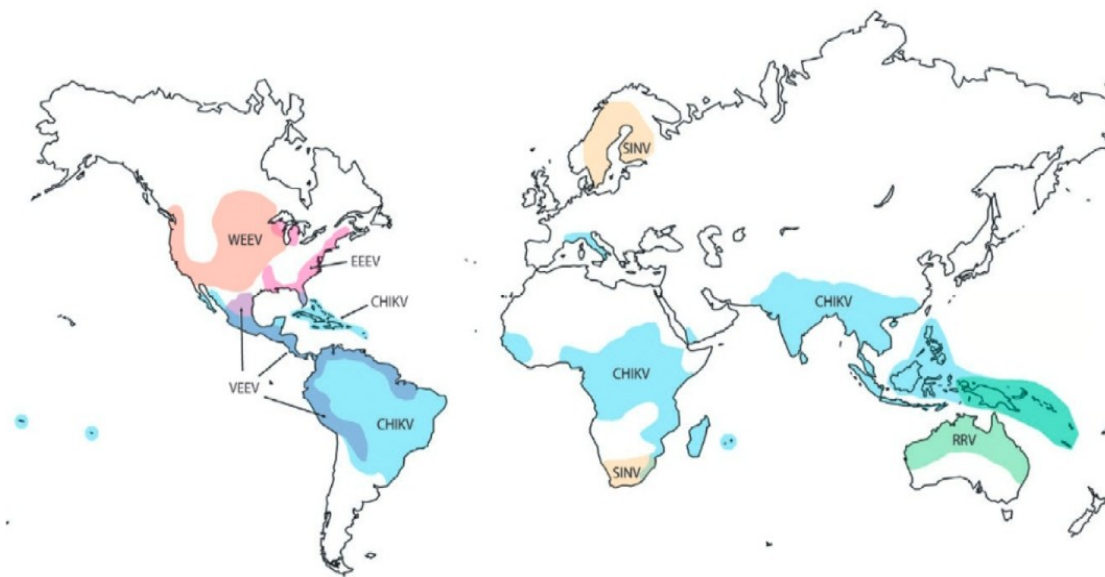
### **3.2 Classification and geographical distribution**

Alphaviruses are part of a widely distributed group of viruses, classified under the *Togaviridae* family. The family includes around 30 members, present in all the continents except Antarctica (figure 9). The genus is divided into seven antigenic

complexes, of which four are medically important types namely: VEEV, EEEV, SFV and WEEV (110). Alphaviruses are further grouped into New World and Old World viruses on the basis of their geographical distribution. New World viruses are widely distributed in Northern, Central as well as Southern America whereas Old World members cover European, African, Asian and Australian continents. There is no clear evidence about the origin of the alphaviruses but on the basis of phylogeny studies it is assumed that they might have originated in New World countries and spread out to the Old World regions (112). Further evolution led to the formation of two major groups SINV and SFV. This divergence of New and Old World might have taken place some 2000-3000 years ago and some migratory birds or animals might have played a role in this dispersion (113). WEEV and EEEV outbreaks are common in American continent especially in different United States regions. They are transmitted by *Culiseta melanura* species of mosquitoes, which help the maintenance of enzootic and epizootic cycle of these viruses (114). Generally, WEEV and EEEV target equines but sporadic cases of human infections are also reported. There are almost eight antigenic subtypes of VEEV having different virulence capabilities. Enzootic strains of VEEV consist of subtypes ID, IE, and II-VI and epizootic of subtypes IAB and IC. The enzootic strains have been isolated in Florida, Mexico, Central America, and South America. Epizootic strains are found mainly in Venezuela, Columbia and Trinidad but also between northern Argentina and Florida (114). Different species of *Aedes* and *Culex* mosquitoes help in the transmission of VEEV into its different hosts. Strains with lesser virulence circulate in a mosquito-rodent/mosquito enzootic cycle. Whereas, highly pathogenic strains, able to cause encephalitis, are able to infect humans crossing the epizootic cycle with the horses (114).

CHIKV was first observed in African regions during 1952, where it was circulating in primates and various forest mosquitoes species. The adaptation to human led to a major epidemics in African and Southern Asian regions. Now the virus has become endemic in southern and southeastern Asia, although recent cases of CHIKV infection are reported in some European countries also (115). MAYV which is antigenically

associated to CHIKV is native from Amazon River Basin. It circulates in primates and hematophagous mosquitoes but can infect humans (116). Ross River virus (RRV), first isolated in 1959 from *Aedes vigilax* mosquitoes, in the Ross River area, in Australia, is found to cause epidemic polyarthrititis mostly in Australia and Pacific Ocean islands (117).

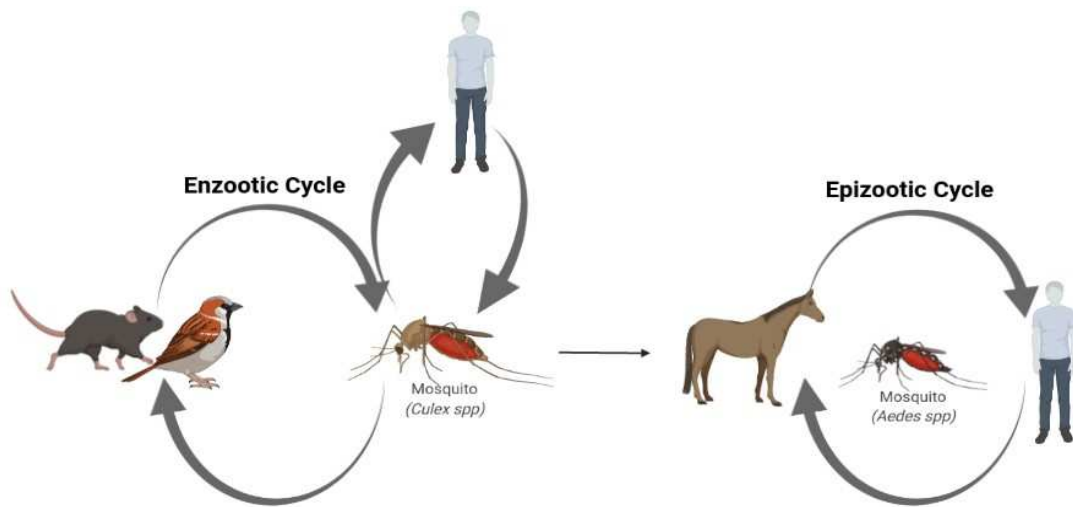


**Figure 9. Geographical distribution of the main human pathological alphaviruses.** CHIKV: chikungunya virus; RRV: Ross River virus; SINV: Sindbis virus; VEEV: Venezuelan equine encephalitis virus, EEEV: eastern equine encephalitis virus; WEEV: Western equine encephalitis virus. The image was adapted from (184).

### 3.3 Clinical manifestation and treatments

Alphaviruses circulates in an enzootic and epizootic cycle in the environment. Arthropod vectors like mosquitoes acquire the infection after biting a viremic host. After viral replication, it is transmitted to other hosts through salivary secretions.

Once inside the body, the virus leads to viremia causing illness in Humans, birds and other mammalian species (figure 12) (113). Inside the principal host, first symptoms are visible after an incubation period of 2-6 days, a period during which the virus tries to increase its copy number. There are further two phases of infection: an acute phase and a chronic phase, lasting from few days to a couple of weeks. In humans and domestic animals, infection with Alphaviruses can lead to encephalitis, polyarthritis, but also hemorrhagic fevers. The symptoms can be correlated to the geographic distribution. Hence, the Old World alphavirus causes myopathies, polyarthritis, and ultimately hemorrhagic fevers. This group includes viruses like the RRV, Barmah Forest virus (BFV), CHIKV or the SINV (118). The New World alphavirus cause encephalitis, vomiting, leukocytosis and miscarriages in humans and domestic animals. This group includes VEEV, EEEV, and WEEV. Old World viruses present rare cases of mortality, whereas New World ones are known to cause fatal encephalitis. In humans, although the death rate is low (less than 1%), up to 15% of individuals infected with “encephalic” alphaviruses can develop long-term or even lifelong neurological sequelae (118). The fever associated to the infection does not stay for long, but severe joint pain, headache and nausea can be suffered for 7-8 days. In addition, although these viruses are normally transmitted by arthropods, laboratory accidents have shown that viruses causing equine encephalitis can be highly infectious by the air (119). Majority of the alphavirus members are pathogenic to humans and can easily be used as weapon for bioterrorism. Clinical diagnosis of the infection in the case of alphaviruses is difficult sometimes, since symptoms can overlap those of rubella or enterovirus infection. making the disease complicated to identify at the initial stages. To date, although different vaccine strategies have been investigated, the only approved vaccine against VEEV is an attenuated strain of the virus, TC-83, only authorized in veterinary medicine (120). In addition, no specific antiviral treatments have been approved. As a result, these viruses represent significant epidemic and epizootic risks.



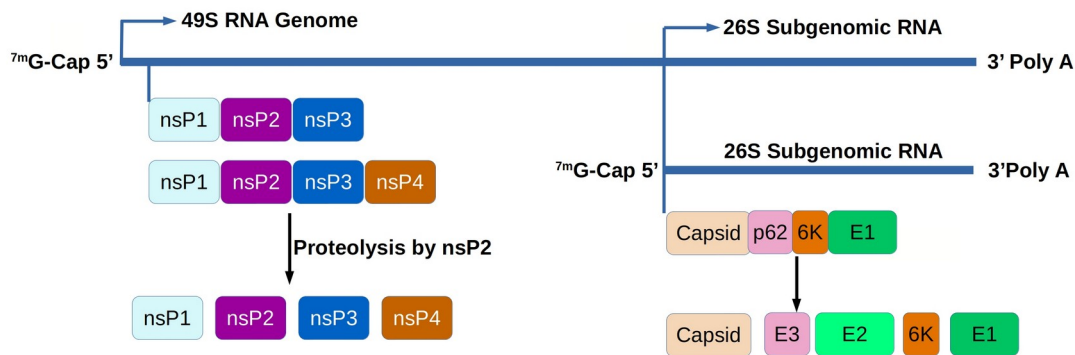
**Figure 10. Enzootic and epizootic/epidemic transmission cycles of Alphavirus.**

The enzootic transmission cycle of alphavirus is maintained among rodents, birds and other vertebrates as reservoirs. Mosquitoes, as *Culex spp*, represent primary vectors. In epizootic cycle, alphavirus is transmitted by mosquito vectors (e.g., *Aedes spp.*) to susceptible amplification hosts such as horses or humans. The image was created in *BioRender.com*.

### 3.4 Genome organization

The (+) RNA genome in alphaviruses is approximately 11.7 kb in length. The genome is capped at the 5' terminus and present a poly-adenylated tail at the 3' terminus with untranslated regions (UTR) at the 5' and 3' ends (figure 11). The overall organization of the alphaviral genome is conserved. The nsPs, forming the replication complex are encoded by the 5' proximal end. Upon cell penetration, the genomic (49S) RNA, is translated onto four nsPs (nsP1, nsP2, nsP3 and nsP4) inside host cytoplasm. Structural proteins are encoded from the last 1/3 through a sgRNA, leading to the production of a capsid protein, three glycosylated proteins (E1, E2, E3) and a small 64 amino-acids glycoprotein 6K) (121). Moreover, alphavirus genomic RNA contains

four highly conserved regions, known as cis-acting conserved sequence elements (CSEs). These CSEs play an important role in the regulation of the viral genome replication (122).



**Figure 11. Genome organization of alphavirus.** Alphavirus genome is 5' capped and 3' polyadenylated. The 49S RNA genome encodes for the nsPs polyprotein precursor, whereas the structural protein precursor is translated from the 26S sgRNA located in the 3' end of the genome. Both precursors are cleaved by viral and cellular proteases into individual functional proteins. The image was created in *BioRender.com*.

### 3.4.1 Life cycle

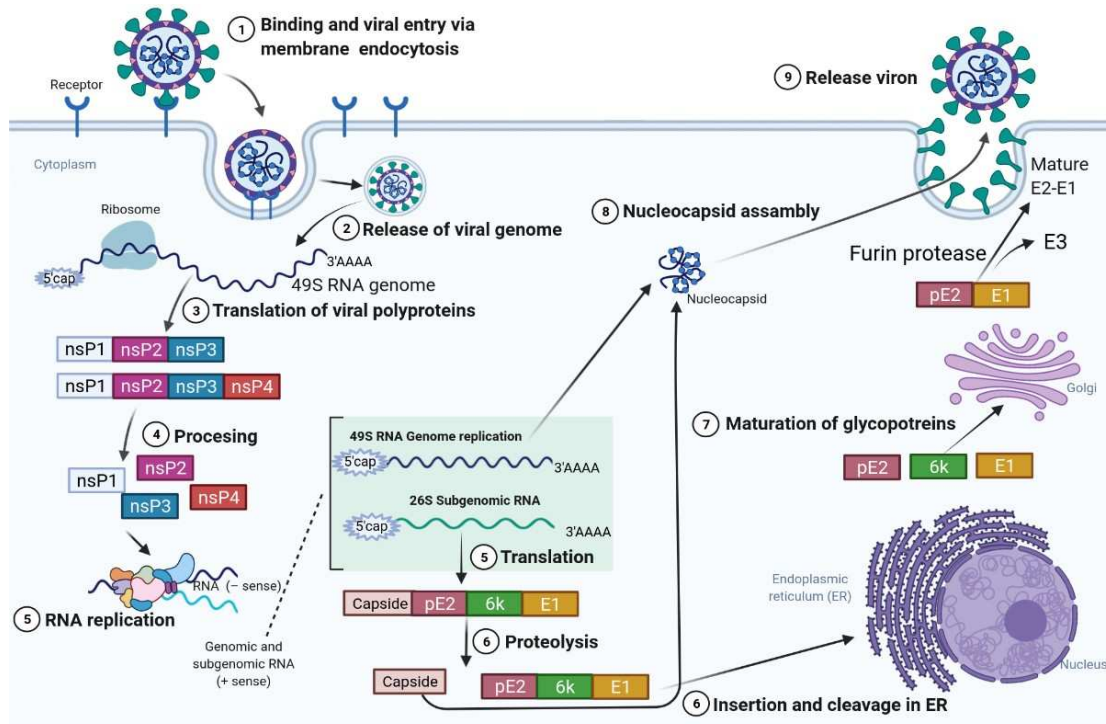
Entry of the virus inside the host cell is considered to be an engagement between viral glycoproteins and some uncharacterized host cell receptors (figure 12). Although the exact mechanism is still not fully elucidated, it seems that the common receptor is laminin-like (found on mosquitoes cell surface) (123). Once bound to the cell receptor, the virus enters by endocytosis, via a clathrin dependant mechanism. With the maturation of this endocytic vesicle, the pH drops and becomes acidic, this change in pH destabilizes the E1-E2 interaction and a fusion loop previously hidden and present at the end of the E1 glycoprotein becomes exposed. This loop is then inserted into the endosomal membrane, allowing the fusion of viral and endosomal membranes. This process leads to the release of the nucleocapsid into the cytoplasm

through a pore formation mechanism (124). Inside the cytoplasm, genomic RNA is released from the capsid, providing a template for replication (figure 12).

### **3.4.2 Virus replication**

Alphaviruses, like other positive-strand RNA viruses, replicate their RNA genome in association with cellular membranes. The 49S genomic RNA once released into the cytoplasm is directly translated with the help of cellular machinery into a polyprotein formed by either nsP123, and 4 or nsP12, and 3 (125). The translation takes place when the polymerase reads through the opal codon between the nsp3 and nsp4. The polyprotein is cleaved into isolated nsP1 to 4. The cleavage is the result of cis and trans activity of the papain like protease nsP2 (125). There are different intermediate stages of the polyprotein cleavage and each cleavage regulates a different step in the viral replication cycle. The polyprotein is first cleaved by cis activity of nsP2 leading to the formation of the nsP123 and nsP4 RdRp enzyme. Then during the late stage of the infection, the polyprotein formed by nsP123 gets further cleaved into nsP1 and nsP23 by the trans activity of nsP2. Finally, nsP23 gets cleaved into nsP2 and nsP3 by the trans activity of the nsP2 protease (126). Replication complex formation occurs on vacuoles derived from cytoplasmic and organelle's membranes. Vacuoles are identified to be the place where the (-) RNA is synthesized with the help of the replication complex formed by nsPs (127). The different types of replication complexes, listed above, are implicated at different replication stages. Hence, nsP1 to 4 combine together to form a complex required for the synthesis of (-) RNA during the early stages of the infection (127). As the infection proceeds, a transient intermediate stage is observed to be formed by nsP1, nsP2, nsP3 and nsp4 which might play a role in the (-) and (+) strand synthesis. This mature form of nsPs also leads to the synthesis of the 26S sgRNA from the subgenomic promoter on the genomic RNA (128). During the later stages of the infection, nsP23 and nsP1 and nsP4 combine to form replication complex leading to the synthesis of (+) RNA (figure 12). All the nsPs play different but critical roles in the replication process.



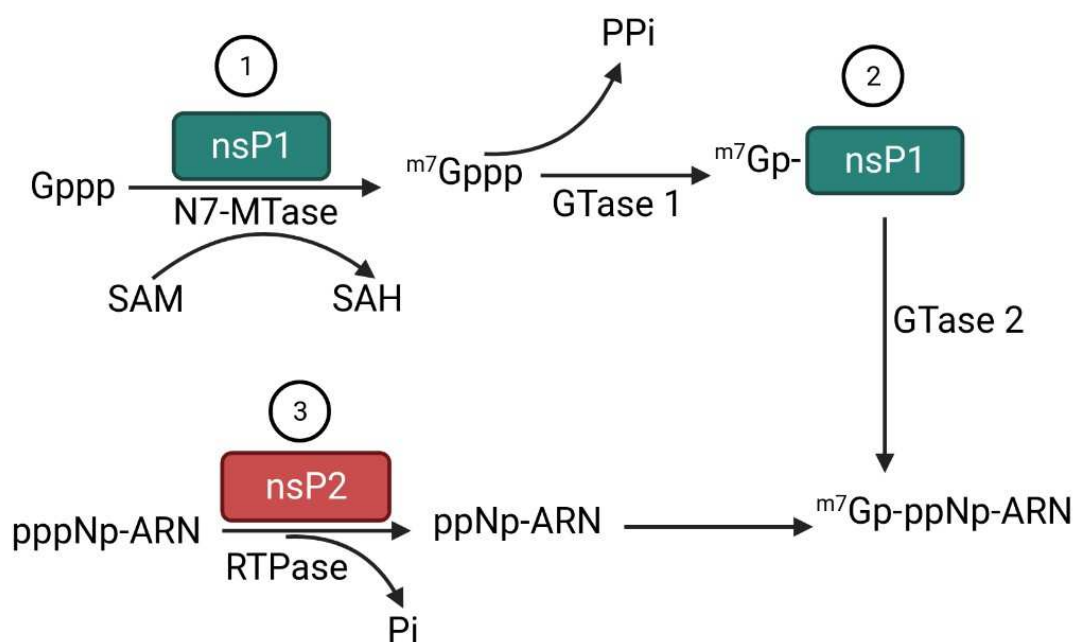


**Figure 12. Alphavirus replication cycle.** (1) Virus entry by receptor-mediated endocytosis. (2) Release of the nucleocapsid into the cytoplasm and the liberation of the viral genome. (3) Translation of nsPs. (4) Differential processing of nsP's. (5) nsPs with the viral RNA and host proteins form the membrane-bound replication complex, required for genome replication and the transcription of (-) RNA strand. The latter serves as a template for synthesis of both the full-length (+) genome and the and 26S sgRNA. (6, 7) Translation of sgRNA and processing of structural polyprotein (capsid, E1, E2, E3 and 6K). (8)(9) Glycoproteins are further modified in the endoplasmic reticulum (ER) and the Golgi. (10), (11) New virus particles are assembled and released by budding from the plasma membrane. The image was created in *BioRender.com*.

### 3.4.3 nsPs: Viral replication complex

#### 3.4.3.1 nsP1

The nsP1 is known as the capping enzyme. It possesses a MTase and GTase activities (figure 13), needed for the capping of the genomic RNA and sg RNA. Unlike other capping enzymes, nsP1 is unique in terms of harboring both capping activities. The reaction catalyzed by nsP1 is said to be “unconventional” (figure 13). The capping starts with the methylation of a GTP molecule at position N7 by the nsP1 MTase activity in the presence of SAM S-adenosylmethionine (SAM) methyl donor to form  $m^7GTP + Sadenosylhomocysteine (SAH)$ . Then, a GTase1 activity hydrolyzes the  $m^7GTP$  molecule and forms a covalent bond  $m^7GMP$ -nsP1 complex. The latter is transferred, through a GTase 2 activity, on the 5' diphosphate end of the viral RNA, forming a cap-0 structure ( $7^mGpppRNA$ ) (26). The removal of the  $\gamma$ -phosphate of viral RNA is insured by the RTPase activity carried by nsP2 protein (129). Alphaviruses lack 2'O MTase activity and only synthesize 0-caps at the 5' end of their RNAs, which makes them susceptible to the recognition by innate immunity restriction factor IFIT1. However, the presence of a stem-loop structure at the 5'UTR reduces the recognition of alphavirus RNAs, by RIG-1 and the IFIT 1/3, see general introduction section (130). Amphipathic helices, present in nsP1, ensure the anchoring of the replication complex to membranes (131). Moreover, cysteine residues 418-420 were found to be palmitoylated which further strengthens the association of alphavirus replication complex to membrane structures (132). nsP1 has also been found to participate directly to the initiation and the elongation of (-) RNA synthesis, through an interaction with nsP4 (133). nsP1 represents a good therapeutic target due to its unique mode of action. Targeted inhibition of this enzyme might reduce viral RNAs capping and make the virus vulnerable to innate immunity.

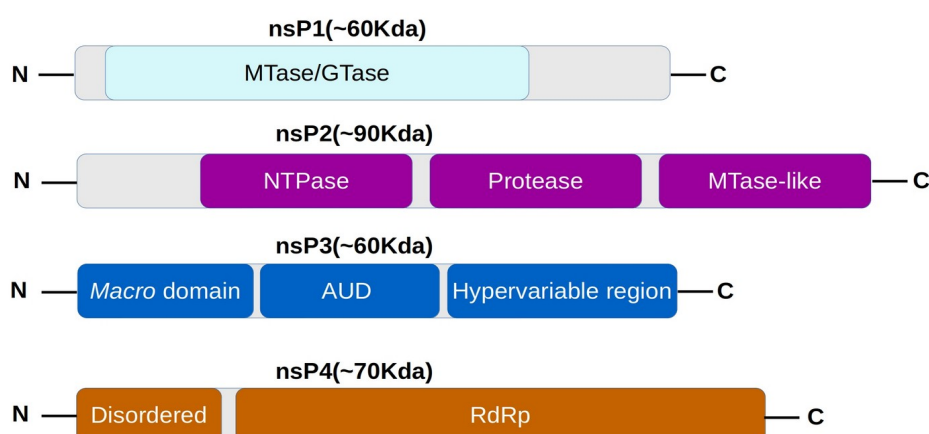


**Figure 13. Alphavirus nsP1 capping reactions.** (1) Using SAM as a methyl donor, nsP1 methylates GTP at position 7 (<sup>7m</sup>GTP). (2) <sup>7m</sup>GTP gets covalently attached to nsP1, forming a <sup>7m</sup>GMP-nsP1 complex. (3) The 5'-triphosphate end of the viral RNA is then converted to 5'-diphosphate with the help of RNA 5'-triphosphatase activity of nsP2. (4) The <sup>7m</sup>GMP-nsP1 is transferred to the 5'-end of the viral RNA, forming cap-0 structure. The image was created in *BioRender.com*.

### 3.4.3.2 nsP2

The alphavirus nsP2 is a multiple domain protein with several enzymatic activities. The N-terminal and C-terminal domains are assumed to possess helicase and protease activities, respectively (figure 14). Protease domain at the C-terminal is involved in the cleavage of the polyprotein nsP1234 at three different positions, i.e nsP1/2, nsP2/3 and nsP3/4 (125). The N terminal domain contain an NTP site and harbours ATPase and GTPase activities. NTPase conserved motifs I (GSGKS) and II (DEAF) beginning from residues 189 and 250 is responsible for NTP binding (134). Sequence of these and other downstream conserved motifs share considerable homology with the motifs found in superfamily 1 of helicases. nsP2 is able to bind RNA molecules

indicating its possible role in the unwinding of RNA during the replication process (135). First crystal structure of the protease domain was elucidated for VEEV, which gave insights about the different structural features of the C-terminal domain. From VEEV nsP2 protease structure, it is established that the C-terminal domain can further be subdivided into two subdomains: one, N-terminal harboring the protease activity and second C-terminal containing a MTase-like domain (figure 14) (136). However, this latter doesn't seem to be functional, suggesting that it must have other non-enzymatic functions. Only a small fraction of nsP2 is incorporated into viral replicative complex. The main fraction of nsP2 is present both in the cytoplasm and the cell nucleus (since it processes an NLS), where it is not associated with nsPs. nsP2 has been shown to antagonize the antiviral response by blocking STAT1 nuclear translocation (137).



**Figure 14. Domain organization of alphavirus nsPs: nsP1, nsP2, nsP3 and nsP4.** Recognized domains are indicated as colored boxes in their relative position. The image was created in *BioRender.com*.

### 3.4.3.3 nsP3

Nsp3 is divided into three domains (figure 14): (i) a *Macro* domain, highly conserved among alphaviruses, coronaviruses, Hepatitis E virus and Rubella virus; (ii) a

conserved alphavirus-unique domain (AUD) with a zinc binding domain (ZBD). Viable mutants of this domain were shown to exhibit defects in early events of RNA replication, likely the formation of the replication complex for (-) RNA synthesis was impeded (138); (iii) a C-terminal hypervariable domain, crucial for virus replication and host partners recruitment (139). This protein has been shown to be a phosphoprotein and majority of the phosphorylated serine and threonine residues are located in the hypervariable region (140). Moreover, nsP3 is an important factor in alphavirus neurovirulence.

#### **3.4.3.4 nsP4**

The C-terminal end of nsP4 constitutes the core RdRp domain, responsible for RNA synthesis in the viral replication complex (figure 14). The N-terminal domain is alphavirus-specific and might be partially disordered structurally. It appears to be important for the interaction with polyprotein P123 in the process of replication complex formation. Possible nsP1-nsP4 interaction was suggested based on the identification of compensation mutations in nsP1 along with substitutions of non-aromatic amino acids at the N-terminal of nsP4 (133). In SINV nsP4, mutation studies of the RdRp domain demonstrated a terminal adenylyltransferase activity, suggesting that the enzyme has a potential role in the maintenance of the 3' poly-A tail of the (+) RNA (141). The nsP4 is the most highly conserved protein in alphaviruses.

#### **3.4.4 Structural proteins involved in assembly and viral release**

Structural proteins are expressed from the 26S sgRNA in a precursor form as Capsid/p62/6k/E1 polyprotein. The polyprotein is processed by host and viral proteases to release individual structural proteins. Major modifications, particularly glycosylations, and processing of the p62 and E3 proteins takes place in the ER. Cleaved capsid proteins recognize a packaging signal located in the 5' region of the

genome, leading to the formation of nucleocapsid. In the final stage of the infection, this nucleocapsid interacts with the E2 at the plasma membrane, through a hydrophobic pocket present on its surface (142). This interaction leads to the attraction of more capsid glycoproteins, resulting in the formation of glycoprotein coat around the nucleocapsid which eventually buds out through the plasma membrane.

### 3.5 Objectives

Alphaviruses are RNA viruses that are capable of causing severe disease and are a significant burden to public health. The viral nsP1 is responsible of MTase and GTase activities, involved in adding the cap structure on viral RNAs. RNA capping machineries are drawing a lot of attention for antiviral drug design. The research component in this part of the thesis are focused on the structure/function study of the various activities carried out by nsP1.

The aim of this work is to understanding the specific contribution of nsP1-residues in drug resistance and the implication of structure and sequence of the 5' conserved stem-loop, in VEEV genome, for the capping process.

#### **Specific goals:**

- (i) Perform a mutational analysis on two nsP1 VEEV positions associated with ribavirin resistance in alphavirus.
- (ii) Determine the implication of the mutated residues in the regulation of the capping reaction.
- (ii) Explore the specificities of the GTase acceptor substrate, in terms of length, RNA secondary structure, and/or sequence of the 5' VEEV genome.

# Results

### 3.6 Article 2

#### ***Structure and sequence requirements for RNA capping at the Venezuelan Equine Encephalitis Virus RNA 5'-end***

**Oney Ortega Granda<sup>a</sup>** , Coralie Valle<sup>a\*</sup> , Ashleigh Shannon<sup>a</sup> , Etienne Decroly<sup>a</sup> , Bruno Canard<sup>a</sup> , Bruno Coutard<sup>b</sup> and Nadia Rabah<sup>a,c,#</sup>

<sup>a</sup> Aix Marseille Université, CNRS, AFMB UMR 7257, Marseille, France

<sup>b</sup> Unité des Virus Emergents (UVE: Aix-Marseille Univ-IRD 190-Inserm, 1207-IHU Méditerranée Infection) Marseille, France

<sup>c</sup> Université de Toulon, 83130 La Garde, France

\*Present address: National Institute for Public Health and Environment (RIVM). Bilthoven, The Netherlands

# Address correspondence to Nadia Rabah, [nadia.rabah@univ-tln.fr](mailto:nadia.rabah@univ-tln.fr)

**Journal of Virology (2021) Jul 12;95(15):e0077721.**





# Structure and Sequence Requirements for RNA Capping at the Venezuelan Equine Encephalitis Virus RNA 5' End

Oney Ortega Granda,<sup>a</sup> Coralie Valle,<sup>a\*</sup> Ashleigh Shannon,<sup>a</sup> Etienne Decroly,<sup>a</sup> Bruno Canard,<sup>a</sup> Bruno Coutard,<sup>b</sup> Nadia Rabah<sup>a,c</sup>

<sup>a</sup>Aix Marseille Université, CNRS, AFMB UMR 7257, Marseille, France

<sup>b</sup>Unité des Virus Emergents (UVE): Aix-Marseille Univ-IRD 190-Inserm, Marseille, France

<sup>c</sup>Université de Toulon, La Garde, France

**ABSTRACT** Venezuelan equine encephalitis virus (VEEV) is a reemerging arthropod-borne virus causing encephalitis in humans and domesticated animals. VEEV possesses a positive single-stranded RNA genome capped at its 5' end. The capping process is performed by the nonstructural protein nsP1, which bears methyl and guanylyltransferase activities. The capping reaction starts with the methylation of GTP. The generated m<sup>7</sup>GTP is complexed to the enzyme to form an m<sup>7</sup>GMP-nsP1 covalent intermediate. The m<sup>7</sup>GMP is then transferred onto the 5'-diphosphate end of the viral RNA. Here, we explore the specificities of the acceptor substrate in terms of length, RNA secondary structure, and/or sequence. Any diphosphate nucleosides but GDP can serve as acceptors of the m<sup>7</sup>GMP to yield m<sup>7</sup>GpppA, m<sup>7</sup>GpppC, or m<sup>7</sup>GpppU. We show that capping is more efficient on small RNA molecules, whereas RNAs longer than 130 nucleotides are barely capped by the enzyme. The structure and sequence of the short, conserved stem-loop, downstream to the cap, is an essential regulatory element for the capping process.

**IMPORTANCE** The emergence, reemergence, and expansion of alphaviruses (genus of the family *Togaviridae*) are a serious public health and epizootic threat. Venezuelan equine encephalitis virus (VEEV) causes encephalitis in human and domesticated animals, with a mortality rate reaching 80% in horses. To date, no efficient vaccine or safe antivirals are available for human use. VEEV nonstructural protein 1 (nsP1) is the viral capping enzyme characteristic of the *Alphavirus* genus. nsP1 catalyzes methyltransferase and guanylyltransferase reactions, representing a good therapeutic target. In the present report, we provide insights into the molecular features and specificities of the cap acceptor substrate for the guanylation reaction.

**KEYWORDS** capping, *Alphavirus*, guanylyltransferase, methyltransferase, RNA

Venezuelan equine encephalitis virus (VEEV) is an encephalitic pathogen transmitted by hematophagous arthropods. The virus circulates throughout the Americas and causes regular outbreaks affecting humans and equids. In horses, the mortality rate is high and can reach up to 80%. In humans, even though the fatality rate is low (<1%), up to 14% of infected individuals can develop long-lasting and even permanent neurological sequelae (1). Over the last few decades, much effort has been devoted to vaccine development. The strategies adopted include the use of live-attenuated, inactivated, chimeric, and various subunit vaccine candidates. While the live-attenuated vaccine, TC-83, is commonly used in Mexico and Colombia, there is still no approved vaccine for human use, urging the development of antivirals (2).

VEEV is a member of the *Alphavirus* genus, which contains 31 species and belongs to the *Togaviridae* family. VEEV possesses a single-stranded positive-sense RNA of approximately 11.5 kb that is polyadenylated and contains a 7-methylguanosine (m<sup>7</sup>G) cap structure (cap 0, m<sup>7</sup>GpppN, where N represents the first transcribed nucleotide of

**Citation** Ortega Granda O, Valle C, Shannon A, Decroly E, Canard B, Coutard B, Rabah N. 2021. Structure and sequence requirements for RNA capping at the Venezuelan equine encephalitis virus RNA 5' end. *J Virol* 95:e00777-21. <https://doi.org/10.1128/JVI.00777-21>.

**Editor** Susana López, Instituto de Biotecnología/UNAM

**Copyright** © 2021 American Society for Microbiology. All Rights Reserved.

Address correspondence to Nadia Rabah, [nadia.rabah@univ-tln.fr](mailto:nadia.rabah@univ-tln.fr).

\* Present address: Coralie Valle, National Institute for Public Health and Environment (RIVM), Bilthoven, The Netherlands.

**Received** 11 May 2021

**Accepted** 14 May 2021

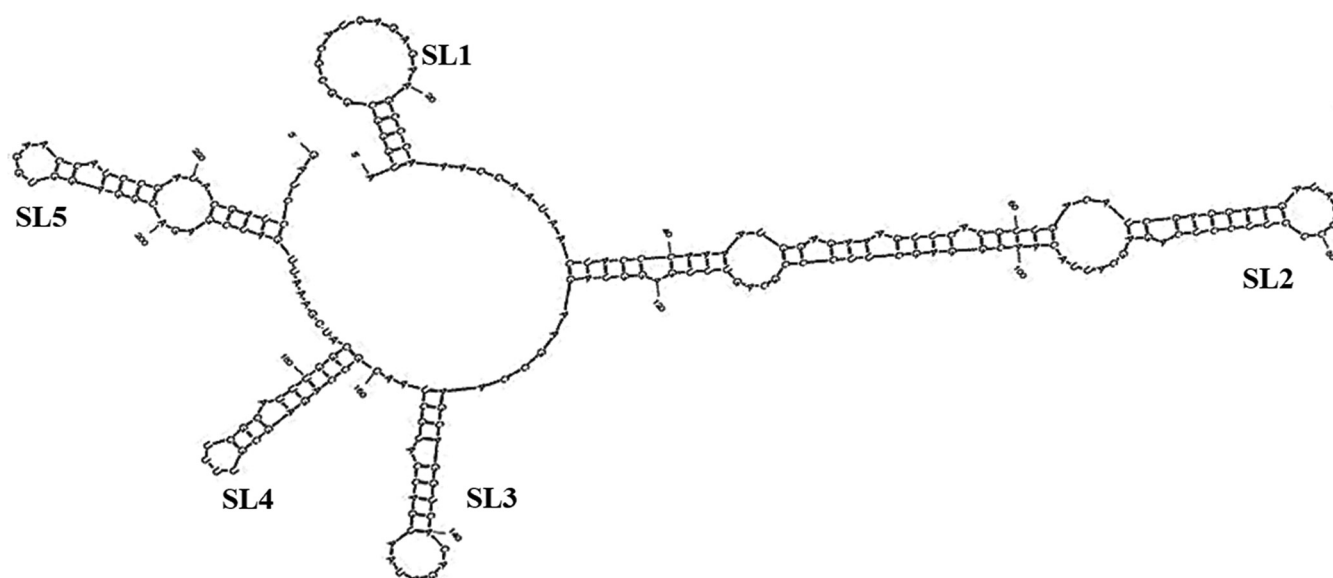
**Accepted manuscript posted online** 19 May 2021

**Published** 12 July 2021

the viral RNA). The genome contains two open reading frames (ORFs) and acts as an mRNA after cell entry. The first ORF leads to the synthesis of nonstructural polyprotein precursors, named P123 and P1234, which are ultimately auto-processed to nsP1 to nsP4, forming the membrane-associated replication/transcription complex in virus-driven lipid structures called spherules. The second ORF encodes the structural polyprotein precursor and is translated from the subgenomic (SG) RNA. This polyprotein is ultimately cleaved to yield E1, E2, E3, C, 6K, and TF structural proteins that drive the formation of capsid shells and glycoprotein spikes, directing the packaging of new viral genomes (3, 4). The capping of both genomic and subgenomic RNAs is mediated by the nsP1, harboring both methyltransferase and guanylyltransferase activities. Jones et al. have recently shown that nsP1 can form membranous dodecamer pore complexes, ensuring the transit between the cytoplasm and the spherule contents (5). The reaction catalyzed by nsP1 involves the following three steps: (i) in the presence of *S*-adenosylmethionine (SAM), GTP is methylated by the methyltransferase activity (MTase) of nsP1 to form m<sup>7</sup>GTP and *S*-adenosylhomocysteine (SAH); (ii) nsP1 guanylyltransferase activity (GTase1) allows the formation a covalent m<sup>7</sup>GMP-nsP1 intermediate, releasing inorganic pyrophosphate (PPi); and (iii) the methylated guanosine is transferred to the ppRNA, whose  $\gamma$ -phosphate is previously removed by the nsP2 protein (6–12). This results in the formation of a type 0 cap structure, methylated at the N7 position only. While the catalytic amino acids in nsP1 have been relatively well characterized, specifics of the RNA substrate required for efficient capping remain poorly described.

It has previously been demonstrated that specific secondary structures present at the 5' region of the viral RNA are essential to the virus. These structures, known as conserved sequence elements (CSE), can greatly vary from one alphavirus species to another, suggesting that they play a key role not only in the regulation of the viral replication cycle but also in host diversity, tissue tropism, and probably disease outcome (13–17). *Cis*-acting elements in the alphavirus genome have been identified in both 5' and 3' untranslated regions (UTRs), but also sprawl through coding regions of viral RNA sequences. Experimental evidences confirm the importance of CSE within the 5' UTR and near the nsP1 start codon for the initiation of viral RNA replication. A 3' CSE of 19 nucleotides upstream of the poly(A) tail was shown to act in concert with the 5' CSE for the regulation of minus- and plus-strand RNA synthesis (15, 17, 18). Similarly, the SG 5' UTR, along with the first capsid protein-coding nucleotides, serves as the SG promoter (19, 20). The role of RNA elements goes beyond RNA synthesis *per se*. Thus, both genomic and SG 5' RNA elements contain translational enhancer sequences (21, 22). RNA structural elements spanning nsP1 and nsP2 sequences were shown to act as genome packaging signals (23–25). Moreover, viral RNA structures can serve as docking sites not only for the viral replication machinery but also for the interaction with host factors (26).

In the VEEV-attenuated strain TC-83, which is the basis of the veterinary vaccine, two key single nucleotide changes were shown to contribute to the attenuated phenotype. One mutation is located in the 5' UTR and corresponds to G3A substitution in the extreme 5' terminus (27, 28). This replacement strongly destabilizes the secondary structure present at the 5' terminus. Interestingly, the mutation results in an increase in the synthesis of genomic RNA, reducing SG RNA synthesis and consequently creating an imbalance in the ratio of subgenomic to genomic RNA (29). Moreover, this secondary structure has additionally been shown to play a crucial role in immune escape by resisting IFIT1 restriction, a product of an interferon-stimulated gene that contributes to the regulation of protein synthesis. While many RNA viruses evade IFIT1 restriction through 2'-*O* methylation of their 5' cap (m<sup>7</sup>GpppN<sub>m</sub>), alphaviruses lack this second methylation. Instead, the secondary structure located in the 5' UTR compensates by altering IFIT1 binding, thereby evading translational inhibition (26). Although the role of secondary structure elements in immune escape and RNA synthesis has thus been established, the possible impact of these structural motifs on VEEV RNA capping has never been addressed. The objective of the present study is therefore to investigate the impact of



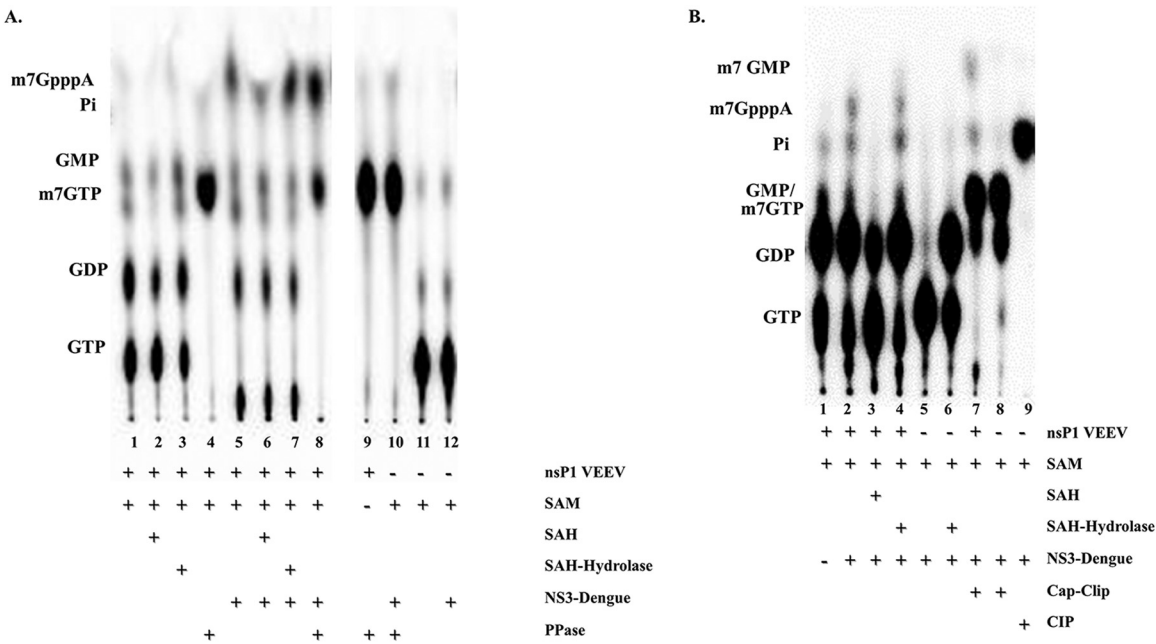
**FIG 1** Mfold-predicted secondary structure of VEEV 5' end genomic RNA. Fold of the first 230 nucleotides showing the 5 stem-loops SL1 to SL5.

RNA length, sequence, and secondary structure on nsP1-RNA binding and efficiency of RNA capping in order to set the minimal acceptor substrate requirement for m<sup>7</sup>GMP transfer.

## RESULTS

In a previous study, we have shown that VEEV-nsP1 is able to cap a synthetic 15-nucleotide-long oligomer mimicking the 5' end of the VEEV genomic RNA (11). In order to deepen our understanding of the molecular motifs potentially modulating capping efficiency, we investigated the effect of 5' RNA length and structure on the cap formation. The computer-predicted secondary structure of the first 230 VEEV RNA nucleotides is presented in Fig. 1. The 5' end of VEEV folds into several stem-loops (SL) similar to those reported for Sindbis virus (SINV) and Semliki Forest virus (SFV) (14, 18, 30). The structure starts with a short stem-loop (SL1) formed by the first 30 nucleotides of the VEEV genome. The second hairpin (SL2) is longer and contains 3 loops. It is comprised of nucleotides 35 to 130 and thus includes part of the nsP1 coding sequence. The latter structure is followed by three short SLs, SL3 to SL5. SL3 and SL4 are believed to form the 51-nucleotide CSE acting as a transcriptional enhancer (16).

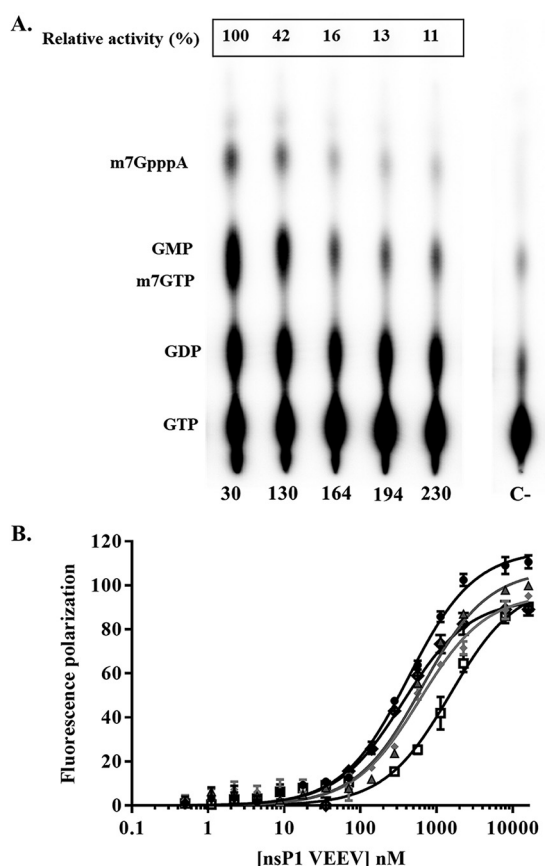
**Capping of VEEV wild-type RNA SL1.** We first characterized the capping parameters of an *in vitro*-transcribed VEEV RNA corresponding to SL1 (Fig. 1). A 5'-triphosphate RNA substrate was incubated with recombinant VEEV nsP1 enzyme (12) in the presence of [ $\alpha$ -<sup>32</sup>P]GTP under various experimental conditions. The sequential capping reaction described in the introduction includes the formation of m<sup>7</sup>GTP, which is then covalently complexed to the enzyme (m<sup>7</sup>GMP-nsP1) before getting transferred onto a  $\gamma$ -dephosphorylated RNA (ppRNA). In cells, this 5'-triphosphatase activity is presumably provided by the helicase domain of nsP2. Here, RNA 5' dephosphorylation was carried out by dengue virus (DENV) nonstructural protein 3 (NS3) (31). Hence, in order to underline the importance of the  $\gamma$ -phosphate dephosphorylation, we used both NS3-treated and untreated RNA. For the characterization of the reaction product, the capped RNA was subsequently treated with P1 endonuclease prior to analysis on thin-layer chromatography (TLC). TLC plates were developed in either (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (Fig. 2A) or LiCl (Fig. 2B), the latter allowing a better resolution of high migration spots. The results are presented in Fig. 2. The first 4 lanes in Fig. 2A, as well as lane 1 in Fig. 2B show the capping reaction on a nondephosphorylated RNA (pppRNA). In the presence of SAM (Fig. 2A and B, lane 1), GTP is methylated to form m<sup>7</sup>GTP. The reaction by-products



**FIG 2** Capping of VEEV WT RNA SL1. RNA was pretreated with DENV NS3 to remove the  $\gamma$ -phosphate. Then, the capping reaction was carried out in the presence of 2  $\mu$ M nsP1, 100  $\mu$ M SAM, and 0.33  $\mu$ M [ $\alpha$ - $P^{32}$ ]GTP. m<sup>7</sup>GpppA caps were digested from RNA with nuclease P1 and resolved on polyethylenimine cellulose thin-layer chromatography (TLC) using (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (A) or LiCl (B) as a mobile phase. The migration of standards are indicated on the left side of each TLC chromatogram. The plates were revealed by autoradiography. The conditions of each capping reaction are summarized below the TLC plate. Data are representative of three independent experiments.

include unused GTP, but also GDP and GMP. However, no m<sup>7</sup>GpppA is detected. It was previously demonstrated that SAH, the by-product of GTP methylation reaction, is required for m<sup>7</sup>GMP-nsP1 complex formation when m<sup>7</sup>GTP is used as the substrate for the capping reaction (7, 11). In the present study, we show that the addition of an exogenous excess of SAH reduces m<sup>7</sup>GTP production (Fig. 2A, lane 2). Moreover, the addition of SAH-hydrolase to the reaction mixture (Fig. 2A, lane 3) promotes the MTase reaction. When the RNA substrate is  $\gamma$ -dephosphorylated, m<sup>7</sup>GpppA is detected (Fig. 2A, lane 5, and Fig. 2B, lane 2). Cap formation is reduced in the presence of an excess of SAH (Fig. 2A, lane 6, and Fig. 2B, lane 3) and slightly increased when SAH-hydrolase is added (Fig. 2A, lane 7, and Fig. 2B, lane 4), in accordance with the production of m<sup>7</sup>GTP observed in lanes 1 to 3 (Fig. 2A). Controls run without enzyme (Fig. 2A, lanes 11 to 12, and Fig. 2B, lane 5) show a main signal corresponding to GTP, which is slightly hydrolyzed to GDP and GMP during the course of the reaction. We note that the addition of SAH-hydrolase alone (Fig. 2B, lane 6) increases GTP hydrolysis to GDP. The addition of a GDPase to the capping reaction did not increase m<sup>7</sup>GTP synthesis (data not shown). In lane 9 (Fig. 2B), the reaction mixture was treated with calf intestinal phosphatase (CIP) to confirm the position of Pi.

Another by-product resulting from the formation of the m<sup>7</sup>GMP-nsP1 complex is PPI, which has previously been shown to inhibit the formation of the m<sup>7</sup>GMP-enzyme covalent link (32, 33). In accordance with this, the addition of inorganic pyrophosphatase (PPase) (Fig. 2A, lanes 4 and 8 to 10) promoted formation of m<sup>7</sup>GTP and/or m<sup>7</sup>GpppA (Fig. 2A, lanes 4 and 8, respectively). In the absence of SAM or nsP1 enzyme (lanes 9 and 10), only GMP can be detected when the PPI hydrolase is present. In order to confirm the identity of the cap, a sample of the reaction, lanes 2 and 5 (Fig. 2B), was treated with Cap-Clip acid pyrophosphatase, an enzyme which hydrolyzes pyrophosphate bonds in m<sup>7</sup>GpppN structures. In the absence of nsP1, only GDP and GMP are detected (Fig. 2B, lane 8). In lane 7, we note that the cap structure is degraded into m<sup>7</sup>GMP, confirming that the product is m<sup>7</sup>GpppA.



**FIG 3** Effect of RNA length on nsP1 binding and capping efficiency. (A) Capping reaction was carried out for VEEV 5' RNAs. m<sup>7</sup>GpppA caps were digested from capped RNA using nuclease P1 and subjected to TLC. The data are representative of three independent experiments. (B) Fluorescence polarization experiment assessing nsP1 binding to 3' Cy5-labeled 30 (●), 130 (Δ), 164 (◆), 194 (□), and 230 (◆)-nucleotide VEEV 5' RNA. Dissociation constants ( $K_d$ ) were determined using Hill slope curve fitting (GraphPad Prism 7 program) and compiled in Table 1. All data points are the means of three wells, and all error bars represent the standard deviation.

**Effect of RNA length on capping efficiency.** Next, we wondered if the other SL structures present at the 5' end could affect the capping reaction and if RNA length might play a role in regulating the capping process. We produced VEEV RNAs from different lengths encompassing the different SLs by *in vitro* transcription, based on the predicted RNA structures (Fig. 1). Five RNAs corresponding to the first 30, 130, 164, 194, and 230 nucleotides were generated. When the five RNAs are used as the substrate for the RNA capping reaction (Fig. 3A), the intensity of m<sup>7</sup>GpppA formation is inversely related to RNA length. A drastic decrease in GTase activity is observed with substrates starting from 130 nucleotides long, with 40% of remained capping compared to the 30-mer RNA. The relative activity drops to 16, 13, and 11% for 164-, 194-, and 230-nucleotide RNA, respectively. These results support that the longer the RNA is, the less efficient the capping synthesis is.

We have previously shown that for short RNAs (15 nucleotides long) mimicking the VEEV 5' end, the phosphorylation state is crucial for the RNA-enzyme interaction (11). In order to evaluate the role of substrate length in RNA recruitment, the affinity of nsP1 for the five RNA substrates was assessed by fluorescence polarization (FP). To that end, we labeled the 3' end of each RNA with Cy5 and performed FP experiments with increasing amounts of enzyme (Fig. 3B). Apparent equilibrium dissociation constant ( $K_d$ ) values are compiled in Table 1. FP results show that the binding constants for 30-, 130-, 164-, 194-, and 230-nucleotide-long RNAs oscillate between 0.36 and 1.3  $\mu$ M. With the exception of the 194-nucleotide-long RNA (SL1 to SL4), which is the least



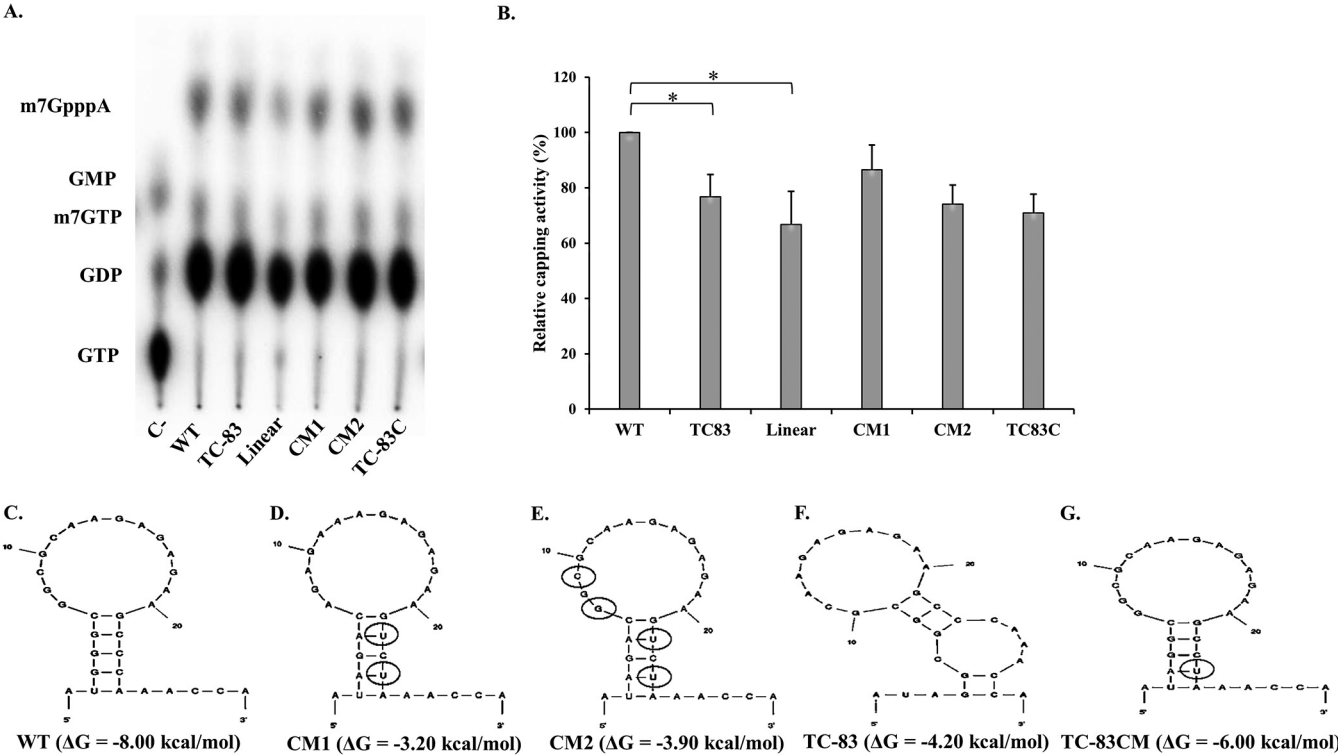
**TABLE 1** Effect of 5' VEEV RNA length on RNA-nsP1 interaction<sup>a</sup>

RNA length (no. of nucleotides)	Apparent $K_d$ ( $\mu$ M)
30	$0.42 \pm 0.03$
130	$0.64 \pm 0.08$
164	$0.60 \pm 0.07$
194	$1.3 \pm 0.18$
230	$0.36 \pm 0.05$

<sup>a</sup>Kinetic parameters were calculated using the SigmaPlot program. Values are expressed as mean  $\pm$  standard deviation with an  $R^2$  value of  $\geq 0.98$ .

preferred RNA substrate for nsP1, no major differences in  $K_d$  are noted within the set of tested RNAs. Comparison of affinity and catalytic data indicate that while nsP1 is still able to bind RNAs longer than 130 nucleotides, the capping process is largely impeded for longer substrates. Hence, weak variations of RNA affinity toward nsP1 are not likely to play a crucial role in the capping efficacy. The presence of SL3 to SL5 is deleterious not only for the formation of the cap structure itself, i.e., m<sup>7</sup>GpppA, but also for the MTase reaction since the signal corresponding to m<sup>7</sup>GTP is very weak. All the tested RNAs share a denominator, SL1, which therefore requires further evaluation in terms of sequence and structure determinant toward capping efficacy.

**Effect of RNA 5' secondary structure on capping efficiency.** The 5'-end SL1 structure has been shown to be important for immune evasion but also the regulation of genomic and subgenomic RNA synthesis (33). To investigate its role in the capping reaction, we linearized SL1 by mutating positions 3, 5, 7, 9, and 11 to adenosines (linear mutant). In order to evaluate if the effect on the capping efficacy can be influenced by changes in either sequence or structure of SL1 (Fig. 4C), we included in the study (i)



**FIG 4** Effect of RNA 5' secondary structure on capping efficiency. (A) We subjected 30-nucleotide oligomers corresponding to the WT, linear, TC-83, and various compensatory mutant sequences (as shown in Fig. 1) to DENV NS3 treatment and then assessed them under capping conditions. The data shown are representative of at least three independent experiments. (B) Relative activity corresponding to means  $\pm$  standard deviation of VEEV nsP1 toward various RNA 5' secondary structures. The relative activity of WT 5' VEEV RNA was considered 100% for comparison with the other reaction conditions. \*,  $P < 0.05$  by  $t$  test. (C and F) Computer-predicted fold of the 5' 30-nucleotide oligomers in WT and TC-83 VEEV RNA, respectively. (D, E, and G) Mfold of compensatory mutants from linear 30-nucleotide oligomers (C and D) and TC-83 (F). The mutations are highlighted with black circles. The computed  $\Delta G$  values for each RNA are mentioned.

**TABLE 2** Effect of SL1 nucleotide mutations on RNA-nsP1 binding kinetics<sup>a</sup>

Type of RNA	Apparent $K_d$ ( $\mu$ M)
WT	0.42 $\pm$ 0.03
Linear	7.53 $\pm$ 2.3
TC-83	4.14 $\pm$ 1.6
CM1	2.69 $\pm$ 0.7
CM2	3.03 $\pm$ 0.7
TC-83CM	2.88 $\pm$ 0.8

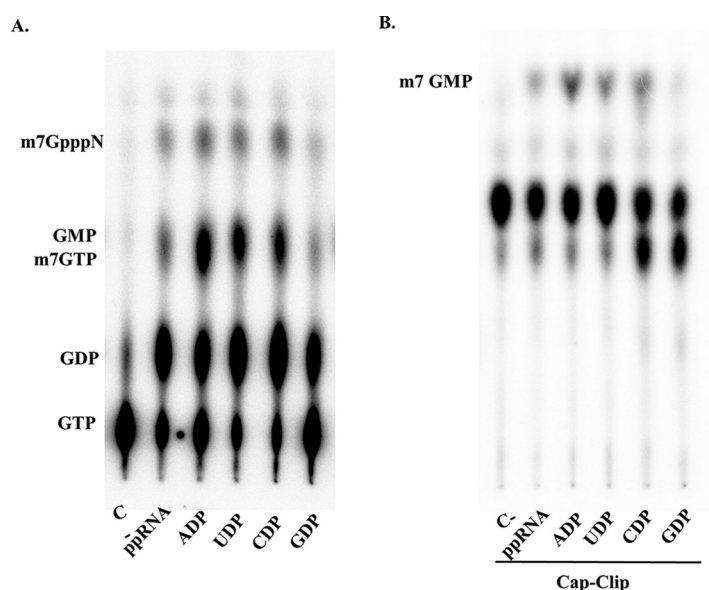
<sup>a</sup>Kinetic parameters were calculated using the SigmaPlot program. Values are expressed as mean  $\pm$  standard deviation with an  $R^2$  value of  $\geq 0.95$ .

two computer-predicted compensatory mutants, CM1 and CM2 (Fig. 4D and E, respectively), that allow the recovery of the SL1 hairpin structure but with different nucleotide composition; (ii) a 30-nucleotide oligomer, the G3A mutation observed in the vaccinal strain TC-83 (26) (Fig. 4F); and (iii) a structural compensatory mutant, TC-83C, allowing the refolding of the TC-83 into a hairpin structure similar to SL1 (Fig. 4G). The set of SL1 RNAs was first tested in the capping reaction. Interestingly, all of them can be capped by nsP1 (Fig. 4A) but with variable efficacy. Relative capping activities were thus calculated and normalized to the total radioactivity in each lane (Fig. 4B). All the mutations leading to a change of the structure or the sequence of the 5' RNA have a negative effect on RNA capping. The TC-83 RNA, mimicking the RNA of the VEEV veterinary vaccine, has more than 20% reduction in cap synthesis. The linear RNA shows more than 30% loss in capping efficiency. Compensatory mutations fail to restore completely RNA capping, although CM1 reaches nearly 85% of the wild-type (WT) activity. We next tested the effect of SL1 sequence/structure modifications on enzyme binding. FP experiments show that WT SL1 has higher polarization values, with a computed apparent  $K_d$  of 0.42  $\mu$ M (Table 2). The lowest affinity is seen for the linear RNA, with an apparent estimated  $K_d$  of 7.5  $\pm$  2.3  $\mu$ M (Table 2). The TC-83 RNA is also significantly impacted in terms of binding compared to the WT substrate, with a  $K_d$  of 4.1  $\mu$ M. All the compensatory mutants, with an SL similar in fold to the WT, partially restore binding affinity, with estimated  $K_d$  values of 2.69, 3.03, and 2.88  $\mu$ M, for CM1, CM2, and TC-83CM, respectively (Table 2).

**Nucleotides as the substrates for N7 cap reaction.** The results obtained so far support the fact that the capping process occurs early, since long RNAs tend to interfere with m<sup>7</sup>GMP transfer on RNA. Moreover, we have previously shown that capping can occur efficiently on 15 mer of 5' VEEV RNA (11) and non-VEEV RNA sequences of comparable size (data not shown). Studies on Bamboo mosaic virus (BaMV), a member of the alpha-like superfamily, have shown that GDP and ADP can also serve as acceptors for m<sup>7</sup>GMP to form m<sup>7</sup>GpppG or m<sup>7</sup>GpppA (33). We therefore conducted a capping experiment with nucleotides as acceptors for m<sup>7</sup>GMP. In the capping reaction mixture containing VEEV-nsP1, [ $\alpha$ -<sup>32</sup>P]GTP, and SAM, RNA was replaced by either GDP, ADP, or UDP (Fig. 5). The cap can be formed efficiently with ADP, UDP, and CDP as observed by the presence of m<sup>7</sup>GpppN (where N represents A, U, or C) (Fig. 5A). Caps corresponding to m<sup>7</sup>GpppU and m<sup>7</sup>GpppC comigrate with the m<sup>7</sup>GpppA signal under our conditions, as described previously (34). When GDP is used as the substrate for m<sup>7</sup>GMP transfer, the signal is very weak (Fig. 5A). A sample of each reaction condition was treated with Cap-Clip acid pyrophosphatase to confirm the cap nature of the observed signal. Following this treatment, the m<sup>7</sup>GpppN signal disappeared. A signal corresponding to m<sup>7</sup>GMP is detected (Fig. 5B), confirming the identity of the cap.

## DISCUSSION

The unconventional capping mechanism of alphavirus nsP1 makes it an attractive drug target. However, exploiting nsP1 drug target potential could rationally be achieved through an extensive and comprehensive characterization of its enzymatic activity. In order to dissect the role of RNA element on VEEV RNA capping, we tested different RNA substrates that



**FIG 5** Nucleotides can act as acceptors for N7 capping reaction. (A) ADP, UDP, CDP, and GDP nucleotides were used in replacement of RNA in the capping reaction described in Materials and Methods. (B) Digestion of m<sup>7</sup>GpppN cap with Cap-Clip acid pyrophosphatase enzyme before TLC migration. The experiment was repeated at least three times. One representative experiment is shown.

were various in length and secondary structure on the capping reaction. The analysis of the extreme 5'-secondary structure SL1 (Fig. 1 and 4C) revealed that this RNA hairpin can be capped efficiently by nsP1, and the capping reaction occurs in concordance with what was previously described in the literature. RNA capping is reduced by an excess of SAH and prompted by the addition of either SAH-hydrolase or PPase. Both SAH and PPi are by-products of the MTase reaction. PPi was shown to play a role in the regulation of cap transfer onto RNA and the regeneration of GTP (35). For eukaryotic guanylyltransferases, the formation of the GMP-protein complex before transfer to RNA was shown to be reversible. When PPi is added in excess to the reaction, GTP cannot be regenerated, and this phenomenon is reversed by the addition of PPase (35). One can speculate that a similar process is occurring in our case, leading to the methylation of all the GTP present in the reaction (Fig. 2A, lanes 4 and 8), thereby enhancing m<sup>7</sup>GpppA formation. It was shown previously that SAH is necessary for m<sup>7</sup>GMP-nsP1 complex formation when m<sup>7</sup>GTP is provided to the enzyme (11, 36). However, when added in excess to the reaction mixture containing GTP as substrate, SAH inhibits GTP methylation (Fig. 2A, lane 2) and, ultimately, m<sup>7</sup>GpppA synthesis (Fig. 2A, lane 6). These data support the proposed *Alphavirus* capping model according to which SAM and SAH share the same binding site (36). Moreover, the results obtained in the presence of SAH-hydrolase (Fig. 2A, lanes 3 and 7) indicate that SAH is not released from the active complex until the end of the capping process; otherwise, the production of m<sup>7</sup>GTP and m<sup>7</sup>GpppA would have been prevented.

We have shown previously that VEEV nsP1 is able to bind di-, tri-, and, to a certain extent, monophosphorylated 15-nucleotide-long VEEV 5' RNAs. However, RNA capping was observed almost exclusively with ppRNA. The faint capping observed on pppRNA led us to postulate, at that time, that either a small proportion of pppRNA was hydrolyzed at the  $\gamma$ -phosphate or the enzyme was able to form m<sup>7</sup>GppppA cap, as reported for D1 subunit of vaccinia virus (VV) capping enzyme mutants or vesicular stomatitis virus (VSV) (11, 37–39). In the present conditions, we failed to detect any significant capping reaction with pppRNA. A contrast between RNA binding capacity and capping efficiency appeared when RNA of different lengths was tested. While nsP1 is able to efficiently bind RNA ranging from 30 to 230 nucleotides, it is barely able to cap RNAs longer than 130 nucleotides (Fig. 3). The rationale behind this experiment relied on several assumptions. First, it was shown in SFV that RNA capping occurs concomitantly with



RNA biosynthesis (40). Second, many studies highlighted that a minimal RNA chain extension is a prerequisite for capping to occur. For example, Tat protein of human immunodeficiency virus (HIV) stimulates capping when the nascent RNA is 19 to 22 nucleotides long (41). In the case of VSV, RNA capping occurs after reaching 30 nucleotides (42). The reported data for BaMV are puzzling. Huang et al. (33) found that efficient capping occurs when the RNA is longer than 50 nucleotides. However, in the same study, they also showed that the BaMV capping enzyme is able to cap efficiently nucleotides, supporting a possible pre- to cotranscriptional capping event, which might relate to RNA synthesis priming, discussed below. Third, in SINV, we know that the ratio of capped versus uncapped RNAs varies between cell types. In the first 9 h of infection, uncapped RNAs are getting encapsidated to form new virions (43). This might suggest that, indeed, the entire genome can be synthesized without being capped. Our data do not support a posttranscriptional RNA capping event, at least for VEEV nsP1. The other option allowing the generation of uncapped genomes would be a decapping event. A recent paper demonstrated that during viral infection, cellular decapping enzymes DCP1 and DCP2 are directed to viral replication sites for the restriction of viral infection (44).

RNA alignment of different *Alphavirus* species suggests that, although the presence of CSEs is conserved, they can vary in sequence. Despite these sequence variations, some regions could fold into similar CSEs, suggesting a conserved role for these elements. Similarly, divergence in RNA fold or sequence might contribute to functional diversity (14). Earlier studies reported that modifications of the sequence or the structure of the SL1 have a negative impact on viral replication and alter the ratio of genomic versus subgenomic RNAs (29, 45). Furthermore, it was described that CSEs are hardly transposable between different *Alphavirus* species. The switch of either sequence or structure had a negative effect on replication (13, 45). Intriguingly, it was confirmed by biophysical techniques that TC-83 stem-loop is less stable than the WT, but this decreased stability did not affect viral RNA translation (29). Concerning the SL1 secondary fold in the capping reaction, even if none of the tested mutants abrogate m<sup>7</sup>GpppA synthesis, our data suggest that both the structure and the sequence of the 5' hairpin are important for an efficient capping process (Fig. 4). Indeed, SL1 sequence and fold disruption might reduce RNA binding up to 17 times (Table 2), leading to a 30% reduction in capping activity. This suggests that SL1 integrity is important for substrate accessibility and enzyme turnover. Moreover, RNA binding is not rate limiting under these conditions. In our experiments, TC-83 RNA has a more than 20% reduction in cap synthesis. This indicates that, in addition to the negative effect on genome replication and viral escape from IFIT 1 restriction (26), the TC-83 mutant could additionally impact viral RNA capping. This difference might create a disbalance in the ratio of capped versus uncapped VEEV RNAs, shown to be crucial not only for viral replication, translation, and immune escape but also for viral infectivity (26, 29, 46). Similarly, the importance of 5' SL structure in capping was highlighted in flaviviruses. A 74-nucleotide double-hairpin structure in the flavivirus genome is crucial for N7 cap methylation, whereas 2'-O methylation can proceed on very short RNAs (47).

In VEEV, the accumulated data and the present study tend to suggest that RNA capping is a very early event (11). Using nucleotides as the substrate for capping showed that, in contrast to BaMV, VEEV nsP1 efficiently caps ADP but not GDP (Fig. 5). GDP was shown to inhibit m<sup>7</sup>GTP synthesis in BaMV (48), which might explain the absence of cap signal in our case. It is worth mentioning that both genomic and SG RNAs of VEEV start with an A. The same seems to apply for other alphaviruses when looking at the reported or putative genomic and SG RNA sequences available in the NCBI database (49, 50). Interestingly, a non-VEEV RNA starting with a G can be capped by nsP1 (data not shown), suggesting that even if GDP is not used as the substrate for m<sup>7</sup>GpppG cap formation, the enzyme is able to cap RNAs starting with G. Another difference with the BaMV study resides in UDP capping. We observe an m<sup>7</sup>GpppU and an m<sup>7</sup>GpppC signals comigrating with m<sup>7</sup>GpppA, as described in reference 34. Recent epitranscriptomic studies revealed that a small proportion of m<sup>7</sup>GpppU- and m<sup>7</sup>GpppU<sub>m</sub>-capped RNAs do exist in eukaryotic cells (51). In the case of Zika, dengue, polio, and hepatitis C

viruses, it was shown that virus infection triggers a modification of epitranscriptomic profiles of the host cell (52). Moreover, m<sup>7</sup>GpppU- and m<sup>7</sup>GpppU<sub>m</sub>-capped RNAs can be translated, and, in general, the nature of the cap analogue (m<sup>7</sup>GpppN), present on the RNA, as well as its methylation status on the N nucleotide, regulate the translation process (53). Although the UDP capping described here is obtained in an *in vitro* setting, one can speculate that it might have a biological significance, potentially for the modification of host cellular mRNA caps in view of innate immunity escape. Finally, different strategies for cap priming have been previously described. Influenza viruses, for example, snatch 10 to 15 cellular mRNA oligomers and use them to prime viral mRNA synthesis (54). In picornaviruses, priming is initiated by a protein-primed mechanism, whereby a small viral protein (VPg) is covalently linked to a UMP via the hydroxyl group of a conserved tyrosine residue. This UMP becomes the first nucleotide of the nascent RNA (55). In L-A yeast virus, it was shown that the cap analogue m<sup>7</sup>GpppG could initiate transcription (56). The latter supports a possible cap analogue synthesis used to prime RNA synthesis, in accordance with our data, which would constitute a novel pathway to achieve cap-protected synthesis of RNA transcripts.

In conclusion, the present study provides insights into VEEV nsP1 substrate preferences for cap structure synthesis. We show that VEEV nsP1 is able to cap RNAs from 1 to at least 130 nucleotides in length and that both the sequence and the structure of the 5' SL1 are important for efficient binding to nsP1 and, therefore, capping efficiency.

## MATERIALS AND METHODS

**Expression and purification of recombinant VEEV nsP1 protein.** The cDNA encoding the VEEV nsP1 (strain P676, amino acids 1 to 535) was codon optimized and cloned into the pET28b (Novagen) vector in fusion with a hexa-histidine coding sequence at its 3' end. The resulting recombinant protein was produced in T7 Express *E. coli* cells (New England Biolabs) after induction of the expression with 0.5 mM isopropyl-β-D-thiogalactopyranoside (Sigma-Aldrich) for 3 h at 17°C as described in references 11 and 12. Briefly, bacteria pellets were lysed by repeated sonication cycles in lysis buffer (20 mM Tris-HCl, pH 7.5, 300 mM NaCl, 5% glycerol, and 5 mM β-mercaptoethanol) supplemented with 20 μg/ml DNase I, 0.25 mg/ml lysozyme, and a tablet of EDTA-free protease inhibitor cocktail (Sigma) per 50 ml of lysate. The soluble material was recovered by centrifugation at 30,000 × *g* for 30 min at 4°C and then subjected to immobilized metal affinity chromatography (IMAC) on a 5-ml HisTrap column (GE Healthcare). Following extensive washings with the lysis buffer supplemented with 40 mM imidazole and 1 M NaCl, the protein was eluted with lysis buffer supplemented with 250 mM imidazole and dialyzed against storage buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, and 50% glycerol) for storage at –20°C. For RNA binding assays, VEEV nsP1 was dialyzed in the same storage buffer with 10% of glycerol and stored at –80°C until use.

***In vitro* RNA transcription.** DNA oligonucleotides were purchased from Eurofins Scientific. Short RNAs 30 nucleotides long were generated from templates of annealed single-stranded DNA (ssDNA) oligonucleotides containing the sequence of T7 ϕ2.5 promoter followed by the 5' sequence of VEEV RNA. Long RNAs >30 nucleotides long were generated using cDNA templates produced by PCR from VEEV strain P676. The forward primers were extended with the T7 ϕ2.5 promoter sequence at the 5' end. The transcription reaction was performed in 40 mM Tris-HCl (pH 7.0), 40 mM MgCl<sub>2</sub>, 2 mM spermidine, 0.01% Triton X-100, and 4% polyethylene glycol (PEG) 8000. RNA synthesis was carried out at 37°C for 4 h in the presence of 8 mM nucleoside triphosphates (NTPs; GE Healthcare), T7 RNA polymerase (0.1 μM), and RNase inhibitor (Ambion). RNA was purified by phenol-chloroform extraction technique and then precipitated with ethanol supplemented with 2.5 M ammonium acetate overnight at 4°C. Long RNAs of >30 nucleotides were further purified on agarose gel after electrophoresis using Nucleospin gel extraction kit-NTC buffer (Macherey-Nagel), as described by manufacturer's protocol. RNA purity was assessed by denaturing acrylamide gel (urea PAGE).

**RNA GTase assay (formation of m<sup>7</sup>GpppRNA).** We pretreated 5'-triphosphate RNAs, generated by *in vitro* transcription, with 1 μM dengue virus (DENV) nonstructural protein 3 (NS3), produced and purified as described elsewhere (31). The γ-phosphate removal was done in 50 mM HEPES (pH 7.5) and 2 mM dithiothreitol (DTT) for 30 min. The enzyme was then heat-inactivated at 65°C for 5 min. GTase reaction was performed with 2 μM nsP1 VEEV in the presence of 10 μCi of [α-P<sup>32</sup>]GTP (3,000 Ci/mmol), 50 mM HEPES (pH 7.5), 10 mM KCl, 2 mM MgCl<sub>2</sub>, 2 mM DTT, and 100 μM SAM for 2 h at 30°C and then stored at –20°C. Capped RNAs were subjected to nuclease P1 (Sigma) digestion in 30 mM sodium acetate (pH 5.3), 5 mM ZnCl<sub>2</sub>, and 50 mM NaCl (2 h, 37°C), followed by proteinase K (NEB) hydrolysis (30 min, 37°C). Digested products were resolved by polyethylenimine cellulose thin-layer chromatography (TLC) (Macherey-Nagel) using 0.45 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> or 1 M LiCl as mobile phase. TLC plates were visualized using Amersham Typhoon phosphor imager (12). Inorganic pyrophosphatase and adenosylhomocysteine-hydrolase (SAH-hydrolase) were purchased from Sigma-Aldrich. Cap-Clip acid pyrophosphatase enzyme was purchased from Cellscript. All commercial enzymes were used as indicated in manufacturers' protocols. AtAPY1-ΔTM GDPase was produced and used as described in Massalski et al. (57). Quantification of m<sup>7</sup>GpppA signal was carried out using the ImageJ program. Statistical analyses were carried out using two-tailed unpaired Student's *t* tests (\*, *P* < 0.05).

**RNA binding assay. (i) RNA labeling.** We incorporated 12.5  $\mu$ M cyanine 5-cytidine-5-phosphate-3-(6-aminoethyl) phosphate (pCpCy5) dye (Jena Bioscience) at the 5' end of *in vitro*-transcribed VEEV RNAs by ligation in the presence of 1 mM ATP and 1 U of T4 ligase 1 (NEB) for 1 h at 37°C. The labeled RNA was purified by ethanol precipitation in the presence of 0.3 M sodium acetate, 1  $\mu$ g/ $\mu$ l glycogen, and 2 volumes of cold ethanol (100%) at  $-80^{\circ}\text{C}$  for 1 h. After the incubation time, the Cy5-labeled RNA was centrifuged for 15 min at 10,000 rpm and washed with ethanol (70%) to eliminate the excess dye.

**(ii) Fluorescence polarization assay.** The binding between the VEEV nsP1 and each RNA was monitored by FP. Each Cy5-labeled RNA strand was mixed with increased concentrations of nsP1 (0.09 to 20  $\mu$ M) and protein storage buffer in a final volume of 20  $\mu$ l. Assays were performed in 384-well opaque microplates (Greiner Bio-One). Fluorescence polarization was measured using PHERAstar FS microplate reader (BMG Labtech) with excitation and emission wavelengths of 590 and 675 nm, respectively. All assays were repeated three times and carried out in triplicate, and a blank control without protein was included. The dissociation constants ( $K_d$ ) were calculated using GraphPad Prism 7 program.

## ACKNOWLEDGMENTS

We thank Iris Steinebrunner for the generous gift of AtAPY1- $\Delta$ TM GDPase.

O.O.G. is a recipient of a Méditerranée Infection Fondation doctoral studentship.

We declare no conflict of interest.

**Author contributions:** Conceptualization, B.Co. and N.R.; Methodology, O.O.G., C.V., A.S., and N.R.; Software, O.O.G. and N.R.; Investigation, O.O.G., B.Co., and N.R.; Resources, B.Co. and B.Ca.; Writing – original draft preparation, O.O.G. and N.R.; Writing – review and editing, O.O.G., C.V., A.S., E.D., B.Ca., B.Co., and N.R.; Project administration, N.R.; Funding acquisition, E.D., B.Ca., B.Co., and N.R. All authors have read and agreed to the published version of the manuscript.

## REFERENCES

1. Ronca SE, Dineley KT, Paessler S. 2016. Neurological sequelae resulting from encephalitic alphavirus infection. *Front Microbiol* 7:959. <https://doi.org/10.3389/fmicb.2016.00959>.
2. Sharma A, Knollmann-Ritschel B. 2019. Current understanding of the molecular basis of Venezuelan equine encephalitis virus pathogenesis and vaccine development. *Viruses* 11:164. <https://doi.org/10.3390/v11020164>.
3. Snyder JE, Kulcsar KA, Schultz KWL, Riley CP, Neary JT, Marr S, Jose J, Griffin DE, Kuhn RJ. 2013. Functional characterization of the alphavirus TF protein. *J Virol* 87:8511–8523. <https://doi.org/10.1128/JVI.00449-13>.
4. Rupp JC, Sokoloski KJ, Gebhart NN, Hardy RW. 2015. Alphavirus RNA synthesis and non-structural protein functions. *J Gen Virol* 96:2483–2500. <https://doi.org/10.1099/jgv.0.000249>.
5. Jones R, Bragagnolo G, Arranz R, Reguera J. 2021. Capping pores of alphavirus nsP1 gate membranous viral replication factories. *Nature* 589:615–619. <https://doi.org/10.1038/s41586-020-3036-8>.
6. Cross RK. 1983. Identification of a unique guanine-7-methyltransferase in Semliki Forest virus (SFV) infected cell extracts. *Virology* 130:452–463. [https://doi.org/10.1016/0042-6822\(83\)90099-5](https://doi.org/10.1016/0042-6822(83)90099-5).
7. Ahola T, Laakkonen P, Vihinen H, Kääriäinen L. 1997. Critical residues of Semliki Forest virus RNA capping enzyme involved in methyltransferase and guanlyltransferase-like activities. *J Virol* 71:392–397. <https://doi.org/10.1128/JVI.71.1.392-397.1997>.
8. Rikonen M, Peränen J, Kääriäinen L. 1994. ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2. *J Virol* 68:5804–5810. <https://doi.org/10.1128/JVI.68.9.5804-5810.1994>.
9. Vasiljeva L, Merits A, Auvinen P, Kääriäinen L. 2000. Identification of a novel function of the alphavirus capping apparatus. RNA 5'-triphosphatase activity of Nsp2. *J Biol Chem* 275:17281–17287. <https://doi.org/10.1074/jbc.M910340199>.
10. Karpe YA, Aher PP, Lole KS. 2011. NTPase and 5'-RNA triphosphatase activities of Chikungunya virus nsP2 protein. *PLoS One* 6:e22336. <https://doi.org/10.1371/journal.pone.0022336>.
11. Li C, Guillén J, Rabah N, Blanjoie A, Debart F, Vasseur J-J, Canard B, Decroly E, Coutard B. 2015. mRNA capping by Venezuelan equine encephalitis virus nsP1: functional characterization and implications for antiviral research. *J Virol* 89:8292–8303. <https://doi.org/10.1128/JVI.00599-15>.
12. Rabah N, Ortega Granda O, Quérat G, Canard B, Decroly E, Coutard B. 2020. Mutations on VEEV nsP1 relate RNA capping efficiency to ribavirin susceptibility. *Antiviral Res* 182:104883. <https://doi.org/10.1016/j.antiviral.2020.104883>.
13. Gorchakov R, Hardy R, Rice CM, Frolov I. 2004. Selection of functional 5' cis-acting elements promoting efficient Sindbis virus genome replication. *J Virol* 78:61–75. <https://doi.org/10.1128/jvi.78.1.61-75.2004>.
14. Kutchko KM, Madden EA, Morrison C, Plante KS, Sanders W, Vincent HA, Cruz Cisneros MC, Long KM, Moorman NJ, Heise MT, Laederach A. 2018. Structural divergence creates new functional features in alphavirus genomes. *Nucleic Acids Res* 46:3657–3670. <https://doi.org/10.1093/nar/gky012>.
15. Niesters HG, Strauss JH. 1990. Mutagenesis of the conserved 51-nucleotide region of Sindbis virus. *J Virol* 64:1639–1647. <https://doi.org/10.1128/JVI.64.4.1639-1647.1990>.
16. Niesters HG, Strauss JH. 1990. Defined mutations in the 5' nontranslated sequence of Sindbis virus RNA. *J Virol* 64:4162–4168. <https://doi.org/10.1128/JVI.64.9.4162-4168.1990>.
17. Kuhn RJ, Hong Z, Strauss JH. 1990. Mutagenesis of the 3' nontranslated region of Sindbis virus RNA. *J Virol* 64:1465–1476. <https://doi.org/10.1128/JVI.64.4.1465-1476.1990>.
18. Fayzulin R, Frolov I. 2004. Changes of the secondary structure of the 5' end of the Sindbis virus genome inhibit virus growth in mosquito cells and lead to accumulation of adaptive mutations. *J Virol* 78:4953–4964. <https://doi.org/10.1128/jvi.78.10.4953-4964.2004>.
19. Levis R, Schlesinger S, Huang HV. 1990. Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. *J Virol* 64:1726–1733. <https://doi.org/10.1128/JVI.64.4.1726-1733.1990>.
20. Wielgosz MM, Raju R, Huang HV. 2001. Sequence requirements for Sindbis virus subgenomic mRNA promoter function in cultured cells. *J Virol* 75:3509–3519. <https://doi.org/10.1128/JVI.75.8.3509-3519.2001>.
21. Nickens DG, Hardy RW. 2008. Structural and functional analyses of stem-loop 1 of the Sindbis virus genome. *Virology* 370:158–172. <https://doi.org/10.1016/j.virol.2007.08.006>.
22. Frolov I, Schlesinger S. 1996. Translation of Sindbis virus mRNA: analysis of sequences downstream of the initiating AUG codon that enhance translation. *J Virol* 70:1182–1190. <https://doi.org/10.1128/JVI.70.2.1182-1190.1996>.
23. Frolova E, Frolov I, Schlesinger S. 1997. Packaging signals in alphaviruses. *J Virol* 71:248–258. <https://doi.org/10.1128/JVI.71.1.248-258.1997>.
24. Volkova E, Gorchakov R, Frolov I. 2006. The efficient packaging of Venezuelan equine encephalitis virus-specific RNAs into viral particles is determined by nsP1-3 synthesis. *Virology* 344:315–327. <https://doi.org/10.1016/j.virol.2005.09.010>.
25. Kim DY, Firth AE, Atasheva S, Frolova EI, Frolov I. 2011. Conservation of a packaging signal and the viral genome RNA packaging mechanism in

- alphavirus evolution. *J Virol* 85:8022–8036. <https://doi.org/10.1128/JVI.00644-11>.
26. Hyde JL, Gardner CL, Kimura T, White JP, Liu G, Trobaugh DW, Huang C, Tonelli M, Paessler S, Takeda K, Klimstra WB, Amarasinghe GK, Diamond MS. 2014. A viral RNA structural element alters host recognition of nonself RNA. *Science* 343:783–787. <https://doi.org/10.1126/science.1248465>.
  27. Johnson BJ, Kinney RM, Kost CL, Trent DW. 1986. Molecular determinants of alphavirus neurovirulence: nucleotide and deduced protein sequence changes during attenuation of Venezuelan equine encephalitis virus. *J Gen Virol* 67:1951–1960. <https://doi.org/10.1099/0022-1317-67-9-1951>.
  28. Kinney RM, Johnson BJ, Welch JB, Tsuchiya KR, Trent DW. 1989. The full-length nucleotide sequences of the virulent Trinidad donkey strain of Venezuelan equine encephalitis virus and its attenuated vaccine derivative, strain TC-83. *Virology* 170:19–30. [https://doi.org/10.1016/0042-6822\(89\)90347-4](https://doi.org/10.1016/0042-6822(89)90347-4).
  29. Kulasegaran-Shylini R, Thivyanathan V, Gorenstein DG, Frolov I. 2009. The 5'UTR-specific mutation in VEEV TC-83 genome has a strong effect on RNA replication and subgenomic RNA synthesis, but not on translation of the encoded proteins. *Virology* 387:211–221. <https://doi.org/10.1016/j.virol.2009.02.027>.
  30. Frolov I, Hardy R, Rice CM. 2001. Cis-acting RNA elements at the 5' end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. *RNA* 7:1638–1651. <https://doi.org/10.1017/S135583820101010X>.
  31. Milhas S, Raux B, Betzi S, Derviaux C, Roche P, Restouin A, Basse M-J, Rebuffet E, Lugari A, Badol M, Kashyap R, Lissitzky J-C, Eydoux C, Hamon V, Gourdell M-E, Combes S, Zimmermann P, Aurrand-Lions M, Roux T, Rogers C, Müller S, Knapp S, Trinquet E, Collette Y, Guillemot J-C, Morelli X. 2016. Protein-protein interaction inhibition (2P2I)-oriented chemical library accelerates hit discovery. *ACS Chem Biol* 11:2140–2148. <https://doi.org/10.1021/acschembio.6b00286>.
  32. Ahola T, Ahlquist P. 1999. Putative RNA capping activities encoded by brome mosaic virus: methylation and covalent binding of guanylate by replicase protein 1a. *J Virol* 73:10061–10069. <https://doi.org/10.1128/JVI.73.12.10061-10069.1999>.
  33. Huang Y-L, Hsu Y-H, Han Y-T, Meng M. 2005. mRNA guanylation catalyzed by the S-adenosylmethionine-dependent guanylyltransferase of Bamboo mosaic virus. *J Biol Chem* 280:13153–13162. <https://doi.org/10.1074/jbc.M412619200>.
  34. Haugland RA, Cline MG. 1980. Post-transcriptional modifications of oat coleoptile ribonucleic acids. 5'-Terminal capping and methylation of internal nucleosides in poly(A)-rich RNA. *Eur J Biochem* 104:271–277. <https://doi.org/10.1111/j.1432-1033.1980.tb04425.x>.
  35. Mizumoto K, Kaziro Y, Lipmann F. 1982. Reaction mechanism of mRNA guanylyltransferase from rat liver: isolation and characterization of a guanylyl-enzyme intermediate. *Proc Natl Acad Sci U S A* 79:1693–1697. <https://doi.org/10.1073/pnas.79.6.1693>.
  36. Huang Y-L, Han Y-T, Chang Y-T, Hsu Y-H, Meng M. 2004. Critical residues for GTP methylation and formation of the covalent m7GMP-enzyme intermediate in the capping enzyme domain of Bamboo mosaic virus. *J Virol* 78:1271–1280. <https://doi.org/10.1128/jvi.78.3.1271-1280.2004>.
  37. Yu L, Martins A, Deng L, Shuman S. 1997. Structure-function analysis of the triphosphatase component of vaccinia virus mRNA capping enzyme. *J Virol* 71:9837–9843. <https://doi.org/10.1128/JVI.71.12.9837-9843.1997>.
  38. Ogino T, Banerjee AK. 2007. Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Mol Cell* 25:85–97. <https://doi.org/10.1016/j.molcel.2006.11.013>.
  39. Ogino T, Banerjee AK. 2008. Formation of guanosine(5')tetraphospho(5') adenosine cap structure by an unconventional mRNA capping enzyme of vesicular stomatitis virus. *J Virol* 82:7729–7734. <https://doi.org/10.1128/JVI.00326-08>.
  40. Cross RK, Gomas PJ. 1981. Concomitant methylation and synthesis in vitro of Semliki Forest virus (SFV) ss RNAs by a fraction from infected cells. *Virology* 114:542–554. [https://doi.org/10.1016/0042-6822\(81\)90234-8](https://doi.org/10.1016/0042-6822(81)90234-8).
  41. Chiu Y-L, Ho CK, Saha N, Schwer B, Shuman S, Rana TM. 2002. Tat stimulates cotranscriptional capping of HIV mRNA. *Mol Cell* 10:585–597. [https://doi.org/10.1016/s1097-2765\(02\)00630-5](https://doi.org/10.1016/s1097-2765(02)00630-5).
  42. Tekes G, Rahmeh AA, Whelan SPJ. 2011. A freeze frame view of vesicular stomatitis virus transcription defines a minimal length of RNA for 5' processing. *PLoS Pathog* 7:e1002073. <https://doi.org/10.1371/journal.ppat.1002073>.
  43. Sokoloski KJ, Haist KC, Morrison TE, Mukhopadhyay S, Hardy RW. 2015. Noncapped alphavirus genomic RNAs and their role during infection. *J Virol* 89:6080–6092. <https://doi.org/10.1128/JVI.00553-15>.
  44. Ng CS, Kasumba DM, Fujita T, Luo H. 2020. Spatio-temporal characterization of the antiviral activity of the XRN1-DCP1/2 aggregation against cytoplasmic RNA viruses to prevent cell death. *Cell Death Differ* 27:2363–2382. <https://doi.org/10.1038/s41418-020-0509-0>.
  45. Kulasegaran-Shylini R, Atasheva S, Gorenstein DG, Frolov I. 2009. Structural and functional elements of the promoter encoded by the 5' untranslated region of the Venezuelan equine encephalitis virus genome. *J Virol* 83:8327–8339. <https://doi.org/10.1128/JVI.00586-09>.
  46. LaPointe AT, Moreno-Contreras J, Sokoloski KJ. 2018. Increasing the capping efficiency of the Sindbis virus nsP1 protein negatively affects viral infection. *mBio* 9:e02342-18. <https://doi.org/10.1128/mBio.02342-18>.
  47. Dong H, Ray D, Ren S, Zhang B, Puig-Basagot F, Takagi Y, Ho CK, Li H, Shi P-Y. 2007. Distinct RNA elements confer specificity to flavivirus RNA cap methylation events. *J Virol* 81:4412–4421. <https://doi.org/10.1128/JVI.02455-06>.
  48. Hu R-H, Lin M-C, Hsu Y-H, Meng M. 2011. Mutational effects of the consensus aromatic residues in the mRNA capping domain of Bamboo mosaic virus on GTP methylation and virus accumulation. *Virology* 411:15–24. <https://doi.org/10.1016/j.virol.2010.12.022>.
  49. Dubin DT, Stollar V, Hsueh CC, Timko K, Guild GM. 1977. Sindbis virus messenger RNA: the 5'-termini and methylated residues of 26 and 42 S RNA. *Virology* 77:457–470. [https://doi.org/10.1016/0042-6822\(77\)90471-8](https://doi.org/10.1016/0042-6822(77)90471-8).
  50. Pettersson RF, Söderlund H, Kääriäinen L. 1980. The nucleotide sequences of the 5'-terminal T1 oligonucleotides of Semliki-Forest-virus 42-S and 26-S RNAs are different. *Eur J Biochem* 105:435–443. <https://doi.org/10.1111/j.1432-1033.1980.tb04518.x>.
  51. Wang J, Alvin Chew BL, Lai Y, Dong H, Xu L, Balamkundu S, Cai WM, Cui L, Liu CF, Fu X-Y, Lin Z, Shi P-Y, Lu TK, Luo D, Jaffrey SR, Dedon PC. 2019. Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA. *Nucleic Acids Res* 47:e130. <https://doi.org/10.1093/nar/gkz751>.
  52. McIntyre W, Netzband R, Bonenfant G, Biegel JM, Miller C, Fuchs G, Henderson E, Arra M, Canki M, Fabris D, Pager CT. 2018. Positive-sense RNA viruses reveal the complexity and dynamics of the cellular and viral epitranscriptomes during infection. *Nucleic Acids Res* 46:5776–5791. <https://doi.org/10.1093/nar/gky029>.
  53. Sikorski PJ, Warminski M, Kubacka D, Ratajczak T, Nowis D, Kowalska J, Jemielity J. 2020. The identity and methylation status of the first transcribed nucleotide in eukaryotic mRNA 5' cap modulates protein expression in living cells. *Nucleic Acids Res* 48:1607–1626. <https://doi.org/10.1093/nar/gkaa032>.
  54. De Vlucht C, Sikora D, Pelchat M. 2018. Insight into influenza: a virus cap-snatching. *Viruses* 10:641. <https://doi.org/10.3390/v10110641>.
  55. Paul AV, Wimmer E. 2015. Initiation of protein-primed picornavirus RNA synthesis. *Virus Res* 206:12–26. <https://doi.org/10.1016/j.virusres.2014.12.028>.
  56. Fujimura T, Esteban R. 2010. Yeast double-stranded RNA virus L-A deliberately synthesizes RNA transcripts with 5'-diphosphate. *J Biol Chem* 285:22911–22918. <https://doi.org/10.1074/jbc.M110.138982>.
  57. Massalski C, Bloch J, Zebisch M, Steinebrunner I. 2015. The biochemical properties of the arabidopsis ecto-nucleoside triphosphate diphosphohydrolase AtAPY1 contradict a direct role in purinergic signaling. *PLoS One* 10:e0115832. <https://doi.org/10.1371/journal.pone.0115832>.

### 3.7 Article 3

#### ***Mutations on VEEV nsP1 relate RNA capping efficiency to ribavirin susceptibility***

Nadia Rabah <sup>A,C,\*</sup>, **Oney Ortega Granda** <sup>A</sup>, Gilles Quérat <sup>B</sup>, Bruno Canard <sup>A</sup>, Etienne Decroly <sup>A</sup>, Bruno Coutard <sup>B,\*</sup>.

<sup>A</sup> Aix Marseille Université, CNRS, AFMB UMR 7257, Marseille, France

<sup>B</sup> Unité des Virus Emergents (UVE: Aix-Marseille Univ-IRD 190-Inserm, 1207-IHU Méditerranée Infection) Marseille, France

<sup>C</sup> Université de Toulon, 83130 La Garde, France

\*Corresponding authors Nadia Rabah and Bruno Coutard

E-mail: bruno.coutard@univ-amu.fr; nadia.rabah@univ-tln.fr

Tel: +33 4 13 73 21 62; +33 4 91 82 55 46

**Antiviral Research 182 (2020) 104883**





## Mutations on VEEV nsP1 relate RNA capping efficiency to ribavirin susceptibility

Nadia Rabah<sup>a,c,\*\*</sup>, Oney Ortega Granda<sup>a</sup>, Gilles Quérat<sup>b</sup>, Bruno Canard<sup>a</sup>, Etienne Decroly<sup>a</sup>, Bruno Coutard<sup>b,\*</sup>

<sup>a</sup> Aix Marseille Université, CNRS, AFMB, UMR, 7257, Marseille, France

<sup>b</sup> Unité des Virus Emergents (UVE: Aix-Marseille Univ-IRD 190-Inserm, 1207-IHU Méditerranée Infection), Marseille, France

<sup>c</sup> Université de Toulon, 83130, La Garde, France

### ABSTRACT

Alphaviruses are arthropod-borne viruses of public health concern. To date no efficient vaccine nor antivirals are available for safe human use. During viral replication the nonstructural protein 1 (nsP1) catalyzes capping of genomic and subgenomic RNAs. The capping reaction is unique to the *Alphavirus* genus. The whole three-step process follows a particular order: (i) transfer of a methyl group from S-adenosyl methionine (SAM) onto a GTP forming <sup>m7</sup>GTP; (ii) guanylation of the enzyme to form a <sup>m7</sup>GMP-nsP1 adduct; (iii) transfer of <sup>m7</sup>GMP onto 5'-diphosphate RNA to yield capped RNA. Specificities of these reactions designate nsP1 as a promising target for antiviral drug development. In the current study we performed a mutational analysis on two nsP1 positions associated with Sindbis virus (SINV) ribavirin resistance in the Venezuelan equine encephalitis virus (VEEV) context through reverse genetics correlated to enzyme assays using purified recombinant VEEV nsP1 proteins. The results demonstrate that the targeted positions are strongly associated to the regulation of the capping reaction by increasing the affinity between GTP and nsP1. Data also show that in VEEV the S21A substitution, naturally occurring in Chikungunya virus (CHIKV), is a hallmark of ribavirin susceptibility. These findings uncover the specific mechanistic contributions of these residues to nsP1-mediated methyl-transfer and guanylation reactions.

### 1. Introduction

Global warming, increased worldwide travel and urbanization, among other causes, promote a significant geographical expansion of many arboviruses vectors, including *Aedes*, *Psorophora* and *Culex* mosquitoes. These arthropods act as vectors for a plethora of causal agents of infectious diseases and thus the emergence and re-emergence of a variety of Alphaviruses is currently observed. Alphaviruses infect diverse hosts including mammals, birds, rodents and salmonids. The genus *Alphavirus* contains 31 species among which 21 were reported to induce pathological symptoms in humans (Weaver et al., 2012). Alphaviruses pathogenic for humans can roughly be divided into two main groups according to the clinical symptoms they trigger. Arthralgic symptoms include fever, cutaneous rash, polyarthralgia and arthritis. They are mainly caused by alphaviruses of the Old World (OW), including Chikungunya virus (CHIKV), O'nyong-nyong virus (ONNV) and Semliki Forest Virus (SFV), with the exception of Mayaro virus (MAYV) which circulates in the Americas. Conversely, encephalitic symptoms are characteristic of viruses restricted to the New World (NW), such as Eastern, Western and Venezuelan equine encephalitis

viruses (EEEV, WEEV and VEEV). They comprise headache, nausea, anorexia and ultimately encephalitis with more than 15% of fatal outcomes. Long-lasting symptoms and the morbidity associated to alphavirus outbreaks make them a serious public health and epizootic threat (Lwande et al., 2015; Zacks and Paessler, 2010). So far, no efficient vaccine nor licenced antivirals are available for safe human use (Abdelnabi et al., 2017). Moreover, the absence of a specific treatment for alphavirus infection limits the therapy, in the case of outbreaks, to the use of medication based mainly on symptoms relieve, such as analgesic, anti-pyretic, anti-inflammatory and immunosuppressing drugs (Sales et al., 2018).

Alphaviruses are enveloped viruses possessing a single positive-stranded genomic RNA. This RNA is first translated by the host cell machinery into polyproteins P123 and P1234 that are subsequently processed by the viral protease yielding proteolytic intermediates and the four non-structural proteins (nsPs) nsP1 to 4. These proteolytic intermediates and the matured nsPs constitute the replication/transcription complexes (RTC) organised in spherules at the plasma membrane. RTC drive the replication of the viral genome but also the transcription and the capping of genomic and subgenomic RNA coding for the

\* Corresponding author.

\*\* Corresponding author. Aix Marseille Université, CNRS, AFMB, UMR, 7257, Marseille, France.

E-mail addresses: [nadia.rabah@univ-tln.fr](mailto:nadia.rabah@univ-tln.fr) (N. Rabah), [bruno.coutard@univ-amu.fr](mailto:bruno.coutard@univ-amu.fr) (B. Coutard).

<https://doi.org/10.1016/j.antiviral.2020.104883>

Received 10 October 2019; Received in revised form 30 June 2020; Accepted 14 July 2020

Available online 1 August 2020

0166-3542/© 2020 Published by Elsevier B.V.

structural proteins. NsP1 is the viral RNA capping enzyme, carrying methyl- and guanylyltransferase activities necessary for the viral replication (Ahola and Kaariainen, 1995) (Ahola et al., 1997) (Li et al., 2015). NsP2 possesses NTPase, RNA triphosphatase and RNA helicase activities (Rikkonen et al., 1994) (Law et al., 2019; Vasiljeva et al., 2000). The C-terminal region of nsP2 has a papain-like protease fold and is implicated in the P1234 processing (Hardy and Strauss, 1989; Shirako and Strauss, 1994). NsP3 contains a Macro domain and a C-terminal hypervariable region. The protein nsP4 corresponds to the viral RNA dependent RNA polymerase (RdRp) responsible for RNA synthesis (Chen et al., 2017).

The unique alphaviral replicative process is attractive for the development of potent and specific inhibitors. Several inhibitor screening efforts showed that an antiviral effect can be achieved through targeting a specific nsP. Purine nucleoside analogues or urea-derivatives can inhibit RNA synthesis in SFV, SINV and CHIKV (Albulescu et al., 2015; Delang et al., 2014; Pohjala et al., 2008) (Uraikova et al., 2017), suggesting a direct effect on the nsP4 RdRp. Similarly, formamide-based cysteine protease inhibitors seem promising in targeting CHIKV and VEEV nsP2 (Das et al., 2016; Hu et al., 2016). The list of active molecules against nsP1 was recently extended with capping inhibitors of different structural classes (Gigante et al., 2014) (Delang et al., 2016) (Feibelman et al., 2018) (Gomez-SanJuan et al., 2018) (Ferreira-Ramos et al., 2019).

In the case of alphaviruses, ribavirin administration to infected patients decreases arthralgia symptoms associated to CHIKV infection (Ravichandran and Manian, 2008). Nevertheless, infection experiments performed in cell lines have evidenced that antiviral effect of ribavirin on alphaviruses is cell-, virus species-, and even virus strain-specific. Hence, CHIKV replication can be inhibited by ribavirin and the antiviral effect is more prominent in hepatoma cells than in lung or kidney cell lines (Franco et al., 2018). In Vero cells, ribavirin has a potent antiviral effect against CHIKV, but also against SINV and Dakar strain of SFV. In contrast, both the Uganda strain of SFV and VEEV species are resistant to ribavirin (Briolant et al., 2004) (De Clercq et al., 1991; Markland et al., 2000). Interestingly, in the same cells, VEEV is sensitive to VX-497, a carbamic acid derivative, targeting IMPDH by uncompetitive inhibition, underlying the high potential of targeting IMPDH in alphavirus infection (Markland et al., 2000).

Sheidel and Stollar isolated SINV mutants resistant to three inhibitors of IMPDH, namely MPA, ribavirin and 2-amino-1,3,4-thiadiazole. The reported mutants had a 3000 fold increased viral load when challenged with the drug compared to the WT. Genomes of resistant mutants carried three non-synonymous mutations in the nsP1 coding sequence, yielding the amino acid substitutions Q21K, S23N and V302M (Scheidel et al., 1987) (Rosenblum et al., 1994). Positions 21 and 23 are in the vicinity of another resistant mutation selected against the capping inhibitors [1,2,3]triazolo[4,5-d]pyrimidin-7(6H)-ones bearing antiviral activity (Delang et al., 2016). Together, these observations suggest that the inhibition of IMPDH affects indirectly viral capping. The nsP1 promotes specific RNA capping in the following order: (i) a methyltransferase reaction (MTase), where GTP + SAM form  $m^7$ GTP and the S-adenosylhomocysteine (SAH) by-product (ii) A first guanylyltransferase reaction (GTase1), where the enzyme forms a covalent complex  $m^7$ GMP-nsP1. (iii) a second guanylyltransferase reaction (GTase2), where the  $m^7$ GMP is transferred onto a 5'-diphosphate RNA (ppRNA), to form the cap structure  $m^7$ GpppRNA (Li et al., 2015).

In this study, we focus on VEEV nsP1 Q19 and S21 residues, which correspond to the SINV ribavirin resistant mutants mentioned above. We have generated different combinations of K, N and A substitutions both in the VEEV reverse genetics system and nsP1 recombinant enzyme in order to elucidate the role of these amino acids in the viral capping and the viral susceptibility/resistance towards compounds regulating the intracellular GTP pool. Our data demonstrate that: (i) By modulating SAM and GTP binding, Q19 is a key residue for the MTase reaction; (ii) S21 is essential for  $m^7$ GMP-nsP1 complex formation during the GTase1 reaction; (iii) The double mutant Q19K-S21N, leads to a 3-fold increase

in GTase2 activity. (iv) S21A substitution, corresponding to the naturally occurring CHIKV polymorphism, is a hallmark of ribavirin susceptibility.

## 2. Materials and methods

### 2.1. Cloning, mutagenesis and expression of VEEV nsP1 proteins

The DNA sequence corresponding to the VEEV nsP1 protein (strain P676, amino acid 1 to 535) was synthesized by GenScript and cloned into the pET28b (Novagen) vector in fusion with a C-terminal hexahistidine tag after codon optimization for bacterial expression. Site-directed mutagenesis was performed by PCR amplification of the WT sequence with primers carrying the desired mutations and using PFU Turbo (Ambion) as described by the manufacturer's protocol. All constructs were confirmed by DNA Sanger sequencing. T7 Express *E. coli* (NEB) were transformed by the VEEV nsP1 expressing plasmids and grown at 37 °C in Terrific Broth medium until an OD<sub>600nm</sub> reached 0.6. Induction of protein expression was carried with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG, Sigma) for 3 h at 17 °C. Cells were harvested by centrifugation at 5000×g for 15 min. The pellets were stored at −80 °C until purification.

### 2.2. Protein purification

Bacteria pellets were thawed on ice, then lysed in 20 mM Tris-HCl pH7.5, 300 mM NaCl, 5% glycerol, 5 mM mercaptoethanol supplemented with 20 µg/ml DNase I, 0.25 mg/ml lysozyme and Complete EDTA-free Protease Inhibitor Cocktail (Sigma). After complete dissolution of the pellet, samples were sonicated and clarified by centrifugation (30 000×g for 30 min at 4 °C). Following the addition of imidazole to a final concentration of 40 mM, soluble fractions were incubated with Ni-Sepharose resin (GE Healthcare; 0.5 ml/l culture) for 1 h at 4 °C, with gentle shaking. Beads were washed 2 times with 5 column volumes (CV) of lysis buffer then 5 CV of wash buffer (20 mM Tris-HCl, pH 7.5, 1 M NaCl, 5% glycerol, 40 mM imidazole). Proteins were eluted in 20 mM Tris-HCl pH 7.5, 300 mM NaCl, 5% glycerol, 250 mM imidazole. Finally, proteins were concentrated using Amicon Ultra (EMD Millipore) ultrafiltration units and dialysed against storage buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 50% glycerol) for storage at −20 °C.

### 2.3. Methyltransferase assay

The transfer of the methyl group from S-adenosyl-[methyl-3H] methionine (SAM [ $H^3$ ]) (PerkinElmer) to GTP or guanylylimidodiphosphate (GIDP) was performed as described previously (Li et al., 2015). Briefly, 5 µM of VEEV nsP1 WT and mutant were incubated in 50 mM Tris pH 7.0, 2 mM DTT, 10 mM KCl, 2 mM MgCl<sub>2</sub>, 2 mM GTP or GIDP, 0.55 µCi (SAM [ $H^3$ ]) (0.33 µM), at 30 °C. The reaction was stopped by loading the samples on DEAE-cellulose filters (PerkinElmer). The filters were washed twice with 10 mM ammonium formate, once with H<sub>2</sub>O, and finally with 95% ethanol. The radioactivity was measured by scintillation counting with SCINT BETAPLATE solution in a MicroBeta 2 counter (PerkinElmer). The kinetics parameters were determined by varying the concentrations of either SAM, GTP or GIDP. Data were analysed using SigmaPlot software.

### 2.4. nsP1 guanylyltransferase assay (GTase1)

The formation of  $m^7$ GMP-nsP1 complex was monitored by incubating 5 µM of each VEEV-nsP1 protein with [ $\alpha$ -P<sup>32</sup>] GTP (3000 Ci/mmol), 100 µM SAM, 10 mM NaCl, 2 mM MgCl<sub>2</sub> and 2 mM DTT in 20 mM HEPES buffer pH 7, at 30 °C for 1 h. The complex was then resolved by electrophoresis on a 12% SDS-PAGE. The radiolabeled material was visualized using Amersham Typhoon phosphor-imager.

## 2.5. *In vitro* transcription

DNA oligonucleotide corresponding to the first 30 nucleotides of VEEV genomic RNA preceded by the phi 2.5 class II promoter was used as a template in an *in vitro* RNA reaction synthesis. Transcription buffer contained 40 mM Tris-HCl (pH 7.0), 40 mM MgCl<sub>2</sub> and 2 mM Spermidine, 0.01% Triton, 4% PEG 8000. The reaction was conducted at 37 °C for 4 h in the presence of 8 mM of NTPs (GE Healthcare), T7 RNA polymerase (0.1 μM) and RNase inhibitor (Ambion). The RNA solution was centrifuged 15 min at 3000×g, treated with 15 min at 37 °C with DNase (Ambion), then subjected to a phenol/chloroform extraction. RNA was subsequently precipitated with ethanol supplemented with 2.5 M of Ammonium acetate O/N at 4 °C. The purified RNA is hereafter named VEEV RNA.

## 2.6. RNA guanylyltransferase assay (GTase2: formation of m<sup>7</sup>GpppRNA)

In order to remove the γ-phosphate of T7 expressed RNA, 5 μM of VEEV RNA was preincubated with 1 μM of Dengue NS3, produced and purified as described previously (Milhas et al., 2016) in 50 mM HEPES (pH 7.5) and 2 mM DTT for 30 min. The reaction was stopped by heat inactivation at 65 °C for 5 min. The generated diphosphate 5' termini VEEV RNA was then incubated with 2 μM of nsP1 VEEV proteins in a reaction mixture containing 10 μCi of [α-P<sup>32</sup>] GTP (3000 Ci/mmol), 50 mM HEPES (pH 7.5), 10 mM KCl, 2 mM MgCl<sub>2</sub>, 2 mM DTT and 100 μM SAM for 2 h at 30 °C. After one freezing/thawing cycle the capped RNAs were digested with 1 U of nuclease P1 (Sigma) in 30 mM sodium acetate (pH 5.3), 5 mM ZnCl<sub>2</sub> and 50 mM NaCl (2 h, 37 °C). At the end of the reaction, proteins were hydrolysed with 1 U of proteinase K (NEB). Digestion products were separated by polyethylenimine cellulose thin-layer chromatography (Macherey-Nagel) and resolved using 0.45 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> as mobile phase. The radiolabeled material was visualized as described above.

## 2.7. Generation of a recombinant VEE virus

Wild type and mutants VEEV were generated by co-transfection in Vero E6 cells of overlapping synthetic (GenScript) molecular clones covering the whole genome using the Infectious Subgenomic Infectious Amplicons (ISA) method as previously described (Aubry et al., 2015). Briefly, Three VEEV fragments were generated. The first 5' end fragment covers the CMV early promoter linked to the first 80 nucleotides of the VEEV genome (strain P676). The second fragment encompasses either WT or mutated nsP1 from nucleotide position 1 to position 1772 of the VEEV genome. The third fragment encompasses Nsp2 to Nsp4 from nucleotide position 1683 to 7583 and the 3' end fragment encompasses the structural proteins and the 3' UTR (7503 to genome end) linked to a poly A tail, a ribozyme site and the SV40 poly A signal. All fragments were linearized by amplification from the GenScript-delivered plasmid using Taq polymerase and directly purified from the amplification reaction. The day before transfection Vero E6 cells were seeded in 96 wells plates at 50000 cells per well. Transfection was carried using Lipofectamine 3000 reagent (Life Technologies) and 100 ng of total DNA per well. Supernatants were collected on day 6 post transfection. Viral genomes were quantified by quantitative RT-PCR of viral RNA extracted from the supernatant in presence of a DNase treatment. Subsequently, viral stocks were grown on Vero cells.

## 2.8. Antiviral assay

Antiviral assays were performed using recombinant VEEV WT, S21A (see above) and Chikungunya virus strain Opy1 (La Réunion Island LR2006\_OPY1; EVAg 001v-EVA83). The amount of each virus and the duration of the assay had initially been calibrated so that the replication is still in the log phase of growth at the time of readout and the cycle threshold (CT) standard deviations of qRT-PCR quantification

(quadruplicate) is below 0.5. Approximate multiplicity of infection (MOI) range from 10<sup>-4</sup> to 10<sup>-3</sup> depending on the strain.

One day prior to infection 5 × 10<sup>4</sup> Vero E6 cells were seeded in 100 μl of medium (supplemented with 2.5% FCS) in each well of a 96-well titer plate. The next day, 8 two-fold serial dilutions of the compounds (beginning at 400 μM final concentration, down to 0.16 μM), in duplicates or triplicates, were added to the cells (25 μl/well, in 2.5% FCS containing medium). Four Virus Control (VC) wells (per virus) were supplemented with 25 μl medium. Fifteen minutes later, 25 μl of a virus mix containing the appropriate amount of viral stock diluted in medium (2.5% FCS) were added to the 96-well plates.

Cells were cultivated for 36–48 h after which 100 μl of the supernatant were collected for viral RNA purification. The supernatants were transferred to 96 well S-Bloc from QIAgen preloaded with VXL mix and extract by the Cador Pathogen 96 QIAcube HT kit run on QIAcube HT automat according to QiaGen protocol. Purified RNAs were eluted in 80 μl of water. Viral RNAs were then quantified by real time one step RT-PCR to determine viral RNA yield using 3.5 μl of RNA and 6.5 μl of RT-PCR mix using standard cycling parameters. The four control wells were replaced by four 2 log dilutions of an appropriate T7-generated RNA standards of known quantities for each viral genome (100 copies to 100 million copies).

Mean inhibition of virus yield is equal to 100 X (mean quantities of viral RNA in VC quadruplicates - mean quantities of viral RNA in drug treated triplicates)/mean quantities of viral RNA in VC. The inhibition values (expressed as percent inhibition, in linear scale) obtained for each drug concentration (expressed in μM, in log scale) are plotted using Kaleidagraph plotting software (Synergy Software) and the best sigmoidal curve, fitting the mean values, is determined by a macro in the software: (Inhibition, Y is given by  $Y = 100 / (1 + (m0/m1)^{m2})$ ). This macro allows to determine the best curve fit and the m1 and m2 parameters, where m1 corresponds to EC<sub>50</sub>.

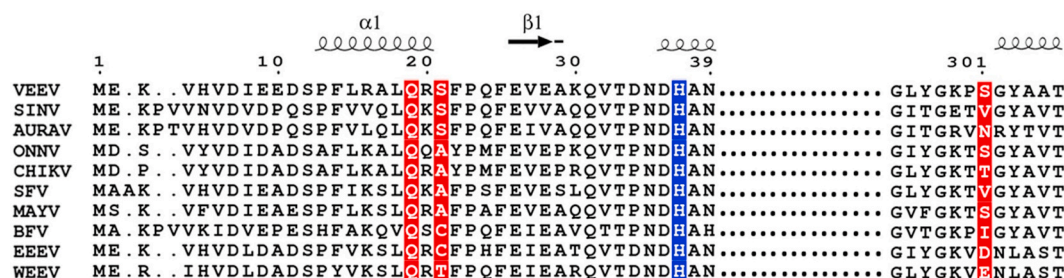
## 3. Results

The aim of this study was to examine the effect of nsP1 mutations both in infected cells and at the enzyme level in order to get mechanistic insights into the alphavirus RNA-capping process. In SINV, three mutations in nsP1 were shown to be associated to ribavirin resistance (Fig. 1): (i) a glutamine (Q) to a lysine (K) in position 21; (ii) a serine (S) to an asparagine (N) at position 23 and (iii) a valine (V) to a methionine (M) in position 302. These substitutions induced a cross-resistance to ribavirin and 2-amino-1,3,4-thiadiazole, both inhibitors of the cellular enzyme IMPDH. These data might suggest that cellular GTP pool balance is a concern in resistance development through a potential nsP1 capping activity modulation. In order to refine the role of these amino acids and their substitutions, we first assessed the molecular context associated to ribavirin resistance by an alignment analysis restricted to the three loci involved in ribavirin-resistance in the SINV model (Fig. 1). Position 21 is occupied by a conserved glutamine (Q) along the human-tropic alphaviruses. In contrast, the amino acid at position 23 is varying following the geographical distribution of human-tropic alphaviruses, with an alanine (A) for OW viruses and with a polar residue serine (S), threonine (T) or cysteine (C)) for NW viruses. In position 302, the amino acid can greatly vary with no obvious relation with the geographical distribution. We therefore decided to focus the study on positions 21 and 23 corresponding to residues 19 and 21 on the VEEV nsP1 sequence, respectively. Hence we mutated Q19 and S21 to various combinations of K, N and A residues, both in the VEEV reverse genetics system, and in the recombinant VEEV nsP1 protein for biochemical characterization, VEEV nsP1 being the biochemical model validated for all the capping steps (Li et al., 2015).

### 3.1. Effect of the mutations on methyltransferase activity (MTase)

To understand the effect of substitutions at positions 19 and 21, we



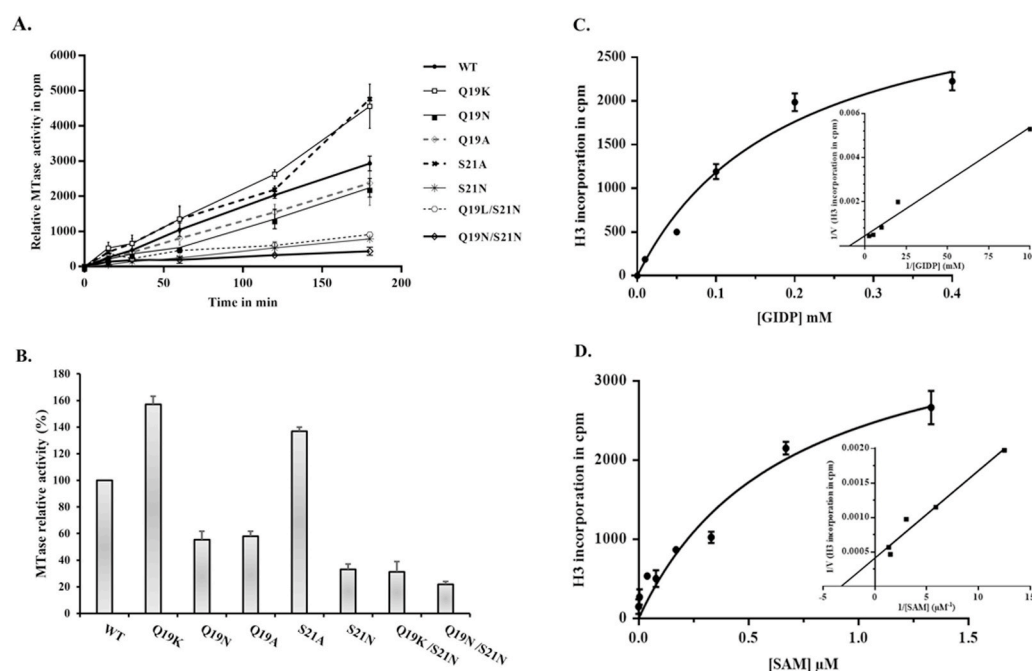


**Fig. 1.** Multiple sequence alignment of alphavirus nsP1 proteins in the vicinity of ribavirin resistance mutations. Sequence alignment was performed using T-Coffee and combined to PHD secondary-structure prediction using the ESPript software. Sequence of VEEV is numbered. Sites of ribavirin resistance mutations are in red boxes. Catalytic histidine is highlighted in the blue box. Black squiggles and arrows represent  $\alpha$  helices and  $\beta$  strands respectively. Dotted lines indicate region with low reliability of secondary structure elements for  $\alpha$  helices and  $\beta$  strands. VEEV, Venezuelan equine encephalitis virus (NC\_001449.1); SINV, Sindbis virus (NC\_001547.1); AURAV, Aura virus (NP\_819010.1); ONNV, O'nyong-nyong virus (NC\_001512.1); CHIKV, Chikungunya virus (MH229986.1); SFV, Semliki Forest virus (NC\_003215.1); MAYV, Mayaro virus (AZM66145.1); BFV, Barmah Forest virus (NC\_001786.1); EEEV, Eastern equine encephalitis virus (KX029319.1); WEEV, Western equine encephalomyelitis virus (NC\_003908.1).

first examined the effect of these mutations on nsP1 methyltransferase activity using time course experiments. For this purpose, the different mutant proteins were incubated with GIDP, a non-hydrolysable analogue of GTP, in the presence of [ $^3$ H] SAM, and the enzymatic activity was determined using a filter binding assay. Fig. 2A shows a time course experiment indicating that under these conditions all mutants are in the linear phase of the reaction before 180 min. Fig. 2B represents a 2 h end-point assay summarizing the resulting effects of the mutations and allowing the comparison of measured MTase activities relative to that of the WT enzyme. It appears clearly that most of the mutants display a decreased activity compared to the WT except Q19K and S21A. Mutants can be divided in two groups. A first group with decreased enzymatic activity compared to the WT including Q19N and Q19A, these latter keep about 50% of the nsP1 MT activity (Fig. 2A and B and Table 1). The other mutants from this group (S21N, Q19K/S21N and Q19N/S21N) show a marked decreased (70–80%) of MT activity. All of them contain the S21N substitution. This observation suggests that activity loss is the hallmark of S21N substitution. Mutants of the second group present an

enhanced MTase activity, with Q19K and S21A single mutations, yielding up to 1.57 time increase of the activity. However, it is noteworthy that the increase of the activity due to Q19K cannot compensate the negative effect of S21N. Altogether these results suggest that the loci on the nsP1 sequence studied here contains residues impacting the methyltransferase activity, relating MPA- and Ribavirin-resistance to this enzyme.

To date no crystal structure of alphavirus nsP1 is available. NsP1 has a predicted Rossmann fold with canonical motifs in its N-terminal domain, and is putatively folded with  $\alpha$ -helices and  $\beta$ -sheets separated by loops probably involved in nucleotide co-factor binding (Ahola et al., 1997) (Ahola and Karlin, 2015). In order to define if the targeted mutations regulate either GTP and/or SAM binding, we determined the apparent  $K_m$  values for SAM and GIDP. Briefly, we quantified the amount of [ $^3$ H] methyl transferred on GIDP at increasing concentration of SAM or GIDP and the  $K_m$  was deduced from Lineweaver-Burk plots (Fig. 2C and D). The kinetic experiments indicate that WT nsP1 binds SAM with an apparent  $K_m$  of 0.5  $\mu$ M (Table 2) which is in the same range



**Fig. 2.** Altered methyl-transfer in mutated nsP1. The activity was monitored by incubating VEEV-nsP1 WT and mutant proteins at 30 °C for varied amounts of time in the presence of 50 mM Tris (pH 7.0), 2 mM DTT, 10 mM KCl, 2 mM MgCl<sub>2</sub>, 2 mM GTP or GIDP, 0.55  $\mu$ Ci (SAM [H<sub>3</sub>]). The reaction product was loaded on DEAE-cellulose filters and the radioactivity quantified by scintigraphy. The mutation effect was investigated either in time course (A) or 2 h end-point experiments (B). (C) Typical Michaelis-Menten and Lineweaver –Burk plot for WT nsP1VEEV. Kinetic parameters were determined using Sigmaplot program and compiled in Table 2.

**Table 1**

Effect of mutations on the capping reactions.

VEEV nsP1	Methyltransferase MTase (%) <sup>a</sup>	m <sup>7</sup> GMP-nsP1 complex GTase 1 (%) <sup>b</sup>	RNA guanylylation GTase 2 (%) <sup>c</sup>	Guanylyltransferase activity GTase 1+GTase 2 (%) <sup>d</sup>
WT (Q19/S21)	100	100	100	100
Q19K	157 ± 6	124	117	145
Q19N	55 ± 6	85	197	167
Q19A	58 ± 3	195	107	209
S21A	137 ± 2	147	136	200
S21N	33 ± 4	191	125	239
Q19K/S21N	31 ± 7	126	207	261
Q19N/S21N	22 ± 2	91	350	318

Note.

<sup>a</sup> m<sup>7</sup>GDP formation refers to the methyltransferase reaction in the presence of 5 μM WT or mutant nsP1, 0.33 μM of SAM [H<sup>3</sup>] and 2 mM of GIDP. Values correspond to means of three different experiment ± standard deviation.

<sup>b</sup> GT1 activity corresponding to the ratio of m<sup>7</sup>GMP-nsP1 complex formation (Fig. 3) vs MT activity, assuming that all the formed m<sup>7</sup>GTP is used in the reaction.

<sup>c</sup> GT2 activity corresponding to the ratio of m<sup>7</sup>GpppA formation (Fig. 4) vs GT1 activity.

<sup>d</sup> Total guanylylation reaction corresponding to the ratio of RNA-guanylylation vs MT activity.

as the Km of other previously characterized recombinant MTases (0.1–8 μM) (Schulz and Rentmeister, 2012; Tomar et al., 2011) (Horiuchi et al., 2013). The Km value of mutant nsP1 varies between 0.14 and 1.5 μM. These values are barely impacted by single mutations except for Q19K showing a ~3-fold increase in SAM binding. Increased SAM binding properties of this mutant correlates with increased MTase activity (Fig. 2A and B). Conversely, the double mutant (Q19K/S21N and Q19N/S21N) show a ~3-fold increased Km, which might explain the reduced MTase activity of these mutants. The Km value of WT nsP1 VEEV for GIDP was also determined (0.27 mM, Fig. 2E). The Km values of mutant proteins vary from 0.1 to 0.6 mM, except for S21A nsP1 which shows a Km at 10 μM, reflecting a 27-fold increased apparent affinity for the GTP analogue compared to that of WT. This dramatic change might explain the enhancement of the corresponding MTase activity observed in Fig. 2D and C. Altogether these results indicate that position 19 and 21 of VEEV nsP1 impact both SAM and GIDP binding. In addition, we observe that mutations increasing the SAM or GIDP binding properties increase the MTase activity of nsP1 whereas those decreasing SAM or GIDP recruitment decrease the MTase activity.

### 3.2. Effect of mutations on nsP1-guanylylation (GTase1)

We next wanted to determine whether the mutants could differentially affect the formation of the m<sup>7</sup>GMP-nsP1 adduct. To test this hypothesis we incubated each mutant in the presence of SAM and [α-<sup>32</sup>P] GTP. During the incubation, the [α-<sup>32</sup>P] GTP is first methylated on its N7 position (MTase reaction) and the guanylylation of nsP1 occurs subsequently (GTase1 reaction). The generated radiolabeled m<sup>7</sup>GMP-nsP1 complex was next separated using SDS-PAGE and the amount of [α-<sup>32</sup>P] GTP bound to nsP1 was quantified. Fig. 3 shows that Q19K and S21A single substitutions result in a 2-fold increase of complex formation compared to that of WT nsP1. This is reminiscent of the increased MTase activity observed above (Fig. 2). The S21N mutation either alone or associated to Q19K or Q19N significantly decreases nsP1-guanylylation

**Table 2**

Effect of mutations on kinetic parameters for MTase reaction.

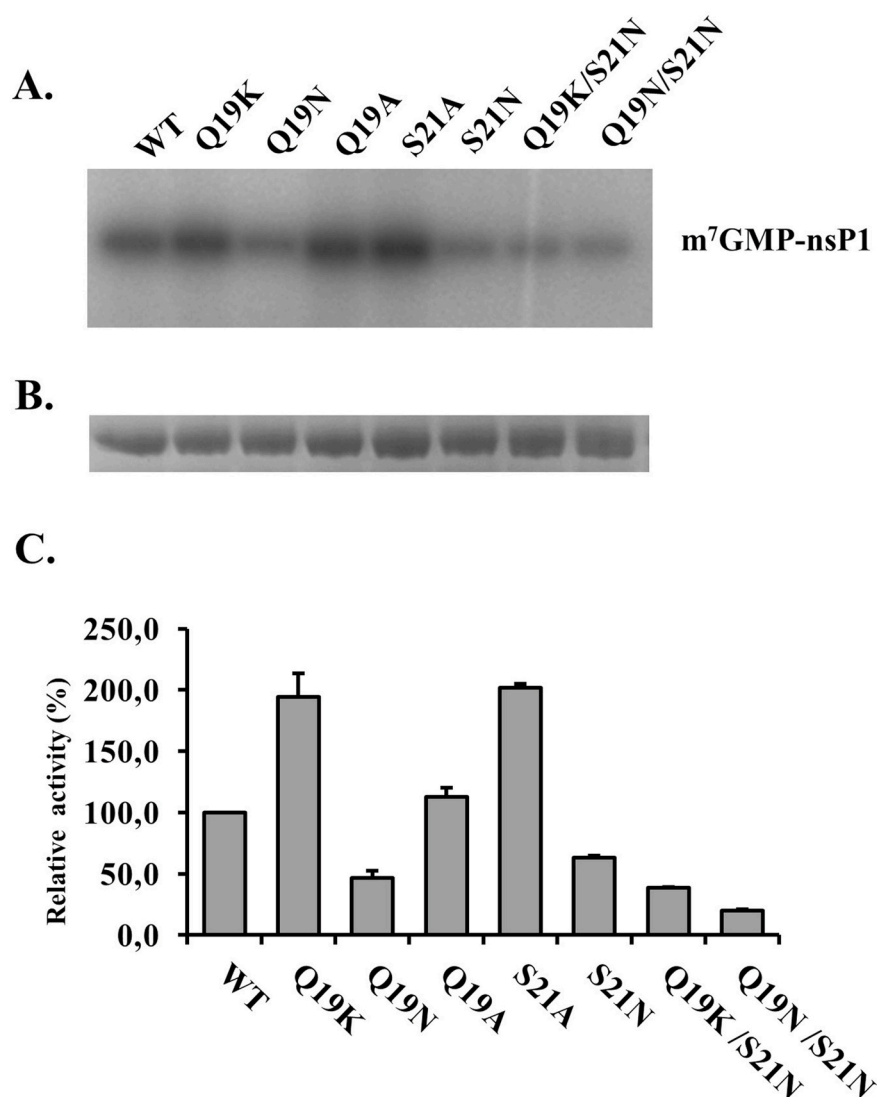
VEEV nsP1	Km (SAM) (μM)	Km (GIDP) (mM)
WT	0.55 ± 0.16	0.27 ± 0.08
Q19K	0.14 ± 0.03	0.1 ± 0.02
Q19N	0.5	0.23 ± 0.1
Q19A	0.75 ± 0.3	0.65 ± 0.2
S21A	0.58 ± 0.1	0.01 ± 0.003
S21N	0.71 ± 0.05	0.2 ± 0.002
Q19K/S21N	1.54 ± 0.4	0.11 ± 0.07
Q19N/S21N	1.3 ± 0.3	0.08 ± 0.02

Note: Kinetic parameters were calculated from Lineweaver –Burk plots using Sigmaplot program.

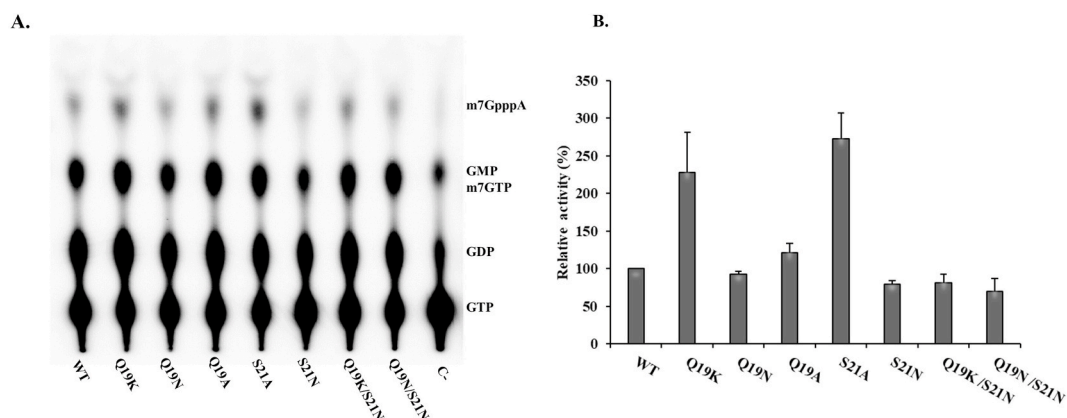
as observed for the MTase reaction. In contrast, the Q19A mutant which is characterized by moderate MTase activity shows an increase in nsP1 guanylylation, since its activity is similar to WT. Indeed, since MTase and nsP1-guanylylation reactions are sequential, this result suggests that the MTase reaction is not rate-limiting for this mutant. Hence, the assessment of the GTase1 reaction *per se* for each individual mutant can be evaluated by calculating the ratio global m<sup>7</sup>GMP-nsP1 complex formation/MTase (Table 1). By doing so one can notice that mutating Q19 to K but not N increases nsP1-guanylylation (25%). In addition we calculated an increase of about 50% for the S21A mutant and almost 100% (2-fold), for Q19A and S21N. Interestingly, it seems that there is no synergistic effect between the two residues for the GTase1 reaction, since the double mutants behave almost identically to the corresponding single Q19 mutant.

### 3.3. Effect of mutations on RNA-guanylyltransferase activity (GTase2)

Subsequently, we explored the effect of amino acid substitutions on the nsP1-mediated RNA-guanylylation reaction leading to the formation of the m<sup>7</sup>GpppRNA cap structure (GTase2 reaction). To assess this reaction, the first 30 nucleotides corresponding to the 5' sequence of VEEV genomic RNA were *in vitro* transcribed and the 5'-triphosphate end of RNA was dephosphorylated to generate a 5'-diphosphate end. We then incubated this RNA with nsP1 mutants, SAM and [α-<sup>32</sup>P] GTP. The reaction products - capped RNAs - were next digested by P1 endonuclease treatment. The released cap structures were separated using thin layer chromatography (TLC), and revealed by autoradiography (Fig. 4). Typically, TLC profiles reveal the presence of all nucleotide moieties present during the different steps of the reaction, namely GTP, GDP, GMP, m<sup>7</sup>GTP and the cap structure m<sup>7</sup>GpppA. For the negative control (without nsP1) the main product observed is GTP, although faint spots corresponding to GDP and GMP are present, generated by GTP hydrolysis. When nsP1 enzymes are present, GTP, GDP and GMP all the above mentioned chemical species are still present (Fig. 4A). In particular GDP is a well-known by-product of the cap synthesis mediated by the Bamboo mosaic virus capping enzyme (Hu et al., 2011; Huang et al., 2004; Lin et al., 2012). The m<sup>7</sup>GTP is the product of the MTase reaction and migrates very close to GMP. Lastly, the m<sup>7</sup>GpppA cap structure corresponds to the final product of the global capping reaction. The intensity of the m<sup>7</sup>GpppA product differs from one mutant to another and reflects the effect of the mutations on the global capping reaction. Fig. 4B points out that the substitution to N of one or both Q19 and S21 residues negatively affects the overall capping reaction. Q19A substitution induces a 20% gain over the WT nsP1 capping activity. Strikingly, Q19K and S21A induce a 2.5-fold increase in capping compared to the WT nsP1. Still, the activity described above corresponds to the global capping activity including MTase, GTase1 and GTase2 activities. When looking at the



**Fig. 3.** Effect of mutations on the formation of m<sup>7</sup>GMP-nsP1 complex. A: nsP1 guanylylation was monitored by SDS-PAGE following m<sup>7</sup>-GMP-nsP1 complex formation using [ $\alpha$ -P<sup>32</sup>] GTP, 100  $\mu$ M cold SAM and 5  $\mu$ M of each nsP1 enzyme. B: Coomassie blue stained gel of the nsP1-guanylylation reaction for protein load normalisation. C: Relative activity of VEEV nsP1 WT and mutant enzymes.



**Fig. 4.** Impact of ribavirin resistant mutation on RNA Guanylyltransferase activity. A: VEEV-RNA guanylylation reaction was carried out in the presence of diphosphate 5' termini VEEV RNA 30 oligomer nucleotides, 2  $\mu$ M of nsP1, 100  $\mu$ M of SAM and 0.33  $\mu$ M of [ $\alpha$ -P<sup>32</sup>] GTP. m<sup>7</sup>GpppA caps were digested from RNA with nuclease P1 and resolved on polyethylenimine cellulose thin-layer chromatography. B: The relative activity was calculated by considering that of WT as 100%. Values correspond to means  $\pm$  standard deviation. See Material and method for detailed procedure.

ratio of GTase2 over GTase1 (Table 1), one can estimate the rate of the GTase2 reaction *per se* for each mutant. Hence, the substitution of one residue leads to a 1.5 to 2-fold rise in GTase2 activity. Q19 and S21 residues appear to act in synergy, since the double mutants are 2.5 to 3-fold more active than the WT. It should be emphasised that the guanylyltransferase activity of nsP1 comprising the formation the m7GMP-nsP1 complex and the transfer to RNA is increased for all the mutants.

### 3.4. Effect of ribavirin on WT and S21A VEEV replication

To investigate the role of Q19 and S21 nsP1 VEEV residues on ribavirin susceptibility, we next introduced the mutations of interest into the VEEV molecular clone using the ISA method in Vero E6 cells (Aubry et al., 2015). Unfortunately, in this system only S21A substitution led to the generation of viable viral particles. The other substitutions yielded no or low-efficiency replicating viruses, insufficient to initiate an antiviral assays in the presence of ribavirin (data not shown). Therefore, the analysis focused on VEEV WT and S21A, using CHIKV as reference. To test the effect of ribavirin on these viruses, the amount of viral RNA in infected cultures containing increased amounts of ribavirin was compared. The dose response curves are presented in Fig. 5 and EC<sub>50</sub> compiled in Table 3. As expected, VEEV WT is not sensitive to ribavirin below 250  $\mu$ M, whereas EC<sub>50</sub> of ribavirin on CHIKV could be determined and is about  $26 \pm 8$   $\mu$ M (Table 3). Interestingly, VEEV S21A became ribavirin sensitive in the assay, with EC<sub>50</sub> of  $60 \pm 13$   $\mu$ M. In the genomic sequence of CHIKV the naturally occurring residue at that position is an A (Fig. 1). This observation suggests that this position might be one of the hallmarks towards ribavirin sensitivity in alphaviruses. To confirm that the antiviral effect of ribavirin is associated to the inhibition of the cellular IMPDH, the experience was repeated using a culture medium supplemented with GMP. Exogenous supply of GMP allows to bypass the IMPDH pathway and partly restores GTP synthesis. In the presence of exogenous GMP, the antiviral effect of ribavirin is partly abrogated for both CHIKV and VEEV S21A, suggesting that

**Table 3**

Antiviral activity of ribavirin on VEEV WT and S21A.

EC50 ( $\mu$ M)	VEEV WT	VEEV S21A	CHIKV OPY1
Ribavirin	>270	$60 \pm 13$	$26 \pm 8$
Ribavirin+ 100 $\mu$ M GMP	>400	>400	$90 \pm 18$

ribavirin effect on both viruses is mediated through IMPDH inhibition.

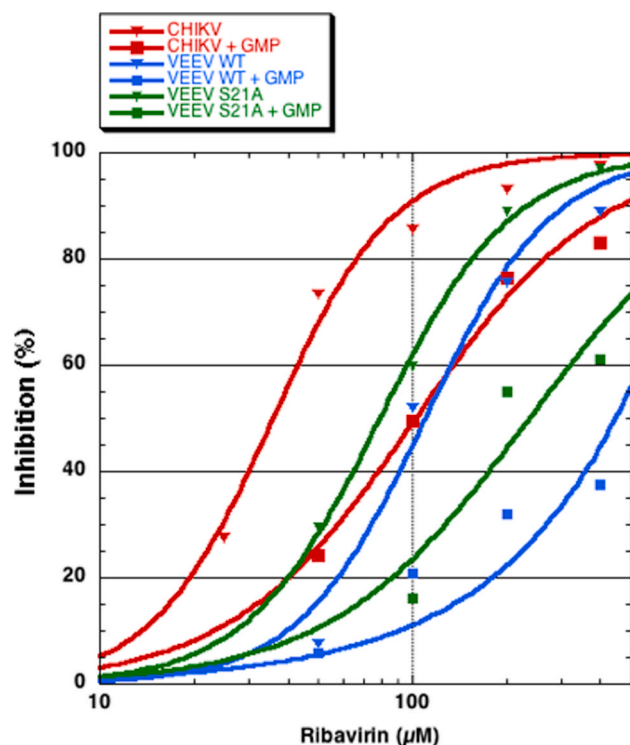
## 4. Discussion

The results presented here support the initial hypothesis according to which the development of resistance following GTP depletion is associated to a modulation of the capping efficiency. Hence, mutation of Q19 to K led to a three fold increase in SAM and GTP affinity, thereby enhancing the first step of the capping pathway, *i.e.*, the MTase reaction. The importance of this specific lysine in the GTP methylation is supported by the fact that substitution of Q19 to N has nearly no effect on SAM and GTP kinetic parameters. It is noteworthy that K in position 19 selected in SINRV ribavirin resistant virus is also naturally observed in fish alphaviruses SPDV and SDV, reflecting possible discrepancies of GTP concentration in the intracellular environment between mammal- and fish-infecting alphaviruses. The only example in the literature where a nearby residue was mutated is the substitution of L18 to E (L19E) in SFV nsP1, which leads to the complete and partial loss of SAM and GTP binding in UV-crosslinking experiments, respectively (Ahola et al., 1997). Concerning position 21, the results suggest that this residue regulates the guanylation reaction. The S21N substitution has a weak MTase activity but a strong nsP1 guanylation activity (GTase1). When acting in synergy with the K in position 19, the RNA guanylation (GTase2) is increased up to 207% vs 117% and 125% for the single mutants. This phenomenon endorses the resistance mutations observed in SINRV, in which upon GTP depletion the adaptive response could be the generation of an enzyme with higher GTP affinity and a stronger cap transfer on RNA. This hypothesis would have however to be validated on SINRV nsP1 model.

Our results point out to a close interconnection between GTP and SAM binding, as evidenced by the kinetic parameters determined for each substrate. This observation is also in agreement with previous studies highlighting that SAH produced during methylation is regulating the downstream steps of the capping (Li et al., 2015). From a structural point of view, these data support the fact that the two binding sites could be close to each other and involve leaning residues in the extreme N-terminus of the MT domain, as illustrated in the structure of Flavivirus virus (ZIKV) NS5 protein (Coutard et al., 2017; Zhao et al., 2017).

In parallel, we show that some polymorphism exists at position 21 in the alphavirus family (Fig. 1). We established that S21A shows an increased sensitivity of VEEV to ribavirin. As the addition of GMP decreases the antiviral effect of ribavirin, we assume that S21 might represent a key residue in ribavirin susceptibility. Together with the increase of the susceptibility towards ribavirin, S21A substitution on recombinant VEEV nsP1 shows a strong MTase activity and a 2-fold increased guanylyltransferase activity. This substitution, which is OW alphavirus specific (Fig. 1) points out that there are differences in the molecular mechanism governing capping reaction between alphavirus members. However, it remains possible that drug resistance/sensitivity could be modulated by compensatory mutations in other nsPs which remain to be determined.

Hence, it is important to bear in mind that drug resistance can be concomitant to mutations in other nsPs. Compensatory mutations are mainly due to the dynamic of interactions between the various nsPs alone or in the context of different polyprotein state in the course of processing, as exemplified by the switch between negative and positive sense RNA synthesis (Lemm et al., 1994). Recent study on Favipiravir, a potent antiviral pyrazine derivative against alphaviruses, led to the isolation of CHIKV nsP4 K291R resistant mutation. This mutant could



**Fig. 5.** Dose response curves of ribavirin on the replication of CHIKV, VEEV WT and VEEV S21A, in presence or absence of GMP.



not survive in the absence of the drug unless compensatory mutations in nsP2 and nsP3 were present (Delang et al., 2014). In SINV, synergic effect of compensatory mutations between nsP1 and nsP4 was observed in methionine depletion resistant mutants (Stollar et al., 2013). Moreover, subjecting VEEV infected cells to cytidine analog,  $\beta$ -D-N4-hydroxycytidine, showed that it has a strong antiviral potential. However, virus developed resistance mutations in nsP4, explained by its role in nucleotide incorporation, but also to a lesser extent into all other nsPs (Urakova et al., 2017).

Interestingly, ribavirin resistance was generally associated to mutations in viral RdRp domains. These mutations could be explained by the triphosphate form of ribavirin acting as an obligate or non-obligate chain terminator for the polymerase, or by the monophosphate form of ribavirin inhibiting IMPDH and inducing the depletion of GTP together with hypermutation by lack of available nucleotide (De Clercq and Li, 2016). However, other mode of action of ribavirin has emerged in certain virus family, with some of them targeting the cap formation. For example, ribavirin 5'-triphosphate was crystallized into the GTP-binding site of the DENV MTase NS5 (Benarroch et al., 2004). Likewise, a direct interaction of ribavirin with vaccinia virus (VV) capping D1 subunit was described. The monophosphate of ribavirin is able to form a RMP-D1 complex leading to the formation of RpppRNA (Bougie and Bisaillon, 2004). The characterization of the L protein of different rhabdoviruses showed that ribavirin diphosphate (RDP) could also bind the GDP polyribonucleotidyltransferase domain and interfere with cap formation only if the concentration of cellular GDP is low (Ogino and Ogino, 2017). Moreover, we and others have shown that 5'-triphosphate ribavirin is not acting directly on nsP1, as demonstrated by mM range  $IC_{50}$  values for VEEV and CHIKV nsP1 MTase (Kaur et al., 2018; Li et al., 2015). A similar  $IC_{50}$  value was obtained for VEEV nsP1 when tested in GTase 2 assay in the presence of RTP (data not shown).

RNA capping is a crucial process for controlling RNA stability, translation and recognition of "self" by host cells. Diphosphate, triphosphate, mis-capped or mis-folded RNA trigger the activation of RIG-I, MDA 5 and toll like receptors (TLR7/8 and TLR3) leading to Interferon secretion and thereby to the restriction of mis-capped viral genomes (Freund et al., 2019; Hyde et al., 2014). This is emphasised in different studies. In reoviruses for example increased infectivity can be associated with improved vRNA capping (Eaton et al., 2017). Likewise, Flaviviruses can boost their guanylation activity under oxidative stress conditions to enhance viral replication (Gullberg et al., 2015). In alphaviruses, the situation appears puzzling. In general, drug resistance mutations observed in alphavirus nsPs (including those studied here) decrease viral infectivity and negatively impact viral replication. This translates into small viral plaques and reduced viral titers in resistant mutants (Scheidel and Stollar, 1991; Urakova et al., 2017). Hence, point mutations leading to an increased nsP1 activity suggest that over-capping has a deleterious effect on the efficiency of viral infection (LaPointe et al., 2018; Stollar et al., 2013). During alphavirus replication, viral RNA population is heterogeneous. In addition to capped viral RNA, the infection leads to the accumulation of mono, di, tri and even non phosphate viral RNA. Interestingly, capped and uncapped RNA can be encapsidated with or without host's translation factors (Sokoloski et al., 2015). Clearly, the balance between capped and uncapped viral RNA is an important issue, whose biological significance is still to be defined. Over-capping might lead to a competition over translational host factors with a negative impact on infectivity. Yet in the case of drug-induced GTP pool depletion, the maintenance of an appropriate ratio of capped/uncapped RNA could be achieved by capping increase.

In conclusion, our results suggest that positions 19 and 21 are key players in the capping mechanism. In the presence of IMPDH inhibitors, RNA capping can be enhanced by mutations of these residues in order to overcome GTP depletions. The nature of the amino acid at position 21 might represent an important sensitivity/resistance marker to be taken into account for the rational design of nsP1 inhibitors.

Note:  $EC_{50}$  was determined as the mean of the results from three

independent experiments (ribavirin alone) and duplicates (Ribavirin + 100  $\mu$ M GMP).

## Acknowledgments

Oney Ortega is a recipient of Méditerranée Infection Fondation studentship.

## References

- Abdelnabi, R., Neyts, J., Delang, L., 2017. Chikungunya virus infections: time to act, time to treat. *Curr Opin Virol* 24, 25–30.
- Ahola, T., Kaariainen, L., 1995. Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. *Proc. Natl. Acad. Sci. U. S. A.* 92, 507–511.
- Ahola, T., Karlin, D.G., 2015. Sequence analysis reveals a conserved extension in the capping enzyme of the alphavirus supergroup, and a homologous domain in nodaviruses. *Biol. Direct* 10, 16.
- Ahola, T., Laakkonen, P., Vihinen, H., Kaariainen, L., 1997. Critical residues of Semliki Forest virus RNA capping enzyme involved in methyltransferase and guanylyltransferase-like activities. *J. Virol.* 71, 392–397.
- Albulescu, I.C., van Hoolwerff, M., Wolters, L.A., Bottaro, E., Nastruzzi, C., Yang, S.C., Tsay, S.C., Hwu, J.R., Snijder, E.J., van Hemert, M.J., 2015. Suramin inhibits chikungunya virus replication through multiple mechanisms. *Antivir. Res.* 121, 39–46.
- Aubry, F., Nougairède, A., de Fabritus, L., Piorkowski, G., Gould, E.A., de Lamballerie, X., 2015. ISA-lation" of single-stranded positive-sense RNA viruses from non-infectious clinical/animal samples. *PLoS One* 10, e0138703.
- Benarroch, D., Egloff, M.P., Mulard, L., Guerreiro, C., Romette, J.L., Canard, B., 2004. A structural basis for the inhibition of the NS5 dengue virus mRNA 2'-O-methyltransferase domain by ribavirin 5'-triphosphate. *J. Biol. Chem.* 279, 35638–35643.
- Bougie, I., Bisaillon, M., 2004. The broad spectrum antiviral nucleoside ribavirin as a substrate for a viral RNA capping enzyme. *J. Biol. Chem.* 279, 22124–22130.
- Briolant, S., Garin, D., Scaramozzino, N., Jouan, A., Crance, J.M., 2004. In vitro inhibition of Chikungunya and Semliki Forest viruses replication by antiviral compounds: synergistic effect of interferon-alpha and ribavirin combination. *Antivir. Res.* 61, 111–117.
- Chen, M.W., Tan, Y.B., Zheng, J., Zhao, Y., Lim, B.T., Cornvik, T., Lescar, J., Ng, L.F.P., Luo, D., 2017. Chikungunya virus nsP4 RNA-dependent RNA polymerase core domain displays detergent-sensitive primer extension and terminal adenylyltransferase activities. *Antivir. Res.* 143, 38–47.
- Coutard, B., Barral, K., Lichiere, J., Selisko, B., Martin, B., Aouadi, W., Lombard, M.O., Debart, F., Vasseur, J.J., Guillemot, J.C., Canard, B., Decroly, E., 2017. Zika virus methyltransferase: structure and functions for drug design perspectives. *J. Virol.* 91, e02202.
- Das, P.K., Puusepp, L., Varghese, F.S., Utt, A., Ahola, T., Kananovich, D.G., Lopp, M., Merits, A., Karelson, M., 2016. Design and validation of novel chikungunya virus protease inhibitors. *Antimicrob. Agents Chemother.* 60, 7382–7395.
- De Clercq, E., Cools, M., Balzarini, J., Snoeck, R., Andrei, G., Hosoya, M., Shigetani, S., Ueda, T., Minakawa, N., Matsuda, A., 1991. Antiviral activities of 5-ethynyl-1-beta-D-ribofuranosylimidazole-4-carboxamide and related compounds. *Antimicrob. Agents Chemother.* 35, 679–684.
- De Clercq, E., Li, G., 2016. Approved antiviral drugs over the past 50 years. *Clin. Microbiol. Rev.* 29, 695–747.
- Delang, L., Li, C., Tas, A., Querat, G., Albulescu, I.C., De Burghgraeve, T., Guerrero, N.A., Gigante, A., Piorkowski, G., Decroly, E., Jochmans, D., Canard, B., Snijder, E.J., Perez-Perez, M.J., van Hemert, M.J., Coutard, B., Leyssen, P., Neyts, J., 2016. The viral capping enzyme nsP1: a novel target for the inhibition of chikungunya virus infection. *Sci. Rep.* 6, 31819.
- Delang, L., Segura Guerrero, N., Tas, A., Querat, G., Pastorino, B., Froeyen, M., Dallmeier, K., Jochmans, D., Herdewijn, P., Bello, F., Snijder, E.J., de Lamballerie, X., Martina, B., Neyts, J., van Hemert, M.J., Leyssen, P., 2014. Mutations in the chikungunya virus non-structural proteins cause resistance to favipiravir (T-705), a broad-spectrum antiviral. *J. Antimicrob. Chemother.* 69, 2770–2784.
- Eaton, H.E., Kobayashi, T., Dermody, T.S., Johnston, R.N., Jais, P.H., Shmulevitz, M., 2017. African swine fever virus NP868R capping enzyme promotes reovirus rescue during reverse genetics by promoting reovirus protein expression, virion assembly, and RNA incorporation into infectious virions. *J. Virol.* 91, e02416-02416.
- Feibelman, K.M., Fuller, B.P., Li, L., LaBarbera, D.V., Geiss, B.J., 2018. Identification of small molecule inhibitors of the Chikungunya virus nsP1 RNA capping enzyme. *Antivir. Res.* 154, 124–131.
- Ferreira-Ramos, A.S., Li, C., Eyedoux, C., Contreras, J.M., Morice, C., Querat, G., Gigante, A., Perez Perez, M.J., Jung, M.L., Canard, B., Guillemot, J.C., Decroly, E., Coutard, B., 2019. Approved drugs screening against the nsP1 capping enzyme of Venezuelan equine encephalitis virus using an immuno-based assay. *Antivir. Res.* 163, 59–69.
- Franco, E.J., Rodriguez, J.L., Pomeroy, J.J., Hanrahan, K.C., Brown, A.N., 2018. The effectiveness of antiviral agents with broad-spectrum activity against chikungunya virus varies between host cell lines. *Antivir. Chem. Chemother.* 26, 2040206618807580.

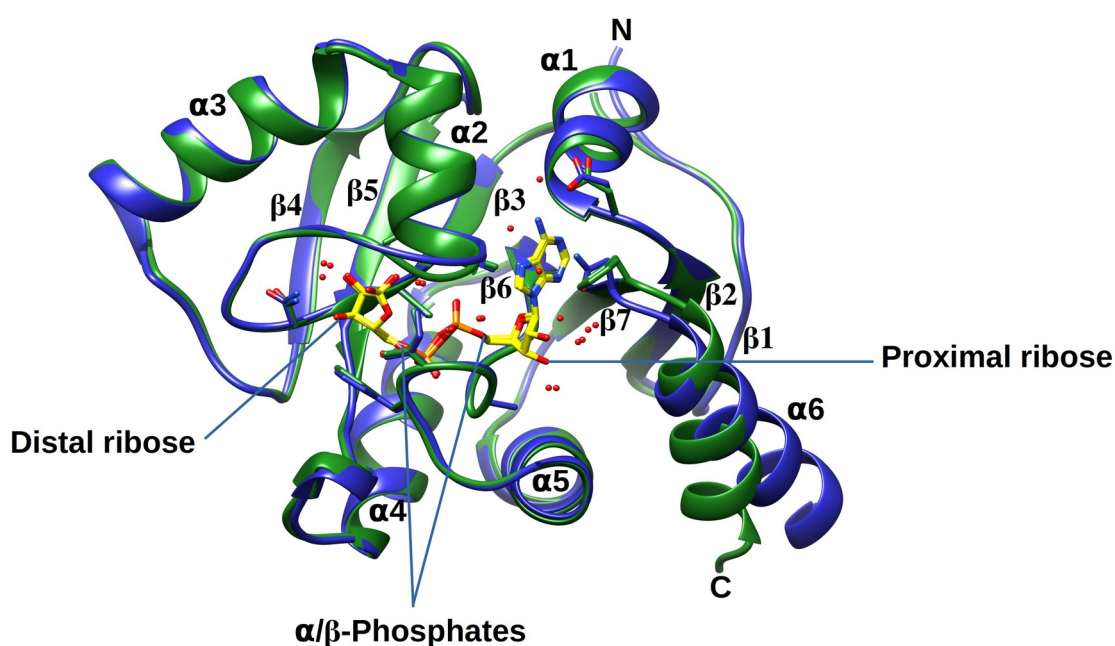
- Freund, I., Buhl, D.K., Boutin, S., Kotter, A., Pichot, F., Marchand, V., Vierbuchen, T., Heine, H., Motorin, Y., Helm, M., Dalpke, A.H., Eigenbrod, T., 2019. 2'-O-methylation within prokaryotic and eukaryotic tRNA inhibits innate immune activation by endosomal Toll-like receptors but does not affect recognition of whole organisms. *RNA* 25, 869–880.
- Gigante, A., Canela, M.D., Delang, L., Priego, E.M., Camarasa, M.J., Querat, G., Neyts, J., Leyssen, P., Perez-Perez, M.J., 2014. Identification of [1,2,3]triazolo[4,5-d]pyrimidin-7(6H)-ones as novel inhibitors of Chikungunya virus replication. *J. Med. Chem.* 57, 4000–4008.
- Gomez-SanJuan, A., Gamo, A.M., Delang, L., Perez-Sanchez, A., Amrun, S.N., Abdelnabi, R., Jacobs, S., Priego, E.M., Camarasa, M.J., Jochmans, D., Leyssen, P., Ng, L.F.P., Querat, G., Neyts, J., Perez-Perez, M.J., 2018. Inhibition of the replication of different strains of chikungunya virus by 3-aryl-[1,2,3]triazolo[4,5-d]pyrimidin-7(6H)-ones. *ACS Infect. Dis.* 4, 605–619.
- Gullberg, R.C., Jordan Steel, J., Moon, S.L., Soltani, E., Geiss, B.J., 2015. Oxidative stress influences positive strand RNA virus genome synthesis and capping. *Virology* 475, 219–229.
- Hardy, W.R., Strauss, J.H., 1989. Processing the nonstructural polyproteins of sindbis virus: nonstructural proteinase is in the C-terminal half of nsP2 and functions both in cis and in trans. *J. Virol.* 63, 4653–4664.
- Horiuchi, K.Y., Eason, M.M., Ferry, J.J., Planck, J.L., Walsh, C.P., Smith, R.F., Howitz, K. T., Ma, H., 2013. Assay development for histone methyltransferases. *Assay Drug Dev. Technol.* 11, 227–236.
- Hu, R.H., Lin, M.C., Hsu, Y.H., Meng, M., 2011. Mutational effects of the consensus aromatic residues in the mRNA capping domain of Bamboo mosaic virus on GTP methylation and virus accumulation. *Virology* 411, 15–24.
- Hu, X., Compton, J.R., Leary, D.H., Olson, M.A., Lee, M.S., Cheung, J., Ye, W., Ferrer, M., Southall, N., Jadhav, A., Morazzani, E.M., Glass, P.J., Marugan, J., Legler, P.M., 2016. Kinetic, mutational, and structural studies of the Venezuelan equine encephalitis virus nonstructural protein 2 cysteine protease. *Biochemistry* 55, 3007–3019.
- Huang, Y.L., Han, Y.T., Chang, Y.T., Hsu, Y.H., Meng, M., 2004. Critical residues for GTP methylation and formation of the covalent m7GMP-enzyme intermediate in the capping enzyme domain of bamboo mosaic virus. *J. Virol.* 78, 1271–1280.
- Hyde, J.L., Gardner, C.L., Kimura, T., White, J.P., Liu, G., Trobaugh, D.W., Huang, C., Tonelli, M., Paessler, S., Takeda, K., Klimstra, W.B., Amarasinghe, G.K., Diamond, M. S., 2014. A viral RNA structural element alters host recognition of nonself RNA. *Science* 343, 783–787.
- Kaur, R., Mudgal, R., Narwal, M., Tomar, S., 2018. Development of an ELISA assay for screening inhibitors against divalent metal ion dependent alphavirus capping enzyme. *Virus Res.* 256, 209–218.
- LaPointe, A.T., Moreno-Contreras, J., Sokolowski, K.J., 2018. Increasing the capping efficiency of the sindbis virus nsP1 protein negatively affects viral infection. *mBio* 9, e02342-02318.
- Law, Y.S., Utt, A., Tan, Y.B., Zheng, J., Wang, S., Chen, M.W., Griffin, P.R., Merits, A., Luo, D., 2019. Structural insights into RNA recognition by the Chikungunya virus nsP2 helicase. *Proc. Natl. Acad. Sci. U. S. A.* 116, 9558–9567.
- Lemm, J.A., Rumenapf, T., Strauss, E.G., Strauss, J.H., Rice, C.M., 1994. Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus- and plus-strand RNA synthesis. *EMBO J.* 13, 2925–2934.
- Li, C., Guillen, J., Rabah, N., Blanjoie, A., Debart, F., Vasseur, J.J., Canard, B., Decroly, E., Coutard, B., 2015. mRNA capping by Venezuelan equine encephalitis virus nsP1: functional characterization and implications for antiviral research. *J. Virol.* 89, 8292–8303.
- Lin, H.Y., Yu, C.Y., Hsu, Y.H., Meng, M., 2012. Functional analysis of the conserved histidine residue of Bamboo mosaic virus capping enzyme in the activity for the formation of the covalent enzyme-m7GMP intermediate. *FEBS Lett.* 586, 2326–2331.
- Lwande, O.W., Obanda, V., Bucht, G., Mosomtai, G., Otieno, V., Ahlm, C., Evander, M., 2015. Global emergence of Alphaviruses that cause arthritis in humans. *Infect. Ecol. Epidemiol.* 5, 29853.
- Markland, W., McQuaid, T.J., Jain, J., Kwong, A.D., 2000. Broad-spectrum antiviral activity of the IMP dehydrogenase inhibitor VX-497: a comparison with ribavirin and demonstration of antiviral additivity with alpha interferon. *Antimicrob. Agents Chemother.* 44, 859–866.
- Milhas, S., Raux, B., Betzi, S., Derviaux, C., Roche, P., Restouin, A., Basse, M.J., Rebuffet, E., Lugari, A., Badol, M., Kashyap, R., Lissitzky, J.C., Eyedoux, C., Hamon, V., Gourd, M.E., Combes, S., Zimmermann, P., Aurrand-Lions, M., Roux, T., Rogers, C., Muller, S., Knapp, S., Trinquet, E., Collette, Y., Guillemot, J.C., Morelli, X., 2016. Protein-protein interaction inhibition (2P2D)-Oriented chemical library accelerates hit discovery. *ACS Chem. Biol.* 11, 2140–2148.
- Ogino, M., Ogino, T., 2017. 5'-Phospho-RNA acceptor specificity of GDP polyribonucleotidyltransferase of vesicular stomatitis virus in mRNA capping. *J. Virol.* 91.
- Pohjala, L., Barai, V., Azhayev, A., Lapinjoki, S., Ahola, T., 2008. A luciferase-based screening method for inhibitors of alphavirus replication applied to nucleoside analogues. *Antivir. Res.* 78, 215–222.
- Ravichandran, R., Manian, M., 2008. Ribavirin therapy for Chikungunya arthritis. *J. Infect. Dev. Ctries* 2, 140–142.
- Rikonen, M., Peranen, J., Kaariainen, L., 1994. ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2. *J. Virol.* 68, 5804–5810.
- Rosenblum, C.I., Scheidel, L.M., Stollar, V., 1994. Mutations in the nsP1 coding sequence of Sindbis virus which restrict viral replication in secondary cultures of chick embryo fibroblasts prepared from aged primary cultures. *Virology* 198, 100–108.
- Sales, G., Barbosa, I.C.P., Canejo Neta, L.M.S., Melo, P.L., Leitao, R.A., Melo, H.M.A., 2018. Treatment of chikungunya chronic arthritis: a systematic review. *Rev. Assoc. Med. Bras.* 64, 63–70, 1992.
- Scheidel, L.M., Durbin, R.K., Stollar, V., 1987. Sindbis virus mutants resistant to mycophenolic acid and ribavirin. *Virology* 158, 1–7.
- Scheidel, L.M., Stollar, V., 1991. Mutations that confer resistance to mycophenolic acid and ribavirin on Sindbis virus map to the nonstructural protein nsP1. *Virology* 181, 490–499.
- Schulz, D., Rentmeister, A., 2012. An enzyme-coupled high-throughput assay for screening RNA methyltransferase activity in *E. coli* cell lysate. *RNA Biol.* 9, 577–586.
- Shirako, Y., Strauss, J.H., 1994. Regulation of Sindbis virus RNA replication: uncapped P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. *J. Virol.* 68, 1874–1885.
- Sokolowski, K.J., Haist, K.C., Morrison, T.E., Mukhopadhyay, S., Hardy, R.W., 2015. Noncapped alphavirus genomic RNAs and their role during infection. *J. Virol.* 89, 6080–6092.
- Stollar, V., Mensah, V., Adams, S., Li, M.L., 2013. Evolution of Sindbis virus with a low-methionine-resistant phenotype is dependent both on a pre-existing mutation and on the methionine concentration in the medium. *PLoS One* 8, e60504.
- Tomar, S., Narwal, M., Harms, E., Smith, J.L., Kuhn, R.J., 2011. Heterologous production, purification and characterization of enzymatically active Sindbis virus nonstructural protein nsP1. *Protein Expr. Purif.* 79, 277–284.
- Urakova, N., Kuznetsova, V., Crossman, D.K., Sokratian, A., Guthrie, D.B., Kolykhalov, A. A., Lockwood, M.A., Natchus, M.G., Crowley, M.R., Painter, G.R., Frolova, E.I., Frolov, I., 2017. beta-D-N(4)-hydroxycytidine is a potent anti-alphavirus compound that induces high level of mutations in viral genome. *J. Virol.* 92, e01965.
- Vasiljeva, L., Merits, A., Auvinen, P., Kaariainen, L., 2000. Identification of a novel function of the alphavirus capping apparatus. RNA 5'-triphosphatase activity of Nsp2. *J. Biol. Chem.* 275, 17281–17287.
- Weaver, S.C., Winegar, R., Manger, I.D., Forrester, N.L., 2012. Alphaviruses: population genetics and determinants of emergence. *Antivir. Res.* 94, 242–257.
- Zacks, M.A., Paessler, S., 2010. Encephalitic alphaviruses. *Vet. Microbiol.* 140, 281–286.
- Zhao, B., Yi, G., Du, F., Chuang, Y.C., Vaughan, R.C., Sankaran, B., Kao, C.C., Li, P., 2017. Structure and function of the Zika virus full-length NS5 protein. *Nat. Commun.* 8, 14762.

## 4 *Macro domain key player in coronavirus infections*

### 4.1 *Macro Domains*

*Macro* domains are widely distributed among life kingdoms. They have been phylogenetically subdivided into six distinct clades: MacroD-type, MacroH2A-type, Macro2-type, Viral-type, PARG (Poly (ADP-ribose) glycohydrolases) and ALC1 (amplified in liver cancer 1)-type. *Macro* domains are found in more than 150 RNA viruses, including *Togaviridae*, *Hepeviridae* and *Coronaviridae*. Most RNA virus *Macro* domains are included into the MacroD-like family. In viruses, they were originally termed X-domains then renamed because of their structural similarity with H2A histone variant called *MacroH2A*. *Macro* domains, which are able to bind various mono-MAR derivatives, including ADP-ribose 1" phosphate (Appr1p), O-acetyl-ADP-ribose and PAR, in a free form or conjugated to protein, are also known as *Mac1*. In addition, SARS-CoVs, MERS-CoV and other members of the  $\beta$ -coronaviruses, contain a non-conserved region with two additional *Macro* domains (*Mac2/Mac3* – formally known as SUD domains), which bind to RNA substrates (143). The *Macro* domain from thermophilic bacterium *Archaeoglobus fulgidus* was the first *Macro* domain which atomic structure was resolved (144). It was shown to bind ADP-ribose with a well-conserved ligand-binding pocket. The *Macro* domain contains approximately 130–190 amino acid residues that form a globular mix of six stranded  $\beta$ -sheet flanked by five  $\alpha$ -helices (figure 15) (145). The conserved cleft that bind ADP-ribose on the surface have been confirmed by several crystal structures and biochemical studies (figure 15). The residues found in  $\beta$ -sheet 2,  $\beta$ -sheet 5,  $\alpha$ -helix 2 and  $\alpha$ -helix 5 coordinated the distal ribose moiety. Also, the residues situated in this region contribute to the backbone contacts with the  $\alpha$ - $\beta$  phosphate of ADP-ribose. The proximal ribose is sandwiched between  $\beta$ -sheet 5,6 and 7 in the C-terminal region. Viral *Macro* domain were shown to bind ADP-ribose derivatives, but also to hydrolyse MARylated (de-MARylation) and, in some cases, PARylated (de-PARylation) substrates (145). Moreover, a phosphatase activity against Appr1p was

reported for alphaviruses, Human coronavirus 229E (HCoV-229E) and SARS-CoV (146,147). Hydrolysis activities are believed to counter the above mentioned PARPs anti-viral effects (see general introduction section). This assumption is supported by many mutagenesis studies of *Macro* domain key residues implicated in ADP-ribose binding, showing a reduction of virus replication and virulence of HEV, alphaviruses and several coronaviruses (148). Hence, MHV *Macro* domain catalytic mutant failed to induce acute hepatitis; and its growth is restricted in culture cells unless IFN receptor knockout cells are used or PARPs inhibitors are added (148). In HCoV-229, the mutant virus becomes susceptible to type I and II IFN and is unable to suppress the activation of ISGs. In mouse adapted SARS-CoV, mutation of ADP-ribose binding pocket renders the virus more susceptible to cytokines, including IFN, TNF and IL-6, and protects mice from lethal infection outcome (148). These data are supported by Heer *et al* stating that SARS-CoV-2 infection alters PARP family gene expression and disrupts NAD<sup>+</sup> biosynthesis (149). The accumulated data highlight the importance of ADP-ribosylation in host-viral conflict and the role of viral *Macro* domains for the viral virulence. Thus, *Macro* domain could be a novel therapeutic target preventing severe coronavirus, alphavirus or HEV-induced disease.





**Figure 15. Structure of SARS-CoV and SARS-CoV-2 *Macro* domain in complex with ADP-ribose molecule.** SARS-CoV (blue) and SARS-CoV-2 (green) structure superposition with ADP-ribose were made with Chimera (150). Secondary structures are labeled and ADP-ribose molecule is shown in yellow sticks with oxygen (red), nitrogen (blue) and bound water (red spheres). Structure of SARS-CoV (PDB: 2FAV) (151) and SARS-CoV-2 (PDB: 6WOJ) (152) *Macro* domains were extracted from PDB.

#### 4.1.1 Viral *Macro* Domains: prominent antiviral targets

The implication of viral *Macro* domains in the modulation of host cell defence and the promotion of viral replication via ADP-ribosylhydrolase activities supports their targeting for anti-viral drug development. Until now research related to the development of inhibitors against viral *Macro* domains remains limited and involves mainly SARS-CoV-2 and CHIKV. Most of the identified inhibitors bind to *Macro* domain in the region of adenosine and diphosphate bond. In the case of CHIKV *Macro* domain inhibitors, several computational and experimental drug design methods have been undertaken. Two 4'-halogenated dihydrorugosaflavonoids (chloro and bromo derivatives) were found to bind in the ADP-ribose region by molecular docking (153). Evaluation of these compound in cell culture revealed that they inhibit viral replication in Vero cells up to 95% and 92%, respectively (153). In another studies Shimizu *et al.* selected 12 molecules out of 48 750 compounds based on the docking scores of ADP-ribose binding of CHIKV *Macro* domain for further *in vitro* analysis (154). Some of the compounds appeared to interfere with virus replication. Although, experimental evidences showing the binding of these compounds to the *Macro* domain remain lacking. In addition, the interactions of flavonoids from natural resources with protein targets as *Macro* domains have been studied by computational tools. Some of them such as hesperetin, naringenin, baicalin and quercetagenin revealed a highest theoretical binding affinity to the CHIKV nsP3 *Macro* domain (155). Further, in another recently study nontoxic and noncarcinogenic flavonoids

were also identified to bind to the CHIKV nsP3 *Macro* by in silico-based pharmacological tests (155). Despite the numerous compounds identified, experimental validation confirming the inhibitory effect is lacking. Among the clinically used antiviral nucleoside analogs, remdesivir was the only metabolite occupying the adenosine site. It has a comparable binding affinity to the ADP-ribose measure by ITC (156). The plasticity of the ADP-ribose binding cleft provides new insights for antiviral drug design.

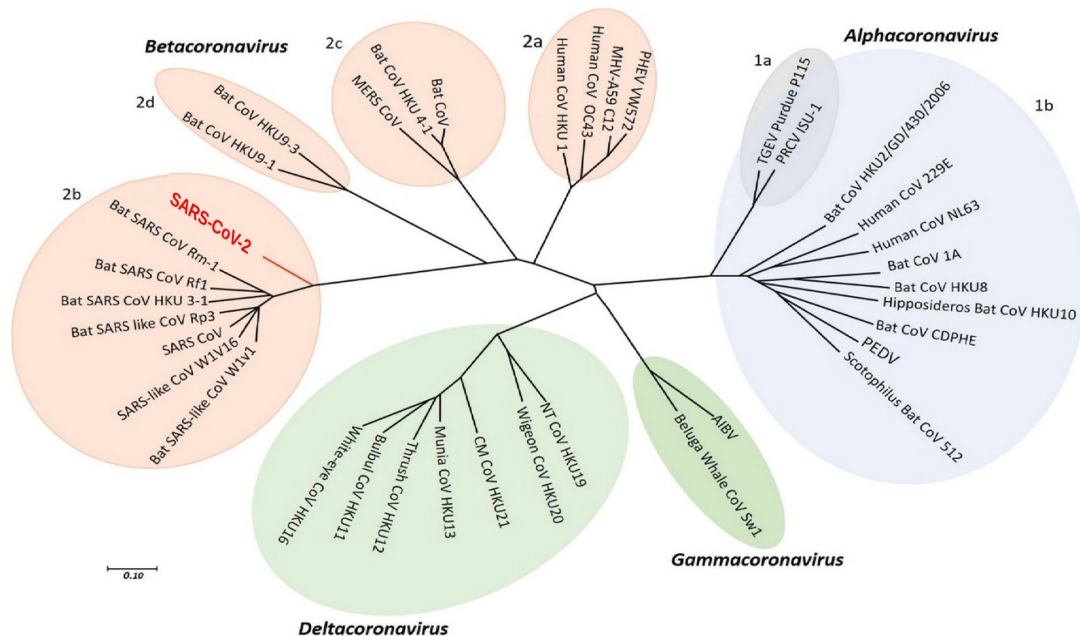
## 4.2 Coronavirus: Classification and epidemiology

Coronavirus (CoVs) are one of the known largest group of viruses belonging to the *Nidovirales* order, including *Coronaviridae*, *Roniviridae*, *Mesoniviridae*, and *Arteriviridae* families. The *Coronaviridae* comprise one of two subfamilies *Coronavirinae* and *Torovirinae*. *Coronavirinae* subfamily is further subdivided into four genera, alpha, beta, gamma and delta CoVs. Genera alpha and beta are able to infect mammalian cells. While, gamma and delta CoVs have affinity towards birds (157). Alpha and beta CoVs are responsible for known human CoVs related diseases. Actually, principal CoVs lineages causing significant illness in human comprise: HCoV-229E, Human coronavirus OC43 (HCoV-OC43), Human coronavirus NL63 (HCoV-NL63), Human coronavirus HKU1 (HCoV-HKU1), SARS-CoV, SARS-CoV-2 and MERS-CoV. These viruses cause infections of the upper and lower respiratory tract. In general, the symptoms of infections range from a mild common cold to a lethal respiratory infection, depending on the virus. The main epidemics affecting humans in recent years have been caused by SARS-CoV and MERS-CoV.

The epidemic of SARS-CoV originated in Guangdong Province in Southern China during November 2002, then it was brought to Hong Kong in February 2003 (158). The disease spread rapidly to Asian, North American and European countries causing atypical pneumonia with a fatality rate of 10%. In June 2012, the MERS-CoV, a new

betacoronavirus (beta-CoV) was isolated from a patient in Saudi Arabia who died of severe pneumonia and renal failure (158). Others, outbreaks of MERS-CoV have been reported in Korea in 2015 with 180 cases, and in 27 other countries, causing 848 deaths in July 2019. At the end of 2019, a novel beta-CoV similar to SARS-CoV emerged in the city of Wuhan, China. This new virus was designated as SARS-CoV-2 causing an outbreak of an unusual viral pneumonia. SARS-CoV-2 shares 79 and 50% genome sequence identity with SARS-CoV and MERS-CoV, respectively (159). Phylogenetic analysis included SARS-CoV-2 in the subgenus Sarbecovirus of the genus beta-CoV along with SARS-CoV and SARS-related coronaviruses (SARSr-CoVs) (figure 16), found in bats. To date, it is still not clear where and when the SARS-CoV-2 first humans were infected, although genetic evidences suggest that it is a natural virus that probably originated in animals (158). This novel coronavirus is responsible for the actual pandemic known as coronavirus disease 2019 (COVID-19). SARS-CoV-2 is a highly transmissible coronavirus with a fast worldwide spread. As of March 2022, 238,229,951 cases have been confirmed, including 4,859,277 cumulative deaths (160). This new virus affects all the ages of the population but the median age of the infections is around 50 years. People having co-morbidities are more likely to develop severe respiratory disease. Mortality rate among COVID-19 patients ranges between 2 and 4%. While mortality rate of SARS-CoV and MERS-CoV reached 15 and 37%, respectively (161).

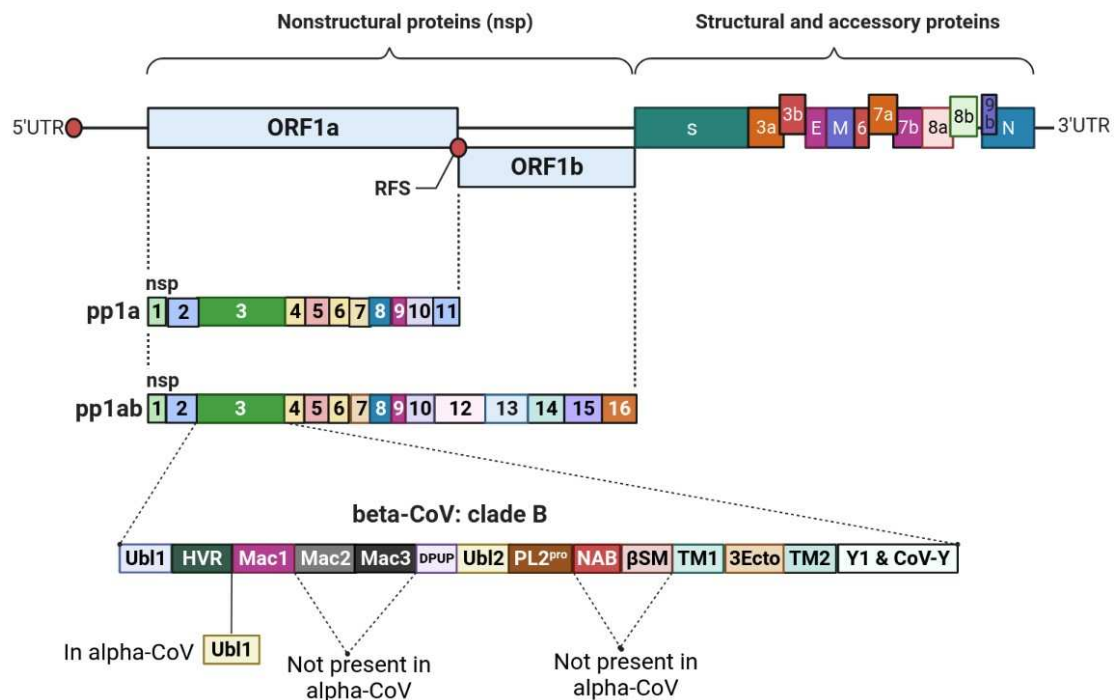
Several serological and molecular diagnostics tools were developed to detect CoVs infections (162). The most useful diagnostic method available actuality for SARS-CoV-2 is the reverse transcription-quantitative polymerase chain reaction (RT-qPCR). The persistence of SARS-CoV-2 outbreak for almost two years represents the most serious threat to health and world economy in the XXI century.



**Figure 16. Coronavirus phylogenetic tree.** Coronaviruses are grouped into four genera: alphacoronavirus (alpha-CoV) (sky blue), betacoronavirus (beta-CoV) (pink), gammacoronavirus (gamma-CoV) (green) and deltacoronavirus (delta-CoV) (light blue). SARS-CoV-2 is highlighted in red. The image was adapted from (163).

#### 4.2.1 Genome organization and replication cycle

Coronaviruses are (+) RNA viruses with exceptionally large genomes of ~30 kb, containing at least six ORFs. The first ORFs (ORF1a/b) is located at the 5' end and represents two-thirds of the whole genome length (figure 17). It encodes the non-structural proteins (nsp), described below. The ORFs located at the 3' end lead to the synthesis of the structural and the accessory proteins, including envelop glycoprotein spike (S), membrane (M) proteins, envelope (E) proteins and the nucleocapsid (N) protein (figure 17) (164).



**Figure 17. Coronavirus genomic organization.** The 5'-capped and 3'-polyadenylated genome contains at least 6 ORFs. ORF1a and ORF1b, encode two polyproteins (pp1a and pp1ab). These polyproteins are cleaved by viral proteases generating 16 non-structural proteins (nsp1–nsp16). A zoom on nsp3 is represented to highlight the multiple domains of this protein. The description of each domain is in the main text. 3'ORFs encode the structural and accessory proteins, mentioned in the text. The image was created in *BioRender.com*.

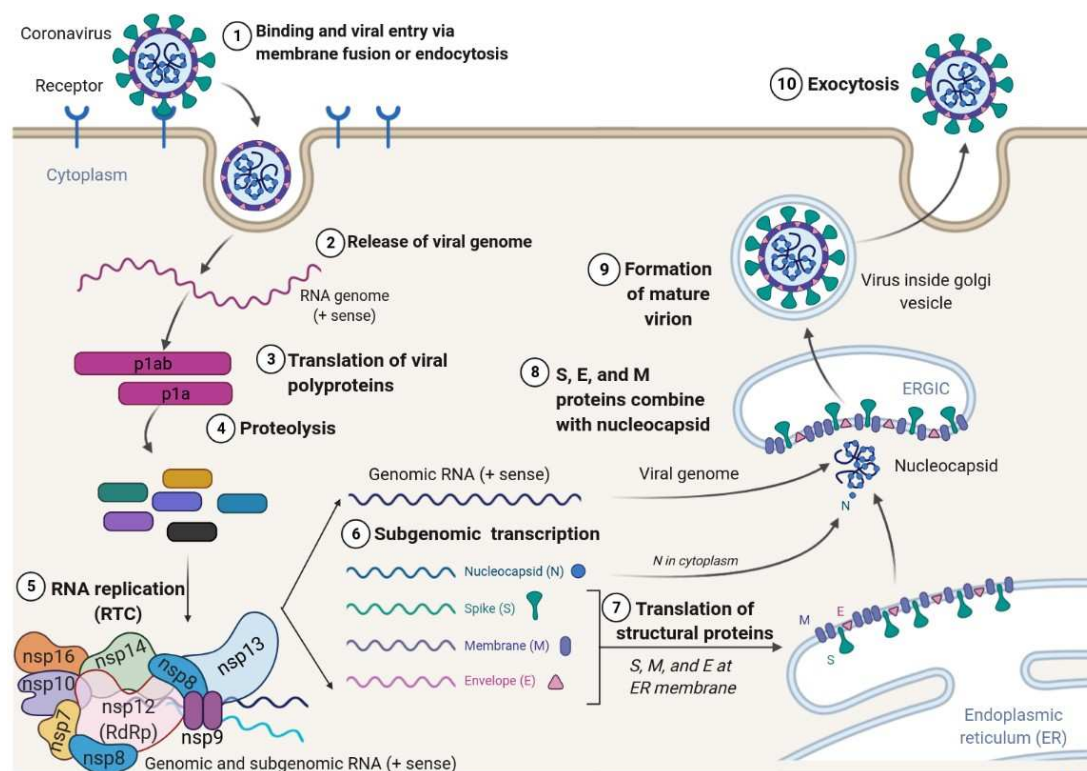
#### 4.2.1.1 Replication cycle

Coronavirus enters the host cell by receptor-mediated endocytosis utilizing a variety of entry receptors (figure 18). The initial attachment of the virion to the host cell is initiated by interactions between the S protein and its receptor. Many coronaviruses utilize aminopeptidase N (APN) as entry receptor. SARS-CoV, SARS-CoV-2 and HCoV-NL63 use angiotensin-converting enzyme 2 (ACE2). MERS-CoV binds to

dipeptidyl-peptidase 4 (DPP4) at the cell surface (164). The binding is mediated through a specific receptor-binding domain (RBD) exposed at the surface of S1 after the proteolytic cleavage of S protein into two sub-units S1 and S2. The S-protein/receptor interaction is the primary determinant for a coronavirus to infect a host species and governs the tissue tropism of the virus. The viral RNA is released into the cells following entry of the virus. Once in the cytoplasm the replication starts with the translation of the 5'-proximal ORFs of the viral genome corresponding to ORF1a and ORF1b. The translation of the ORFs by the cellular machinery results in the synthesis of two large replicase polyproteins: pp1a and pp1ab. Synthesis of pp1ab, a C-terminally extended form of pp1a, involves a ribosomal frameshift signal (RFS) occurring near the 3'end of ORF1a. This frameshift acts like a regulator to balance the abundance of replicase enzymes compared to other proteins and regulates the stoichiometry of nsPs (165). The two polyproteins pp1a and pp1ab then undergo proteolytic maturation allowing the production of 16 non-structural proteins. This process is driven by 1 or 2 papain-like protease domains (PL1<sup>pro</sup>) and a chymotrypsin-like enzyme also known as main protease (3CL<sup>pro</sup> or M<sup>pro</sup>) that reside in nsp3 and nsp5, respectively (165).

After polyprotein maturation, most of the nsps assemble to form the replication-transcription complex (RTC). The RTC uses the genomic (+) RNA as a template for replication of the new genomic RNA and transcription of sgRNA through genomic and sg (-) RNA intermediates. CoV RNA synthesis alternates between (-) and (+) RNAs similarly to all positive strand RNA viruses. Viral RNA intermediates, including dsRNA, have been identified in double-membrane vesicles (DMVs), which seem to be either completely closed or partially open and are connected to the ER. DMVs seem to be the only place where viral RNA synthesis occurs despite the presence of other conceivable replication organelles such as convoluted membranes and single-membrane vesicles. The nsp12, the viral RdRp, is the centrepiece of the RTC, along with its co-factors nsp7 and nsp8. The complex also includes nsp13, its helicase activity unwinds dsRNA in a 5' to 3' direction, and nsp14, which can excise

mismatching nucleotides incorporated in the RNA chain, since it harbours an exonuclease activity (166). Two essential modification are introduced on both sg and genomic RNA to mimic cellular mRNAs: a cap-1 structure at the 5' end and a polyadenylated tail at the 3' end. The capping mechanism involves the RNA triphosphatase (RTPase) and the GTase activity of nsp13; the MTase activity of nsp14 (N7-MTase) and 2'O-MTase of nsp16 (167). Following sgRNA synthesis, structural and accessory proteins are translated in the cytoplasm. The structural proteins S, M and E are inserted into the membrane of the ER. They move along the secretory pathway to the ER-Golgi intermediate compartment (ERGIC), where assembly of the new viral particle takes place. The viral genome is encapsidated by the N protein and buds across membranes of the ERGIC carrying E proteins to form mature virions (168). Following assembly, virions are transported to the cell surface in vesicles and released into the extracellular medium by exocytosis (figure 18). Recently, it has been suggested that budding of SARS-CoV-2 and MHV may also occur via lysosomal trafficking (186).





**Figure 18. Replication cycle of coronavirus.** The different steps (1 to 10) of the viral cycle are represented. See text for details. The image was created in *BioRender.com*.

#### 4.2.1.2 nsp3

The nsp3 is the largest multi-domain protein of coronavirus, with a molecular weight of 200 kDa. The organization of nsp3 varies within the CoV genera. Through bioinformatics analysis, it has been ‘split’ into at least 8 domains and two transmembrane conserved regions. At the N-terminus of nsp3 is located the ubiquitin-like domain 1 (Ubl1) and the Glu-rich acidic region also called nsp3a. These regions exist in all CoVs, despite a low (~15%) amino-acid sequence identity (143). Ubl1 is involved in the binding to ssRNA and the interaction with the N protein. A second ubiquitin-like subdomain (Ubl2) is located within nsp3. The exact functional role of the Ubl2 domain is not clear despite its conservation among the CoVs compared to Ubl1. A HVR links these two domains to the PL1<sup>pro</sup> in the case of the alpha-CoVs and in clade A of beta-CoVs. However, all other CoVs possess other papain-like protease 2 domain (PL2<sup>pro</sup>) (169). Both papain-like protease domains are responsible for the cleavage of the polyproteins 1a/1ab to release nsp1, nsp2, and nsp3 from the N-terminal region. The PL2<sup>pro</sup> was reported also to possess deubiquitinating and deISGylating activities. The deubiquitylation activity of the MHV nsp3 protein has been shown to be involved in the inactivation of TBK-1 preventing IRF3 translocation to the nucleus and activation of the IFN response (170). A conserved *Macro* domain (Mac1) is present between HVR and PL1<sup>pro</sup> domain in all the nsp3 of coronaviruses. Mac1 is an ADP-ribose binding domain and the function of this domain were described above. Mac1 is followed by a “SARS-unique domain” (SUD) (171). This region includes two *Macro*-like domains: the *Macro* domain II and III (Mac2 and Mac3) as well as a subdomain called: Domain Preceding Ubl2 and PL2 pro (DPUP). A region corresponding to parts of SUD was also found in MERS-CoV and MHV



(143). Mac2 and Mac3 has been show to be essential for SARS-CoV and SARS-CoV-2 replication/transcription and translation. Most of the functions of Mac2 and 3 are related to RNA binding. It has been proposed that the SUD domain regulates the antiviral response of the host cell by binding to RNA or DNA carrying G-quadruplex sequences (143). In beta-CoV, two domains tail PL1<sup>pro</sup>, the nucleic-acid binding (NAB) and the beta-CoVs-specific marker ( $\beta$ SM) domains. They are implicated in the binding to ssRNA and unwinding of dsDNA, respectively. The C-terminal region of nsp3 includes: two transmembrane regions (TM1 and TM2), an ectodomain (3Ecto), a Y1 and a CoV-Y domains. The TM1 and TM2 allow nsp3 to pass the ER membrane twice. It is thought that transmembrane regions and 3Ecto domain are important for PL2<sup>pro</sup> cleavage of the nsp3-4 polyprotein (172). The Y1 domain is a conserved domain among the *Nidovirales* order while the CoV-Y is conserved only in CoVs. Both domains are located in ER on the cytosolic side. Currently, the role of these two domains are not clear but were shown to be important for the binding to nsp4. Hence, nsp3 protein plays several important roles during CoVs infection by interaction with viral nsps and with the RNA to form the RTC. Furthermore, nsp3 antagonize the host innate immune response by counteracting different PMTs such as de-MAR/PARYlation, deISGylation or deubiquitination.

### 4.3 Coronavirus: treatments and vaccines.

To date, there is no registered effective antiviral treatment for any affecting human coronavirus. The outbreak of COVID-19 pandemic and the SARS-CoV-2 spread worldwide at an exponential rate led to the development of various drugs. Many inhibitory molecules from other viruses have been redirected against SARS-CoV-2. In addition, molecules in development against SARS-CoV and MERS-CoV have been also tested. One exception of the authorized drug against SARS-CoV-2 is remdesivir, an inhibitor of RdRp developed against Ebola virus. Remdesivir have been shown to inhibit MERS-CoV and SARS-CoV-2 polymerases *in vitro* (173,174). However, only

modest clinical efficacy and no impact on the survival rate of COVID-19 patients was reported. Other clinical approved antiviral drug tested against SARS-CoV-2 is the combination of Lopinavir and ritonavir. Both potent inhibitors of the human immunodeficiency virus (HIV) protease (175). Ritonavir is also an inhibitor of Lopinavir metabolism and therefore increases its half-life, which justifies their combination. Although HIV protease and SARS-CoV 3CL protease do not belong to the same family, a previous study showed that Lopinavir partially inhibits 3CL protease *in vitro* (176). In clinical trials, testing multiple molecules as lopinavir/ritonavir combination was found to inhibit SARS-CoV replication in cell culture, reducing patient symptoms. However, a study, using this combination, does not suggest any improvement in hospitalized patients with severe COVID-19 disease (175). Moreover, favipiravir, a purine analogue nucleotide, which inhibits the polymerases of various RNA viruses was also tested on SARS-CoV-2 infected patients. This study shown better results over lopiravir/ritonavir in the presence of inhaled IFN (177). Other small molecules with a broad-spectrum inhibitory action on viral entry of several viruses is umifenovir. Although its mechanism of action on SARS-CoV-2 entry is unknown, a study showed an improvement in patient's condition (178). The list of molecules tested against SARS-CoV-2 is constantly growing, but until now there is no effective approved treatment eliminating the virus or stopping the symptoms of the disease.

Thus far, driven forward by the COVID-19 pandemic, vaccine development has by far outpaced antiviral drug discovery. The majority of vaccine development strategies is based on using the S protein as an antigen to elicit a potent neutralizing antibody response. Antibodies generated against S protein in recovered SARS patients are immunodominant and long-lasting in humans and animals (179). Approaches previously used to develop CoV vaccines included DNA plasmids, nanoparticles, virus-like particles, viral vector preparations using adenovirus or vaccinia virus as platforms encoding viral antigens, chemically inactivated virus and live -attenuated virus (180). In 2020, innovative solutions using mRNA technology resulted in the

fast-track development and production of FDA-approved vaccines (181). Although, no vaccines against SARS and MERS have been approved and most of them did not progress beyond Phase I clinical trials. Four vaccines have been approved for use against COVID-19 by the European Medicines Agency (EMA). These include mRNA vaccines encoding the S protein of SARS-CoV-2 from Pfizer/BioNTech and Moderna, and adenovirus-vector vaccines encoding the SARS-COV-2 spike glycoprotein from AstraZeneca and Janssen/Johnson & Johnson. Based on the results of countries with advanced vaccination programs, like Israel and the United Kingdom, vaccination clearly reduces severe disease and the number of hospitalizations. However, the impact on preventing transmission between humans and re-infection of vaccinated people is still under evaluation. It is also too early to estimate how long last the protection derived from vaccination. Moreover, the circulation of new SARS-CoV-2 variants worldwide, carrying multiple mutations in the S protein, may affect the vaccine efficacy. Social and economic factors including compliance by the public and cost-effective production/distribution play an important role in successful vaccination campaigns, especially to achieve the threshold necessary to obtain herd-immunity. Moreover, not all individuals can receive vaccination such as young children, immunocompromised patients and other risk groups. Antiviral therapy can provide treatment of illness at the onset of symptoms, and/or when vaccination-induced protection is incomplete or inefficient. Antiviral drugs can be designed to target different viral components with broad-spectrum activity against multiple CoVs. The developed antiviral could be administered at the onset of a future outbreak of a novel pathogenic human CoV. Therefore, for a number of reasons, it remains critical to pursue antiviral drug development for CoVs

## 4.4 Objectives

The clinical outcome upon SARS-CoV-2 infection can vary from silent infection to mild, severe and ultimately lethal COVID-19. The detrimental progression of

COVID-19 is attributed to a hyper-inflammatory reaction called cytokine storm, leading to an increase in circulating levels of pro-inflammatory cytokines. The low efficiency of re-positioned drugs as anti-COVID treatment underlines the urge of discovery of novel therapeutic agents as well as other specific coronavirus drug targets. This would be considerably facilitated through the characterization and the accumulation of knowledge on coronavirus proteins. Previous studies pinpointed *Macro* domain as a good therapeutic target. Based on published evidence we propose that SARS-CoV-2 *Macro* domain is a valid anti-viral target against COVID -19. Functional characterization of SARS-CoV-2 *Macro* domain will define specific molecular determinants, crucial for antiviral drug design.

The aims of this part of the study is to design and test specific *Macro* domain ligands and inhibitors, determine their mechanism of action (ADP-ribose binding and /or hydrolysis) through a functional and biochemical characterization of SARS-CoV2 *Macro* domain.

The development of this project will rely on the following goals:

- i. Conduct a mutagenesis study targeting specifically the divergent residues between SARS-CoV and SARS-CoV-2, in order to pinpoint differences in their catalytic function in terms of ligand binding and substrate hydrolysis.
- ii. Select potential inhibitory molecules on the basis of ADP-ribose pocket occupancy.
- iii. Evaluate the effect of selected molecules on SARS-CoV and SARS CoV-2 *Macro* domains.

# **Results**

## 4.5 Article 4

### ***Macro domain residue F156: a hallmark of SARS-CoV-2 de-MARylation specificity***

**Oney Ortega Granda**<sup>a</sup>, Karine Alvarez<sup>a</sup>, Maria Mate<sup>a</sup>, Bruno Canard<sup>a,c</sup>, François Ferron<sup>a,c</sup> and **Nadia Rabah**<sup>a,b,#</sup>

<sup>a</sup> Aix Marseille Université, CNRS, AFMB UMR 7257, Marseille, France

<sup>b</sup> Université de Toulon, 83130 La Garde, France

<sup>c</sup> European Virus Bioinformatics Center, Jena, Germany, 07743.

# Address correspondence to Nadia Rabah, [nadia.rabah@univ-tln.fr](mailto:nadia.rabah@univ-tln.fr)

***In preparation***

**Abstract**

The Coronavirus disease 2019 (COVID-2019) pandemic currently provokes a global health and economic crisis due to the generalized spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). SARS-CoV-2 is a large, enveloped and positive sense single stranded RNA virus. Its genome encodes 16 non-structural proteins (nsp 1-16), forming a large replication complex anchored into the membrane. The largest protein of this complex is nsp3, a multi-domain protein that contains a well conserved *Macro* domain (also called X domain or ADP-ribose phosphatase domain). The *Macro* domain can bind to mono-ADP-ribose (MAR) and poly-ADP-ribose (PAR) in their free form or conjugated to a protein or RNA substrates. *Macro* domains also carry de-MARylation and de-PARylation activities, implicated in the inflammation process and the regulation of innate immunity. Herein, we report a mutagenesis study focusing on SARS-CoV-2 F156 and SARS-CoV N157 residues, stipulated important for ADP-ribose orientation within the binding cleft. Our data suggest that the exchange of these residues or their substitution to alanine slightly influence ADP-ribose binding, but drastically impact *Macro* domain de-MARylation activity.

**Introduction**

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is the virus associated with the current pandemic of coronavirus disease (COVID-2019). SARS-CoV-2 is a member of the family *Coronaviridae* and the subfamily *Coronavirinae* [1]. SARS-CoV-2 belongs to  $\beta$ -coronaviruses together with previously identified SARS-CoV and MERS-CoV. All of these  $\beta$ -coronaviruses have been associated with human fatal diseases [2]. SARS-CoV is the agent of SARS that caused 774 deaths and 8096 confirmed cases from 2002 to 2003 [3]. MERS-CoV emerged in 2012 spreading in the Middle East with 2494 positive cases and 858 deaths [4]. Although the mortality rate of SARS-CoV-2 (<2%) is low compared to SARS-CoV (9.6%) and MERS (35%), it spreads faster among humans [5]. Rapidly evolving vaccine strategies seem

effective, however they might be constantly hampered by the emergence of new variants [6].

SARS-CoV-2 is a positive-sense single-stranded RNA (+ssRNA) virus with a length of 29.9 kb [7] and encodes two large open reading frames (ORF1a and ORF1ab), the latter generated by a frame-shift close to the ORF1a end. They lead to the production of two large polypeptides subsequently processed to 16 non-structural proteins nsp1 to 16. The last third of the genome encodes viral structural proteins: spike (S), envelope (E), membrane (M), nucleoprotein (N) and accessory proteins. The translation of this last part of the genome (ORFs) occurs via RNA intermediates or subgenomic RNAs (sgRNA). Structural proteins will ensure genomic packaging, encapsidation, virion formation and release [8]. Amongst nsps produced by coronaviruses, nsp3 is the largest multi-domain protein with an average molecular mass of about 200 kD. Essential component of the replication/transcription complex (RTC), its architecture is however not absolutely conserved within CoV genera, due to duplication or the deletion of some domains. In addition to two conserved transmembrane regions (TM1 and TM2), eight regions remains conserved : the ubiquitin-like domain 1 (Ubl1), the hypervariable region (Glu-rich or acidic domain), at least one *Macro* domain (also named “X domain”), the ubiquitin-like domain 2 (Ubl2), the papain-like protease 2 (PL2<sup>pro</sup>), a zinc-finger domain (ectodomain), the Y1 and CoV-Y domains of unknown functions (PMID: 29128390). In SARS-CoV-2, three tandem *Macro* domains (Mac1 to Mac3) were described . *Macro* domains Mac2 and Mac3 interact with nucleic acids, whereas (Mac1) is involved ADP-ribose binding and hydrolysis [9].

ADP-ribosylation is a ubiquitous post-translational modification (PTM) affecting protein activity, interactions, ubiquitination and degradation. The reaction is driven by ADP-ribosyl transferases (ARDTs) and leads to the addition of one (MAR) or multiple/poly (PAR) ADP-ribose moieties onto charged amino acid residues, mostly aspartate or glutamate. Several groups of ARDTs were described but the best



characterized are poly-ADP-ribosylpolymerases (PARPs), which use nicotinamide adenine dinucleotide (NAD<sup>+</sup>) as a substrate[10]. PARPs family comprises 17 members in humans. Their ADP-ribosylation activities are implicated in DNA repair, chromatin remodelling, transcriptional regulation, cell signalling, inflammation and immune response [11]. PARPs 1, 10, 12, 13 and 14 are involved in the regulation of innate immunity and were shown to have an anti-viral effect *via*: (i) promoting viral proteins proteasome degradation; (ii) inhibiting translation machinery; (iii) inhibiting viral replication; (iv) stimulating the formation of stress granules; and (v) inducing interferon signalling and interferon stimulated genes (ISGs) [12,13]. However, ADP-ribosylation is a reversible PTM. Two enzyme families are able to hydrolyse ADP-ribosyl linkages: ADP-ribosyl hydrolases (ARHs) and *Macro* domains.

*Macro* domains are widely distributed among life kingdoms and are encoded by positive strand RNA viruses including *Togaviridae*, *Hepeviridae* and *Coronaviridae*, as mentioned above [14,15]. These domains are able to bind various mono-ADP-ribose derivatives, including ADP-ribose 1'' phosphate (Appr1p), O-acetyl-ADP-ribose and PAR, in a free form or conjugated to protein or RNA substrates [16,17]. Viral *Macro* domain were shown to bind ADP-ribose derivatives, but also to hydrolyse MARYlated (de-MARYlation) and, in some cases, PARYlated (de-PARYlation) substrates [18–22]. Moreover, a phosphatase activity against Appr1p was reported for alphaviruses, Human coronavirus 229E (HCoV-229E) and SARS-CoV [23–26]. Hydrolysis activities are believed to counter the above mentioned PARPs anti-viral effects. This assumption is supported by many *Macro* domain mutagenesis studies of key residues implicated in ADP ribose binding, which show a reduction of virus replication and virulence of Hepatitis E virus (HEV), alphaviruses and several coronaviruses [27]. Hence, murine hepatitis virus (MHV)'s *Macro* domain catalytic mutants failed to induce acute hepatitis and its growth is restricted in culture cells unless IFN receptor knockout cells are used or PARPs inhibitors are added [28,29]. In HCoV-229, the mutant virus becomes susceptible to type I and II IFN and is unable to suppress the activation of ISGs [30]. In mouse adapted SARS-CoV, mutation of ADP-

ribose binding pocket renders the virus more susceptible to cytokines, including IFN, TNF and IL-6, and protects mice from lethal infection outcome [18]. These data are supported by Heer *et al* [31] stating that SARS-CoV-2 infection alters PARP family gene expression and disrupts NAD<sup>+</sup> biosynthesis. The accumulated data highlight the importance of ADP-ribosylation in host-viral conflict, thus making *Macro* domain a suitable druggable target.

Amino-acid sequence comparison of nsp's domains between SARS-CoV-2 and SARS-CoV reveals that the SARS-CoV-2 *Macro* domain shares 71 % identity with SARS-CoV, hence being the most divergent amongst nsp [32,33]. The crystal structure of SARS-CoV-2 *Macro* domain was solved in its free form (apo) and complexed to various ligands including 2-(N-morpholino) ethanesulfonic acid (MES), AMP and ADP-ribose. The reported structures follow the classical architecture described for *Macro* domains, with 7  $\beta$ -sheets sandwiched between two layers of  $\alpha$ -helices [32–35]. Structural comparison studies between SARS-CoV-2 and SARS-CoV *Macro* domains pinpointed the residue F156 in SARS-CoV-2, corresponding to an N157 in SARS-CoV. The divergence at this position suggest a considerate change in the orientation of ADP-ribose adenine group in the binding pocket between the *Macro* domains of these two viruses, which might impact ADP-ribose binding and/or hydrolysis kinetics.

In the present study, we focus on SARS-CoV-2 F156 and SARS-CoV N157 to evaluate the impact on the stability and activity of these residues on the *Macro* domain and their relation with their natural ligand.

## **Results and discussions**

### **ADP-ribose binding coordination**

The overall structure of the CoV *Macro* domain consisted of six  $\alpha$ -helices and one seven-stranded  $\beta$ -sheet. The  $\beta$ -sheet ( $\beta$ 1– $\beta$ 2– $\beta$ 7– $\beta$ 6– $\beta$ 3– $\beta$ 5– $\beta$ 4) is topped and bottomed by 3  $\alpha$ -helices ( $\alpha$ 1,  $\alpha$ 2, and  $\alpha$ 3 and  $\alpha$ 4,  $\alpha$ 5, and  $\alpha$ 6 respectively) thus

delimiting a groove where the ligand binds (Figure 1). Comparative protein sequence analysis of SARS-CoV and SARS-CoV-2 *Macro* domains (Figure 1A) reveals that most of the differences are located at the extremities of the domain namely the  $\alpha 1$  helix and  $\alpha 1$ - $\beta 3$  loop at the N-terminus, and the  $\alpha 6$  helix at the C-terminal. Divergent residues between the two close related viruses represent ~30 % of the sequence in accordance to previous reports [32,33]. The structural basis of *Macro* domain and ADP-ribose interactions have been characterized [23,24,32–35]. These studies pointed out that the ADP-ribose molecule is partially buried in a chief hydrophobic cleft encompassing, in the case of SARS-CoV-2, the C-terminal end of  $\beta$  strands  $\beta 3$ ,  $\beta 5$ ,  $\beta 6$ ,  $\beta 7$  and two loops,  $\beta 3$ - $\alpha 2$  and  $\beta 6$ - $\alpha 5$  (Figure 1B and C). The cleft accommodating ADP-ribose creates four contact zones. The first forms a hydrophobic patch between the residues I23, V49, P125, V155 and F156 (all conserved except the later), which interacts with adenine and points it toward the polar D at position 22. Residues interacting directly with ADP-ribose are well conserved among various *Macro* domains. The catalytic residue D22 correspond to D23 in SARS-CoV (Figure 1E) and is present in other *Coronavirinae* members, many alphaviruses and non viral *Macro* domains [27,36–38]. Mutagenesis studies showed that these catalytic residues are crucial for ADP-ribosyl-hydrolase and ADP-ribose-1'' phosphatase activities, they drastically impact viral replication and virulence (reviewed in [27]). The second, a conserved glycine rich stretch (residues 46-48) accommodates the diphosphate moiety of ADP-ribose. Phosphate groups connecting adenosine moiety to the distal ribose interact with various backbone amide nitrogen atoms of residues within the loop  $\beta 3$ - $\alpha 2$  and  $\beta 6$ - $\alpha 5$  region. The  $\alpha$ -phosphate group forms oxygen-hydrogen bonds with V49 and I131 while  $\beta$ -phosphate with S128, G130 and F132 (Figure 1 D and E). Distal ribose fits tightly into the pocket formed by the nitrogen bonds between the backbone amide of G48, G 46 and N 40 with the ribose-1,2 and 3'' oxygen atom. The third, implicates F132 and I131, which stabilize the proximal ribose. The forth, involving L136, A154 and D157, supports the distal ribose *via* water-hydrogen bonding [32–35] .

In SARS-CoV-2 the adenosine moiety is partially covered by F156, which interacts edge-to-face with the aromatic ring system (Figure 1C and D). This residue corresponds to N157 in SARS-CoV (Figure 1E). The proximity of phenylalanine to the adenosine ring is only observed in SARS-CoV-2 among *betacoronaviruses*, even if other hydrophobic residues are present at that position. These data suggest the possible involvement of F156 in the differential alignment of ADP-ribose adenine group in the binding pocket of SARS-CoV-2.

### **Expression and Purification of recombinant *Macro* domain proteins**

In order to understand the influence of F 156 on ADP-ribose binding and hydrolysis kinetics, we generated several substitutions leading to: (i) the exchange of F 156 and N157 between *Macro* domains of SARS-CoV-2 and SARS-CoV, respectively; (ii) the substitution of the target residues to alanine. Recombinant *Macro* domain proteins, with a N-terminal 6 x His tag, of 172 amino acids (aa) and 168 aa, corresponding to SARS-CoV and SARS-CoV-2 respectively, were successfully expressed in *E. coli* (Figure 2A) and purified. Bands of about 20kDa were observed in soluble purification fractions, and enriched throughout the purification process (Figure 2B). Despite being a 20.8 kDa protein, the *Macro* domain of SARS-CoV-2 migrates slightly higher than that of SARS-CoV (20.9 kDa). The difference in the gel migration could be related to the difference in SDS adsorption of these two proteins. Besides, SARS-CoV-2 shows slight differences in purification buffer preferences as compared to SARS-CoV.

First, SARS-CoV-2 was significantly lost during the washing steps at the imidazole concentration used for SARS-CoV (60 mM). Thus, imidazole concentration was decreased to 30 mM. In contrast, salt concentration had to be increased to eliminate more tightly bound contaminants. After two purification steps (IMAC and gel filtration), proteins of interest were pure, as evidenced by SDS-PAGE after Coomassie blue staining (Figure 2C). Identities of WT SARS-CoV-2 and SARS-CoV recombinant *Macro* domains were confirmed by Matrix Assisted Laser Desorption Ionization - Time of Flight (MALDI-TOF) mass spectrometry.

### **CoV *Macro* domains ADP-ribose binding assessment**

Previous studies investigated *Macro* domain binding parameters towards ADP ribose using ITC for various human pathogenic coronaviruses, including SARS-CoV and SARS-CoV-2. The reported  $K_d$  values oscillate between 10 and 24  $\mu\text{M}$  for SARS-CoV, and between 10 and 17  $\mu\text{M}$  for SARS-CoV-2 [23,32,33,36]. The differences being due to slight variations in experimental protocols and constructs design. In order to accurately evaluate the characteristics of generated mutants, we started by setting the references values for our WT *Macro* domain proteins. Hence, dot blot assays were conducted using automodified hPARP-1 and increased amounts of WT *Macro* domains. PAR generated by hPARP-1 was able to bind to SARS-CoV and SARS-CoV-2 WT proteins, and detected using an anti-PAR antibody, in a concentration dependent manner (Figure 3A, lower panel). The quantified binding signal displays drastic differences, up to 75 %, in PAR binding of SARS-CoV regarding SARS-CoV-2 at various *Macro* domain concentrations (Figure 3A, upper panel). Next, we determined the dissociation constants and thermodynamic parameters of ADP-ribose by ITC (Figure 3B). The results are compiled in Table 1. In our hands computed  $K_d$  values for SARS-CoV and SARS-CoV-2 correspond to  $5.9 \pm 2.9 \mu\text{M}$  and  $11.5 \pm 5.7 \mu\text{M}$ , respectively.

To determine if the presence F156, SARS-CoV-2 could cause hindrance, preventing ADP-ribose binding into the pocket,  $K_d$  values were measured for the different mutants (Table 1). Substitution of F156N in SARS-CoV-2 and N157F in SARS-CoV affected slightly ADP-ribose binding affinities (Table1). Thus, computed values being  $6.17 \pm 1.66 \mu\text{M}$  for SARS-CoV N157F and  $9.17 \pm 1.14 \mu\text{M}$  for SARS-CoV-2 F156N. Alanine replacement in SARS-CoV N157A led to a 2 fold increase in  $K_d$  value ( $13.5 \pm 1.67 \mu\text{M}$ ) compared with the WT protein. Conversely, SARS-CoV-2 F156A substitution had minor repercussions in ADP-ribose affinity, with a  $K_d$  of  $9.8 \pm 3.43 \mu\text{M}$ . Hence, even SARS-CoV-2 F156N substitution drives the mutant toward SARS-

CoV K<sub>d</sub> value, single amino acid substitution is not sufficient, in this context, for a complete mimicking of SARS-CoV ADP ribose binding characteristics. Interestingly, the simultaneous mutation of SARS-CoV-2 V24I/E25Q/F156N, mimicking MERS-CoV sequence, generated a *Macro* domain with K<sub>d</sub> value similar to that obtain for MERS-CoV (~3  $\mu$ M) [36]. In SARS-CoV-2 the proximity of V24, E25, and F156 to D22 may impact the strength of hydrogen bonds forming, which could impact the side chain orientation of ADP-ribose adenine group. These data support the necessity of a synergistic effect between the molecular determinants involved in ADP ribose binding (Figure 1 D and E).

### **Altered de-MARylation in CoV *Macro* domains mutants**

Viral *Macro* domains possess MAR and PAR hydrolytic activity leading to the removal of ADP ribose derivatives from MARylated and PARylated proteins [39–41]. To assess the effect of CoV *Macro* domain mutation on protein de-MARylation, catalytic PARP3 protein was used as a substrate in time course hydrolysis reactions (Figure 4). MARylated PARP3 was incubated in the presence of various CoV *Macro* domains mutants. After the indicated time points, the reactions were subjected to mPAGE. Western blot using the anti-MAR reagent allows de-MARylation assessment by the removal of the MAR signal (Figure 4A, upper panel). CoV *Macro* domain load was controlled by Coomassie blue staining (Figure 4A, lower panel). Load controls for PARP3, corresponding to Ponceau membrane staining are shown in (Figure 4A, middle panel). Auto-MARylated PARP3 without CoV *Macro* domain served as a control for zero hydrolysis. Bands intensities were quantified and fitted in nonlinear regression curves (Figure 4B).

SARS-CoV-2 and SARS-CoV *Macro* domain are able to remove almost completely MAR from PARP3 within 30 minutes. Substrate decay is more important in SARS-CoV-2 (70-90 % loss after 5-10 minutes) compared to SARS-CoV (30-70 %) (Figure 4B), in accordance with recently reported data [35]. The effect of *Macro* domain mutations on PARP3 de-MARylation is intriguing. Hence, SARS-CoV-2 F156A

mutant shows a complete loss in hydrolysis activity for the first 10 min. MAR hydrolysis reaches hardly 40 % of WT value after 30 min. Alanine substitution in SARS-CoV has also a negative effect on de-MARylation as observed for SARS-CoV N157A mutant, at 10 min de-MARylation activity is reduced by half. Substitution of the conserved D22 and D23 by alanine in SARS-CoV-2 and SARS-CoV, added to the study as negative controls for hydrolysis activity, abolish totally the de-MARylation activity, corroborating the crucial role of this aspartic acid in the catalysis. Intriguingly, SARS-CoV N157F and SARS-CoV-2 F156N mutants behaved as the mimicked parental *Macro* domain. Hence, SARS-CoV N157F mutant shows an increased hydrolysis activity similar to that of SARS-CoV-2; whereas SARS-CoV-2 F156N MAR hydrolysis was impeded, reaching SARS-CoV values. The aromatic structure of phenylalanine in the hydrophobic cavity creates a stacking interaction with the adenine moiety, thus stabilizing the ADP-ribose in the groove of SARS-CoV-2 *Macro* domain.

To link the observed activities to a potential modification of mutant proteins thermostability, CoV *Macro* domain sequences were subjected to HoTMuSiC tool, to evaluate the changes in melting temperature ( $\Delta T_m$ ) under point mutations, on the basis of its experimental 3D structure [42]. The obtained predictions, listed in Table 2, corroborate de-MARylation activity results. Thus, SARS-CoV-2 F156N has a negative  $\Delta T_m$  (-3.77) value, causing a destabilization of the protein; while SARS-CoV N157F displays a  $\Delta T_m$  of +1.01, favoring thermostability. Alanine substitution impact negatively protein stability with computed  $\Delta T_m$  values of - 3.39 and -0.49 for SARS-CoV-2 F156A and SARS-CoV N157A, respectively. These data suggest that F156, in SARS-CoV-2, and its corresponding N157 residue in SARS-CoV are key players in *Macro* domain de-MARylation activity.

Despite a high similarity in terms of sequence and function among viral *Macro* domains, notable differences in their affinity for ADP-ribose and catalytic activity do exist, affecting the state of ADP-ribosylation substrates in cell. Many studies linked viral *Macro* domain ADP-ribose hydrolase activity to viral pathogenesis. Catalytic

mutations targeting residues D23, H38, N41, H46 and G131 in SARS-CoV (Figure 2E) and equivalent positions in other RNA viruses, were shown deleterious for ADR-ribose binding/hydrolysis activity, PARP activation, and viral virulence [22,23,28,32,34]. As an example, in the case of SARS-CoV N41 catalytic residue, the *Macro* domain mutant N41A is devoid of MAR-hydrolase activity and elicits more significant IFN, ISGs and pro-inflammatory cytokines production than the wild type virus in infected mice [18]. In MHV, mutation of the catalytic D residue (D16, numbered D1329 in ORF1a polyprotein) to A has a severe impact on virus replication and the mutant virus is defective in blocking IFN production. The phenomenon is emphasized when this mutation is combined to the substitution N30A (numbered N1347A in ORF1a polyprotein). The later corresponds to N 40 in SARS-CoV-2, involved in distal ribose interaction[43]. A recent study reports the impact of SARS-CoV-2 N40 mutation on PARP9 activation and IFN signalling [44], highlighting the importance of this residue in immune escape *via* STAT signalling.

In contrast, residues outside the adenine binding pocket are barely investigated. Yet, studies on Chikungunya virus (CHIKV) *Macro* domain showed that the mutation of Y114, corresponding to F132 in SARS-CoV-2 and situated in the stretch accommodating the diphosphate moiety of ADP-ribose, impaired MAR hydrolysis activity but not ADP-ribose binding, hampering though viral replication[22]. Interestingly, F132L replacement in SARS-CoV-2 impeded MAR hydrolysis activity, as in CHIKV, but F156L substitution didn't show a drastic effect on ADR-ribose hydrolysis from autoribosylated PARP14 WWE-CAT [45]. This data supports our finding, emphasizing the importance of the hydrophobic patch at that position in SARS-Cov-2 and highlights the fact that many questions remain regarding the role of non-conserved and/ or non-catalytic residues.

Finally, mutational analysis conducted in the present paper led to the decoupling ADP-ribose binding affinities and MAR activity for of SARS-CoV and of SARS-CoV-2 *Macro* domains. Moreover, it pinpointed the importance of F157 residue in



SARS-CoV-2 and N156 residue in SARS-CoV-2 for MAR hydrolysis, broadening our knowledge on *Macro* domains function, a prerequisite for anti-viral drug design.

### ***Material and methods***

#### **Expression and Purification of the SARS-CoVs *Macro* domain**

The cDNA encoding nsp3 sequence of SARS-CoV (residues 182 to 355, GenBank #AY291315) and SARS-CoV-2 (residues 207 to 375, NCBI accession YP\_009725299.1) were codon optimized and cloned into the pET28 (Novagen)-TWIST Bioscience. The recombinant proteins were expressed in the competent *E. coli* (C41 (DE3) plys). A detailed protocol on expression and purification of SARS-CoV is already reported [46]. The expression of SARS-CoV-2 recombinant protein was done in Turbo Broth medium (Cat#0104 AthernaES). The induction was carried with 50µM of IPTG (O.D<sub>600 nm</sub> of 0.6) at 25°C for 12-14 hours. At the end of the incubation time, the cultures were centrifuged at 9000 rpm at 4°C for 30 min, and the pellets were kept at -80°C until purification.

The purification of *Macro* domain of SARS-CoV was done following the conditions previously reported [46]. The bacterial pellet from *Macro* domain of SARS-CoV-2 was resuspended in lysis buffer (50 mM Tris-HCl, pH 8, 150 mM NaCl, 20 mM MgSO<sub>4</sub>, 5 % Glycerol, 20 µg/mL DNase, 0.25mg/mL lysozyme and 1 mM PMSF) at 4°C for 1 hour. The cell lysate was sonicated (amplitude- 40 %, 4x 10s on /5s off) and the supernatant was collected. Immobilized metal affinity chromatography (IMAC) purification in batch was performed using Ni-NTA column (GE Healthcare). The filtered supernatant was loaded into the column per-equilibrated with binding buffer (50 mM Tris-HCl, pH 8, 150 mM NaCl, 5 % Glycerol and 10mM Imidazole). The column was then washed with 10 column volume (CV) of wash buffer (50 mM Tris-HCl, pH 8, 150mM NaCl, 5 % Glycerol and 30 mM Imidazole) followed by second wash with high salt concentration wash buffer supplemented with 1 M NaCl. Recombinant protein was eluted with elution buffer (50 mM Tris-HCl, pH 8, 150 mM NaCl, 5 % Glycerol and 200 mM Imidazole). The fractions containing the purified protein were then pooled and dialysed against GF buffer (50 mM Tris-HCl pH 8, 150

mM NaCl, 5 % glycerol). In addition, a second purification step of size-exclusion chromatography were performed by injecting the protein into a Superdex 75 (GE Healthcare) equilibrated with GF buffer. The eluted protein was concentrated and stored at -80°C. Mutants *Macro* domains were expressed and purified following the same conditions as the corresponding wild-type recombinant proteins.

### **Poly (ADP-ribose) synthesis**

PARylated protein was synthesized by auto-ADP-ribosylation of human PARP-1 (hPARP-1) (Sigma) in 300 µL reaction volume in the presence of 100 mM Tris pH 8, 10 mM MgCl<sub>2</sub>, 2 mM DTT, 0.8 units hPARP-1 and 250 µM NAD<sup>+</sup>. The reaction was carried at room temperature (RT) for 2 h under moderate agitation. PARylation was stopped by diluting the reaction in 20mL of dot blot buffer (150 mM NaCl and 0.05 % Tween). The diluted reactions were directly used for the *Macro* domain binding assay.

### **Binding assay**

Binding affinity of ADP-ribose to SARS-CoV and SARS-CoV-2 was performed in a nitrocellulose membrane (Schleicher & Schuell). The binding reactions were carried out by spotting serial dilutions (from 250 to 1.9 pmol) of recombinant *Macro* domain proteins on nitrocellulose membrane using Minifold II dot blot apparatus (Schleicher & Schuell). Bovine serum albumin (BSA) was used as negative control for the assay. The membrane was blocked with dot blot buffer supplemented with 4% skimmed milk, for 1 h, at RT. The blocked membranes were incubated with auto-ADP-ribosylated hPARP-1 for 1h at RT under constant agitation. The unbound material was removed by three extensive washes with dot blot buffer. The primary antibody anti-PAR binding reagent (Sigma) was used diluted 1:1500 in dot blot buffer with 1 % non-fat milk for 2 h, at RT. The secondary antibody, anti-rabbit IgG (Dako) /, was diluted 1:2000 in dot blot buffer and incubated with the membrane for 1 h, at RT. The membrane was washed three times in dot blot buffer after each anti-body. Immunoreactive signals were revealed using ECL reagent (Cat. # 170–5061, Bio-Rad)

and visualized using Amersham™ ImageQuant™ 800 Imager system. Images were analyzed and quantified using ImageJ software [47]. Each experiment was repeated three times. Band's intensity was normalized to the total protein load.

### **Isothermal titration calorimetry (ITC)**

ITC experiments were performed using VP-ITC (MicroCal Inc., Northampton, MA) at 20°C. Purified recombinant SARS-CoV and SARS-CoV-2 WT and mutant proteins were used at 150 µM. The interaction was assessed in GF buffer, using 75 to 2000 µM of ADP-ribose as injected ligand. The curves were generated from the integrated data of injection peaks plotted against ADP-ribose to protein molar ratio. All data were fit in Origin 7.0 (MicroCal), provided by the manufacturer. The dissociation constant (K<sub>d</sub>), stoichiometry (N) and thermodynamic parameters (enthalpy  $\Delta H$  and entropy  $\Delta S$ ) were determined.

### **De-MARylation assay**

MARylated PARP3 was obtained by incubating 10 µM PARP3 in 25 mM Tris pH 8, 100 mM NaCl, 2 mM DTT, and 200 µM NAD<sup>+</sup> for 30 minutes at 30°C. De-MARylation activity was assessed by incubating 1 µM of MARylated PARP3 with 150 nM of recombinant *Macro* domains at 37°C for a period of 0 to 30 min in the same buffer. The reaction was stopped by adding mPAGE™ 4X LDS Millipore sample buffer and heating at 95°C for 5 min. Autoribosylated PARP3 protein without any *Macro* domain served as negative control. Reactions were loaded on a 4 to 20 % mPAGE bis-tris gels precasted gel. After migration, protein bands were transferred to Immobilon-PSQ PVDF membrane (#170–4159, MERCK) overnight at 4°C. The membranes were blocked with 4 % non-fat milk in TBS-Tween buffer for 1 h. The primary antibody anti-MAR reagent was used at a final dilution of 1/2500 in TBS-T with 1 % non-fat milk for 1h. The membrane was then processed as described above in the dot blot binding assay. Quantification was done using ImageJ program. The results were normalized to the value of the negative control.

**Author Contributions:** Conceptualization, O.O.G.; F.F. and N.R.; methodology, O.O.G. and M.M.; software, O.O.G.; F.F. and N.R.; investigation, O.O.G.; F.F. and N.R.; resources, B.C., K.A. and N.R.; writing original draft preparation, O.O.G. and N.R.; writing review and editing, O.O.G.; K.A.; M.M.; D.C.; F.F. and N.R.; project All authors have read and agreed to the published version of the manuscript.

**Funding:** O.O.G. is a recipient of Méditerranée Infection Fondation doctoral studentship.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- 1 Pal M, Berhanu G, Desalegn C & Kandi V (2020) Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update. *Cureus*.
- 2 Lu L, Zhong W, Bian Z, Li Z, Zhang K, Liang B, Zhong Y, Hu M, Lin L, Liu J, Lin X, Huang Y, Jiang J, Yang X, Zhang X & Huang Z (2020) A comparison of mortality-related risk factors of COVID-19, SARS, and MERS: A systematic review and meta-analysis. *Journal of Infection*, S0163445320304606.
- 3 Cherry JamesD (2004) The chronology of the 2002–2003 SARS mini pandemic. *Paediatric Respiratory Reviews* 5, 262–269.
- 4 Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, Wang W, Song H, Huang B, Zhu N, Bi Y, Ma X, Zhan F, Wang L, Hu T, Zhou H, Hu Z, Zhou W, Zhao L, Chen J, Meng Y, Wang J, Lin Y, Yuan J, Xie Z, Ma J, Liu WJ, Wang D, Xu W, Holmes EC, Gao GF, Wu G, Chen W, Shi W & Tan W (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *The Lancet* 395, 565–574.
- 5 Fani M, Teimoori A & Ghafari S (2020) Comparison of the COVID-2019 (SARS-CoV-2) pathogenesis with SARS-CoV and MERS-CoV infections. *Future Virology* 15, 317–323.
- 6 Kim JH, Marks F & Clemens JD (2021) Looking beyond COVID-19 vaccine phase 3 trials. *Nature Medicine* 27, 205–211.
- 7 Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, Hu Y, Tao Z-W, Tian J-H, Pei Y-Y, Yuan M-L, Zhang Y-L, Dai F-H, Liu Y, Wang Q-M, Zheng J-J, Xu L, Holmes EC & Zhang Y-Z (2020) A new coronavirus associated with human respiratory disease in China. *Nature* 579, 265–269.
- 8 Kirtipal N, Bharadwaj S & Kang SG (2020) From SARS to SARS-CoV-2, insights on structure, pathogenicity and immunity aspects of pandemic human coronaviruses. *Infect Genet Evol* 85, 104502.
- 9 Tan J, Vonnrhein C, Smart OS, Bricogne G, Bollati M, Kusov Y, Hansen G, Mesters JR, Schmidt CL & Hilgenfeld R (2009) The SARS-Unique Domain (SUD) of SARS Coronavirus Contains Two Macrod domains That Bind G-Quadruplexes. *PLoS Pathog* 5, e1000428.
- 10 Alhammad YMO & Fehr AR (2020) The Viral Macrod domain Counters Host Antiviral ADP-Ribosylation. *Viruses* 12, 384.
- 11 Brady PN, Goel A & Johnson MA (2018) Poly(ADP-Ribose) Polymerases in Host-Pathogen Interactions, Inflammation, and Immunity. *Microbiol Mol Biol Rev* 83.
- 12 Wang Y, Tang Q, Maul GG & Yuan Y (2006) Kaposi's sarcoma-associated herpesvirus ori-Lyt-dependent DNA replication: dual role of replication and transcription activator. *J Virol* 80, 12171–12186.
- 13 Atasheva S, Akhrymuk M, Frolova EI & Frolov I (2012) New PARP gene with an anti-alphavirus function. *J Virol* 86, 8147–8160.
- 14 Lei J, Kusov Y & Hilgenfeld R (2018) Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. *Antiviral Research* 149, 58–74.

- 15 Rack JGM, Perina D & Ahel I (2016) Macrodomains: Structure, Function, Evolution, and Catalytic Activities. *Annu Rev Biochem* **85**, 431–454.
- 16 Munnur D & Ahel I (2017) Reversible mono-ADP-ribosylation of DNA breaks. *FEBS J* **284**, 4002–4016.
- 17 Munnur D, Bartlett E, Mikolčević P, Kirby IT, Matthias Rack JG, Mikoč A, Cohen MS & Ahel I (2019) Reversible ADP-ribosylation of RNA. *Nucleic Acids Research* **47**, 5658.
- 18 Fehr AR, Channappanavar R, Jankevicius G, Fett C, Zhao J, Athmer J, Meyerholz DK, Ahel I & Perlman S (2016) The Conserved Coronavirus Macrodomain Promotes Virulence and Suppresses the Innate Immune Response during Severe Acute Respiratory Syndrome Coronavirus Infection. *mBio* **7**, e01721-16.
- 19 Egloff M-P, Malet H, Putics A, Heinonen M, Dutartre H, Frangeul A, Gruez A, Campanacci V, Cambillau C, Ziebuhr J, Ahola T & Canard B (2006) Structural and Functional Basis for ADP-Ribose and Poly(ADP-Ribose) Binding by Viral Macro Domains. *Journal of Virology* **80**, 8493–8502.
- 20 Li C, Debing Y, Jankevicius G, Neyts J, Ahel I, Coutard B & Canard B (2016) Viral Macro Domains Reverse Protein ADP-Ribosylation. *J Virol* **90**, 8478.
- 21 Malet H, Dalle K, Brémond N, Tocque F, Blangy S, Campanacci V, Coutard B, Grisel S, Lichère J, Lantéz V, Cambillau C, Canard B & Egloff M (2006) Expression, purification and crystallization of the SARS-CoV macro domain. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**, 405–408.
- 22 McPherson RL, Abraham R, Sreekumar E, Ong S-E, Cheng S-J, Baxter VK, Kistemaker HAV, Filippov DV, Griffin DE & Leung AKL (2017) ADP-ribosylhydrolase activity of Chikungunya virus macrodomain is critical for virus replication and virulence. *Proc Natl Acad Sci U S A* **114**, 1666–1671.
- 23 Egloff M-P, Malet H, Putics A, Heinonen M, Dutartre H, Frangeul A, Gruez A, Campanacci V, Cambillau C, Ziebuhr J, Ahola T & Canard B (2006) Structural and functional basis for ADP-ribose and poly(ADP-ribose) binding by viral macro domains. *J Virol* **80**, 8493–8502.
- 24 Malet H, Coutard B, Jamal S, Dutartre H, Papageorgiou N, Neuvonen M, Ahola T, Forrester N, Gould EA, Lafitte D, Ferron F, Lescar J, Gorbalenya AE, de Lamballerie X & Canard B (2009) The Crystal Structures of Chikungunya and Venezuelan Equine Encephalitis Virus nsP3 Macro Domains Define a Conserved Adenosine Binding Pocket. *J Virol* **83**, 6534–6545.
- 25 Putics Á, Filipowicz W, Hall J, Gorbalenya AE & Ziebuhr J (2005) ADP-Ribose-1"-Monophosphatase: a Conserved Coronavirus Enzyme That Is Dispensable for Viral Replication in Tissue Culture. *Journal of Virology* **79**, 12721–12731.
- 26 Saikatendu KS, Joseph JS, Subramanian V, Clayton T, Griffith M, Moy K, Velasquez J, Neuman BW, Buchmeier MJ, Stevens RC & Kuhn P (2005) Structural Basis of Severe Acute Respiratory Syndrome Coronavirus ADP-Ribose-1"-Phosphate Dephosphorylation by a Conserved Domain of nsP3. *Structure* **13**, 1665–1675.

- 27 Fehr AR, Jankevicius G, Ahel I & Perlman S (2018) Viral Macrodomains: Unique Mediators of Viral Replication and Pathogenesis. *Trends in Microbiology* **26**, 598–610.
- 28 Grunewald ME, Chen Y, Kuny C, Maejima T, Lease R, Ferraris D, Aikawa M, Sullivan CS, Perlman S & Fehr AR (2019) The coronavirus macrodomain is required to prevent PARP-mediated inhibition of virus replication and enhancement of IFN expression. *PLoS Pathog* **15**.
- 29 Eriksson KK, Cervantes-Barragán L, Ludewig B & Thiel V (2008) Mouse Hepatitis Virus Liver Pathology Is Dependent on ADP-Ribose-1''-Phosphatase, a Viral Function Conserved in the Alpha-Like Supergroup. *J Virol* **82**, 12325–12334.
- 30 Kuri T, Eriksson KK, Putics A, Züst R, Snijder EJ, Davidson AD, Siddell SG, Thiel V, Ziebuhr J & Weber F (2011) The ADP-ribose-1''-monophosphatase domains of severe acute respiratory syndrome coronavirus and human coronavirus 229E mediate resistance to antiviral interferon responses. *Journal of General Virology*, **92**, 1899–1905.
- 31 Heer CD, Sanderson DJ, Voth LS, Alhammad YMO, Schmidt MS, Trammell SAJ, Perlman S, Cohen MS, Fehr AR & Brenner C (2020) Coronavirus infection and PARP expression dysregulate the NAD metabolome: An actionable component of innate immunity. *J Biol Chem* **295**, 17986–17996.
- 32 Lin M-H, Chang S-C, Chiu Y-C, Jiang B-C, Wu T-H & Hsu C-H (2020) Structural, Biophysical, and Biochemical Elucidation of the SARS-CoV-2 Nonstructural Protein 3 Macro Domain. *ACS Infect Dis*.
- 33 Frick DN, Viridi RS, Vuksanovic N, Dahal N & Silvaggi NR (2020) Molecular Basis for ADP-Ribose Binding to the Mac1 Domain of SARS-CoV-2 nsp3. *Biochemistry*.
- 34 Michalska K, Kim Y, Jedrzejczak R, Maltseva NI, Stols L, Endres M & Joachimiak A (2020) Crystal structures of SARS-CoV-2 ADP-ribose phosphatase: from the apo form to ligand complexes. *IUCrJ* **7**, 814–824.
- 35 Alhammad YMO, Kashipathy MM, Roy A, Gagné J-P, McDonald P, Gao P, Nonfoux L, Battaile KP, Johnson DK, Holmstrom ED, Poirier GG, Lovell S & Fehr AR (2021) The SARS-CoV-2 Conserved Macrodomain Is a Mono-ADP-Ribosylhydrolase. *Journal of Virology* **95**.
- 36 Cho C-C, Lin M-H, Chuang C-Y & Hsu C-H (2016) Macro Domain from Middle East Respiratory Syndrome Coronavirus (MERS-CoV) Is an Efficient ADP-ribose Binding Module: CRYSTAL STRUCTURE AND BIOCHEMICAL STUDIES. *J Biol Chem* **291**, 4894–4902.
- 37 Allen MD, Buckle AM, Cordell SC, Löwe J & Bycroft M (2003) The Crystal Structure of AF1521 a Protein from *Archaeoglobus fulgidus* with Homology to the Non-histone Domain of MacroH2A. *Journal of Molecular Biology* **330**, 503–511.
- 38 Rack JGM, Zorzini V, Zhu Z, Schuller M, Ahel D & Ahel I Viral macrodomains: a structural and evolutionary assessment of the pharmacological potential. *Open Biology* **10**, 200237.



- 39 Rosenthal F, Feijs KLH, Frugier E, Bonalli M, Forst AH, Imhof R, Winkler HC, Fischer D, Caflisch A, Hassa PO, Lüscher B & Hottiger MO (2013) Macrodomain-containing proteins are new mono-ADP-ribosylhydrolases. *Nat Struct Mol Biol* **20**, 502–507.
- 40 Eckei L, Krieg S, Bütepage M, Lehmann A, Gross A, Lippok B, Grimm AR, Kümmerer BM, Rossetti G, Lüscher B & Verheugd P (2017) The conserved macrodomains of the non-structural proteins of Chikungunya virus and other pathogenic positive strand RNA viruses function as mono-ADP-ribosylhydrolases. *Sci Rep* **7**.
- 41 Jankevicius G, Hassler M, Golia B, Rybin V, Zacharias M, Timinszky G & Ladurner AG (2013) A family of macrodomain proteins reverses cellular mono-ADP-ribosylation. *Nat Struct Mol Biol* **20**, 508–514.
- 42 Pucci F, Kwasigroch JM & Rooman M (2020) Protein Thermal Stability Engineering Using HoTMuSiC. *Methods Mol Biol* **2112**, 59–73.
- 43 Voth LS, O'Connor JJ, Kerr CM, Doerger E, Schwarting N, Sperstad P, Johnson DK & Fehr AR (2021) Unique Mutations in the Murine Hepatitis Virus Macrodomain Differentially Attenuate Virus Replication, Indicating Multiple Roles for the Macrodomain in Coronavirus Replication. *J Virol* **95**, e0076621.
- 44 Russo LC, Tomasin R, Matos IA, Manucci AC, Sowa ST, Dale K, Caldecott KW, Lehtiö L, Schechtman D, Meotti FC, Bruni-Cardoso A & Hoch NC (2021) The SARS-CoV-2 Nsp3 macrodomain reverses PARP9/DTX3L-dependent ADP-ribosylation induced by interferon signalling. *bioRxiv*, 2021.04.06.438552.
- 45 Rack JGM, Zorzini V, Zhu Z, Schuller M, Ahel D & Ahel I (2020) Viral macrodomains: a structural and evolutionary assessment of the pharmacological potential. *Open Biol* **10**, 200237.
- 46 Malet H, Dalle K, Brémond N, Tocque F, Blangy S, Campanacci V, Coutard B, Grisel S, Lichère J, Lantéz V, Cambillau C, Canard B & Egloff M-P (2006) Expression, purification and crystallization of the SARS-CoV macro domain. *Acta Crystallogr F Struct Biol Cryst Commun* **62**, 405–408.
- 47 Schneider CA, Rasband WS & Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**, 671–675.
- 48 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC & Ferrin TE (2004) UCSF Chimera?A visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.
- 49 Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320–324.
- 50 Laskowski RA & Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778–2786.



**Tables:****Table 1.** Thermodynamic parameters for CoV *Macro* domains determined by ITC

CoV <i>Macro</i> domain	Kd( $\mu$ M)	N	$\Delta H$ (kcal/mol)
SARS-CoV	5.9 $\pm$ 2.4	0.945 $\pm$ 0.012	-0.001 $\pm$ 0.21
SARS-CoV-2	11.5 $\pm$ 5.7	0.92 $\pm$ 0.008	-0.001 $\pm$ 0.13
SARS-CoV-N157F	6.17 $\pm$ 1.66	0.67 $\pm$ 0.008	-9.528 $\pm$ 0.17
SARS-CoV-2 F156N	9.17 $\pm$ 1.14	0.784 $\pm$ 0.012	-8.037 $\pm$ 0.2
SARS-CoV N157A	13.5 $\pm$ 1.67	1.05 $\pm$ 0.003	-9.352 $\pm$ 0.043
SARS-CoV-2 F156A	9.8 $\pm$ 3.43	1.28 $\pm$ 0.05	-7.225 $\pm$ 0.37

Kd (dissociation constant)

N (stoichiometry)

 $\Delta H$  (enthalpy)**Table 2.** Changes in melting temperature for CoV *Macro* domain mutants determined by HoTMuSiC

CoV <i>Macro</i> domain mutants	$\Delta T_m$ (K)
SARS-CoV-N157F	1.01
SARS-CoV-2 F156N	-3.33
SARS-CoV N157A	-0.39
SARS-CoV-2 F156A	-3.39

**Figure Legends****Figure 1: Sequence alignment and structure superposition of SARS-CoV with SARS-CoV-2 *Macro* domain proteins.**

(A) Alignments comparison of SARS-CoV and SARS-CoV-2 *Macro* domains. Secondary-structure elements are labelled for SARS-CoV-2. Secondary structure elements are represented with squiggles for  $\alpha$ -helixes,  $\eta$  ( $3_{10}$ -helix) and  $\beta$ -strands are displayed as arrows and strict  $\beta$ -turns as TT. Residues with strict identity are marked in red box, whereas residues considered as highly similar are dyed in red and framed in blue. Sequences and 3D structure of SARS-CoV (PDB: 2FAV) and SARS-CoV-2 (PDB: 6WOJ) *Macro* domains were extracted from PDB. Superposition were made with Chimera[48] and the alignment using the ESPript [49]. (B) SARS-CoV (blue) and SARS-CoV-2 (green) structure superposition with ADP-ribose. The secondary structures are labeled and ADP-ribose molecule is shown in yellow sticks with oxygens (red), nitrogens (blue) and bound water (red spheres). (C) An expanded close-up view of the the ADP-ribose binding groove superposition highlights the amino acids with major divergence N 157 in SARS-CoV and F156 in SARS-CoV-2. ADP-ribose coordinates with their interacting amino acids in the binding cleft of SARS-2 and SARS-CoV-2 (D) and SARS-CoV (D). Chemical structure of ADP-ribose and corresponding amino acids are exposed according to stick and balls model. Interaction of covalent bonds of ADP-ribose are shown in purple and of amino acid residues in brown. Hydrogen bonds formed between residues and ligand are represented as green dashed lines with the bond length as numeric numbers. Surrounding residues, in the hydrophobic pocket, in contact with ADP-ribose, are displayed as red eyelash symbols. Diagrams were generated by using LigPlot+ [50].

**Figure 2: Expression and purification of recombinant SARS-CoV and SARS-CoV-2 *Macro* domains.**

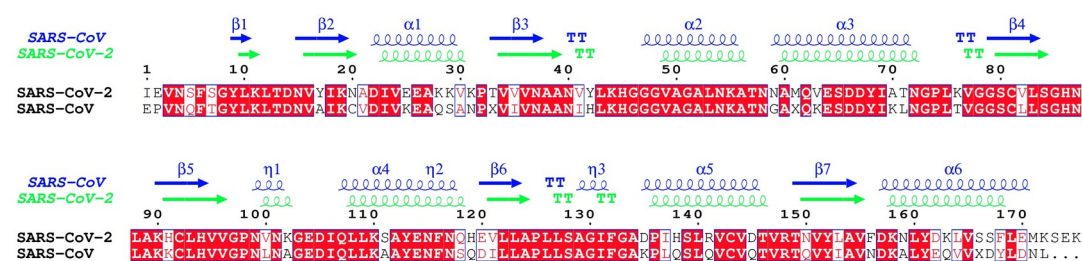
(A) Western blot probing the expressed recombinant proteins insoluble (IF) and soluble (SF) expression fraction. (B) SDS-PAGE of bacterially expressed SARS-CoV-2 *Macro* domain following IMAC purification (E1-9: elution fractions; W: wash; UB: unbound material). (C) SDS-PAGE of SARS-CoV-2 *Macro* domain elution fraction (GF1 to 5) after Gel filtration chromatography.

**Figure 3: ADP-ribose binding to SARS-CoV and SARS-CoV-2 recombinant *Macro* domains.** (A) Lower panel: Dot blot and graphic representation of hPARP1 binding to SARS-CoV and SARS-CoV-2. BSA is used as a negative control for binding. Upper Panel: Binding signal quantification using ImageJ software Average and standard deviation were computed using the GraphPad Prism program. (B) ITC analysis of SARS-CoV and SARS-CoV-2 titration with ADP-ribose.

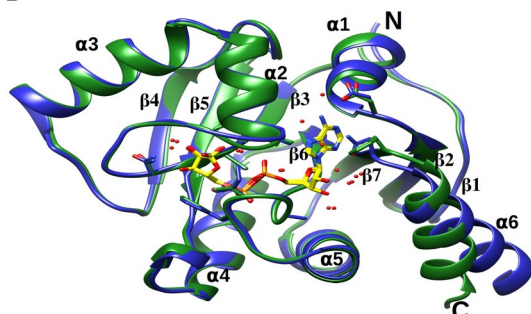
**Figure 4: de-MARylation of automodified PARP3 by CoV *Macro* domains.** (A) Western blot with anti-MAR agent assessing MAR hydrolysis from PARP3 (Upper panel). Ponceau red staining of WB membrane for PARP3 protein load control (Middle Panel). mPAGE Coomassie blue staining for *Macro* domain protein load control. The results are representative of three independent experiments. (B) Bands intensity were quantified using ImageJ software and fitted to a nonlinear regression curve. Error bars indicate standard deviation.

Figure 1

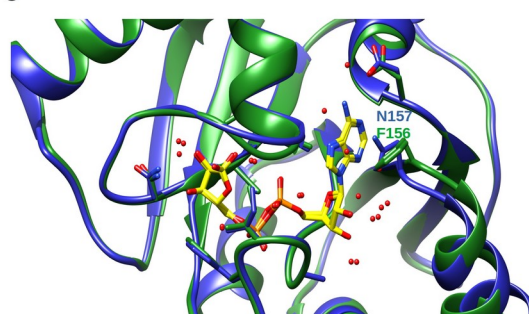
A



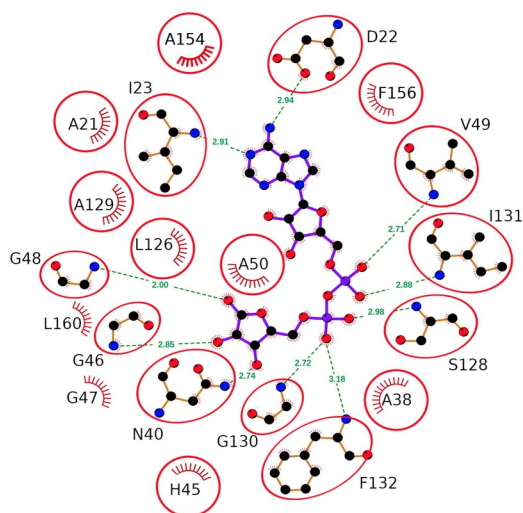
B



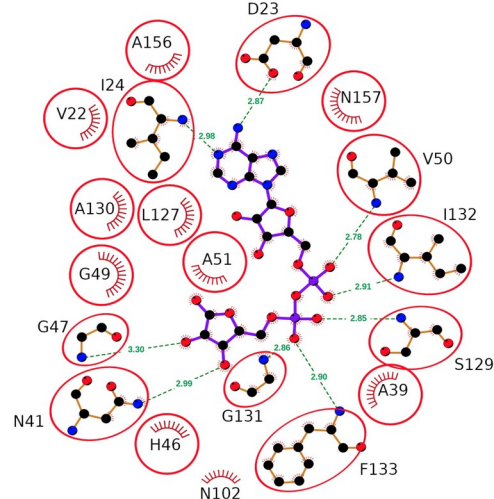
C



D



E



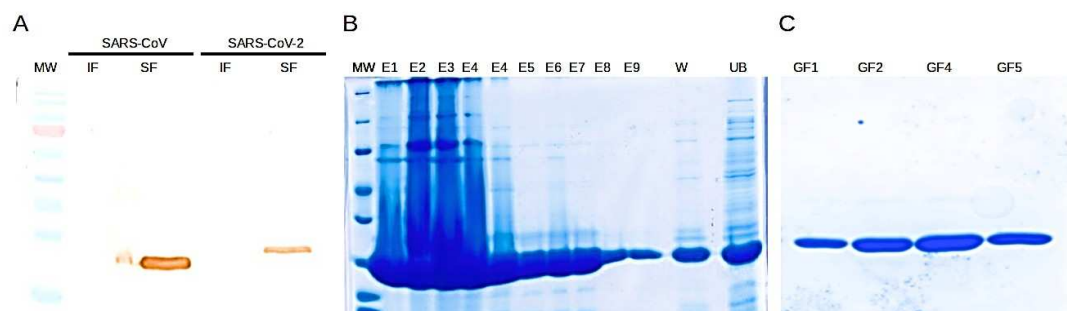
**Figure 2**

Figure 3

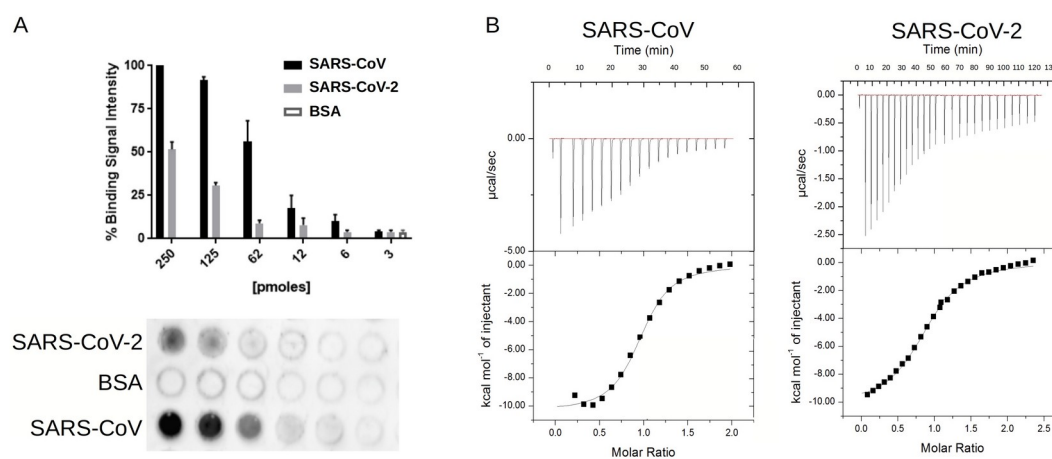
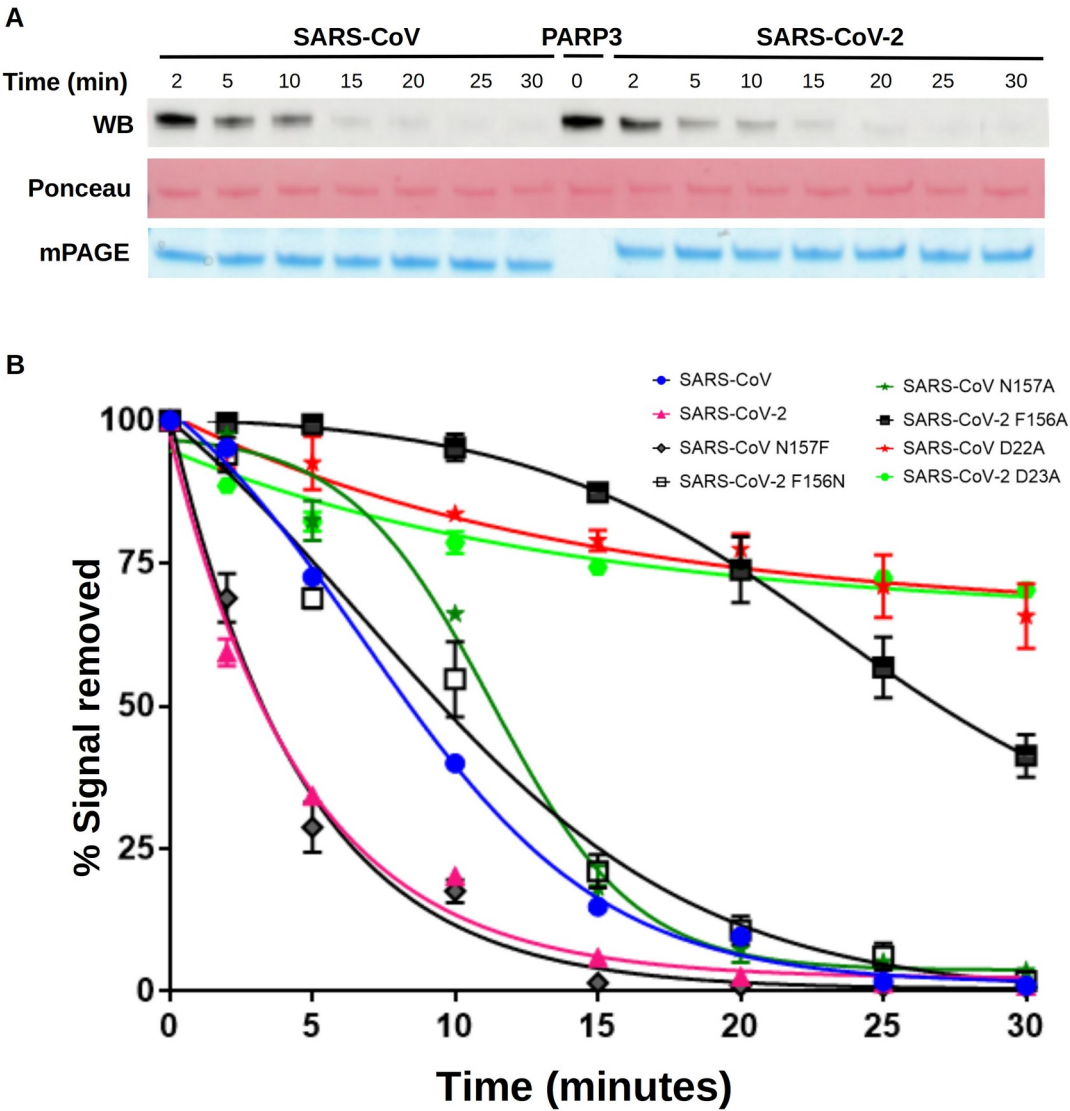


Figure 4



## 4.6 Article 5

### *MOPS derivatives as a scaffold for Macro domain targeted inhibition*

**Oney Ortega Granda<sup>a</sup>**, Karine Alvarez<sup>a</sup>, Mikael Feracci<sup>a</sup>, Bruno Canard<sup>a</sup>, François Ferron<sup>a</sup>,  
and Nadia Rabah<sup>a,b,#</sup>

<sup>a</sup>Aix Marseille Université, CNRS, AFMB UMR 7257, Marseille, France

<sup>d</sup>Université de Toulon, 83130 La Garde, France

# Address correspondence to Nadia Rabah, [nadia.rabah@univ-tln.fr](mailto:nadia.rabah@univ-tln.fr)

*In preparation*



**Abstract**

Coronavirus disease 2019 (COVID-2019) pandemic is an ongoing global health and economic crisis caused by human infection with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The SARS-CoV-2 and SARS-CoV genomes encode for 16 non-structural proteins (nsps). They drive virus replication and participate in viral evasion from the host immune response. Among coronavirus nsps, nsp3 contains a protein module termed *Macro* domain, which was shown to carry an IFN antagonist activity thereby interfering with host innate immunity response (1,2). This domain is able to bind and hydrolyse ADP-ribose derivatives. The hydrolysis activity is correlated to immune escape. *Macro* domains are involved in the regulation of a variety of physiological processes and represent valuable therapeutic targets. Based on the available structural data of the SARS-CoV-2 *Macro* domain, a selected set of MOPS analogues were subjected to molecular docking in the ADP-ribose pocket. In order to screen the selected molecules an immune-enzymatic assay was developed based on the inhibition of recombinant *Macro* domain-ADP-ribose complex formation of SARS-CoV and SARS-CoV-2. Among the tested analogues, MOPSO and CAPSO display the best results in terms of ADP-ribose-binding inhibition. Characterization of these molecules in the ADP-ribose pocket reveal the potential interaction with residues involved in the coordination of ADP-ribose molecule. Therefore, the results suggest the possible use of these molecules as a scaffold for the design of *Macro* domain specific inhibitors.

**Introduction**

*Macro* domains are conserved domains present in nonstructural proteins of several viral families, including *Coronaviridae*, *Togaviridae* and *Hepeviridae*. *Macro* domains contain approximately 170 amino, adopting a defined structure of a central  $\beta$  sheet flanked by  $\alpha$  helices (3,4). *Macro* domains can bind to mono-ADP-ribose (MAR) and poly-ADP-ribose (PAR) in their free form or conjugated to a protein or

RNA substrates. The domain carries hydrolase activities, including a phosphatase activity, against ADP-ribose 1" phosphate, but also de-MARylation and de-PARylation activities. ADP-ribosylation is a ubiquitous post-translational modification (PTM) affecting protein activity, interactions, ubiquitination and degradation. The reaction is driven by ADP-ribosyl transferases (ARDTs) and leads to the addition of MAR or PAR ADP-ribose moieties to charged amino acid residues, mostly aspartates or glutamates. Several groups of ARDTs have been described but the best characterized are poly-ADP-ribosyl polymerases (PARPs), which use nicotinamide adenine dinucleotide (NAD<sup>+</sup>) as a substrate (5). ADP-ribosylation is a crucial regulatory modification implicated in the inflammation process and the regulation of innate immunity. ADP-ribosylation contributes to the establishment of an anti-viral response by inducing type I interferons (IFN) and inhibiting viral translation and replication. The hydrolysis activities of *Macro* domains are believed to counter the above described anti-viral effect of PARPs. This assumption is supported by several mutagenesis studies on viral *Macro* domains, showing that their de-MARylation and de-PARylation activities counter ADP-ribosylation in host cells thereby promoting viral evasion. For example, mutagenesis of *Macro* domain key residues shows a reduction of virus replication and virulence in the case of hepatitis E virus (HEV), alphaviruses and several coronaviruses (2). Mouse hepatitis E virus (MHV) with a catalytic mutation of the *Macro* domain failed to induce acute hepatitis (6-8). Whereas the cellular de-MARylation and de-PARylation substrates of viral *Macro* domains, in the course of infection, are still to be identified, the accumulated data highlight the importance of ADP-ribosylation and hydrolysis in the host-viral conflict.

Complete conservation of *Macro* domains across the *Coronavirinae* subfamily suggests an important role of this domain in the lifecycle of coronaviruses. Based on phylogenetic clustering *Coronavirinae* can be further subdivided into 4 genera: the alpha, beta, gamma, and delta coronaviruses (CoVs). They can cause mild to severe diseases in different mammalian species including agriculture animals and humans.

CoVs are large, enveloped and positive sense single stranded RNA viruses. The genome is ~30 kb long and encodes two large overlapping open reading frames (ORF1a and ORF1ab). Their translation leads to the production of two large polypeptides subsequently processed to 16 non-structural proteins nsp1 to 16, forming the replication-transcription complex (RTC). The largest component of the RTC is the multi-domain protein nsp3. It contains several functional modules among which is the *Macro* domain. The last third of produces several sgRNA, of different lengths, and encodes viral structural proteins: spike (S), envelope (E), membrane (M), nucleoprotein (N); as well as accessory proteins (9).

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the cause of coronavirus disease 2019 (COVID-2019) currently causing the ongoing global health and economic crisis. The World Health Organization (WHO) estimates over 200 million people were tested positive for SARS-CoV-2 infection and more than 5 million persons already died, as of March 2022 (10). The fast development of vaccine candidates and their approval brings a lot of hope for the return to normality. However, emergence of sequence variations in SARS-CoV-2 strains challenges the vaccination efforts, and potentially creates a time lag in adapting vaccines to variant neutralization resistance (11). Thus, different anti-viral strategies should be combined and the development of antivirals against SARS-CoV 2 should not be neglected. Numerous clinical trials testing potential antiviral drugs have been launched since the beginning of the pandemic (12), in particular repurposed drugs. However, no effective treatment seems to emerge (15). Thus, it becomes urgent to pinpoint specific coronavirus druggable targets and develop new-generation of antivirals. *Macro* domain targeting is still at a preliminary ‘docking studies’ stage, lacking experimental data supporting their effective inhibition. Although, flavonoids derivatives, identified by molecular docking, seem promising for the inhibition of the alphavirus *Macro* domain (14,15). A recent report presented the crystal structure of the SARS-CoV-2 *Macro* domain in complex with different nucleosides and nucleotides (16). Among the

clinically used antiviral nucleoside analogs, remdesivir was the only metabolite (GS-441524) which occupied the adenosine site in the *Macro* domain (16).

*Macro* domain inhibitors could be of great therapeutic value against SARS-CoV infections. In this study we evaluate the potential inhibitory effect of MOPS analogues against SARS-CoV and SARS-CoV 2 *Macro* domains. The inhibition of ADP-ribose binding is monitored by thermal shift assay (TSA), enzyme-linked immunosorbent assay (ELISA) , and a dot-blot assay, the latter using anti-poly-ADP-ribose reagent antibodies. Molecular docking of selected MOPS analogues into the ADP-ribose binding site was used to study the molecular basis of the observed inhibition.

### ***Material and methods***

#### **MOPS analogues**

3-(N-morpholino) propanesulfonic acid (MOPS) (Ref.EU0034 EUROMEDEX), 3-Morpholino-2-hydroxypropanesulfonic acid (MOPSO) (Ref.M8389-25G Sigma-Aldrich), 3-(Cyclohexylamino)-1-propanesulfonic acid (CAPS) (Ref.C2632-25G Sigma-Aldrich), 3-(Cyclohexylamino)-2-hydroxy-1-propanesulfonic acid (CAPSO) (Ref.C2278-25G Sigma-Aldrich), 2-(Cyclohexylamino) ethanesulfonic acid (EPPS) (Ref.10213672 Thermo Fischer), 4-(Piperidin-4-yl) butanoic acid hydrochloride (4-PBAK) (Ref.540838-5G Sigma-Aldrich), [(2-Hydroxy-1,1-bis(hydroxymethyl)ethyl)amino]-1-propanesulfonic acid (TAPS)(Ref. T5130-25G Sigma-Aldrich) and 2-(Cyclohexylamino) ethanesulfonic acid (CHEZ)(Ref.29311-10G Sigma-Aldrich).

#### **Immuno-enzymatic assay**

Purified recombinant SARS-CoV and SARS-CoV-2 *Macro* domains were coated on a 96-well microtiter ELISA plates (MTPs, costar EIA/RIA Plates, Thermo Scientific

Inc., USA) overnight at 4°C. The plates then were blocked 1h at 37°C with coating buffer (0.05 M carbonate-bicarbonate buffer pH 9.6), supplemented with 2% of BSA. Purified poly-ADP-ribose (100 nM PAR, Trevigen, cat # 4336-100-01) or autoribosylated hPARP1 (PAR-hPARP1) (Sigma) diluted in PBS/Tween) were added to the plate and incubated 2h at room temperature (RT) with shaking. Anti-poly-ADP-ribose binding reagent (Millipore, cat# MABE1031) diluted 1:2000 in blocking buffer was added for 1h at RT. The second antibody anti-rabbit IgG (Dako) diluted 1:2000 was added to each well, incubated 1h at 37°C. After each step, the plates were washed three times with PBS-Tween (pH 7.4, 0.01 M phosphate buffer, 0.0027 M potassium chloride, 0.137 M sodium chloride and 0.1% tween 20). The PAR–antibody complex was revealed by the addition of o-phenylenediamine solution. The reaction was stopped with 3 M sulfuric acid. Optical densities of formed complexes were measured at 492 nm in the microplate reader (TECAN). BSA was employed as a negative protein control for ADP-ribose binding.

### **Immuno-enzymatic assay of PAR-binding inhibition**

The inhibition of ADP-ribose binding was assessed by adding increased concentrations (50-1000  $\mu$ M) of the corresponding MOPS analogue to the coated *Macro* domains for 2h at RT. Remdesivir metabolite AT-9045 was used as a positive control for ADP-ribose binding inhibition. Signal corresponding to the PAR-antibody complex was measured as described before. All experiments were performed in triplicate and the data was normalized to the value of PAR-binding without inhibitor. Results are presented as means  $\pm$  standard deviation.

### **Inhibition PAR-binding assay**

Inhibition of ADP-ribose binding to SARS-CoV and SARS-CoV-2 *Macro* domains by selected compounds was performed using dot blot assay. The assay was carried out by spotting recombinant SARS-CoV and SARS-CoV-2 *Macro* domain proteins, at various dilutions (750, 500, 250, 125, 62 and 32 pmoles), on a nitrocellulose

membrane (Schleicher & Schuell) using Minifold II dot blot apparatus (Schleicher & Schuell). Bovine serum albumin (BSA) was used as negative control for binding assessment. The membrane was blocked for 1 h at RT with dot blot buffer (10 mM Tris pH 8, 150 mM NaCl and 0.05% of Tween 20), supplemented with 4% skimmed milk. The blocked membranes were incubated with 50 to 1000  $\mu$ M of each tested compound selected by ELISA and TSA for 1h at RT under agitation. The membrane was then subjected to three washes in dot blot buffer, then incubated with PAR-hPARP1 for 1h at RT under agitation. The unbound material was removed by three extensive washes with dot blot buffer. The membrane was next incubated for 2 h, at RT, with the primary antibody, anti-poly ADP-ribose binding reagent (Sigma) diluted 1500 times in dot blot buffer with 1 % non-fat milk. The secondary antibody, anti-rabbit IgG (Dako), diluted up to 1:2000 was incubated with the membrane for 1 h, at RT. After three washes, immunoreactive signals were revealed using ECL reagent (Cat. # 170–5061, Bio-Rad) and visualized using the Amersham™ ImageQuant™ 800 Imager (Cytiva). Images were analyzed and quantified using ImageJ software (17). Each experiment was repeated three times. The intensities of the dots were normalized to the total protein load incubated with auto-ADP-ribosylated hPARP-1 alone.

### **Thermal shift assay (TSA)**

Protein thermal shift assays were performed to determine the stability of SARS-CoV and SARS-CoV-2 *Macro* domains in presence of the evaluated compounds. *Macro* domain proteins were diluted in TSA buffer (20mM HEPES pH 8 and 150mM NaCl) to a final concentration of 2  $\mu$ M. The test was performed in MicroAmp® Fast Optical 96-well reaction plates from Applied Biosystems. The compounds were added at increased concentrations from 50 to 1000  $\mu$ M. Protein thermal shift dye kit (Applied biosystems, Thermo Fisher Scientific) was used diluted in water, according to manufacturer's instructions. After mixing all components in the wells, the plate was sealed and put in the 7500 FastReal-Time PCR System (Applied Biosystems). The reaction was started by a first incubation step at 25°C for 2 min. The temperature was

then increased with a ramp rate of 0.5°C per min to 95°C. Normalized melting curves were obtained with GraphPad Prism. The melting temperatures ( $T_m$ ) were calculated from the inflection point of the melting curves.

### **Virtual molecular docking of the MOPS analogues to the SARS-CoV *Macro* domain**

Molecular docking of MOPS analogues into the ADP-ribose pocket of the SARS-CoV *Macro* domains was carried out using AutoDock Vina (18). A present MOPS ligand and water molecules were removed from the structure of the SARS-CoV *Macro* domain. The docking grid was focused on the center of the ADP-ribose pocket. Structure data files for MOPS and its analogues were extracted from the PubChem database. To rank the docking results, the binding energies ( $\Delta G$ ) and the precision pose on the MOPS place were considered. Docking poses were visualized and analyzed with Chimera.

### ***Results and discussion***

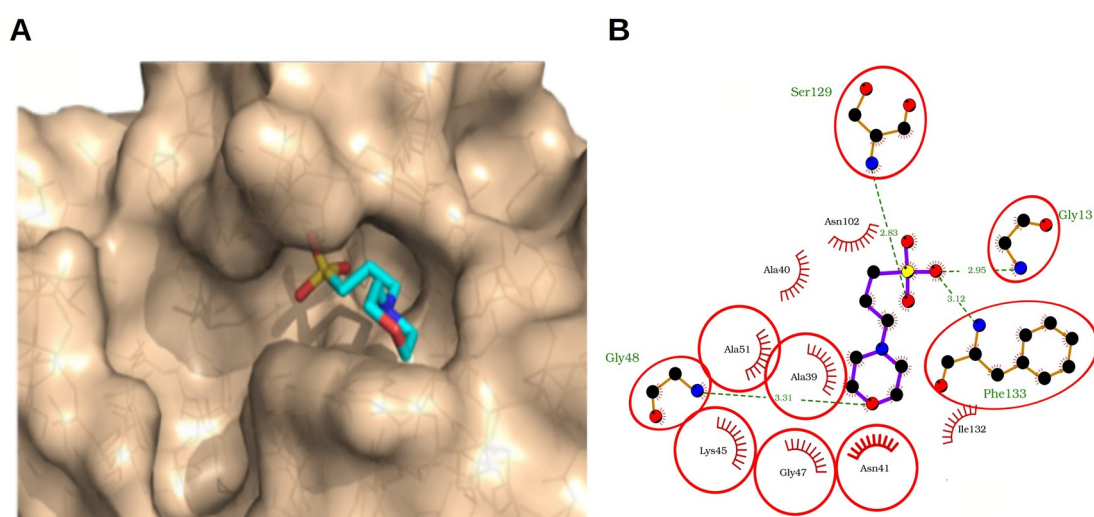
Structural studies on viral *Macro* domains have opened interesting research avenues for the development of antiviral tools. The crystal structure of SARS-CoV-2 *Macro* domain was solved in its free form (apo) and complexed to various ligands including 2-(N-morpholino) ethanesulfonic acid (MES), AMP and ADP-ribose. The reported structures follow the classical architecture described for *Macro* domains, with a  $\beta$ -sheet made of 7  $\beta$ -strands sandwiched between two layers of  $\alpha$ -helixes (19–21) and these data were supported by NMR studies (17). These studies defined that the cleft accommodating ADP-ribose provides four contact zones. The first forms a hydrophobic patch of residues Ile23, Val49, Pro125, Val155 and Phe156, which interacts with adenine pointing it toward the polar Asp at position 22. The second, composed of three conserved Gly (residues 46–48) houses the diphosphate bridge. The third, implicating Phe 132 and Ile 131, stabilizes the proximal ribose. The forth,

involving Leu 136, Ala 154 and Asp 157, supports the distal ribose *via* water-hydrogen bonding. Crystallization studies of SARS-CoV *Macro* domain, done by our group, revealed a MOPS molecule positioned in the ADP-ribose pocket (Figure 1A). The 2-N-morpholine ring and the sulfonic acid of MOPS take the place of the distal ribose and the distal phosphate, respectively. The morpholinic ring and the sulfonic acid of MOPS form hydrogen bonds with residues such as Gly48, Ser129, Gly131 and Phe133 (Figure. 5B, C, E, F) (Figure 1B), which are involved in the interaction with the diphosphate of ADP-ribose (19–21,27,30,31). In addition, residues Asn41, Gly41 and Ile132, forming direct contact with the distal ribose, establish also hydrophobic association with MOPS structure (*Ortega et al* in preparation). These structural observations prompted us to hypothesize that MOPS could be used as a scaffold to design *Macro* domain inhibitors. Therefore, MOPS analogues with various substitutions in the morpholine ring and the connecting alkyl chain (Table 1) were docked to the ADP-ribose pocket and compared to the original placement of MOPS. The superposition of free and ligand bound structures reveal that the fold is flexible and adapts to the bound ligands, namely MOPS, MOPSO, CAP, CAPSO, EPPS, 4-PBAK, TAPS and CHEZ. Hence, these molecules were selected as potential antagonist candidates for ADP-ribose binding. Most of these molecules are structurally related and contain a morpholine ring and a sulfonic acid group.

Some studies highlighted the inhibitory effect of sulfonic acids buffers on enzyme activity. It has been shown that MOPS reduces up to 40% the activity of bovine adrenal tyrosine hydroxylase (22). Moreover, it is able to inhibit the activity of polyester hydrolases from the cutinase family at high concentrations (23). 2-(N-morpholino) ethanesulfonic acid (MES) can also have inhibitory proprieties. It can notably inhibit metallo- $\beta$ -lactamase from *Bacteroides fragilis* (24). The resolution of the crystal structure of CRN-4 from DEDDh family of exonucleases in complex with MES revealed a unique inhibition mechanism, inducing the shift out of the general base from the active site upon MES binding (25). Recently, the crystal structure of SARS-CoV-2 *Macro* domain in complex with MES revealed that it takes position in



the ADP-ribose cleft. The interaction involves hydrogen bonding from Asn40 to the morpholine oxygen and the residues interacting with the proximal phosphate accommodate the sulfonic acid (19,26). In the same study, the 2 morpholinoethanesulfonyl moiety, the major fragment of MES, was incorporated into the linker of poly (ADP-ribose) glycohydrolase (PARG) inhibitors to target both adenosyl and ribose pockets of the SARS-CoV-2 *Macro* domain. This reconfiguration of PARG inhibitor fragments allowed the sulfone functional group to strength the residues backbone contacts with  $\alpha$  and  $\beta$ -phosphates of ADP-ribose. Hence, the flexible accommodation of MOPS analogues in the ADP-ribose binding pocket, supports its use as a skeleton for *Macro* domain inhibitors design.



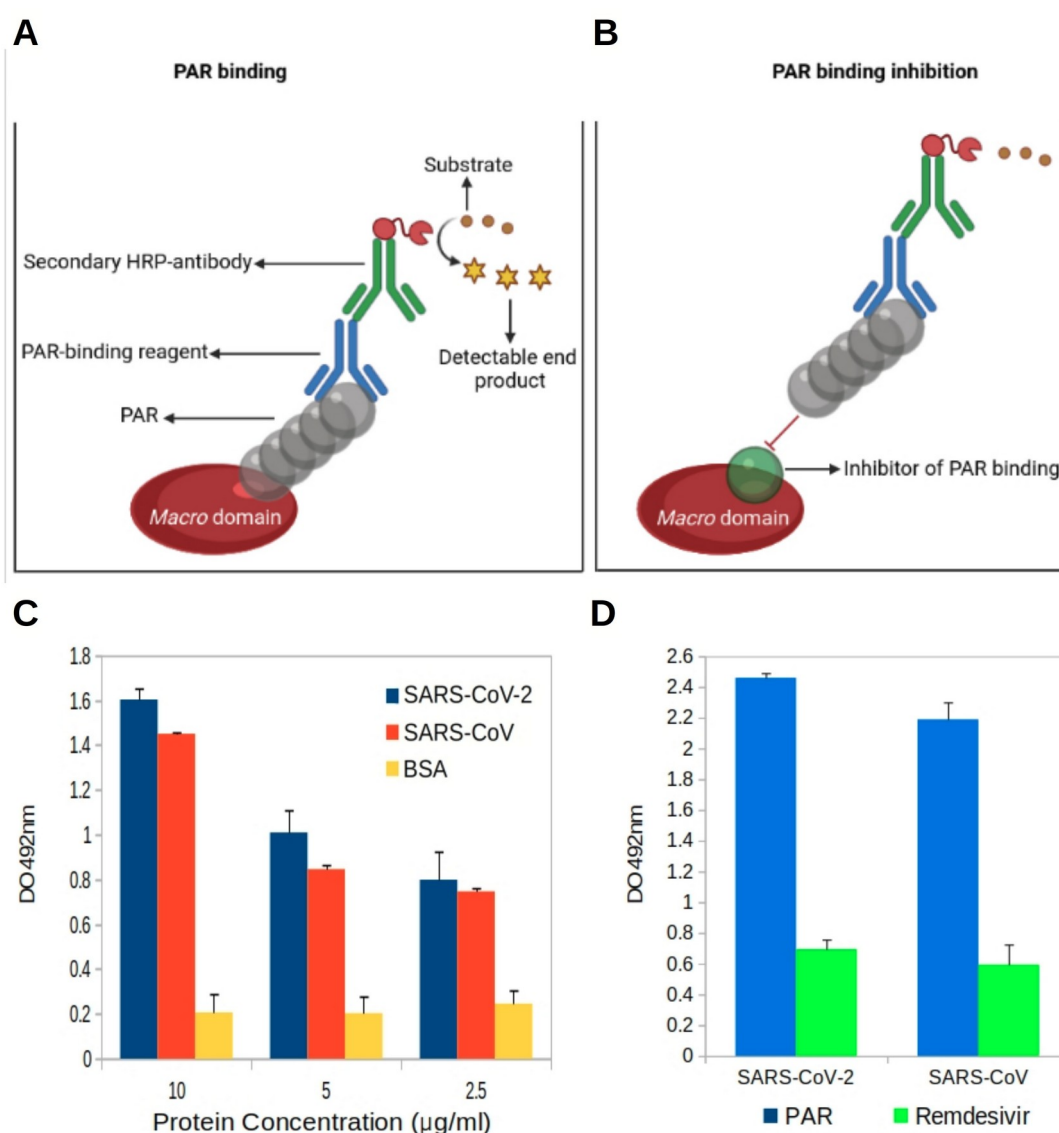
**Figure 1: SARS-CoV *Macro* domain in complex with MOPS.** (A) Molecular surface model of the SARS-CoV *Macro* domain in complex with MOPS, (Ferron and Morin unpublished results). (B) MOPS interaction with the backbone residues in the of SARS-CoV *Macro* domains ADP-ribose binding pocket. Chemical structure of MOPS and corresponding amino acids are exposed according to stick and balls model. The covalent bonds of the ligand are shown in purple and that of the amino acid residues in brown. The crescents with the bristles represent hydrophobic interactions. The hydrophobic atoms are in black. Hydrogen bonds formed between residues and ligand are represented as green dashed lines with the bond length as

numeric numbers. Surrounding residues, in the hydrophobic pocket, in contact with the molecule are displayed as red eyelash symbols. Diagrams were generated by employing LigPlot+ (27).

**ELISA can be exploited to screen PAR-*Macro* domain complex inhibitors at a large scale.**

Recently, an ELISA was developed to test the ADP-ribosylhydrolase activity of SARS-CoV-2 *Macro* domain (28). Here, an ELISA, based on the immune detection of PAR (in a free or protein bound form), was developed to screen rapidly inhibitors affecting ADP-ribose binding to *Macro* domain.

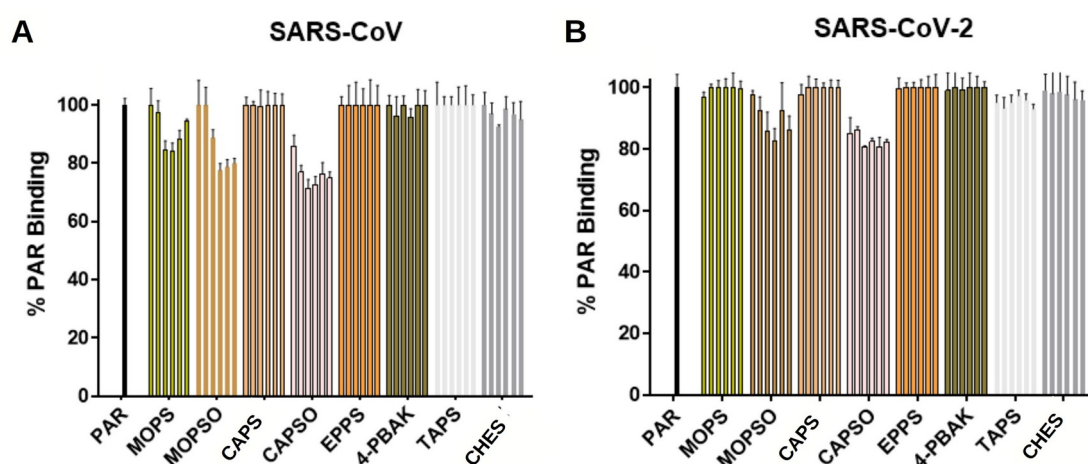
The ELISA is performed by coating the *Macro* domain protein on an ELISA plate. After the blocking, the coated *Macro* domain is saturated with fixed concentrations of either free PAR or PAR-hPARP 1 (Figure 2A). Inhibition assessment is based on the ability of MOPS analogues to occupy the ADP-ribose binding site in the *Macro* domain. Hence, in the presence of an inhibitor molecule, the access to the ADP-ribose binding pocket is denied. There is no PAR binding and no signal is recorded (Figure 2B). The assay validation was performed using Remdesivir as a positive control for inhibition (Figure 2 C, D). Hence, the first step (Figure 2A) sets the maximum value for PAR binding (Figure 2C), after subtraction of the non-specific binding achieved with BSA. The optimal condition for plate coating being set at 10µg/ml of *Macro* domain protein for free- PAR (Figure 2C) and PAR-hPARP 1 (data not shown). When Remdesivir was added to the reaction, an 75% of reduction in PAR binding was observed (Figure 2D). This result confirms previously reported data (26) and validate our ELISA binding assay.



**Figure 2. ELISA for ADP-ribose inhibitors high scale pre-screen.** (A), (B) Cartoon representation of the ELISA assay. (A) Direct detection of PAR binding to *Macro* domain using anti-PAR agent. (B) Inhibition of PAR binding to *Macro* domain in presence of an inhibitor molecule (C) ELISA conditions optimisation, (D) Assay validation in the presence of Remdesivir.

Once the key assay parameters of PAR binding were determined and the assay validated, using a known *Macro* domain inhibitor, the selected MOPS analogs were

tested. For that end, coated *Macro* domains of SARS-CoV and SARS-CoV-2 were incubated with the various MOPS analog molecules, namely MOPS, MOPSO, CAP, CAPSO, EPPS, 4-PBAK, TAPS, CHEZ, before PAR addition (Figure 2B). After PAR addition, the assay was pursued as previously described. The results of data analysis are presented as percentage of PAR binding (Figure 3). One can notice that CAPS, EPPS, 4-PBAK, TAPS, CHES have no effect on PAR binding in the case of SARS-CoV (Figure 3A) and SARS-CoV-2 (Figure 3B). However, MOPS, MOPSO and CAPSO inhibit partially PAR binding at various concentration tested. In the case of SARS-CoV *Macro* domain, inhibition of PAR binding starts at 100 $\mu$ M of the tested molecule. At 250 $\mu$ M of MOPS, MOPSO lead to a reduction of PAR binding of 25%. CAPSO shows a stronger inhibitory effect, since it impedes PAR binding starting from 50  $\mu$ M. Maximal inhibition, up to 30 %, is reached at 250 $\mu$ M (Figure 3A). SARS-CoV-2 *Macro* domain shows a slightly different profile, since it is not sensitive to MOPS even at high concentrations. MOPSO and CAPSO, however, are able to hamper PAR binding. The highest inhibition, up to 25%, is achieved with CAPSO starting from 250 $\mu$ M (Figure 3B).

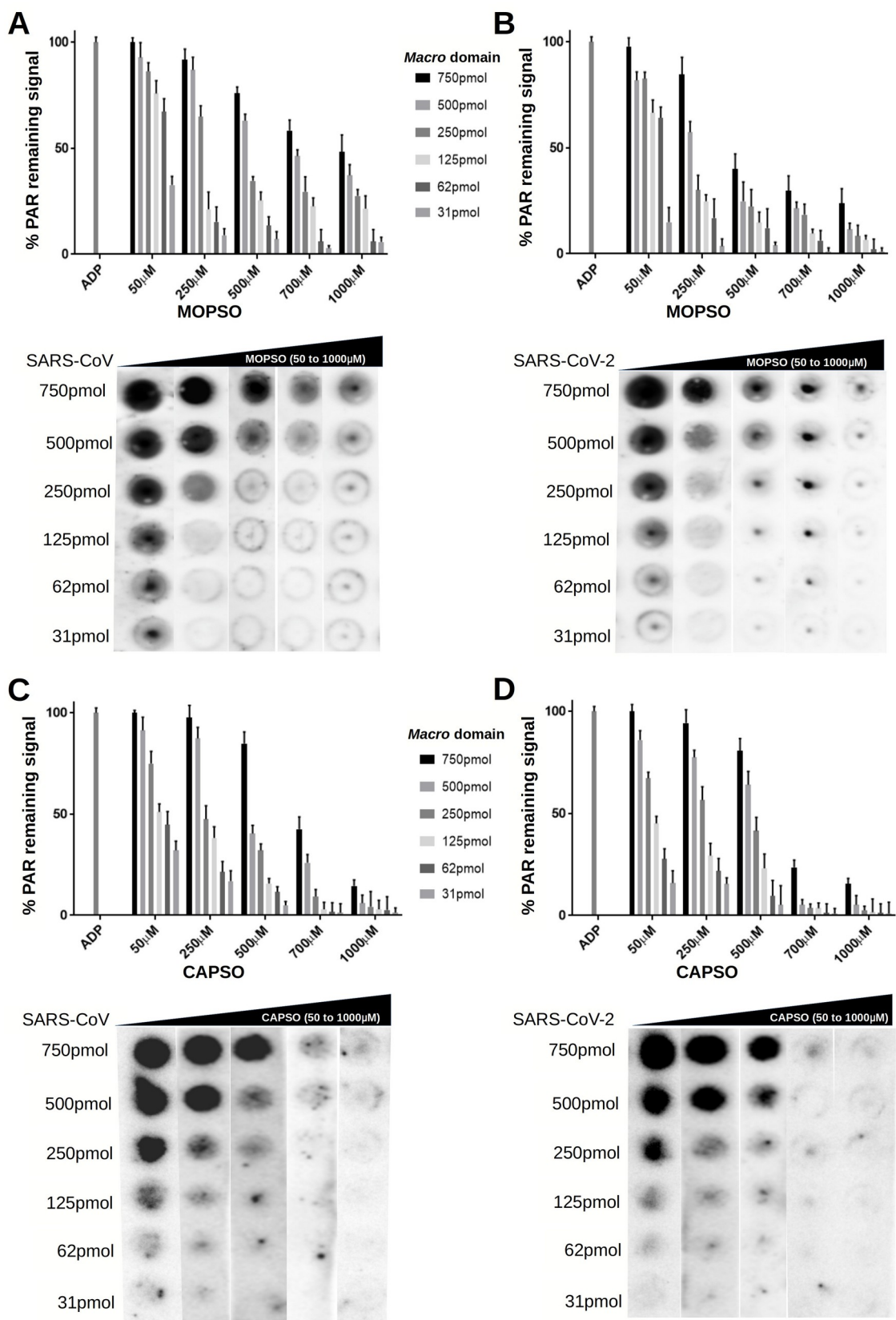


**Figure 3. MOPS analogs screen of ADP-ribose binding inhibition towards *Macro* domains of SARS-CoV and SARS-CoV-2.** MOPS analogues were incubated at increased contractions (50 to 1000 $\mu$ M) with coated *Macro* domains before PAR

addition. Results are presented as percentage of PAR binding compared to negative control (PAR) (without analog addition). (A) SARS-CoV. (B) SARS-CoV-2.

### **MOPS analogs inhibit PAR binding in the context of auto-ribosylated hPARP1**

In order to validate and confirm the inhibitory effect of MOPSO and CAPSO on SARS-CoV and SARS-CoV-2 *Macro* domains, a classical dot blot assay was performed (*Ortega et al* in preparation). Accordingly, *Macro* domains were spotted into a nitrocellulose membrane at serial dilutions and incubated with increased concentrations of MOPSO or CAPSO before the addition of PAR-hPARP1. Following PAR-hPARP1 addition, signals on the membrane were revealed using anti-PAR-binding reagent (Figure 4). When SARS-CoV and SARS-CoV-2 *Macro* domain proteins are spotted at less than 250 pmol, PAR-hPARP1 binding is totally inhibited, even at low MOPSO and CAPSO concentrations. Starting from 250 pmol, a dose dependent inhibition of PAR-hPARP1 is observed. In the case of MOPSO, the inhibitory effect is more prominent with SARS-CoV-2 *Macro* compared to SARS-CoV, with 63% and 30% inhibition, respectively, at 250µM of the inhibitor molecule (Figure 4 A, B). CAPSO, however, seem to inhibit equally both *Macro* domains. At 500µM of CAPSO, the remaining signal ranges from 16% to 20% at 500pmol of protein, for SARS-CoV and SARS-CoV-2, respectively (Figure 4 C, D). An evident drop in signal intensity with increased concentration of CAPSO and MOPSO for both *Macro* domains. This decrease is correlated to the signal decrease observe in the ELISA assay for PAR binding. Thereby, these data validate ELISA results and reinforce its use the pre-screen of potential *Macro* domains inhibitors.

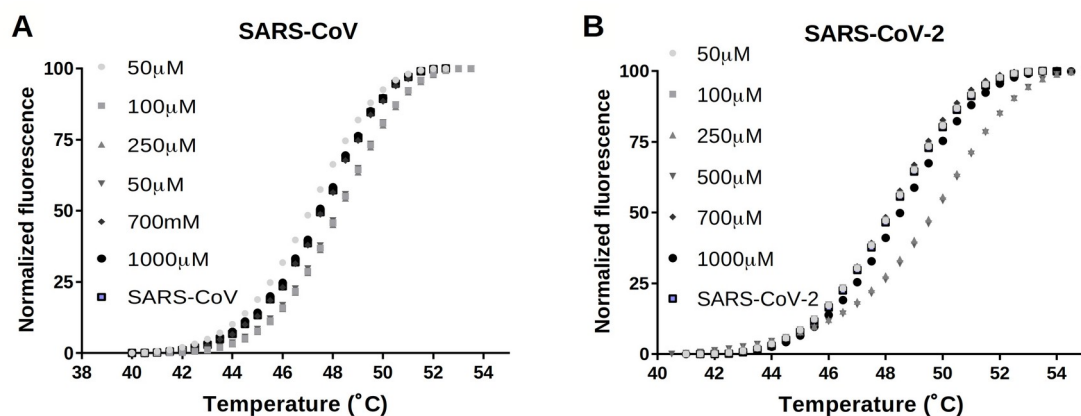


**Figure 4. MOPSO and CAPSO inhibit PAR-hPARP1 binding to *Macro* domains of SARS-CoV and SARS-CoV-2.** Serial dilutions of *Macro* domain proteins were spotted on the dot blot membrane and incubated with increased concentrations of MOPSO or CAPSO before the addition of PAR-hPARP1. (A, B) MOPSO; (C, D) CAPSO. Lower panels: Dot blot and graphic representation of PAR-hPARP1 binding to SARS-CoV (A and C) and SARS-CoV-2 (B and D). Upper Panels: Binding signal quantification, done using ImageJ software. Average and standard deviation were computed using the GraphPad Prism program. Quantitative values were normalized to the binding of PAR-hPARP1 for different dilution of SARS-CoV and SARS-CoV-2 *Macro* domains without any MOPS analog. The data are presented as the percent of PAR-hPARP1 binding values obtained from three independent experiments.

#### **MOPSO and CAPSO stabilize SARS-CoV and SARS-CoV-2 *Macro* domains**

In order to evaluate a possible effect of MOPS analogs on *Macro* domain protein integrity, TSA experiments were carried out (Figure 5). Hence, different dilutions of SARS-CoV and SARS-CoV-2 *Macro* domain were mixed with MOPS analogues at increased concentration (50 to 1000 $\mu$ M). In the absence of inhibitors, the recorded melting temperatures are 47°C and 48°C, for SARS-CoV and SARS-CoV-2 *Macro* domains, respectively. Among all the analogs tested, only MOPSO and CAPSO shift the melting temperature ( $\Delta T_m=1^\circ\text{C}$ ) of both *Macro* domains. In SARS-CoV, the stabilization by CAPSO occurs between 100 $\mu$ M and 500 $\mu$ M (Figure 5A), while in SARS-CoV-2 the stabilization is observed at 250 $\mu$ M and 500 $\mu$ M (Figure 5B). In case of MOPSO the shift of the melting temperature occurs between 250 $\mu$ M and 700 $\mu$ M for SARS-CoV and SARS-CoV-2 (data not shown). Thermal stability of both *Macro* domains increases with increasing analog concentrations until 500  $\mu$ M of MOPSO and 700  $\mu$ M of CAPSO. After these concentrations, melting temperature decreases due to a destabilization effect of MOPSO and CAPSO, consistent with previously reported data (29). TSA results confirm that MOPS analogs penetrate into the ADP-ribose pocket without undermining protein integrity.





**Figure 5. Effect of CAPSO on the thermal stability of SARS-CoV and SARS-CoV-2 *Macro* domains.** Normalized melting profiles of SARS-CoV and SARS-CoV-2 *Macro* domains in presence of increasing concentrations of CAPSO are recorded by TSA. (A) SARS-CoV (B) SARS-CoV-2.

### Modeling of MAPSO and CAPSO molecules into the *Macro* domain ADP-ribose binding site

The structure of MOPSO and CAPSO are similar to that of CAPS and MOPS, except that they possess an additional hydroxyl (OH) group (Table 1). The results obtained with these two compounds are different from the non-hydroxylated analogs. Thus, the presence of the OH group seems to be important in the binding of these molecules to the ADP-ribose binding pocket. In order to ascertain their position in ADP-ribose binding pocket, molecular docking experiments were performed for SARS-CoV *Macro* domain by positioning MOPSO and CAPSO in a similar arrangement as the MOPS ligand described above (Figure 6). Several R and S enantiomers of MOPSO and CAPSO were docked to determine their interacting amino acid residues. The



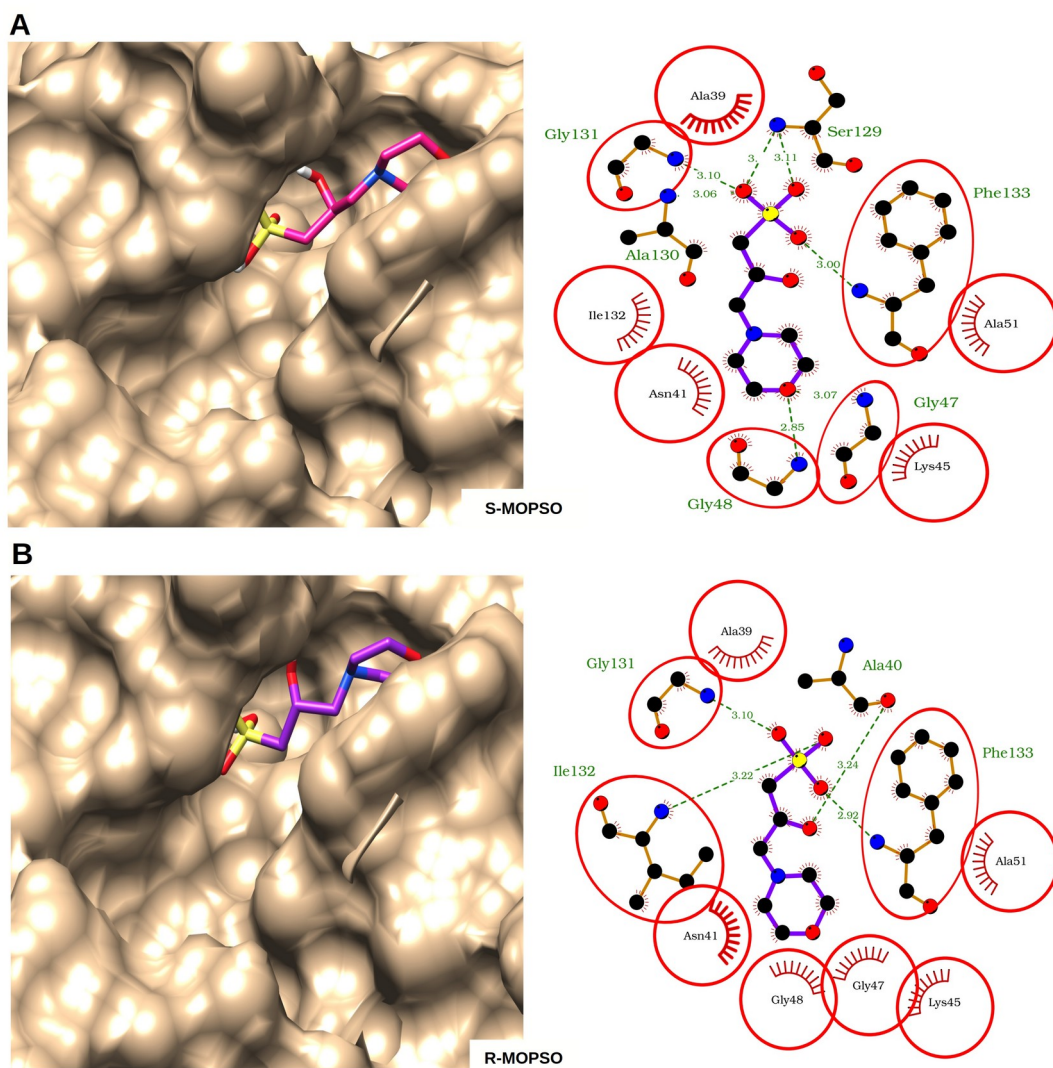
structures of protein-ligand were selected according to binding energies ( $\Delta G$ ) and best fitting compared to MOPS molecule.

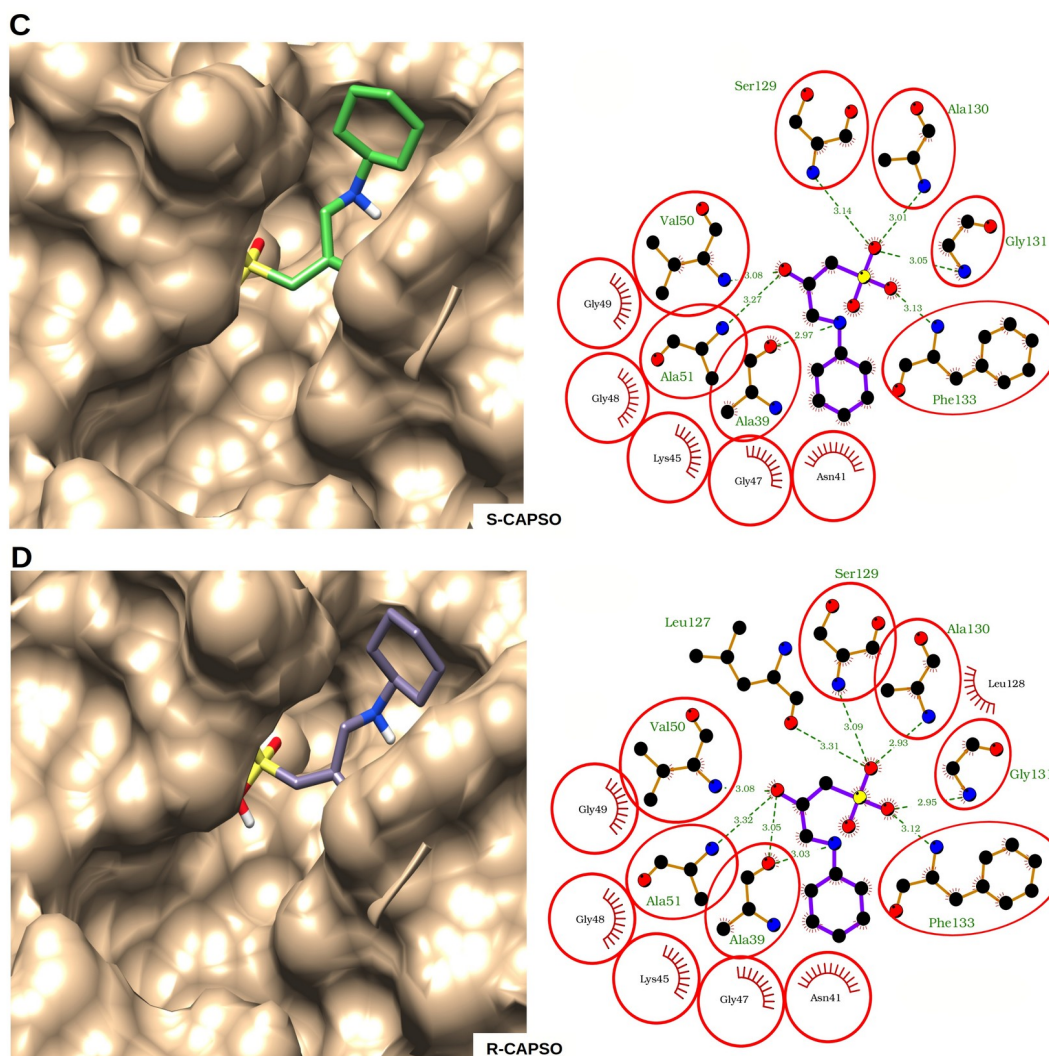
Analyses of the amino acid interaction revealed that MOPSO and CAPSO present most potentially directly contacts with the residues among the molecule in comparison with MOPS (Figure 6). Both MOPSO and CAPSO molecules form hydrogen bonds with residues fixing the ADP-ribose in the binding pocket (Figure 6A, B, C, D).

The sulfonic acid of S-MOPSO form hydrogen bonds with Ser129, Gly131, Phe133 and Ile132 (Figure 6A) that interact with the  $\alpha/\beta$ -phosphate of ADP-ribose molecule (19–21,28,30,31). The other part of the molecule, the morpholinic ring of S-MOPSO, interact with the conserved glycine rich stretch (residues 47-48) (Figure 6A), that accommodates the diphosphate moiety of ADP-ribose (Figure 6A). The hydroxyl group in R-MOPSO is involved in hydrophobic contacts with Ala39 and Ala51 (Figure 6B), that also establish hydrophobic contacts with ADP-ribose molecule (*Ortega et al* in preparation).

In the case of CAPSO, in both the R and S conformation, the sulfonic acid also forms hydrogen bonds with Ser129, Gly131, Phe133 (Figure 5C, D). The hydroxyl group is involved in hydrophobic contacts with Val50 (Figure 5C, D) that interact with the  $\alpha$ -phosphate of ADP-ribose molecule in the SARS-CoV and SARS-CoV-2 binding cleft (21,30). In addition, hydroxyl group interact with Ala39 and Ala51 as R-MOPSO (Figure 6C, D). The hydroxyl group could also interacts with water molecules clusters around the binding cleft, enriching the water-hydrogen bonding with the binding pocket residues. Moreover, MOPSO and CAPSO establish some hydrophobic contact with non-ligand residues involved in hydrogen bonds formation with ADP-ribose, both in SARS-CoV and SARS-CoV-2 *Macro* domains (Figure 6). Thereby, the better direct contact with the residues in ADP-ribose binding pocket and the additional interactions brought by the hydroxyl group could explain the experimental results obtained with CAPSO and MOPSO molecules. These modeling results support all the

more the use of MOPSO and CAPSO as a template for the development *Macro* domain inhibitors.



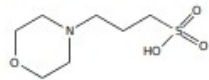
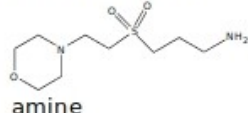
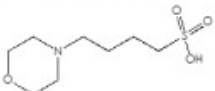
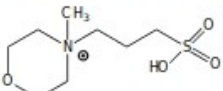
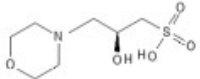
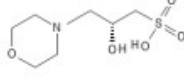
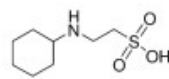
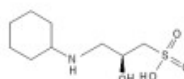
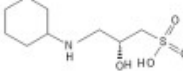
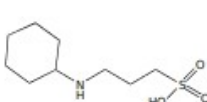
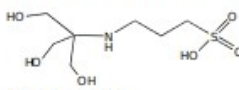
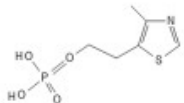
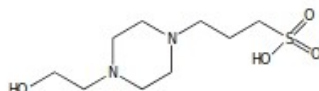


**Figure 6. Molecular docking and amino acids interaction in SARS-CoV Macro domain.** Docking structures (left panels) and potentially amino acids interaction (right panels) in the ADP-ribose binding cleft. (A) S-MOPSO. (B) R-MOPSO. (C) S-CAPSO and (D) R-CAPSO. Docking was performed by positioning each analog at the MOPS place using AutoDock Vina (18). Chemical structure of S/R-MOPSO or S/R-CAPSO and corresponding amino acids are exposed according to stick and balls model. Interaction of covalent bonds of ligands are shown in purple and of amino acid residues in brown. The crescents with the bristles represent hydrophobic interactions.

The hydrophobic atoms are colored in black. Hydrogen bonds formed between residues and ligand are represented as green dashed lines with the bond length as numeric numbers. Surrounding residues, in the hydrophobic pocket, in contact with the molecule are displayed as red eyelash symbols. Diagrams were generated by LigPlot+ (27).

Finally, in the present study, we developed an efficient low cost ELISA assay for the pre-screen of *Macro* domain inhibitors. Based on the presence of a MOPS molecule in the 3D structure of SARS-CoV *Macro* domain, we investigated the inhibitory potential of MOPS analogs on SARS-CoV and SARS-CoV-2 *Macro* domains. We found that MOPSO and CAPSO display an inhibitory effect toward SARS-CoV and SARS-CoV-2 *Macro* domain, prompting the use of MOPS as a scaffold for inhibitor design.

**Table 1.** MOPS analogues

Family		Buffers/ structure		
<b>Morpholin</b>				
MOPS (3-(N-morpholino) propanesulfonic acid)	3-[2-(morpholin-4-yl) ethane sulfonyl]propan-1-	MOBS 4-Morpholine butanesulfonic acid	Morpholinium, 4- methyl-4-(3- sulfopropyl)-, inner salt	
				
MOPSO R 3-Morpholino-2- hydroxypropanesulfon ic acid	MOPSO S			
				
<b>Cyclohexylamino</b>				
CHES bioultra 2-(Cyclohexylamino) ethanesulfonic acid	CAPSO R 3-(Cyclohexylamino)- 2-hydroxy-1- propanesulfonic acid	CAPSO S	CAPS 3-(Cyclohexylamino)- 1-propanesulfonic acid	
				
<b>TRIS</b>				
TAPS [(2-Hydroxy-1,1- bis(hydroxymethyl)ethyl)amino]-1- propanesulfonic acid	4-Methyl-5-hydroxyethyl lthiazole phosphate			
				
<b>Piperazin</b>		<b>Piperidin</b>		
EPPS 3-(4-(2-Hydroxyethyl) piperazin-1-yl) propane-1-sulfonic acid	4-(Piperidin-4-yl) butanoic			
				



**References**

1. Matsuyama T, Kubli SP, Yoshinaga SK, Pfeffer K, Mak TW. An aberrant STAT pathway is central to COVID-19. *Cell Death Differ.* 2020 Dec;27(12):3209–25.
2. Fehr AR, Jankevicius G, Ahel I, Perlman S. Viral Macrodomains: Unique Mediators of Viral Replication and Pathogenesis. *Trends in Microbiology.* 2018 Jul;26(7):598–610.
3. Han W, Li X, Fu X. The macro domain protein family: Structure, functions, and their potential therapeutic implications. *Mutation Research/Reviews in Mutation Research.* 2011 May;727(3):86–103.
4. Rack JGM, Perina D, Ahel I. Macrodomains: Structure, Function, Evolution, and Catalytic Activities. *Annu Rev Biochem.* 2016 Jun 2;85(1):431–54.
5. Fehr AR, Singh SA, Kerr CM, Mukai S, Higashi H, Aikawa M. The impact of PARPs and ADP-ribosylation on inflammation and host-pathogen interactions. *Genes Dev.* 2020 Mar 1;34(5–6):341–59.
6. Grunewald ME, Chen Y, Kuny C, Maejima T, Lease R, Ferraris D, et al. The coronavirus macrodomain is required to prevent PARP-mediated inhibition of virus replication and enhancement of IFN expression. *PLOS Pathogens.* 2019 May 16;15(5):e1007756.
7. Kuri T, Eriksson KK, Putics A, Züst R, Snijder EJ, Davidson AD, et al. The ADP-ribose-1''-monophosphatase domains of severe acute respiratory syndrome coronavirus and human coronavirus 229E mediate resistance to antiviral interferon responses. *Journal of General Virology.* 2011 Aug 1;92(8):1899–905.
8. Fehr AR, Channappanavar R, Jankevicius G, Fett C, Zhao J, Athmer J, et al. The Conserved Coronavirus Macrodomain Promotes Virulence and Suppresses the Innate Immune Response during Severe Acute Respiratory Syndrome Coronavirus Infection. Denison MR, editor. *mBio.* 2016 Diciembre;7(6):e01721-16.
9. Kirtipal N, Bharadwaj S, Kang SG. From SARS to SARS-CoV-2, insights on structure, pathogenicity and immunity aspects of pandemic human coronaviruses. *Infect Genet Evol.* 2020 Nov;85:104502.
10. WHO Coronavirus (COVID-19) Dashboard [Internet]. [cited 2021 Oct 13]. Available from: <https://covid19.who.int>
11. Kim JH, Marks F, Clemens JD. Looking beyond COVID-19 vaccine phase 3 trials. *Nat Med.* 2021 Feb;27(2):205–11.

12. Valle C, Martin B, Touret F, Shannon A, Canard B, Guillemot J-C, et al. Drugs against SARS-CoV-2: What do we know about their mode of action? *Rev Med Virol.* 2020 Nov;30(6):1–10.
13. Repurposed Antiviral Drugs for Covid-19 — Interim WHO Solidarity Trial Results. *New England Journal of Medicine.* 2021 Feb 11;384(6):497–511.
14. Puranik NV, Rani R, Singh VA, Tomar S, Puntambekar HM, Srivastava P. Evaluation of the Antiviral Potential of Halogenated Dihydrorugosaflavonoids and Molecular Modeling with nsP3 Protein of Chikungunya Virus (CHIKV). *ACS Omega.* 2019 Dec 3;4(23):20335–45.
15. Babar Z, Khan M, Zahra M, Anwar M, Noor K, Hashmi HF, et al. Drug similarity and structure-based screening of medicinal compounds to target macrodomain-I from SARS-CoV-2 to rescue the host immune system: a molecular dynamics study. *J Biomol Struct Dyn.* 2020 Sep 8;1–15.
16. Ni X, Schröder M, Olieric V, Sharpe ME, Hernandez-Olmos V, Proschak E, et al. Structural Insights into Plasticity and Discovery of Remdesivir Metabolite GS-441524 Binding in SARS-CoV-2 Macrodomain. *ACS Med Chem Lett.* 2021 Apr 8;12(4):603–9.
17. Cantini F, Banci L, Altincekic N, Bains JK, Dhamotharan K, Fuks C, et al. <sup>1</sup>H, <sup>13</sup>C, and <sup>15</sup>N backbone chemical shift assignments of the apo and the ADP-ribose bound forms of the macrodomain of SARS-CoV-2 non-structural protein 3b. *Biomol NMR Assign.* 2020;14(2):339–46.
18. Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems | Wiley [Internet]. Wiley.com. [cited 2021 Nov 14]. Available from: <https://www.wiley.com/en-us/Enzyme+Kinetics+%3A+Behavior+and+Analysis+of+Rapid+Equilibrium+and+Steady+State+Enzyme+Systems-p-9780471303091>
19. Michalska K, Kim Y, Jedrzejczak R, Maltseva NI, Stols L, Endres M, et al. Crystal structures of SARS-CoV-2 ADP-ribose phosphatase: from the apo form to ligand complexes. *IUCrJ.* 2020 Sep 1;7(Pt 5):814–24.
20. Lin M-H, Chang S-C, Chiu Y-C, Jiang B-C, Wu T-H, Hsu C-H. Structural, Biophysical, and Biochemical Elucidation of the SARS-CoV-2 Nonstructural Protein 3 Macro Domain. *ACS Infect Dis.* 2020 Nov 13;6(11):2970–8.
21. Frick DN, Viridi RS, Vuksanovic N, Dahal N, Silvaggi NR. Molecular Basis for ADP-Ribose Binding to the Mac1 Domain of SARS-CoV-2 nsp3. *Biochemistry.* 2020 Jul 21;59(28):2608–15.

22. Vigny A, Henry JP. Bovine adrenal tyrosine hydroxylase: comparative study of native and proteolyzed enzyme, and their interaction with anions. *J Neurochem.* 1981 Feb;36(2):483–9.
23. Schmidt J, Wei R, Oeser T, Belisário-Ferrari MR, Barth M, Then J, et al. Effect of Tris, MOPS, and phosphate buffers on the hydrolysis of polyethylene terephthalate films by polyester hydrolases. *FEBS Open Bio.* 2016 Sep;6(9):919–27.
24. Fitzgerald PMD, Wu JK, Toney JH. Unanticipated Inhibition of the Metallo- $\beta$ -lactamase from *Bacteroides fragilis* by 4-Morpholineethanesulfonic Acid (MES): A Crystallographic Study at 1.85-Å Resolution <sup>‡</sup>. *Biochemistry.* 1998 May;37(19):6791–800.
25. Huang K-W, Hsu K-C, Chu L-Y, Yang J-M, Yuan HS, Hsiao Y-Y. Identification of Inhibitors for the DEDDh Family of Exonucleases and a Unique Inhibition Mechanism by Crystal Structure Analysis of CRN-4 Bound with 2-Morpholin-4-ylethanesulfonate (MES). *J Med Chem.* 2016 Sep 8;59(17):8019–29.
26. Ni X, Schröder M, Olieric V, Sharpe ME, Hernandez-Olmos V, Proschak E, et al. Structural Insights into Plasticity and Discovery of Remdesivir Metabolite GS-441524 Binding in SARS-CoV-2 Macrodomain. *ACS Med Chem Lett.* 2021 Apr 8;12(4):603–9.
27. Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model.* 2011 Oct 24;51(10):2778–86.
28. Alhammad YMO, Kashipathy MM, Roy A, Gagné J-P, McDonald P, Gao P, et al. The SARS-CoV-2 Conserved Macrodomain Is a Mono-ADP-Ribosylhydrolase. Gallagher T, editor. *J Virol.* 2020 Nov 6;95(3):e01969-20, /jvi/95/3/JVI.01969-20.atom.
29. Taha M, Gupta BS, Khoiroh I, Lee M-J. Interactions of Biological Buffers with Macromolecules: The Ubiquitous “Smart” Polymer PNIPAM and the Biological Buffers MES, MOPS, and MOPSO. *Macromolecules.* 2011 Nov 8;44(21):8575–89.
30. Egloff M-P, Malet H, Putics A, Heinonen M, Dutartre H, Frangeul A, et al. Structural and Functional Basis for ADP-Ribose and Poly(ADP-Ribose) Binding by Viral Macro Domains. *Journal of Virology.* 2006 Sep 1;80(17):8493–502.
31. Malet H, Coutard B, Jamal S, Dutartre H, Papageorgiou N, Neuvonen M, et al. The Crystal Structures of Chikungunya and Venezuelan Equine Encephalitis Virus nsP3 Macro Domains Define a Conserved Adenosine Binding Pocket. *Journal of Virology.* 2009 Jul 1;83(13):6534–45.



## 5 General Conclusions and perspectives

This thesis aimed to identify and characterize (+) RNA viruses nsPs involved in viral escape to develop antiviral tools. During the replication of these viruses, the accumulation of intermediate dsRNAs trigger the antiviral response in the infected cells. These dsRNA intermediates act as a signal to activate different receptors as PKR, inhibiting protein translation, and OAS, activating RNase L. Despite all the antiviral strategies, the virus has developed several tools to counteract the immune response. nsPs play a key role in this interference by interacting specifically with host factors. Despite the growing interest in the replication complex of emerging viruses in recent years, many questions remain unanswered about the function of different nsP's domains. Biochemical and functional characterization of nsPs are the step forward in understanding the multiple mechanisms and interactions involved in the viral infection. The study of molecular basis of the interactions between RNA viruses and host cells is a field that needs urgent attention.

Acknowledging this soon will make it possible to merge antiviral approaches, directly by inhibiting viral enzymes, and indirectly, by modulating/adequately stimulating the innate immunity in the infected cells. In this context, HEV was the chosen model to study viral escape carried out through the activation of OAS pathway. In alphaviruses, we studied the mechanism of cap structure addition, which is unique in these viruses and poses to be an attractive therapeutic target. In addition, coronavirus was the selected model to investigate the role of the *Macro* domain in viral evasion via ADP ribosylation.

HEV, belonging to the alphavirus-like super family, is now recognized as a global health threat around the globe. HEV causes acute and chronic hepatitis in humans. The HEV ORF1 encodes the viral nsps, essential for RNA replication and crucial for

the down regulation of innate immune response, via the inhibition of IFN production. Bioinformatic analysis allowed us to identify the presence of putative 2-5 PDE conserved motif in the ORF1, which could be involved in the inhibition of OAS/RNase L pathway. The IFN inducible OAS/RNase L system relies on the synthesis of short 2-5 oligonucleotides (2-5A) by OAS that activate the latent cellular RNase L. The latter degrades viral and cellular RNAs which then restricts viral infection. Here, we report  $\beta$ -IFN inhibition in response to PDE HEV ORF1 domains transfection in Huh7.5 cells. The ORF1 domains -XD, XD-HVD, HEL and RdRp reduce  $\beta$ -IFN in comparison to the control. Intriguingly, a high reduction in  $\beta$ -IFN was achieved by the presence of more than one domain. This reduction was accentuated when XD or XD-HVD domains were present. In this work, we tried to determine the implication of these motifs in the inactivation of the innate immune response against HEV. Biochemical results showed that 2-5 A can be degraded only when incubated in the presence of XD-HVD and HEL, leading to the generation of the possible tri- and mono-phosphorylated 2-5 oligo-dimers. The formation of these two compounds can be due to PDE activity, leading to the cleavage of the 2-5A bound, associated to the NTPase activity of HEL. These results suggest that the combination of these domains could play an important role in the virus escape, because the 2-5 oligodimers are not able to activate RNase L. Mutagenesis studies are underway to validate and complete the obtained results. This work will help consolidating the role of HEV ORF1 proteins in viral escape.

As a long-term perspective of this study, it will be interesting to evaluate the implication of these motifs in functions related to RNA metabolism. Taking in to account that viral PDE activities are thought to derive from ancestral activities of the family of proteins related to RNA ligases (LigT-like), found in bacteria and archaea, and involved in the metabolism and maturation of RNAs. This suggests that the Hx(T/S)x motif may present a binding site to RNA and / or nucleotides (ATP, 2-5A, NADP + and ADP derivatives ribose). These interactions could also play an important role in evading the innate immunity.

Alphaviruses are one of the three families of the arthropod-borne viruses (arboviruses) that currently cause human disease. The alphavirus nsP1 carries MTase and GTase activities, involved in the addition of the cap structure of viral RNA. Resistance mutations reported in nsP1 following virus challenging with Inosine-5'-monophosphate dehydrogenase (IMPDH) inhibitors, such as ribavirin, highlighted the role of non-catalytic residues in the cap structure addition process (182). Our study explained that the resistance mechanism is based on a change in affinity to GTP, the substrate of MTase reaction. Furthermore, reverse genetic experiments have revealed residues indicative of a possible susceptibility to ribavirin. Our results represent an important contribution to the understanding of possible resistance mechanisms, crucial for the future rational design of nsP1 inhibitors.

In parallel, we elucidated the role of 5' RNA secondary structures of the VEEV genome in the interaction with nsP1 and the cap structure addition (183). Secondary structures at the 5'-end of the genomic RNA of alphaviruses play a critical role in the regulation of the synthesis of genomic and sgRNAs. In addition, the presence of a hairpin structure at the 5'-end of the genome has been shown to play a crucial role in the innate immunity escape mediated by restriction factor- IFIT1. The latter being involved in the inhibition of viral RNA (vRNA) translation. Our findings showed that nsP1 VEEV effectively caps RNA ranging from 1 to 130 nucleotides. RNA length of more than 130 nts do not represent a good substrate for nsP1 activity. Moreover, we showed that both structure and sequence of 5'-hairpin of the genome are crucial for the binding to nsP1. This suggests that the attenuated phenotype of strain TC-83 could be accentuated by a deficiency in the cap structure acquisition along with the other factors.

Coronaviruses are one of the known largest group of viruses belonging to the order *Nidovirales* that typically affect the respiratory tract of mammals, including humans. Before the current ongoing COVID-19 pandemic, researchers studying CoVs encompassed a small community. This changed when SARS-CoV-2 emerged and the COVID-19 was declared a public health emergency by WHO. The rapid development

of vaccines and their approval provided a glimmer of hope for the world to return to normal. However, the emergence of SARS-CoV-2 variants might cause a resistance to vaccines. An orthogonal and complementary approach to vaccine strategies lies in the development of antiviral drugs against COVID-19. This requires the rapid identification of antiviral targets and the development of specific therapeutic agents.

Data from the literature tend to designate the *Macro* domain of SARS-CoV-2 as a good therapeutic target. However, to be effective, an antiviral agent it must be specific and selective. The only way to bring these two prerequisites together is to broaden fundamental knowledge about the drug target. But also to delimit the differences and possible redundancies between the modules that are maintained within a virus family. The aims of this part of the thesis was based on the biochemical and functional characterization of SARS-CoV-2 *Macro* domain, but also to develop and screen specific inhibitors that are interfering with the binding of ADP-ribose and/or hydrolysis of its derivatives.

The structural data, published recently, described the *Macro* domain folding of SARS-CoV-2 (184). Here we investigated the role of phenylalanine residue 156 (F156) in the *Macro* domain of SARS-CoV-2 and its corresponding asparagine 157 (N157) in the *Macro* domain of SARS-CoV. The mutational analysis of the mentioned residues - F156 and N157 implied only slight variations in ADP-ribose binding. However, these mutations had a drastic impact on the *Macro* domain de-MARylation activity in SARS-CoV-2 and SARS-CoV. Besides, the effect of MOPS analogues in the ADP-ribose binding pocket of SARS-CoV and SARS-CoV-2 *Macro* domain were investigated. Two MOPS analogues, MOPSO and CAPSO were identified as potential candidates for highly specific *Macro* domain inhibitor design. An ELISA was developed to pre-screen the *Macro* domain inhibitors. This new technique allows a rapid screen of a large numbers of samples and conditions at once.

The future perspectives regarding to *Macro* domains will be focused on the design, on the basis of structural data, of small molecules with inhibitory potential. These

molecules will be tested, on the recombinant SARS-CoV and SARS-CoV-2 *Macro* domains, in terms of inhibition of binding and hydrolysis of ADP-ribose derivatives. Active candidates are to be validated in the cellular system. Further, mutagenesis studies will need to be performed to target specific divergent residues between SARS-CoV and SARS-CoV-2 with the objective of highlighting differences between the two viruses in terms of ligand binding and hydrolysis activities. Evaluating the effect of these mutations on susceptibility to IFN and the activation of ISGs in infected cells will increase the knowledge of the specificity of each virus.

Finally, the identification and understanding of viral activities involved in the activation of cellular defence could allow the identification of targets that can be exploited for the generation of specific antiviral agents.

## 6 References

1. Viruses and Human Disease - 2nd Edition [Internet]. [cited 2021 Sep 29]. Available from: <https://www.elsevier.com/books/viruses-and-human-disease/strauss/978-0-12-373741-0>
2. Salonen A, Ahola T, Kääriäinen L. Viral RNA Replication in Association with Cellular Membranes. *Membrane Trafficking in Viral Replication*. 2005;285:139–73.
3. Trask OJ. Nuclear Factor Kappa B (NF- $\kappa$ B) Translocation Assay Development and Validation for High Content Screening. In: Markossian S, Grossman A, Brimacombe K, Arkin M, Auld D, Austin CP, et al., editors. *Assay Guidance Manual* [Internet]. Bethesda (MD): Eli Lilly & Company and the National Center for Advancing Translational Sciences; 2004 [cited 2021 Dec 22]. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK100914/>
4. Melchjorsen J. Learning from the Messengers: Innate Sensing of Viruses and Cytokine Regulation of Immunity—Clues for Treatments and Vaccines. *Viruses*. 2013 Jan 31;5(2):470–527.
5. Bowie AG, Unterholzner L. Viral evasion and subversion of pattern-recognition receptor signalling. *Nat Rev Immunol*. 2008 Dec;8(12):911–22.
6. Li Y, Banerjee S, Goldstein SA, Dong B, Gaughan C, Rath S, et al. Ribonuclease L mediates the cell-lethal phenotype of double-stranded RNA editing enzyme ADAR1 deficiency in a human cell line. Nilsen TW, editor. *eLife*. 2017 Mar 31;6:e25687.
7. 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.
8. Goodfellow I, Chaudhry Y, Gioldasi I, Gerondopoulos A, Natoni A, Labrie L, et al. Calicivirus translation initiation requires an interaction between VPg and eIF4E. *EMBO Rep*. 2005 Oct;6(10):968–72.
9. Cougot N, van Dijk E, Babajko S, Séraphin B. ‘Cap-tabolism.’ *Trends in Biochemical Sciences*. 2004 Aug 1;29(8):436–44.
10. Borden KLB. The eukaryotic translation initiation factor eIF4E wears a “cap” for many occasions. *Translation (Austin)*. 2016 Aug 10;4(2):e1220899.
11. 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;42(5):3228–45.
12. Coutard B, Barral K, Lichi re J, Selisko B, Martin B, Aouadi W, et al. Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. *J Virol.* 2017 Mar 1;91(5):e02202-16.
  13. Ringiard 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.
  14. Tanaka N, Nakanishi M, Kusakabe Y, Goto Y, Kitade Y, Nakamura KT. Structural basis for recognition of 2',5'-linked oligoadenylates by human ribonuclease L. *EMBO J.* 2004 Oct 13;23(20):3929–38.
  15. Drappier M, Michiels T. Inhibition of the OAS/RNase L pathway by viruses. *Current Opinion in Virology.* 2015 Dec;15:19–26.
  16. Han J-Q, Barton DJ. Activation and evasion of the antiviral 2'-5' oligoadenylate synthetase/ribonuclease L pathway by hepatitis C virus mRNA. *RNA.* 2002 Apr;8(4):512–25.
  17. Thornbrough JM, Jha BK, Yount B, Goldstein SA, Li Y, Elliott R, et al. Middle East Respiratory Syndrome Coronavirus NS4b Protein Inhibits Host RNase L Activation. *mBio.* 2016 Mar 29;7(2):e00258.
  18. S nchez-Tacuba L, Rojas M, Arias CF, L pez S. Rotavirus Controls Activation of the 2'-5'-Oligoadenylate Synthetase/RNase L Pathway Using at Least Two Distinct Mechanisms. *J Virol.* 2015 Sep 23;89(23):12145–53.
  19. Gupte R, Liu Z, Kraus WL. PARPs and ADP-ribosylation: recent advances linking molecular functions to biological outcomes. *Genes Dev.* 2017 Jan 15;31(2):101–26.
  20. Vyas S, Chesarone-Cataldo M, Todorova T, Huang Y-H, Chang P. A systematic analysis of the PARP protein family identifies new functions critical for cell physiology. *Nat Commun.* 2013;4:2240.
  21. Fehr AR, Singh SA, Kerr CM, Mukai S, Higashi H, Aikawa M. The impact of PARPs and ADP-ribosylation on inflammation and host-pathogen interactions. *Genes Dev.* 2020 Mar 1;34(5–6):341–59.
  22. Boehi F. Interplay between ADP-ribosyltransferases and essential cell signaling pathways controls cellular responses. 2021;22.



23. Iwata H, Goettsch C, Sharma A, Ricchiuto P, Goh WWB, Halu A, et al. PARP9 and PARP14 cross-regulate macrophage activation via STAT1 ADP-ribosylation. *Nat Commun*. 2016 Oct 31;7:12849.
24. Liu S-Y, Sanchez DJ, Aliyari R, Lu S, Cheng G. Systematic identification of type I and type II interferon-induced antiviral factors. *Proc Natl Acad Sci U S A*. 2012 Mar 13;109(11):4239–44.
25. Atasheva S, Akhrymuk M, Frolova EI, Frolov I. New PARP gene with an anti-alphavirus function. *J Virol*. 2012 Aug;86(15):8147–60.
26. Atasheva S, Frolova EI, Frolov I. Interferon-stimulated poly(ADP-Ribose) polymerases are potent inhibitors of cellular translation and virus replication. *J Virol*. 2014 Feb;88(4):2116–30.
27. Li L, Zhao H, Liu P, Li C, Quanquin N, Ji X, et al. PARP12 suppresses Zika virus infection through PARP-dependent degradation of NS1 and NS3 viral proteins. *Sci Signal*. 2018 Jun 19;11(535):eaas9332.
28. Jankevicius G, Hassler M, Golia B, Rybin V, Zacharias M, Timinszky G, et al. A family of macrodomain proteins reverses cellular mono-ADP-ribosylation. *Nat Struct Mol Biol*. 2013 Apr;20(4):508–14.
29. Aggarwal R. Hepatitis E: Historical, contemporary and future perspectives. *J Gastroenterol Hepatol*. 2011 Jan;26 Suppl 1:72–82.
30. TEO C-G. HISTORICAL REVIEW: Fatal outbreaks of jaundice in pregnancy and the epidemic history of hepatitis E. *Epidemiology and Infection*. 2012;140(5):767–87.
31. Smith DB, Simmonds P, Members Of The International Committee On The Taxonomy Of Viruses Study Group null, Jameel S, Emerson SU, Harrison TJ, et al. Consensus proposals for classification of the family Hepeviridae. *J Gen Virol*. 2014 Oct;95(Pt 10):2223–32.
32. Smith DB, Simmonds P, Izopet J, Oliveira-Filho EF, Ulrich RG, Johne R, et al. Proposed reference sequences for hepatitis E virus subtypes. *J Gen Virol*. 2016 Mar;97(3):537–42.
33. Sun P, Lin S, He S, Zhou E-M, Zhao Q. Avian Hepatitis E Virus: With the Trend of Genotypes and Host Expansion. *Frontiers in Microbiology*. 2019;10:1696.
34. Smith DB, Purdy MA, Simmonds P. Genetic variability and the classification of hepatitis E virus. *J Virol*. 2013 Apr;87(8):4161–9.

35. Fierro NA, Realpe M, Meraz-Medina T, Roman S, Panduro A. Hepatitis E virus: An ancient hidden enemy in Latin America. *World J Gastroenterol*. 2016 Feb 21;22(7):2271–83.
36. Realpe-Quintero M, Viera-Segura O, Fierro NA. Hepatitis E Virus: Still an Enigma in Mexico. *Annals of Hepatology*. 2018 Jul 1;17(4):544–6.
37. Wang B, Akanbi OA, Harms D, Adesina O, Osundare FA, Naidoo D, et al. A new hepatitis E virus genotype 2 strain identified from an outbreak in Nigeria, 2017. *Virology Journal*. 2018 Oct 23;15(1):163.
38. Nan Y, Zhang Y-J. Molecular Biology and Infection of Hepatitis E Virus. *Frontiers in Microbiology*. 2016;7:1419.
39. Rendon J, Hoyos MC, di Filippo D, Cortes-Mancera F, Mantilla C, Velasquez MM, et al. Hepatitis E Virus Genotype 3 in Colombia: Survey in Patients with Clinical Diagnosis of Viral Hepatitis. *PLoS One*. 2016;11(2):e0148417.
40. Melgaço JG, Gardinali NR, de Mello V da M, Leal M, Lewis-Ximenez LL, Pinto MA. Hepatitis E: Update on Prevention and Control. *Biomed Res Int*. 2018;2018:5769201.
41. Purcell RH, Emerson SU. Hepatitis E: an emerging awareness of an old disease. *J Hepatol*. 2008 Mar;48(3):494–503.
42. Mirazo S, Ramos N, Mainardi V, Gerona S, Arbiza J. Transmission, diagnosis, and management of hepatitis E: an update. *HMER*. 2014 Jun 3;6:45–59.
43. Yugo DM, Meng X-J. Hepatitis E Virus: Foodborne, Waterborne and Zoonotic Transmission. *Int J Environ Res Public Health*. 2013 Oct;10(10):4507–33.
44. Ren F, Zhao C, Wang L, Wang Z, Gong X, Song M, et al. Hepatitis E virus seroprevalence and molecular study among blood donors in China. *Transfusion*. 2014 Mar;54(3 Pt 2):910–7.
45. Khuroo MS, Kamili S, Khuroo MS. Clinical course and duration of viremia in vertically transmitted hepatitis E virus (HEV) infection in babies born to HEV-infected mothers. *J Viral Hepat*. 2009 Jul;16(7):519–23.
46. Hewitt PE, Ijaz S, Brailsford SR, Brett R, Dicks S, Haywood B, et al. Hepatitis E virus in blood components: a prevalence and transmission study in southeast England. *Lancet*. 2014 Nov 15;384(9956):1766–73.
47. Domanović D, Tedder R, Blümel J, Zaaier H, Gallian P, Niederhauser C, et al. Hepatitis E and blood donation safety in selected European countries: a shift to screening? *Euro Surveill*. 2017 Apr 20;22(16):30514.

48. Dalton HR, Kamar N, Izopet J. Hepatitis E in developed countries: current status and future perspectives. *Future Microbiol.* 2014;9(12):1361–72.
49. Clemente-Casares P, Ramos-Romero C, Ramirez-Gonzalez E, Mas A. Hepatitis E Virus in Industrialized Countries: The Silent Threat. *BioMed Research International.* 2016 Dec 14;2016:e9838041.
50. Yu C, Engle RE, Bryan JP, Emerson SU, Purcell RH. Detection of immunoglobulin M antibodies to hepatitis E virus by class capture enzyme immunoassay. *Clin Diagn Lab Immunol.* 2003 Jul;10(4):579–86.
51. Hepatitis E [Internet]. [cited 2021 Sep 13]. Available from: <https://www.who.int/news-room/fact-sheets/detail/hepatitis-e>
52. Krawczynski K, Meng X-J, Rybczynska J. Pathogenetic elements of hepatitis E and animal models of HEV infection. *Virus Res.* 2011 Oct;161(1):78–83.
53. Xu B, Yu H-B, Hui W, He J-L, Wei L-L, Wang Z, et al. Clinical features and risk factors of acute hepatitis E with severe jaundice. *World J Gastroenterol.* 2012 Dec 28;18(48):7279–84.
54. Wedemeyer H, Pischke S, Manns MP. Pathogenesis and treatment of hepatitis e virus infection. *Gastroenterology.* 2012 May;142(6):1388-1397.e1.
55. Faber M, Willrich N, Schemmerer M, Rauh C, Kuhnert R, Stark K, et al. Hepatitis E virus seroprevalence, seroincidence and seroreversion in the German adult population. *J Viral Hepat.* 2018 Jun;25(6):752–8.
56. Navaneethan U, Mohajer MA, Shata MT. Hepatitis E and Pregnancy-Understanding the pathogenesis. *Liver Int.* 2008 Nov;28(9):1190–9.
57. Tsarev SA, Emerson SU, Reyes GR, Tsareva TS, Legters LJ, Malik IA, et al. Characterization of a prototype strain of hepatitis E virus. *Proc Natl Acad Sci U S A.* 1992 Jan 15;89(2):559–63.
58. Graff J, Zhou Y-H, Torian U, Nguyen H, St Claire M, Yu C, et al. Mutations within potential glycosylation sites in the capsid protein of hepatitis E virus prevent the formation of infectious virus particles. *J Virol.* 2008 Feb;82(3):1185–94.
59. Nair VP, Anang S, Subramani C, Madhvi A, Bakshi K, Srivastava A, et al. Endoplasmic Reticulum Stress Induced Synthesis of a Novel Viral Factor Mediates Efficient Replication of Genotype-1 Hepatitis E Virus. *PLoS Pathog.* 2016 Apr 1;12(4):e1005521.
60. LeDesma R, Nimgaonkar I, Ploss A. Hepatitis E Virus Replication. *Viruses.* 2019 Aug 6;11(8):E719.

61. Parvez MK. The hepatitis E virus nonstructural polyprotein. *Future Microbiol.* 2017 Aug;12:915–24.
62. Osterman A, Stellberger T, Gebhardt A, Kurz M, Friedel CC, Uetz P, et al. The Hepatitis E virus intraviral interactome. *Sci Rep.* 2015 Oct 14;5:13872.
63. Sehgal D, Thomas S, Chakraborty M, Jameel S. Expression and processing of the Hepatitis E virus ORF1 nonstructural polyprotein. *Virology Journal.* 2006 May 26;3(1):38.
64. Ansari IH, Nanda SK, Durgapal H, Agrawal S, Mohanty SK, Gupta D, et al. Cloning, sequencing, and expression of the hepatitis E virus (HEV) nonstructural open reading frame 1 (ORF1). *J Med Virol.* 2000 Mar;60(3):275–83.
65. Saraswat S, Chaudhary M, Sehgal D. Hepatitis E Virus Cysteine Protease Has Papain Like Properties Validated by in silico Modeling and Cell-Free Inhibition Assays. *Frontiers in Cellular and Infection Microbiology.* 2020;9:478.
66. Magden J, Takeda N, Li T, Auvinen P, Ahola T, Miyamura T, et al. Virus-Specific mRNA Capping Enzyme Encoded by Hepatitis E Virus. *J Virol.* 2001 Jul;75(14):6249–55.
67. Ahola T, Kääriäinen L. Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. *Proc Natl Acad Sci U S A.* 1995 Jan 17;92(2):507–11.
68. Magden J, Takeda N, Li T, Auvinen P, Ahola T, Miyamura T, et al. Virus-specific mRNA capping enzyme encoded by hepatitis E virus. *J Virol.* 2001 Jul 1;75(14):6249–55.
69. Emerson SU, Nguyen H, Graff J, Stephany DA, Brockington A, Purcell RH. In Vitro Replication of Hepatitis E Virus (HEV) Genomes and of an HEV Replicon Expressing Green Fluorescent Protein. *J Virol.* 2004 May;78(9):4838–46.
70. Borkakoti J, Ahmed G, Rai A, Kar P. Report of novel H105R, D29N, V27A mutations in the methyltransferase region of the HEV genome in patients with acute liver failure. *Journal of Clinical Virology.* 2017 Jun 1;91:1–4.
71. Ahola T, Karlin DG. Sequence analysis reveals a conserved extension in the capping enzyme of the alphavirus supergroup, and a homologous domain in nodaviruses. *Biology Direct.* 2015 Apr 11;10(1):16.
72. Rabah N, Coutard B, Canard B. Les protéines non structurales des Alphavirus : rôle dans la réplication et l'interaction du virus avec la cellule hôte. *Virologie (Montrouge).* 2013 Feb 1;17(1):31–45.

73. Milligan G, Parenti M, Magee AI. The dynamic role of palmitoylation in signal transduction. *Trends Biochem Sci.* 1995 May;20(5):181–7.
74. Parvez MK. Mutational analysis of hepatitis E virus ORF1 “Y-domain”: Effects on RNA replication and virion infectivity. *World J Gastroenterol.* 2017 Jan 28;23(4):590–602.
75. Nan Y, Yu Y, Ma Z, Khattar SK, Fredericksen B, Zhang Y-J. Hepatitis E Virus Inhibits Type I Interferon Induction by ORF1 Products. *Journal of Virology.* 2014 Oct 15;88(20):11924–32.
76. Karpe YA, Lole KS. Deubiquitination activity associated with hepatitis E virus putative papain-like cysteine protease. *J Gen Virol.* 2011 Sep;92(Pt 9):2088–92.
77. Purdy MA, Lara J, Khudyakov YE. The hepatitis E virus polyproline region is involved in viral adaptation. *PLoS One.* 2012;7(4):e35974.
78. Koonin EV, Gorbalenya AE, Purdy MA, Rozanov MN, Reyes GR, Bradley DW. Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of an additional group of positive-strand RNA plant and animal viruses. *Proc Natl Acad Sci U S A.* 1992 Sep 1;89(17):8259–63.
79. Pudupakam RS, Huang YW, Opriessnig T, Halbur PG, Pierson FW, Meng XJ. Deletions of the hypervariable region (HVR) in open reading frame 1 of hepatitis E virus do not abolish virus infectivity: evidence for attenuation of HVR deletion mutants in vivo. *J Virol.* 2009 Jan;83(1):384–95.
80. Lhomme S, Abravanel F, Dubois M, Sandres-Saune K, Mansuy J-M, Rostaing L, et al. Characterization of the polyproline region of the hepatitis E virus in immunocompromised patients. *J Virol.* 2014 Oct;88(20):12017–25.
81. Diefenbach J, Bürkle A. Introduction to poly(ADP-ribose) metabolism. *Cell Mol Life Sci.* 2005 Apr;62(7–8):721–30.
82. Vikram T, Kumar P. Analysis of Hepatitis E virus (HEV) X-domain structural model. *Bioinformation.* 2018 Jul 31;14(7):398–403.
83. Li C, Debing Y, Jankevicius G, Neyts J, Ahel I, Coutard B, et al. Viral Macro Domains Reverse Protein ADP-Ribosylation. Perlman S, editor. *J Virol.* 2016 Oct 1;90(19):8478.
84. Parvez MK. The hepatitis E virus ORF1 “X-domain” residues form a putative macrodomain protein/Appr-1”-pase catalytic-site, critical for viral RNA replication. *Gene.* 2015 Jul 15;566(1):47–53.

- 
85. Borkakoti J, Ahmed G, Kar P. Report of a novel C1483W mutation in the hepatitis E virus polymerase in patients with acute liver failure. *Infection, Genetics and Evolution*. 2016 Oct;44:51–4.
  86. Ojha NK, Lole KS. Hepatitis E virus ORF1 encoded macro domain protein interacts with light chain subunit of human ferritin and inhibits its secretion. *Mol Cell Biochem*. 2016 Jun;417(1–2):75–85.
  87. Devhare P, Sharma K, Mhaindarkar V, Arankalle V, Lole K. Analysis of helicase domain mutations in the hepatitis E virus derived from patients with fulminant hepatic failure: Effects on enzymatic activities and virus replication. *Virus Res*. 2014 May 12;184:103–10.
  88. Cao D, Ni Y-Y, Meng X-J. Substitution of amino acid residue V1213 in the helicase domain of the genotype 3 hepatitis E virus reduces virus replication. *Virology Journal*. 2018 Feb 8;15(1):32.
  89. Smith DB, Simmonds P. Hepatitis E virus and fulminant hepatitis—a virus or host-specific pathology? *Liver Int*. 2015 Apr;35(4):1334–40.
  90. Koonin EV. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *Journal of General Virology*. 1991 Sep 1;72(9):2197–206.
  91. Agrawal S, Gupta D, Panda SK. The 3' end of hepatitis E virus (HEV) genome binds specifically to the viral RNA-dependent RNA polymerase (RdRp). *Virology*. 2001 Mar 30;282(1):87–101.
  92. Mahilkar S, Paingankar MS, Lole KS. Hepatitis E virus RNA-dependent RNA polymerase: RNA template specificities, recruitment and synthesis. *J Gen Virol*. 2016 Sep;97(9):2231–42.
  93. Todt D, Gisa A, Radonic A, Nitsche A, Behrendt P, Suneetha PV, et al. In vivo evidence for ribavirin-induced mutagenesis of the hepatitis E virus genome. *Gut*. 2016 Oct;65(10):1733–43.
  94. Debing Y, Gisa A, Dallmeier K, Pischke S, Bremer B, Manns M, et al. A Mutation in the Hepatitis E Virus RNA Polymerase Promotes Its Replication and Associates With Ribavirin Treatment Failure in Organ Transplant Recipients. *Gastroenterology*. 2014 Nov 1;147(5):1008-1011.e7.
  95. Shiota T, Li T-C, Yoshizaki S, Kato T, Wakita T, Ishii K. The hepatitis E virus capsid C-terminal region is essential for the viral life cycle: implication for viral genome encapsidation and particle stabilization. *J Virol*. 2013 May;87(10):6031–6.

- 
96. Lenggenger D, Gouttenoire J, Malehmir M, Bawohl M, Honcharova-Biletska H, Kreutzer S, et al. Visualization of hepatitis E virus RNA and proteins in the human liver. *J Hepatol*. 2017 Sep;67(3):471–9.
  97. Tang Z-M, Tang M, Zhao M, Wen G-P, Yang F, Cai W, et al. A novel linear neutralizing epitope of hepatitis E virus. *Vaccine*. 2015 Jul 9;33(30):3504–11.
  98. Shrestha MP, Scott RM, Joshi DM, Mammen MP, Thapa GB, Thapa N, et al. Safety and Efficacy of a Recombinant Hepatitis E Vaccine. *New England Journal of Medicine*. 2007 Mar 1;356(9):895–903.
  99. Yamada K, Takahashi M, Hoshino Y, Takahashi H, Ichiyama K, Nagashima S, et al. ORF3 protein of hepatitis E virus is essential for virion release from infected cells. *Journal of General Virology*. 2009 Aug 1;90(8):1880–91.
  100. Semiletov IA, Dement'eva EV null, Iashina TL, Favorov MO, Shibnev VA. [Synthesis and antigenic activity of peptides from the ORF3 protein sequence of hepatitis E virus]. *Bioorg Khim*. 1995 Feb;21(2):156–7.
  101. Subramani C, Nair VP, Anang S, Mandal SD, Pareek M, Kaushik N, et al. Host-Virus Protein Interaction Network Reveals the Involvement of Multiple Host Processes in the Life Cycle of Hepatitis E Virus. *mSystems*. 2018 Feb;3(1):e00135-17.
  102. Primadharsini PP, Nagashima S, Okamoto H. Genetic Variability and Evolution of Hepatitis E Virus. *Viruses*. 2019 May 18;11(5):456.
  103. Yadav KK, Boley PA, Fritts Z, Kenney SP. Ectopic Expression of Genotype 1 Hepatitis E Virus ORF4 Increases Genotype 3 HEV Viral Replication in Cell Culture. *Viruses*. 2021 Jan 7;13(1):75.
  104. Affeldt P, Di Cristanziano V, Grundmann F, Wirtz M, Kaiser R, Benzing T, et al. Monitoring of hepatitis E virus RNA during treatment for chronic hepatitis E virus infection after renal transplantation. *Immun Inflamm Dis*. 2021 Jun;9(2):513–20.
  105. Cao Y, Bing Z, Guan S, Zhang Z, Wang X. Development of new hepatitis E vaccines. *Hum Vaccin Immunother*. 2018 Jun 18;14(9):2254–62.
  106. Zhang X, Wei M, Pan H, Lin Z, Wang K, Weng Z, et al. Robust manufacturing and comprehensive characterization of recombinant hepatitis E virus-like particles in Hecolin®. *Vaccine*. 2014 Jul 7;32(32):4039–50.
  107. Ma H, Song X, Harrison TJ, Li R, Huang G, Zhang H, et al. Immunogenicity and efficacy of a bacterially expressed HEV ORF3 peptide, assessed by experimental infection of primates. *Arch Virol*. 2009;154(10):1641–8.



108. Naik S, Aggarwal R, Naik SR, Dwivedi S, Talwar S, Tyagi SK, et al. Evidence for activation of cellular immune responses in patients with acute hepatitis E. *Indian J Gastroenterol*. 2002 Aug;21(4):149–52.
109. Weaver SC, Ferro C, Barrera R, Boshell J, Navarro J-C. Venezuelan equine encephalitis. *Annu Rev Entomol*. 2004;49:141–74.
110. Azar SR, Campos RK, Bergren NA, Camargos VN, Rossi SL. Epidemic Alphaviruses: Ecology, Emergence and Outbreaks. *Microorganisms*. 2020 Aug 1;8(8):1167.
111. McLoughlin MF, Graham DA. Alphavirus infections in salmonids – a review. *Journal of Fish Diseases*. 2007;30(9):511–31.
112. Forrester NL, Palacios G, Tesh RB, Savji N, Guzman H, Sherman M, et al. Genome-Scale Phylogeny of the Alphavirus Genus Suggests a Marine Origin. *J Virol*. 2012 Mar;86(5):2729–38.
113. Fields BN, Knipe DM, Howley PM, Griffin DE. *Fields virology*. Philadelphia: Lippincott Williams & Wilkins; 2001.
114. Crosby B, Crespo ME. Venezuelan Equine Encephalitis. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 [cited 2021 Oct 7]. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK559332/>
115. Staples JE, Breiman RF, Powers AM. Chikungunya Fever: An Epidemiological Review of a Re-Emerging Infectious Disease. *Clinical Infectious Diseases*. 2009 Sep 15;49(6):942–8.
116. Ganjian N, Riviere-Cinamond A. Mayaro virus in Latin America and the Caribbean. *Revista Panamericana de Salud Pública*. 2020 Feb 11;44:1.
117. Yu W, Mengersen K, Dale P, Mackenzie JS, Toloo G (Sam), Wang X, et al. Epidemiologic Patterns of Ross River Virus Disease in Queensland, Australia, 2001–2011. *Am J Trop Med Hyg*. 2014 Jul 2;91(1):109–18.
118. Lwande OW, Obanda V, Bucht G, Mosomtai G, Otieno V, Ahlm C, et al. Global emergence of Alphaviruses that cause arthritis in humans. *Infect Ecol Epidemiol*. 2015;5:29853.
119. Pedrosa PBS. Viral infections in workers in hospital and research laboratory settings: a comparative review of infection modes and respective biosafety aspects. *International Journal of Infectious Diseases*. 2011;11.
120. Stromberg ZR, Fischer W, Bradfute SB, Kubicek-Sutherland JZ, Hraber P. Vaccine Advances against Venezuelan, Eastern, and Western Equine Encephalitis Viruses. *Vaccines*. 2020 Jun 3;8(2):273.



121. Schmaljohn AL, McClain D. Alphaviruses (Togaviridae) and Flaviviruses (Flaviviridae). In: Baron S, editor. Medical Microbiology [Internet]. 4th ed. Galveston (TX): University of Texas Medical Branch at Galveston; 1996 [cited 2021 Oct 7]. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK7633/>
122. Pardigon N, Strauss JH. Mosquito homolog of the La autoantigen binds to Sindbis virus RNA. *Journal of Virology*. 1996 Feb;70(2):1173–81.
123. Leung JY-S, Ng MM-L, Chu JJH. Replication of Alphaviruses: A Review on the Entry Process of Alphaviruses into Cells. *Adv Virol*. 2011;2011:249640.
124. Holmes AC, Basore K, Fremont DH, Diamond MS. A molecular understanding of alphavirus entry. *PLOS Pathogens*. 2020 Oct 22;16(10):e1008876.
125. Shirako Y, Strauss JH. Regulation of Sindbis virus RNA replication: uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. *Journal of Virology*. 1994 Mar;68(3):1874–85.
126. Rupp JC, Sokoloski KJ, Gebhart NN, Hardy RW. Alphavirus RNA synthesis and non-structural protein functions. *J Gen Virol*. 2015 Sep;96(Pt 9):2483–500.
127. Spuul P, Balistreri G, Hellström K, Golubtsov AV, Jokitalo E, Ahola T. Assembly of Alphavirus Replication Complexes from RNA and Protein Components in a Novel trans-Replication System in Mammalian Cells ▽. *J Virol*. 2011 May;85(10):4739–51.
128. Rice CM. Examples of expression systems based on animal RNA viruses: Alphaviruses and influenza virus. *Current Opinion in Biotechnology*. 1992 Oct;3(5):523–32.
129. Vasiljeva L, Merits A, Auvinen P, Kääriäinen L. Identification of a novel function of the alphavirus capping apparatus. RNA 5'-triphosphatase activity of Nsp2. *J Biol Chem*. 2000 Jun 9;275(23):17281–7.
130. Hyde JL, Gardner CL, Kimura T, White JP, Liu G, Trobaugh DW, et al. A viral RNA structural element alters host recognition of nonself RNA. *Science*. 2014 Feb 14;343(6172):783–7.
131. Jones R, Bragagnolo G, Arranz R, Reguera J. Capping pores of alphavirus nsP1 gate membranous viral replication factories. *Nature*. 2021 Jan;589(7843):615–9.
132. Spuul P, Salonen A, Merits A, Jokitalo E, Kääriäinen L, Ahola T. Role of the Amphipathic Peptide of Semliki Forest Virus Replicase Protein nsP1 in Membrane Association and Virus Replication. *J Virol*. 2007 Jan;81(2):872–83.

- 
133. Shirako Y, Strauss EG, Strauss JH. Suppressor mutations that allow sindbis virus RNA polymerase to function with nonaromatic amino acids at the N-terminus: evidence for interaction between nsP1 and nsP4 in minus-strand RNA synthesis. *Virology*. 2000 Oct 10;276(1):148–60.
134. Sawicki DL, Perri S, Polo JM, Sawicki SG. Role for nsP2 Proteins in the Cessation of Alphavirus Minus-Strand Synthesis by Host Cells. *J Virol*. 2006 Jan;80(1):360–71.
135. Lain S, Riechmann JL, Garcíá JA. RNA helicase: a novel activity associated with a protein encoded by a positive strand RNA virus. *Nucl Acids Res*. 1990;18(23):7003–6.
136. Russo AT, White MA, Watowich SJ. The Crystal Structure of the Venezuelan Equine Encephalitis Alphavirus nsP2 Protease. *Structure*. 2006 Sep 1;14(9):1449–58.
137. Fros JJ, van der Maten E, Vlak JM, Pijlman GP. The C-Terminal Domain of Chikungunya Virus nsP2 Independently Governs Viral RNA Replication, Cytopathicity, and Inhibition of Interferon Signaling. *J Virol*. 2013 Sep;87(18):10394–400.
138. Wang Y, Sawicki S, Sawicki D. Alphavirus nsP3 functions to form replication complexes transcribing negative-strand RNA. *Journal of virology*. 1994 Nov 1;68:6466–75.
139. Götte B, Liu L, McInerney GM. The Enigmatic Alphavirus Non-Structural Protein 3 (nsP3) Revealing Its Secrets at Last. *Viruses*. 2018 Feb 28;10(3):105.
140. Gao Y, Goonawardane N, Ward J, Tuplin A, Harris M. Multiple roles of the non-structural protein 3 (nsP3) alphavirus unique domain (AUD) during Chikungunya virus genome replication and transcription. *PLOS Pathogens*. 2019 Jan 22;15(1):e1007239.
141. Tomar S, Hardy RW, Smith JL, Kuhn RJ. Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. *J Virol*. 2006 Oct;80(20):9962–9.
142. Lopez S, Yao JS, Kuhn RJ, Strauss EG, Strauss JH. Nucleocapsid-glycoprotein interactions required for assembly of alphaviruses. *J Virol*. 1994 Mar;68(3):1316–23.
143. Lei J, Kusov Y, Hilgenfeld R. Nsp3 of coronaviruses: Structures and functions of a large multi-domain protein. *Antiviral Research*. 2018 Jan;149:58–74.
144. Allen MD, Buckle AM, Cordell SC, Löwe J, Bycroft M. The Crystal Structure of AF1521 a Protein from *Archaeoglobus fulgidus* with Homology to the Non-

- histone Domain of MacroH2A. *Journal of Molecular Biology*. 2003 Jul;330(3):503–11.
145. Han W, Li X, Fu X. The macro domain protein family: Structure, functions, and their potential therapeutic implications. *Mutation Research/Reviews in Mutation Research*. 2011 May;727(3):86–103.
  146. Putics A, Filipowicz W, Hall J, Gorbalenya AE, Ziebuhr J. ADP-Ribose-1"-Monophosphatase: a Conserved Coronavirus Enzyme That Is Dispensable for Viral Replication in Tissue Culture. *Journal of Virology*. 2005 Oct 15;79(20):12721–31.
  147. Saikatendu KS, Joseph JS, Subramanian V, Clayton T, Griffith M, Moy K, et al. Structural Basis of Severe Acute Respiratory Syndrome Coronavirus ADP-Ribose-1"-Phosphate Dephosphorylation by a Conserved Domain of nsP3. *Structure*. 2005 Nov;13(11):1665–75.
  148. Fehr AR, Jankevicius G, Ahel I, Perlman S. Viral Macrodomains: Unique Mediators of Viral Replication and Pathogenesis. *Trends in Microbiology*. 2018 Jul;26(7):598–610.
  149. Heer CD, Sanderson DJ, Voth LS, Alhammad YMO, Schmidt MS, Trammell SAJ, et al. Coronavirus infection and PARP expression dysregulate the NAD metabolome: An actionable component of innate immunity. *J Biol Chem*. 2020 Dec 25;295(52):17986–96.
  150. 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.
  151. Bank RPD. RCSB PDB - 2FAV: Crystal structure of SARS macro domain in complex with ADP-ribose at 1.8 Å resolution [Internet]. [cited 2021 Dec 14]. Available from: <https://www.rcsb.org/structure/2FAV>
  152. Bank RPD. RCSB PDB - 6WOJ: Structure of the SARS-CoV-2 macrodomain (NSP3) in complex with ADP-ribose [Internet]. [cited 2021 Dec 14]. Available from: <https://www.rcsb.org/structure/6woj>
  153. Puranik NV, Rani R, Singh VA, Tomar S, Puntambekar HM, Srivastava P. Evaluation of the Antiviral Potential of Halogenated Dihydroflavonoids and Molecular Modeling with nsP3 Protein of Chikungunya Virus (CHIKV). *ACS Omega*. 2019 Dec 3;4(23):20335–45.
  154. Shimizu JF, Martins DOS, McPhillie MJ, Roberts GC, Zothner C, Merits A, et al. Is the ADP ribose site of the Chikungunya virus NSP3 Macro domain a target for antiviral approaches? *Acta Tropica*. 2020 Jul 1;207:105490.

- 
155. Fu W, Yao H, Bütepage M, Zhao Q, Lüscher B, Li J. The search for inhibitors of macrodomains for targeting the readers and erasers of mono-ADP-ribosylation. *Drug Discovery Today*. 2021 Nov;26(11):2547–58.
156. Ni X, Schröder M, Olieric V, Sharpe ME, Hernandez-Olmos V, Proschak E, et al. Structural Insights into Plasticity and Discovery of Remdesivir Metabolite GS-441524 Binding in SARS-CoV-2 Macrodomain. *ACS Med Chem Lett*. 2021 Apr 8;12(4):603–9.
157. Wilde AH, Snijder EJ, Kikkert M, van Hemert MJ. Host Factors in Coronavirus Replication. *Curr Top Microbiol Immunol*. 2018;419:1–42.
158. Chathappady House NN, Palissery S, Sebastian H. Corona Viruses: A Review on SARS, MERS and COVID-19. *Microbiol Insights*. 2021;14:11786361211002480.
159. Abdelghany TM, Ganash M, Bakri MM, Qanash H, Al-Rajhi AMH, Elhussieny NI. SARS-CoV-2, the other face to SARS-CoV and MERS-CoV: Future predictions. *Biomedical Journal*. 2021 Feb;44(1):86–93.
160. WHO Coronavirus (COVID-19) Dashboard [Internet]. [cited 2021 Oct 13]. Available from: <https://covid19.who.int>
161. Abdelghany TM, Ganash M, Bakri MM, Qanash H, Al-Rajhi AMH, Elhussieny NI. SARS-CoV-2, the other face to SARS-CoV and MERS-CoV: Future predictions. *Biomed J*. 2021 Mar;44(1):86–93.
162. Tahamtan A, Ardebili A. Real-time RT-PCR in COVID-19 detection: issues affecting the results. *Expert Rev Mol Diagn*. 2020 May;20(5):453–4.
163. Shereen MA, Khan S, Kazmi A, Bashir N, Siddique R. COVID-19 infection: Emergence, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research*. 2020 Jul 1;24:91–8.
164. Fehr AR, Perlman S. Coronaviruses: An Overview of Their Replication and Pathogenesis. *Coronaviruses*. 2015 Feb 12;1282:1–23.
165. Plant EP, Rakauskaitė R, Taylor DR, Dinman JD. Achieving a Golden Mean: Mechanisms by Which Coronaviruses Ensure Synthesis of the Correct Stoichiometric Ratios of Viral Proteins. *J Virol*. 2010 May;84(9):4330–40.
166. Ma Y, Wu L, Shaw N, Gao Y, Wang J, Sun Y, et al. Structural basis and functional analysis of the SARS coronavirus nsp14–nsp10 complex. *Proc Natl Acad Sci USA*. 2015 Jul 28;112(30):9436–41.
167. Krafcikova P, Silhan J, Nencka R, Boura E. Structural analysis of the SARS-CoV-2 methyltransferase complex involved in RNA cap creation bound to sinefungin. *Nat Commun*. 2020 Dec;11(1):3717.

- 
168. Ruch TR, Machamer CE. The Coronavirus E Protein: Assembly and Beyond. *Viruses*. 2012 Mar 8;4(3):363–82.
169. Wojdyla JA, Manolaridis I, van Kasteren PB, Kikkert M, Snijder EJ, Gorbalenya AE, et al. Papain-like protease 1 from transmissible gastroenteritis virus: crystal structure and enzymatic activity toward viral and cellular substrates. *J Virol*. 2010 Oct 1;84(19):10063–73.
170. Wang G, Chen G, Zheng D, Cheng G, Tang H. PLP2 of Mouse Hepatitis Virus A59 (MHV-A59) Targets TBK1 to Negatively Regulate Cellular Type I Interferon Signaling Pathway. *PLoS One*. 2011 Feb 18;6(2):e17192.
171. Lei J, Ma-Lauer Y, Han Y, Thoms M, Buschauer R, Jores J, et al. The SARS-unique domain (SUD) of SARS-CoV and SARS-CoV-2 interacts with human Paip1 to enhance viral RNA translation. *EMBO J*. 2021 Jun 1;40(11):e102277.
172. Harcourt BH, Jukneliene D, Kanjanahaluethai A, Bechill J, Severson KM, Smith CM, et al. Identification of severe acute respiratory syndrome coronavirus replicase products and characterization of papain-like protease activity. *J Virol*. 2004 Dec;78(24):13600–12.
173. Gordon CJ, Tchesnokov EP, Feng JY, Porter DP, Götte M. The antiviral compound remdesivir potently inhibits RNA-dependent RNA polymerase from Middle East respiratory syndrome coronavirus. *J Biol Chem*. 2020 Apr 10;295(15):4773–9.
174. Gordon CJ, Tchesnokov EP, Woolner E, Perry JK, Feng JY, Porter DP, et al. Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase from severe acute respiratory syndrome coronavirus 2 with high potency. *J Biol Chem*. 2020 May 15;295(20):6785–97.
175. Chen P-L, Lee N-Y, Cia C-T, Ko W-C, Hsueh P-R. A Review of Treatment of Coronavirus Disease 2019 (COVID-19): Therapeutic Repurposing and Unmet Clinical Needs. *Frontiers in Pharmacology*. 2020;11:1782.
176. Wu C-Y, Jan J-T, Ma S-H, Kuo C-J, Juan H-F, Cheng Y-SE, et al. Small molecules targeting severe acute respiratory syndrome human coronavirus. *PNAS*. 2004 Jul 6;101(27):10012–7.
177. Furuta Y, Komeno T, Nakamura T. Favipiravir (T-705), a broad spectrum inhibitor of viral RNA polymerase. *Proc Jpn Acad Ser B Phys Biol Sci*. 2017;93(7):449–63.
178. Wang Z, Yang B, Li Q, Wen L, Zhang R. Clinical Features of 69 Cases With Coronavirus Disease 2019 in Wuhan, China. *Clin Infect Dis*. 2020 Jul 28;71(15):769–77.

- 
179. Zhao S, Zhang H, Yang X, Zhang H, Chen Y, Zhan Y, et al. Identification of potent human neutralizing antibodies against SARS-CoV-2 implications for development of therapeutics and prophylactics. *Nat Commun*. 2021 Aug 9;12(1):4887.
180. Li Y-D, Chi W-Y, Su J-H, Ferrall L, Hung C-F, Wu T-C. Coronavirus vaccine development: from SARS and MERS to COVID-19. *Journal of Biomedical Science*. 2020 Dec 20;27(1):104.
181. Krammer F. SARS-CoV-2 vaccines in development. *Nature*. 2020 Oct;586(7830):516–27.
182. Rabah N, Ortega Granda O, Quérat G, Canard B, Decroly E, Coutard B. Mutations on VEEV nsP1 relate RNA capping efficiency to ribavirin susceptibility. *Antiviral Res*. 2020 Oct;182:104883.
183. Granda OO, Valle C, Shannon A, Decroly E, Canard B, Coutard B, et al. Structure and Sequence Requirements for RNA Capping at the Venezuelan Equine Encephalitis Virus RNA 5' End. *Journal of Virology* [Internet]. 2021 May 19 [cited 2021 Dec 14]; Available from: <https://journals.asm.org/doi/abs/10.1128/JVI.00777-21>
184. Alhammad YMO, Kashipathy MM, Roy A, Gagné J-P, McDonald P, Gao P, et al. The SARS-CoV-2 Conserved Macrodome Is a Mono-ADP-Ribosylhydrolase. Gallagher T, editor. *J Virol*. 2020 Nov 6;95(3):e01969-20, /jvi/95/3/JVI.01969-20.atom.
185. Hepatitis E [Internet]. [cited 2022 Jan 5]. Available from: <https://www.who.int/news-room/fact-sheets/detail/hepatitis-e>
186. V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat Rev Microbiol*. 2021 Mar;19(3):155–70.