

UNIVERSITE D'AIX-MARSEILLE
ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA
SANTÉ (ED62)

Thèse présentée pour obtenir le grade universitaire de
Docteur en science

Discipline : Pathologies humaines
Spécialité : Maladies infectieuses

Présentée et défendue publiquement le 10 juillet 2020 par

Coralie VALLE

Les méthyltransférases des *Mononegavirales*,
une cible antivirale attractive

Réalisée au sein de l'équipe « *Réplicases Virales : Structure Mécanisme et Drug-Design* »
Laboratoire Architecture et Fonction des Macromolécules Biologiques (AFMB)
Sous la direction du **Docteur Étienne DECROLY**

Membres du jury de thèse :

Dr. Marie GALLOUX
Pr. Hervé BOURHY
Pr. Bruno COUTARD
Dr. Lucie ÉTIENNE
Dr. Mathieu MATEO
Dr. Étienne DECROLY

Rapporteur
Rapporteur
Examineur
Examineur
Examineur
Directeur de thèse

« La réussite appartient à tout le monde. C'est au travail d'équipe qu'en revient le mérite »

Franck Piccard.

REMERCIEMENTS

Tout d'abord, j'aimerais remercier les membres du jury d'avoir accepté d'évaluer mon travail de thèse, et pour l'intérêt porté à l'égard de mes travaux de recherche en s'engageant à être rapporteurs : Dr. Marie Galloux, Pr Hervé Bourhy et examinateurs : Pr Bruno Coutard, Dr. Lucie Etienne et Dr. Mathieu Mateo.

Je tiens à remercier Yves Bournes, directeur de l'AFMB, de m'avoir accueillie au sein de son unité de recherche.

J'ai eu la chance pendant mes trois années et demie de thèse d'être si bien entourée. J'ai énormément appris à vos côtés et je n'y serai arrivée sans vous, il est temps pour moi d'utiliser ces quelques lignes pour vous remercier.

En premier lieu, je tiens à remercier chaleureusement mon directeur de thèse Étienne Decroly, de m'avoir donné l'opportunité de réaliser cette thèse. Merci pour la confiance que tu m'as accordée tout au long de ce projet. Ce fut un réel plaisir de travailler à tes côtés, j'ai énormément appris scientifiquement et personnellement. J'ai eu l'opportunité de travailler sur de nombreux projets, sur lesquels j'ai pu pleinement m'exprimer et je t'en suis reconnaissante. Enfin merci pour ton soutien, je garde dans un coin de ma tête cette histoire du psychiatre dans l'ascenseur.

Je remercie infiniment Bruno Coutard, merci pour ton expertise sans faille, ta disponibilité, ta patience et pour avoir répondu à toutes mes questions avec la plus grande pédagogie. Merci pour ta bienveillance et ton soutien dans les moments plus difficiles. Enfin merci de m'avoir donné toutes les cartes pour devenir, dixit Barbara : « *une étudiante en thèse autonome, responsable et réfléchie !* ».

Je tiens à remercier tous les membres de l'équipe BC qui ont contribué à égayer mon quotidien.

Merci Bruno Canard, pour ta bienveillance et tes conseils. Le café du matin avec Nadia est devenu mon moment sacré pour bien commencer la journée !

Shipovnik Nadia ! et merci de m'avoir soutenue, écoutée, conseillée et aidée à solutionner de nombreux problèmes. Les discussions sciences, tracas paillasses, (mais pas que) ont alimenté mes réflexions et m'ont permis d'avancer. Merci de m'avoir donné l'opportunité d'enseigner, j'en garde un très bon souvenir. Ta bonne humeur contagieuse et le chai vont vraiment me manquer.

Barbara, ce fut un réel plaisir de partager mon bureau avec toi. Merci pour ta bonne humeur quotidienne, pour ton aide, ta bienveillance et ton écoute. Et merci de me garder les plaques de verre en cas d'urgence émotionnelle.

Le matin rime avec Cécilia, merci pour nos discussions débriefing, ton écoute et tes conseils avisés. « J'ai retenu la leçon ;-) » !

Je tiens à remercier Jean-Claude Guillemot pour les encouragements et toute la bienveillance qu'il a eu à mon égard.

Merci Véro F, pour ta disponibilité, ton aide, ta bonne humeur et pour les précultures du dimanche ;-) Avec Barbara vous gérez les petits tracas de tout le monde avec professionnalisme et bienveillance, MERCI !

Merci François, pour tes encouragements et pour l'important travail cristallographique que tu as réalisé et qui a permis de résoudre la structure 6YU8. Tu m'as permis d'ajouter la cerise sur le gâteau, et je t'en suis reconnaissante.

Je remercie mes compagnons de galère, Oney et son vocabulaire français si riche, Priscilla E merci pour ton soutien et ta bonne humeur, Ash « you're the best ! » je reste impressionnée par ton parcours, merci pour l'organisation des after-lab, et merci pour toujours avoir du consommable sur ta paillasse. Alice, Bhawna, Priscilla S, je vous souhaite le meilleur pour la suite.

Je voudrais également remercier Claire, « ma Claire » avec moi depuis le début. Merci pour ta bienveillance, ton soutien sans faille, tes conseils et ta franchise. C'est sans oublier nos midis marche (rapide) au combien nécessaire pour repartir du bon pied et relativiser face aux aléas de la vie.

Parce que ce projet n'aurait pu aussi bien aboutir sans l'aide précieuse de Véronique Z. Merci de m'avoir permis d'ajouter la cerise sur le gâteau ! Merci d'avoir cru en ce projet structure (souvent même pour deux). J'ai été ravie de partager ce travail avec toi. Les cristaux naissent au petit matin et c'était avec grand plaisir que je me rendais au labo quand tout le monde était encore endormi.

Je tiens à remercier Baptiste, avec qui j'ai apprécié travailler et partager les réussites des projets. Merci pour le temps que tu as accordé à la relecture de ce manuscrit. Merci pour ton accueil à mon arrivée au labo et pour ton soutien dans la dernière ligne droite de cette aventure.

Je voudrais remercier mon « *best bureau Eva* » : Christine, Philippe, Sarah, Maria et Bruno avec qui j'ai eu la chance de partager quotidiennement mon bureau pendant les premiers 18 mois de ma thèse. Merci pour votre bonne humeur, votre bienveillance, votre soutien, et pour m'avoir aidé à surmonter certains aléas. Je remercie Sarah, qui en plus du bureau me supportait à la paillasse. Merci d'avoir répondu à toutes les questions d'une novice en production de protéines, ton aide m'a été précieuse. Merci Philippe pour ton écoute, ton soutien et tes boutades qui permettent de dédramatiser. Milles mercis Christine pour avoir toujours répondu présente à mes « Need to talk ! », merci pour ton écoute et tes conseils qui autour d'un verre sont tellement appréciables.

Merci à Renaud et Fabien pour la partie criblage haut débit. Et merci, Fabien je garde de notre coin paillasse avec Sarah et Laurie un excellent souvenir.

Patricia, Brigitte, Denis, merci pour votre disponibilité, votre aide. Merci Brigitte pour les gâteaux, chocolats que tu me mettais de côté et pour le café au petit matin.

Polo, Chantal et Fabienne qui rendent la vie au labo plus facile. Polo, quel plaisir d'entendre ton accent chantant.

Je voudrais remercier Isabelle Leparc-Goffart, qui a encadré l'ensemble de mon parcours universitaire. Il y a 10ans je réalisais mon premier stage dans son équipe. Stage qui a marqué un tournant et m'a donné l'envie de poursuivre dans la recherche. Merci infiniment à Yvan, qui a rendu cette aventure possible et avec qui je prends toujours autant plaisir à discuter épidémio.

J'ai des collègues au top mais également une famille extra. Merci à mes parents qui m'ont toujours encouragée et poussée à poursuivre mes projets. Je ne serai jamais arrivée si loin sans votre soutien. Merci mes sœurs : mes doublons, Morgane et Mélanie, avec qui j'avance depuis toujours, et Ambre qui me comprend si bien, m'épaule et me soutien. À Clémence mon rayon de soleil, sensible et malicieux. Merci à ma seconde maman et Man qui sont toujours là pour moi, pour nous. À ma Camille, ce bébé qui égayait toutes mes fins de semaine pendant ces années thèse et qui me comble chaque jour de bonheur.

Je souhaite remercier la famille que l'on choisit, Victoire à mes côtés depuis tant années, exilée outre atlantique mais toujours là pour moi (allez reviens maintenant la blague a assez duré !). Mathieu, pour ton écoute et tes conseils toujours accompagnés d'un bon vin. Merci Cécile d'être toujours là pour moi, je te souhaite le meilleur pour cette nouvelle aventure qui t'attend. Je remercie aussi mes coloc, Kim, Max, Mathieu, Florent, Pablo avec qui j'ai appris les joies de la vie en communauté et à prendre sur moi ;-). Je garde le meilleur souvenir de ces années.

Je réserve ces dernières lignes pour Romain, mon pilier, mon repère, mon confident depuis tant d'année. Merci de m'aider à garder le cap, promis les études c'est fini !

Table des matières

LISTE DES ABRÉVIATIONS.....	13
LISTE DES FIGURES.....	15
INTRODUCTION GÉNÉRALE.....	17
1. Les virus.....	15
1.1. Définition	15
1.2. Classification.....	16
1.3. Les virus à ARN négatif.....	17
2. Généralités sur le virus de la rage	19
2.1. Historique : Louis Pasteur contre la rage	19
2.2. État de la rage dans le monde.....	20
2.3. Physiopathologie du virus de la rage	21
2.4. Le cycle viral.....	22
2.5. Prise en charge et traitements.....	23
3. Généralités sur le virus Ebola	24
3.1. Épidémiologie	24
3.1.1 Émergence du premier filovirus	24
3.1.2. Les épidémies de fièvre hémorragique à filovirus.....	26
3.1.2.1. Identification de nouvelles espèces.....	26
3.1.3. Flambée à virus Ebola de 2013 à 2016.....	27
3.1.4. La République Démocratique du Congo face à Ebola	27
3.2. Transmission du virus Ebola	28
3.3. Manifestations cliniques.....	30
3.4. Pathogénicité	30
3.4.1. Tropisme cellulaire	31
3.4.2 Le cycle viral	32

3.5. Prise en charge et traitements.....	33
4. Bases moléculaires de la réplication des <i>Mononegavirales</i>	36
4.1. Organisation génomique générale des mononégavirus.....	36
4.2. Cycle de réplication et transcription chez les <i>Mononegavirales</i>	37
4.2.1. Régulation de la réplication/transcription.....	38
4.2 Les protéines impliquées dans la réplication/transcription virale	39
4.2.1. La nucléoprotéine	39
5.2.1.1. La nucléoprotéine du virus de la rage	40
4.2.1.2. La nucléoprotéine du virus Ebola	42
4.2.2. La phosphoprotéine	44
4.2.2.1. La phosphoprotéine du virus de la rage	44
4.2.2.2. La phosphoprotéine du virus Ebola (VP35).....	46
4.2.3. La protéine L des mononégavirus	48
4.2.3.1. L'extrémité N-terminale de la protéine L	50
4.2.3.2. L'extrémité C-terminale de la protéine L	52
4.2.3.3. L'interaction de la protéine L et son cofacteur P.....	53
5. Mécanisme de coiffage chez les eucaryotes et les virus	56
5.1. La structure coiffe	56
5.2. La synthèse de la structure coiffe chez les eucaryotes	57
5.3. L'ajout de la structure coiffe chez les virus	58
5.3.1. Les voies de synthèse de la coiffe non-conventionnelles	60
5.3.1.1. Cas des <i>Alphavirus</i>	60
5.3.1.2. Cas des <i>Mononegavirales</i>	61
5.4. Rôles de la structure coiffe.....	62
6. Immunité innée de l'hôte dans un contexte d'infection virale.....	63
6.1. Les récepteurs de l'immunité innée	63
6.1.1. Les récepteurs de type Toll (TLRs).....	64

6.1.2. Les récepteurs cytoplasmiques de type RIG-I (RLRs).....	64
6.1.2.1. RIG-I.....	65
6.1.2.2. MDA5	66
6.1.2.3. LPG2.....	66
6.2. La cascade de signalisation vers la synthèse d'interférons	67
6.3. La réponse à l'IFN.....	69
7. La subversion des défenses de l'hôte : vue d'ensemble	72
OBJECTIFS DU PROJET DE THÈSE.....	76
RÉSULTATS	78
1. Etude des domaines C-terminaux de la protéine L du virus de la rage.....	78
1.1. Etude bio-informatique	78
1.2. Stratégies d'expression en système bactérien	80
1.2.2. Criblage haut débit des conditions de culture.....	80
1.2.3. Stratégies d'expression en système bactérien.....	82
1.3. Stratégie d'expression en système baculovirus	85
1.3.2. Caractérisation de l'activité MTase	86
1.4. Matériel et méthodes	88
1.4.1. Stratégie de clonage.....	88
1.4.2. Expression et purification domaine CD-MTase-CTD en système baculovirus	90
1.4.2. Test d'activité MTase	91
1.4.3. Chromatographie sur couche mince (TLC).....	92
2. Etude des domaines C-terminaux de la protéine L du virus Soudan	93
2.1. Introduction vers une caractérisation fonctionnelle	93
Article 1 : Ebola virus L protein harbors a new enzymatic activity involved in the internal methylation of RNAs	93
2.2. Caractérisation du domaine CTD de la protéine L du virus Soudan.....	99

Article 2: The C-terminal domain of the Sudan ebolavirus L protein is essential for RNA binding and methylation.	99
2.3. Etude structure/fonction du domaine MTase de la protéine L du virus Soudan .	129
Article 3: New insights into structural features of Ebolavirus methyltransferase activities (<i>Draft</i>).	129
3. Les MTases comme cibles antivirales	164
Article 4: Design, Synthesis and Discovery of N,N'-Carbazoyl-aryl-urea Inhibitors of Zika NS5 Methyltransferase and Virus Replication.	164
CONCLUSION GÉNÉRALE ET PERSPECTIVES	175
RÉFÉRENCES BIBLIOGRAPHIQUES	181
ANNEXES	229

LISTE DES ABRÉVIATIONS

AC : anticorps
ADN : acide désoxyribonucléique
ARN : acide ribonucléique
ARNm : ARN messenger
CARD : caspase activation and recruitment domain
CD : connecteur domaine
cryo-ME : cryo-microscopie électronique
CTD : domaine C-terminal
ELISA : enzyme-linked immunosorbent assay
G : glycoprotéine
GP : glycoprotéine
GTPase : guanosine 5' triphosphatase
hMPV : métapneumovirus humain
hRSV : virus respiratoire syncytial humain
IFN : interféron
IKK : inhibitor of NF- κ B (I κ B) kinase ϵ
IPTG : isopropyl β -D-1-thiogalactopyranoside
ISG : IFN-stimulated genes
L : large protéine L
Le : leader
LeARN : ARN leader
LGP2 : laboratory of genetics and physiology 2
M : protéine de matrice
MARV : virus Marburg
MAVS : mitochondrial antiviral signaling
MDA5 : melanoma differentiation-associated gene 5
ME : microscopie électronique
MeV : virus de la rougeole
MT : méthytransférase
MTase : méthyltransférase
MyD88 : myeloid differentiation primary response 88
N-ARN : nucléocapside

NLR : Nod-like receptors
NP : nucléoprotéine
NPBP : nucleoprotein binding domain
NS : protéines non structurales
NTD : domaine N-terminal
N-term : N-terminal
OMS : Organisation mondiale de la santé
P : phosphoprotéine
PAMP : pathogen-associated molecular patterns
pe : post-exposition
Pi : phosphate inorganique
PMSF : fluorure de phénylméthylsulfonyl
PPi : pyrophosphate
ppp-RNA : ARN 5'-triphosphate
pp-RNA : ARN 5'-diphosphate
PRNTase : polyribonucléotidyltransférase
RABV : virus de la rage
RDC : République démocratique du Congo
RdRp : ARN polymérase ARN-dépendante
RE : réticulum endoplasmique
RESTV : virus Reston
RIG-I : retinoic acid-inducible gene I
RNP : complexe ribonucléoprotéique
RTPase : triphosphatase
RT-qPCR : reverse transcription-polymerase chain reaction
rVSV : virus de stomatite vésiculaire recombinant
SAM : S-adénosylméthionine
SER389 : sérine position 389
SNC : système nerveux central
SUDV : *Sudan ebolavirus*
TBK-1 : tankbinding kinase 1
TLR : Toll-like receptors
Tr : trailer
TRAF : TNF receptor-associated factors

LISTE DES FIGURES

Figure 1 : La classification de Baltimore	16
Figure 2 : Classification des virus de l'ordre des Mononegavirales	18
Figure 3 : la rage dans le monde.....	20
Figure 4 : Le cycle viral du virus de la rage	22
Figure 5 : Cycle de transmission du virus Ebola.....	29
Figure 6 : Le cycle viral du virus Ebola	32
Figure 7 : Organisation génomique des Mononegavirales	36
Figure 8 : Réplication et transcription des Mononegavirales (modèle VSV)	37
Figure 9 : La nucléoprotéine (N) du virus de la rage	41
Figure 10 : La nucléoprotéine du virus Ebola.	42
Figure 11 : La phosphoprotéine du virus de la rage	44
Figure 12 : Organisation de VP35	46
Figure 13 : Structure de la protéine L des Mononegavirales (modèle VSV)	48
Figure 14 : La protéine L du virus de la rage	50
Figure 15 : Extrémité C-terminale de la protéine L de hMPV	52
Figure 16 : Structure du complexe L/P chez les rhabdovirus.....	54
Figure 17 : La structure coiffe des ARNs messagers	56
Figure 18 : Formation de la structure coiffe chez les eucaryotes	57
Figure 19 : Modifications 5' des ARNm : les stratégies déployées par les virus.....	59
Figure 20 : Synthèse de la structure coiffe chez les virus du genre Alphavirus.....	60
Figure 21 : Synthèse de la structure coiffe chez les Mononegavirales.....	61
Figure 22 : Les récepteurs cytoplasmiques de type RIG-I (RLR).....	65
Figure 23 : Réponse immunitaire innée antivirale	68
Figure 24 : La réponse aux IFNs	70
Figure 25 : Subversion des défenses antivirales de l'hôte	72
Figure 26 : Représentation des différentes constructions pour l'expression de protéines recombinantes.....	79
Figure 27 : Stratégie de criblage haut débit des conditions d'expression	81
Figure 28 : Résumé des diverses conditions testées	82
Figure 29 : Analyse expression et production du domaine MTase-CTD.....	85
Figure 30 : Purification du domaine CD-MTase-CTD de RABV.....	86
Figure 31 : Caractérisation de l'activité MTase du domaine CD-MTase-CTD	86

Figure 32 : PCR d'assemblage	88
Figure 33 : Stratégie de clonage	89
Figure 34 : Test d'activité MTase.	91
Figure 35 : Chromatographie sur couche mince (TLC)	92
Figure 36 : Modèle des remaniements structuraux de la protéine L	177
Figure 37 : Rôles potentiels des méthylations de l'ARN chez les filovirus	179

INTRODUCTION GÉNÉRALE

1. Les virus

1.1. Définition

Le terme « virus » a été proposé en 1953 par André Lwoff (1). Les virus sont des agents infectieux capables d'infecter de nombreuses cellules vivantes telles que : les bactéries, les mammifères, les plantes, les reptiles, les insectes ou encore les champignons. Par opposition aux bactéries, les virus constituent des parasites obligatoires souvent de petite taille (diamètre < 200 nm). Ces entités biologiques partagent les caractéristiques suivantes : (i) ils possèdent un seul type d'acide nucléique (ADN ou ARN), (ii) ils se répliquent à partir de leur propre acide nucléique, (iii) leur croissance dépend de leur hôte. Quelques années plus tard, André Lwoff complète cette définition, ajoutant que les virus utilisent les ribosomes de la cellule hôte pour se multiplier (2). En 2003, la découverte d'un virus géant (le Mimivirus) remet en cause ce paradigme. En effet, de par sa taille (> 0,7 µm), la complexité de son génome (plus de 1200 gènes) et la présence d'ADN et d'ARN dans les particules virales, le Mimivirus invalide certains éléments de la définition donnée par André Lwoff (3,4). En 2008, D. Raoult et P. Forterre proposent de revoir la définition d'un virus, proposant de diviser le monde biologique en 2 groupes d'organismes : les organismes codant pour le ribosome tels que les bactéries, les eucaryotes et les archées et les organismes codant pour une capsid. Ainsi, il en découle une redéfinition du virus comme « *un organisme codant la capsid, composé de protéines et d'acides nucléiques, s'auto-assemblant dans une nucléocapsid et utilisant un organisme codant le ribosome pour mener à bien son cycle de vie* » (5).

1.2. Classification

En 1962, A. Lwoff, R. Horne et P. Tournier proposent un schéma pour la classification de tous les virus, indépendamment de l'origine virale et de la spécificité de l'hôte (6). Les virus sont classés selon un système hiérarchique linnéen : Ordre (*-virales*), Famille (*-viridae*), sous-famille (*-virinae*), genre (*-virus*) et espèce. Cette classification est basée sur 4 critères suivant les propriétés de la particule virale :

1. la nature de l'acide nucléique ;
2. la symétrie de la capside ;
3. la présence ou l'absence d'une enveloppe ;
4. les dimensions du virion et de la capside.

Le Comité international sur la taxonomie des virus est chargé de développer, d'affiner et de maintenir une taxonomie universelle des virus.

En 1971, David Baltimore met au point un second schéma de classification complémentaire prenant en compte la nature de l'acide nucléique viral et le mode d'expression de ce dernier (7,8). La classification de Baltimore comprend 7 groupes décrits dans la **Figure 1**. Les travaux de thèse présentés dans ce manuscrit ont été menés sur deux virus à ARN simple brin de polarité négative appartenant à l'ordre des *Mononegavirales*.

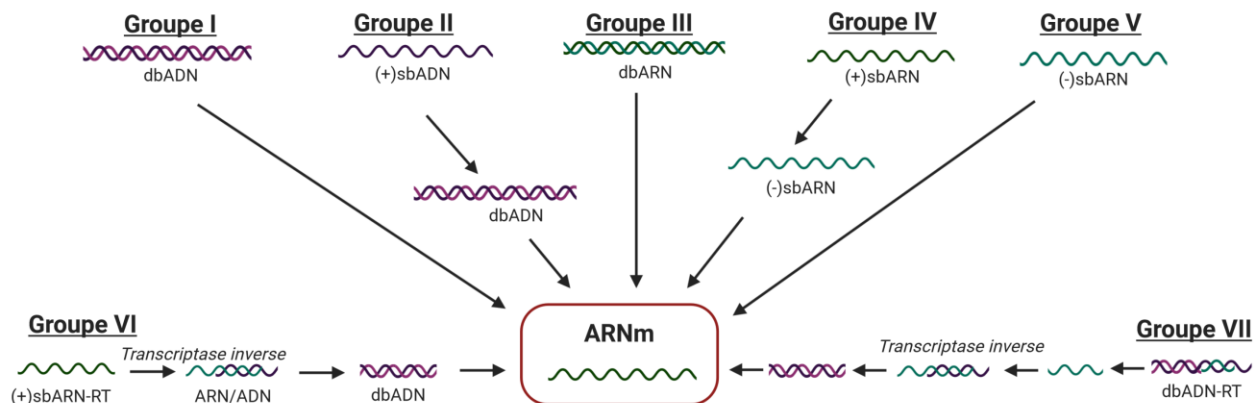


Figure 1 - La classification de Baltimore

La classification de Baltimore, basée sur la nature de l'acide nucléique et sur leur polarité, comprend 7 groupes définis selon la voie de synthèse de l'ARNm à partir du génome viral. Groupe I: virus à ADN double brin (exemples: Adénovirus et le virus de l'Herpès); Groupe II : virus à ADN simple brin de polarité positive utilisant comme intermédiaire de transcription un ADN double brin (exemples : Parvovirus); Groupe III : virus à ARN double brin (exemples : rotavirus, réovirus); Groupe IV : ARN simple brin de polarité positive utilisant un ARN simple brin de polarité négative comme intermédiaire de transcription (exemples : flavivirus, alphavirus, poliovirus); Groupe V : virus à ARN simple brin de polarité négative (exemples : virus de la rage, virus Ebola, la rougeole); Groupe VI : virus à ARN simple brin de polarité positive utilisant l'ADN comme intermédiaire de la réplication via la transcriptase inverse (exemple : VIH); Groupe VII : virus à ADN double brin passant par un intermédiaire simple brin d'ARN (exemple : virus de l'hépatite B).

1.3. Les virus à ARN négatif

L'ordre des *Mononegavirales*, créé en 1991, regroupe des virus à ARN simple brin (sb), linéaire, non-segmenté et de polarité négative. A l'origine, l'ordre était constitué de 3 grandes familles : les filovirus (*Filoviridae*), les rhabdovirus (*Rhabdoviridae*) et les paramyxovirus (*Paramyxoviridae*) (comprenant 2 sous-familles *Paramyxovirinae*, et *Pneumovirinae*) (9). Au cours des dernières années, l'ordre des *Mononegavirales* s'est étendu et comprend à ce jour 11 familles : bornavirus (*Bornaviridae*), mymonavirus (*Mymonaviridae*), nyamivirus (*Nyamiviridae*), artovirus (*Artoviridae*), lispivirus (*Lispiviridae*), pneumovirus (*Pneumoviridae*), xinmovirus (*Xinmoviridae*), sunvirus (*Sunviridae*) et les 3 familles historiques, précédemment nommées (10,11). Chacune des familles est subdivisée en genres puis en espèces (**Fig. 2**). L'ordre des *Mononegavirales* comprend d'important pathogènes humains, tels que le virus de la rougeole (MeV), le virus de la rage (RABV), le métapneumovirus humain (hMPV), le virus des oreillons (MuV), le virus respiratoire syncytial humain (hRSV) et le virus Ebola. Les travaux suivants ont porté sur 2 de ces différents pathogènes humains : le virus de la rage et le virus Ebola. RABV appartient à la famille des *Rhabdoviridae* et au genre *Lyssavirus* (**Fig. 2**). Le virus Ebola appartient à la famille des *Filoviridae* qui est subdivisée en six genres (*Cuevavirus*, *Marburgvirus*, *Ebolavirus*, *Striavirus*, *Thamnovirus* et *Dianlovirus*). Le genre *Ebolavirus* comprend d'importantes espèces pathogènes pour l'Homme comme le virus Soudan (SUDV) (12,13).

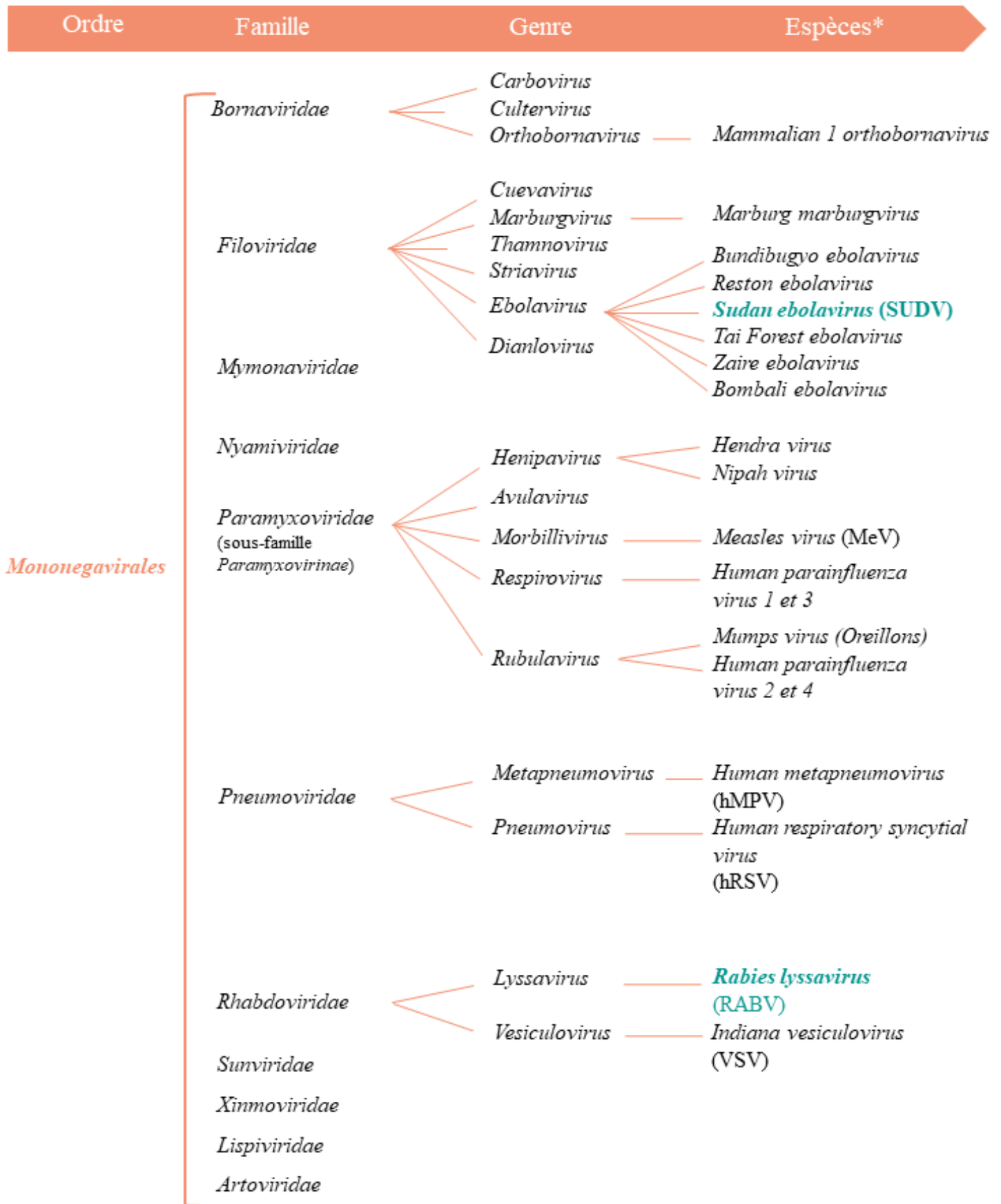


Figure 2 - Classification des virus de l'ordre des *Mononegavirales*

L'ordre des *Mononegavirales* regroupe des virus à ARN simple brin non segmenté de polarité négative. Cet ordre comprend 11 familles. Chaque famille est subdivisée en genres puis en espèces. *Cette figure ne présente pas l'ensemble des espèces, mais celles présentant un intérêt médical. Les virus étudiés dans ces travaux sont indiqués en vert.

2. Généralités sur le virus de la rage

2.1. Historique : Louis Pasteur contre la rage

En 1879, Louis Victor Galtier, professeur à l'école vétérinaire de Lyon, publie les premiers travaux sur la rage. Ses recherches, menées entre 1879 et 1881, ont montré que la rage pouvait être transmise au lapin par injection de salive de chiens enragés. Ses travaux sur les moutons ont par ailleurs mis en évidence que l'injection de salive en intraveineuse conférait une immunité sans que ces derniers ne développent la maladie (14). Louis Pasteur prend connaissance des travaux de Louis Victor Galtier et débute des recherches sur la rage en 1880 (15). En 1881, Louis Pasteur et ses collaborateurs, Emile Roux, Charles Chamberland et Louis Thuillier, mettent en évidence le tropisme du virus pour le système nerveux et l'intérêt de l'inoculation intracérébrale. Louis Pasteur va ensuite rechercher un moyen efficace de diminuer la virulence d'une souche. Dans un premier, le passage en série du virus par inoculation intracérébrale sur lapin, permet à L. Pasteur d'isoler un virus dit « fixe » ayant des caractéristiques constantes (temps d'incubation, pouvoir pathogène). Le traitement par dessiccation et potasse caustique (hydroxyde de potassium) de prélèvements, post-mortem, de moelle de lapin inoculé par la souche « fixe », permet l'obtention de souches atténuées (« *méthode des moelles* »). Pasteur montre qu'entre 2 et 9 jours de dessiccation, la virulence décroît progressivement et après 9 jours toute virulence a disparu. L'inoculation de souches de virulence de plus en plus croissante permet le déclenchement d'une réponse immunitaire : le vaccin canin antirabique est ainsi obtenu (16). En juillet 1885, Joseph Meister, jeune garçon âgé de 9 ans mordu par un chien enragé, se présente au laboratoire de L. Pasteur. L'enfant reçoit plusieurs injections de moelle de lapin rabique de virulence croissante et ne développe pas la maladie après injection de l'échantillon de moelle le plus virulent. Le traitement du jeune garçon marque le premier succès de vaccination antirabique chez l'Homme (17) et introduit la notion de vaccin thérapeutique.

2.2. État de la rage dans le monde

La situation de la rage dans le monde est très variable d'un pays à un autre et évolue continuellement (**Fig. 3**).

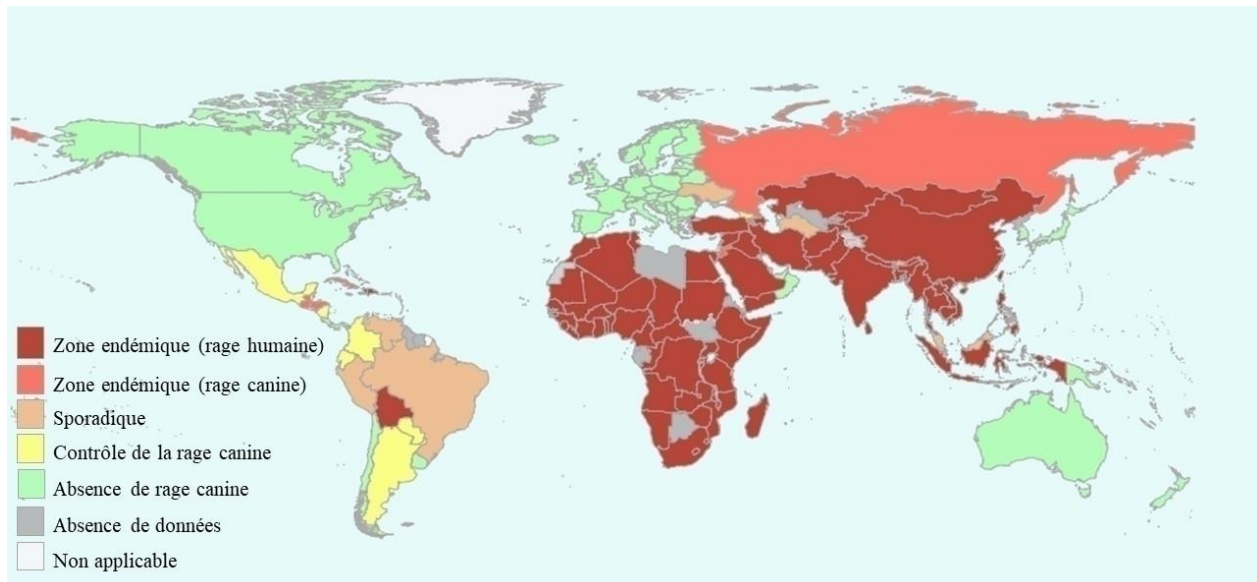


Figure 3 - La rage dans le monde

Carte représentant la répartition de la rage humaine et canine dans le monde. Bien que la rage ait été éradiquée dans 80 pays, elle continue d'être présente dans 70 pays, et tue chaque année près de 59 000 personnes dans le monde.

(apps.who.int/neglected_diseases/ntddata/rabies/rabies.html).

L'Afrique et l'Asie comptent le plus de cas de rage humaine, avec un nombre de décès estimé à plus de 59 000 par an (18). Le chien, principal vecteur sur ces continents, est responsable de 99% des contaminations humaines (19). Cependant, d'autres animaux peuvent être porteurs du virus rabique. L'apparition d'épidémies sporadiques au sein des populations animales reste donc imprévisible, soulignant la nécessité de mieux comprendre le mode de transmission et de propagation du virus dans les hôtes réservoirs. En Europe, les campagnes vaccinales des chiens domestiques ainsi que l'élimination des chiens errants avaient permis à de nombreux pays de l'Europe de l'Ouest de rester indemnes de toute contamination rabique. Malgré cela, à la fin de la seconde guerre mondiale, le virus rabique s'est adapté au renard, entraînant sa propagation dans de nombreux pays. Dès la fin des années 70, l'utilisation de vaccins antirabiques (virus de la vaccine recombinant) contenus dans des appâts a permis la diminution drastique du nombre de cas de rage (20,21). En France, depuis 1924, aucun cas de rage humaine autochtone a été rapporté (22). Les rares cas de décès déclarés en France sont des cas importés qui ont contracté la rage au cours d'un voyage dans un pays endémique. Au

cours du siècle dernier, un certain nombre de pays ont mis en œuvre des programmes de vaccination à grande échelle qui ont permis l'élimination de la rage canine, réduisant considérablement l'incidence de la maladie chez l'Homme. Ainsi, la rage a pu être éradiquée des États-Unis, du Canada, de l'Europe occidentale, du Japon, de la Malaisie et de Hong Kong (23,24). Le risque de réémergence dans ces pays est principalement lié à l'importation non contrôlée d'animaux infectés (25). En décembre 2015, l'OMS, l'Organisation mondiale de la santé animale (OIE), l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO) et l'Alliance mondiale pour la lutte contre la rage (GARC) se sont fixés pour objectif de ramener à zéro le nombre de décès humains dus à la rage d'ici 2030 (26). Pour répondre à cet objectif, la surveillance et le contrôle de la rage animal représentent une mesure clef vers la diminution du nombre de décès lié à la rage humaine. La lutte contre ce virus nécessite l'action concertée de différents partenaires, médicaux, vétérinaires mais également éducatifs, pour la formation du personnel soignant et sensibiliser les populations des zones endémiques.

2.3. Physiopathologie du virus de la rage

Le virus de la rage est un virus neurotrope qui attaque le système nerveux central (SNC). Le terme rage est dérivé du latin *rabere* qui signifie « être fou ». La maladie évolue vers une mort inévitable lorsqu'elle est cliniquement déclarée. La transmission du virus est, dans la plupart des cas, liée à une morsure d'un chien infecté, mais également par léchage de plaies. Le virus est retrouvé dans la salive de l'animal infecté, trois jours avant le début des symptômes neurologiques et l'animal décède environ 15 jours suivant la déclaration des symptômes. Après pénétration dans l'organisme, le virus se multiplie localement avant de pénétrer le système nerveux périphérique. Le virus remonte ensuite le long de la moelle épinière, via un transport axonal, pour rejoindre le cerveau. Un fois qu'il a atteint le SNC, le virus est disséminé dans l'organisme touchant plusieurs organes comme les glandes salivaires, l'œil, le foie ou encore le pancréas (27,28). Chez l'Homme, les caractéristiques cliniques associées à cette infection peuvent être divisées en 5 phases : la période d'incubation, le prodrome, la phase neurologique aiguë, le coma et la mort. La période d'incubation d'un virus neurotrope peut varier de 1 à 3 mois (29). Dans le cas de la rage, une morsure au niveau de la tête, du cou, de la face est généralement associée à une période d'incubation très courte pouvant être réduite à 7 jours. Le prodrome correspond à l'apparition des symptômes et marque la fin de la période d'incubation. Cette phase correspond à la migration du virus de la périphérie vers le SNC, provoquant une inflammation de la moelle épinière et une atteinte neuronale. La maladie peut évoluer soit en forme furieuse soit en forme paralytique. La forme

furieuse est caractérisée par une hyperactivité, des troubles neuropsychiatriques (hallucinations, délire, insomnie), une hydrophobie et des spasmes. Le décès lié à un arrêt cardiorespiratoire survient en quelques jours. L'évolution vers une forme paralytique est moins spectaculaire que la forme furieuse. Les muscles se paralysent progressivement jusqu'au coma puis décès du patient (30–32).

2.4. Le cycle viral

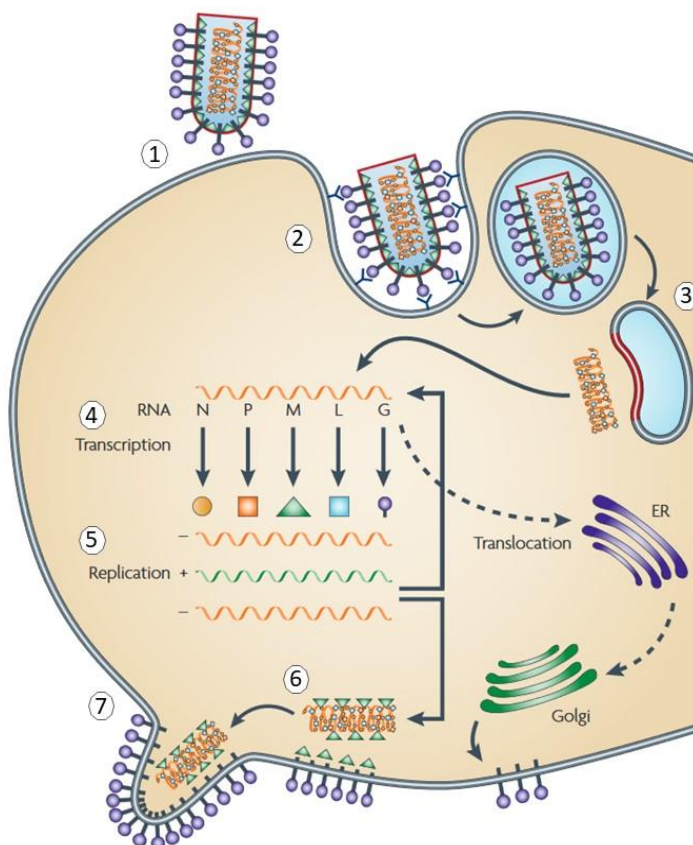


Figure 4 - Le cycle viral du virus de la rage

Le cycle viral du virus de la rage est strictement cytoplasmique et divisé en plusieurs étapes. 1) Le virus s'attache à la surface de la cellule hôte via sa glycoprotéine de surface G. 2) Un phénomène d'endocytose permet l'entrée de la particule virale dans la cellule. 3) L'acidification du compartiment endosomal permet la libération du génome viral dans le cytoplasme, s'en suit les étapes de 4) transcription et 5) réplication, qui permettent la multiplication virale. 6) De nouvelles particules virales sont formées et 7) sont libérées par bourgeonnement dans le milieu extracellulaire.

Issue de M.J. Schnell et al., 2010

Le virus de la rage est transmis à l'Homme suite à une morsure puis migre vers le SNC. Le cycle viral du virus de la rage est strictement cytoplasmique et se divise en 3 grandes phases. (i) L'entrée du virus dans la cellule hôte. Cette phase débute par la fixation de la particule virale via sa glycoprotéine (G) au récepteur cellulaire. À ce jour, plusieurs récepteurs, tels que le récepteur 2 du glutamate métabotrope mGluR2 (« metabotropic glutamate receptors »), la glycoprotéine NCAM (« *Neural cell adhesion molecule* »), les récepteurs nicotiques de l'acétylcholine (AChR), le récepteur du facteur de survie et de différenciation des neurones p75NTR (« *p75 neurotrophin receptor* »), les glucides et les

gangliosides, ont été identifiés (33). Suite à cette fixation, le virus pénètre dans la cellule par un mécanisme d'endocytose. L'acidification de l'endosome entraîne un changement de conformation de la glycoprotéine (G), facilitant la fusion de la membrane de la particule virale et de l'endosome (34). La fusion des membranes permet la libération de la nucléocapside (RNP) dans le cytosol (35) (**Fig. 4 - Etapes 1 à 3**). (ii) La multiplication virale regroupe les étapes de transcription/traduction et de réplication du génome viral (**Fig. 4 - Etapes 4 et 5**). La transcription a lieu dans une «usine de virus» spécialisée, les corps de Negri qui correspondent à des inclusions cytoplasmiques retrouvées dans les cellules infectées par RABV (36). La protéine G est transloquée dans le réticulum endoplasmique (RE) avant de transiter vers l'appareil de Golgi, représentant une étape indispensable dans le processus de bourgeonnement. Les autres protéines virales sont, quant à elles, traduites à proximité des corps d'inclusion (37). (iii) Libération des particules virales. Cette dernière phase permet la sortie du virus de la cellule hôte. Suite à la réplication, les virions sont assemblés avec l'ARN génomique et les protéines virales puis libérés à la surface de la membrane de la cellule hôte via un phénomène de bourgeonnement (38) (**Fig.4 : étapes 6 et 7**).

2.5. Prise en charge et traitements

Bien que le vaccin antirabique ait été mis au point il y a plus de 130 ans par Louis Pasteur, il n'existe à ce jour aucun moyen de traiter cette infection une fois les symptômes déclarés. L'enjeu des traitements actuels est de prévenir le développement de la maladie. Une prise en charge rapide réduit considérablement la survenue des symptômes, pouvant ainsi réduire le risque de mortalité. La prophylaxie post-exposition (pe) a pour objectif de neutraliser le virus avant qu'il ne pénètre le système nerveux, avec une fenêtre d'action ne dépassant pas 2h à 3h pe. La prise en charge pe comprend un nettoyage profond des plaies, l'immunisation passive et la vaccination active (39). Le nettoyage des plaies est une étape critique dans la prévention de la survenue de la maladie. En effet, il a été montré *in vivo*, que le lavage des plaies dans les 3h pe réduisait considérablement la propagation du virus (40). L'immunisation passive consiste à administrer un cocktail d'immunoglobulines antirabiques permettant de fournir des anticorps capables de neutraliser le virus localement (41). Les anticorps (AC) ne pouvant pas franchir la barrière hémato-encéphalique, ils ne sont plus efficaces une fois que le virus a atteint le SNC. Des doses vaccinales sont également administrées. Les personnes non vaccinées contre la rage reçoivent 5 doses contre 2 de rappel pour les personnes vaccinées (42). Une vaccination préexposition est recommandée pour les personnes à risque d'être en contact avec l'agent infectieux, comme les employés de laboratoire, les vétérinaires, les

personnes voyageant dans les zones où la rage circule. L'administration de trois doses vaccinales est réalisée à jour 0, 3 et 28, avec un taux de séroconversion estimé à 98,2% (43).

3. Généralités sur le virus Ebola

3.1. Épidémiologie

3.1.1 Émergence du premier filovirus

En 1967, des cas de fièvre hémorragique inexplicée ont été déclarés en Allemagne (Marburg) et en Yougoslavie suscitant un intérêt considérable. Les symptômes associés étaient de violents maux de tête, une fièvre élevée, des douleurs articulaires et des hémorragies. La majorité des patients, travaillant pour le groupe pharmaceutique Behring, avaient été en contact direct avec des échantillons (sang, organes ou cultures cellulaires) de singes verts (*Cercopithecus aethiops*) importés d'Ouganda. Le virus a pu être isolé à partir de ces échantillons et baptisé Marburg (44). Depuis sa découverte, le virus Marburg (MARV) fut à l'origine de plus d'une dizaine d'épidémies principalement en Afrique de l'Est, avec une première apparition du virus sur le continent Africain en 1975 (45) (**Tableau 1**).

Années	Pays	Souche Virale	Cas humains	Décès reportés
1976 (Juin-Novembre)	Soudan	SUDV	284	151
1976 (Septembre-Octobre)	RDC	EBOV	318	280
1976 (Novembre)	Angleterre	SUDV	1	0
1977 (Juin)	RDC	EBOV	1	1
1979 (Juillet-Octobre)	Soudan	SUDV	34	22
1989-1990	Philippines et US	RESTV	7 (asymptomatiques)	0
1992	Italie	RESTV	0	0
1994	Gabon	EBOV	52	31
1994	Côte d'ivoire	TAFV	1	0
1995	RDC	EBOV	315	250
1996 (Janvier-Avril)	Gabon	EBOV	37	21
1996-1997 (Juillet-Janvier)			60	45
1996	Afrique du Sud	EBOV	2	1
1996	US et Philippines	RESTV	0	0
1996	Russie	EBOV	1	1
2000-2001	Ouganda	SUDV	425	224
2001-2002 (Octobre-Mars)	Gabon et Congo	EBOV	124	97
2002-2003 (Décembre-Avril)	Congo	EBOV	143	128
2003 (Novembre-Décembre)			35	29
2004	Soudan	SUDV	17	7
2004	Russie	EBOV	1	1
2005	Congo	EBOV	12	10
2007	RDC	EBOV	264	186
2007-2008 (Décembre-Janvier)	Ouganda	BDBV	131	42
2008	Philippines	RESTV	6 (asymptomatiques)	0
2008-2009 (Décembre-Février)	RDC	EBOV	32	15
2011 (Mai)	Ouganda	SUDV	1	1
2012 (Juillet-Août)			11	4
2012 (Août-Octobre)	RDC	BDBV	36	13
2012-2013	Ouganda	SUDV	6	3
2013-2016	Plusieurs pays	EBOV	28 616	11 310
2014 (Août-Novembre)	RDC	EBOV	69	49
2015	Philippines	RESTV	0	0
2017 (Mai-Juillet)	RDC	EBOV	8	4
2018 (Mai-Juillet)			54	33
2018-2020 (Août-Avril)			3461*	2279*

Tableau 1 - Les épidémies de fièvre hémorragique à filovirus

EBOV : virus Ebola; RESTV: virus Reston; SUDV: virus Soudan; BDBV: virus Bundibugyo; TAFV: virus Taï forest (Tableau issu de *Languon et al.* 2019)

*Situation en avril 2020

3.1.2. Les épidémies de fièvre hémorragique à filovirus

3.1.2.1. Identification de nouvelles espèces

Depuis la découverte de MARV, plus d'une trentaine d'épidémies de fièvre hémorragique à filovirus ont été recensées par l'OMS (**Tableau 1**). Les deux premières, en 1976, touchèrent simultanément la République Démocratique du Congo (RDC) et 4 villes du Soudan (Nzara, Maridi, Tembura, and Juba) (46). Une étude sérologique a permis d'identifier le foyer épidémique : une usine de coton à Nzara peuplée de chauves-souris (46,47). Dans la même année, un cas de contamination a été rapporté dans un laboratoire en Angleterre suite à la manipulation d'échantillons biologiques en provenance du Soudan (48). Les épidémies au Soudan et en RDC ont permis l'isolement de deux nouvelles souches virales appartenant à un second genre, *Ebolavirus*, et nommées respectivement virus Soudan (*Sudan ebolavirus*, SUDV) et virus Ebola (*Zaire ebolavirus*, EBOV), liés mais non identiques (49). Les espèces virales SUDV et EBOV représentent aujourd'hui les souches circulantes les plus virulentes.

Le virus Ebola réémerge en Virginie et en Pennsylvanie, après une période de latence de près de 10 ans. Entre 1989 et 1990, la mort d'un grand nombre de singes liée à une infection à filovirus a été observée faisant suite à l'importation de macaques en provenance des Philippines au centre de primatologie de Reston (USA). Une nouvelle espèce, nommée virus Reston (*Reston ebolavirus*, RESTV) a pu ainsi être isolée (50,51). Un épisode d'infection à RESTV a de nouveau été à l'origine d'épidémies aux Philippines (52), en 2008, chez des cochons. Aucun des éleveurs porcins ne présentaient des symptômes de maladie à virus Ebola bien que des anticorps aient pu être détectés dans leurs séras, mettant en évidence une absence de pathogénicité de RESTV pour l'Homme (53). En 1994, le sérum d'un ethnologue contaminé suite à l'autopsie d'un chimpanzé en Côte d'Ivoire, a permis d'isoler une nouvelle espèce nommée virus Taï Forest (*Tai Forest ebolavirus*, TAFV) (54). Le virus Bundibugyo (*Bundibugyo ebolavirus*, BDBV) a, quant à lui, pu être identifié suite à une épidémie qui toucha Bundibugyo en 2007 (55).

En 2002, des colonies de chauves-souris (*Miniopterus schreibersii*) ont subi une mortalité massive dans des grottes de certains pays européens (en France, en Espagne et au Portugal) (56). Les techniques de séquençage haut débit ont permis de caractériser un nouveau genre de filovirus nommé virus Lloviu (*Lloviu cuevavirus*, LLOV) (57). En 2018, une nouvelle espèce du genre *Ebolavirus* a été identifiée en Sierra Leone, et nommée virus Bombali (*Bombali ebolavirus*, BOMV). Cette souche a été détectée chez des chauve-souris retrouvées dans les

habitations suggérant un risque de transmission à l'Homme. Cependant, la pathogénicité de cette nouvelle souche pour l'Homme reste à démontrer (58). Dans la même année, un filovirus de chiroptère, phylogénétiquement distinct d'EBOV et MARV, appartenant au genre *Dianlovirus*, a été identifié et nommé virus Mënglà (*Mengla dianlovirus*, MLAV) (59). Plus récemment, deux nouvelles espèces ont été décrites chez les poissons, *Xilang striavirus* (XILV) et *Huangjiao thamnovirus* (HUVJ), ajoutant deux nouveaux genres de virus à la famille des filovirus (11,60) (**Tableau 1**).

3.1.3. Flambée à virus Ebola de 2013 à 2016

Une importante épidémie à virus Ebola toucha plusieurs pays d'Afrique de l'Ouest, entre Décembre 2013 et Mars 2016. Elle représente l'une des épidémies les plus meurtrières avec plus de 28,000 cas et 11,000 décès recensés par l'OMS (61). Le cas index, un enfant d'un village de Guinée, a été infecté par des chauves-souris insectivores (62), s'ensuivit plusieurs cas de diarrhée hémorragique et une alerte sur une maladie non identifiée est lancée. Des échantillons sont envoyés à l'Institut Pasteur (Paris) afin de statuer sur l'agent pathogène mis en cause. L'épidémie atteint rapidement les pays limitrophes, continuant de décimer les populations (63), et devenant rapidement incontrôlable notamment en raison de la faiblesse des systèmes de surveillance et du manque d'infrastructures de santé publique. Le caractère inattendu et la rapide progression de l'épidémie ont mis en exergue la nécessité de mettre au point de nouvelles techniques de diagnostic rapides, sensibles, spécifiques pouvant être réalisées sur place sans besoin matériel. En Août 2014, l'épidémie a été déclarée comme une urgence de santé publique de portée internationale (64). Des efforts internationaux de lutte contre l'épidémie ont permis de réduire la transmission du virus Ebola, d'améliorer la capacité de détection des nouveaux foyers et d'endiguer l'épidémie.

3.1.4. La République Démocratique du Congo face à Ebola

Depuis 1976, la RDC a dû faire face à une dizaine d'épidémies à virus Ebola, dont deux entre 2018 et 2020 (65) (**Tableau 1**). La première débuta le 8 mai à Bikoro. La mise en place de mesures de contrôle intensives et le lancement d'une campagne de vaccination, ont permis de mettre fin à cette flambée le 24 juillet 2018 (66). Le 1er Août, quelques jours après l'annonce de la fin de la première, la seconde épidémie débuta dans le nord-ouest de la RDC à plus de 1000 km de la première (67). En juin 2019, des premiers cas d'infection à virus Ebola ont été rapportés en dehors de la RDC. Cet événement a conduit, le 17 juillet 2019, le comité d'urgence du règlement sanitaire international sur les maladies virales à virus Ebola (IHREC)

a déclaré l'épidémie comme étant une urgence de santé publique de portée internationale (68). Les efforts internationaux ont permis de lutter contre ce virus. Cependant, l'épidémie à virus Ebola en RDC reste une urgence mondiale et les efforts de lutte continuent. Alors qu'en Mars 2020, le dernier patient quittait le centre de de soin, un nouveau cas est rapporté en avril 2020.

Les épidémies récentes touchant la RDC soulignent la nécessité de mettre au point des traitements antiviraux et des vaccins efficaces pour traiter les patients et les cas contacts afin de lutter contre la propagation de ce virus. Cependant, la recherche sur le virus Ebola reste limitée, notamment du fait que ce virus doit être cultivé dans un environnement de haute sécurité (BSL4). Le manque de moyens techniques, la détérioration des centres de soin anti-Ebola et le contexte politique et culturel conduisant le plus souvent à une méfiance de la population face aux traitements ont mis à mal l'efficacité de la riposte et contribué à la propagation du virus. Le comportement individuel et les coutumes locales jouent un rôle important dans le déroulement des épidémies et la propagation du virus. Le contexte d'épidémie à virus Ebola a montré l'importance d'inclure les sciences humaines et sociales dans les stratégies de réponse conduites par les systèmes de santé afin de mener une riposte efficace et éclairée.

3.2. Transmission du virus Ebola

Le virus Ebola est une zoonose, se transmettant à l'Homme, via contact direct avec des animaux infectés comme les chauves-souris frugivores et les primates non-humains principalement lors de la consommation de viandes de brousse. Le virus peut ensuite être facilement transmis d'un individu à l'autre (**Fig. 5**).

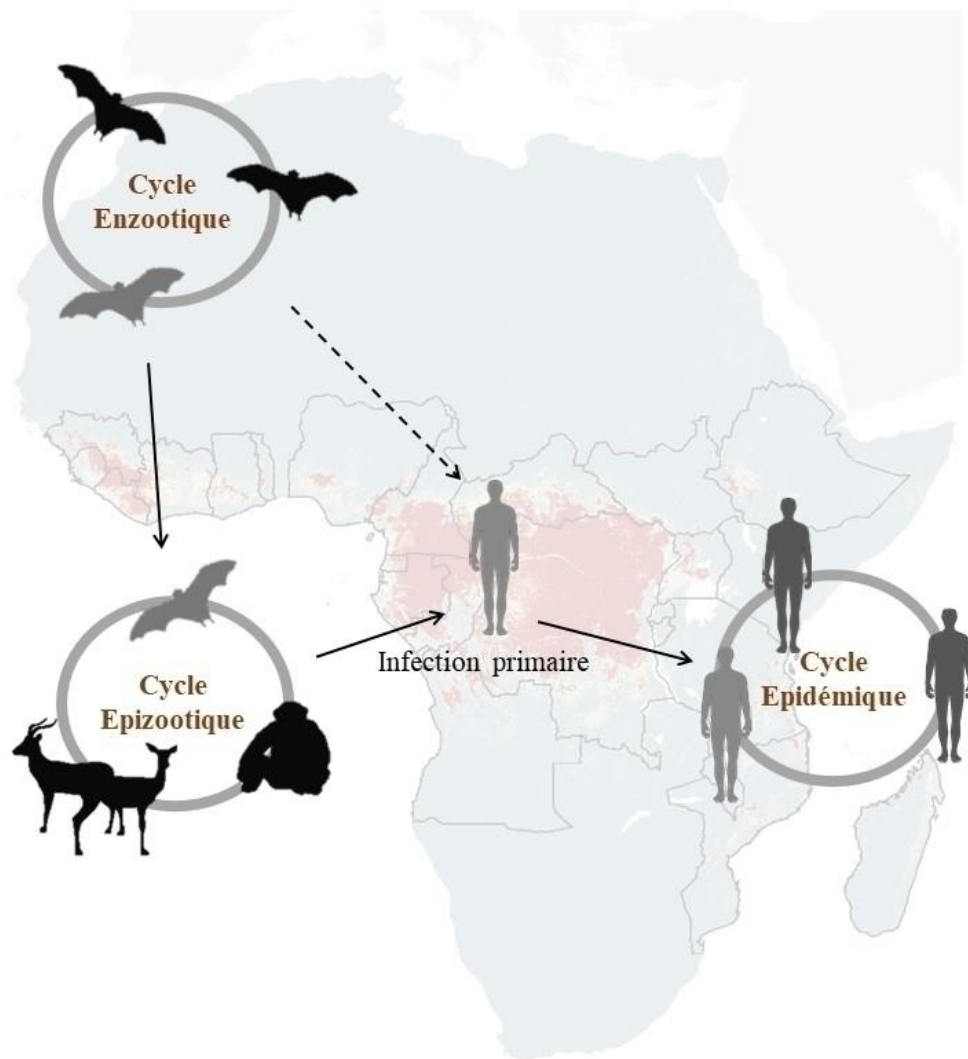


Figure 5 - Cycle de transmission du virus Ebola

Les chauves-souris frugivores sont considérées comme les réservoirs naturels du virus Ebola. Elles permettent la dissémination du virus et le maintien du cycle viral. La transmission du virus au sein du cycle enzootique entre chauves-souris est encore mal connu. L'épizootie liée au virus Ebola apparaît de façon sporadique. Les chauves-souris infectées entre en contact avec de animaux de la faune sauvage tels que les gorilles, les antilopes des forêt et d'autres singes, pouvant entraîner des épidémies de grande ampleur. L'Homme représente un hôte accidentel et l'infection primaire est le plus souvent liée à la chasse et la consommation de viande de brousse. La transmission interhumaine passe par le contact direct avec une personnes infectées ou le défunt lors de rites funéraires.

La transmission du virus Ebola d'Homme à Homme se fait, dans la plupart des cas, à la suite d'un contact direct avec une personne infectée ou le défunt lors des rites funéraires (69). La voie d'entrée du virus dans le corps se fait par un contact direct entre la peau lésée (plaies, abrasions cutanée) ou les muqueuses (nasale, génitale, rectale et parentérale) d'un sujet sain et les fluides et tissus corporels infectés comme le sang, les sécrétions nasales, sécrétions orales ou encore les sécrétions génitales d'une personne infectée (70–74). L'Homme comme les

grands singes sont considérés comme des hôtes opportunistes et non comme des espèces réservoirs (75). Bien que beaucoup d'efforts aient été déployés pour identifier les réservoirs naturels le sujet reste controversé. Pendant longtemps, les rongeurs (76) et les chauves-souris (77) ont été considérés comme des réservoirs potentiels et l'isolement de MARV chez la *Rousettus aegyptiacus* conforta l'idée de la chauve-souris comme espèce réservoir des filovirus (78). Cependant, l'existence d'autres espèces susceptibles d'être infectées par un filovirus reste envisageable, notamment après la découverte de RESTV chez les porcs aux Philippines (53), et des hôtes potentiellement amplificateurs ou réservoirs secondaires pourraient jouer un rôle au cours des épidémies.

3.3. Manifestations cliniques

Après une période d'incubation de 2 à 21 jours, l'infection à virus Ebola débute par des signes cliniques non spécifiques tels qu'une forte fièvre, des frissons, des malaises et des myalgies (douleurs musculaires). Après quelques jours, ces symptômes sont généralement suivis d'une atteinte multi-viscérale observée à différents niveaux : systémique (prostration), gastro-intestinal (anorexie, nausée, vomissement, douleur abdominale, diarrhée), respiratoire (douleur thoracique, essoufflement, toux, écoulement nasal), vasculaire (injection conjonctivale, hypotension, œdème) et neurologiques (maux de tête, confusion, coma) (79). Des pertes sanguines massives, principalement au niveau du tractus gastro-intestinal, peuvent être observées. Les patients décèdent, généralement entre 6 à 16 jours après l'apparition des premiers symptômes, des suites d'une défaillance multi-organes (80). Pour les survivants la convalescence est longue et des troubles sur le long terme tels qu'une hépatite récurrente, une myélite, une perte de cheveux prolongée, une psychose et une uvéite peuvent être observés (81). Une récente étude épidémiologique a mis en évidence un risque de mortalité 5 fois plus élevé chez les personnes ayant survécus à l'infection à virus Ebola au cours de la première année de rémission, le plus souvent lié à une insuffisance rénale. De plus, cette étude montre que le risque est fonction de l'âge, du sexe et du temps passé dans les systèmes de soin (82). Dans un contexte post-épidémie, cette étude indicative pourrait permettre de mettre en place un système de suivi et une prise en charge adaptée des survivants de l'infection à virus Ebola.

3.4. Pathogénicité

La virulence du virus Ebola chez l'Homme est variable et dépend de la souche considérée. Les infections par EBOV ont les taux de létalité les plus élevés (60 à 90%), suivies de celles par SUDV (40 à 60%) (81). L'infection à virus Ebola entraîne une nécrose coagulante de

divers organes tels que le foie, les reins, les poumons, la rate et les testicules (83). Suite à la transmission du virus, le décès est lié à un dysfonctionnement de la réponse immunitaire, une défaillance multiviscérale et des troubles vasculaires et de la coagulation (84).

3.4.1. Tropisme cellulaire

Le virus Ebola peut infecter un large éventail de types cellulaires. Les analyses de tissus de patients ou de primates non-humains infectés montrent que le virus Ebola a un tropisme pour les monocytes, les macrophages, les cellules dendritiques, les cellules endothéliales, les fibroblastes, les hépatocytes, les cellules corticales surrénaliennes et plusieurs types de cellules épithéliales (72,83,85–88). Au stade précoce de la maladie, le virus infecte principalement les monocytes, les macrophages et les cellules dendritiques. Ces cellules semblent jouer un rôle dans la dissémination du virus. En effet, le virus se propage depuis le site d'infection initial via ces dernières jusqu'aux ganglions lymphatiques régionaux dans un premier temps puis vers d'autres organes (89). Les cellules des glandes surrénales et les hépatocytes semblent être des cibles de choix pour les filovirus, jouant un rôle important dans la pathogénèse de la maladie (85,87). Chez les patients ayant succombés à l'infection mais également chez les primates non-humains infectés expérimentalement, il a été observé une nécrose de la rate, du thymus et des ganglions lymphatiques (85). La libération de cytokines, de chimiokines et d'autres médiateurs pro-inflammatoires par les cellules présentatrices d'antigènes entraîne une défaillance multiviscérale accompagnée d'un choc septique, conduisant au décès de la personne infectée (90–92).

3.4.2 Le cycle viral

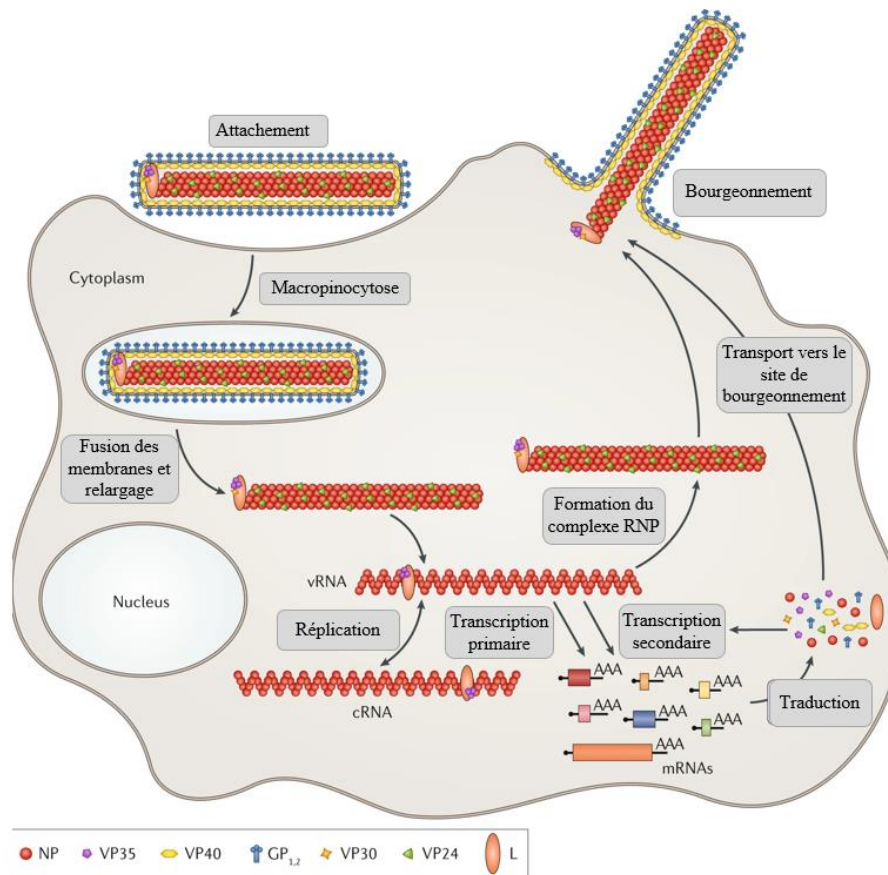


Figure 6 - Le cycle viral du virus Ebola

Le cycle viral du virus Ebola est strictement cytoplasmique et divisé en plusieurs étapes. 1) Le virus s'attache à la surface de la cellule hôte via sa glycoprotéine de surface GP, 2) et entre dans la cellule via un mécanisme de macropinocytose, 3) l'acidification du compartiment endolysosomal permet la libération du génome viral dans le cytoplasme, s'en suit les étapes de 4) transcription et 5) réplication qui permettent la multiplication virale. 6) De nouvelles particules virales sont formées et 7) seront libérées par bourgeonnement dans le milieu extracellulaire.

Issue de T. Hoennen et al., 2019

Les particules virales des filovirus présentent une forme allongée et filamenteuse. Cette forme caractéristique est à l'origine de leur nom, *Filoviridae*, du latin *filum*, qui signifie fil. L'ensemble des 7 protéines virales sont contenues dans les virions. Le génome est encapsidé par la nucléoprotéine (NP), associé à la protéine L, son co-facteur la protéine VP35 et le facteur d'activation de la transcription VP30 (93), formant le complexe ribonucléoprotéique (RNP). Ce complexe est enfermé par une enveloppe extérieure, dérivée de la membrane plasmique de la cellule hôte, et tapissée de la protéine matricielle VP40, responsable de la morphogénèse et le bourgeonnement des particules virales (94). Enfin, la glycoprotéine (GP), composée de deux sous-unités GP1 et GP2, est ancrée sous forme trimérique dans l'enveloppe extérieure (95).

Le cycle viral des filovirus est classique des virus à ARN de polarité négative se répliquant dans le cytoplasme et se compose de trois grandes étapes (**Fig. 6**). (i) L'entrée du virus débute par l'attachement du virion à la surface de la cellule hôte, assurée par la GP. En effet, les GP interagissent avec des leptines de type-C (CLRs), DC-SIGN et L-SIGN exprimés à la surface des cellules dendritiques et favorisent la fusion des membranes virale et cellulaire permettant ainsi l'entrée de la particule virale. Une fois la particule virale attachée à la cellule hôte, elle est internalisée suivant un mécanisme de macropinocytose et se retrouve dans le compartiment endolysosomal (96,97). L'acidification de l'endosome et le clivage de la GP par des cathepsines cellulaires déclenchant la fusion entre les membranes virale et de l'endosome (98,99). Cette fusion des membranes permet le relargage du complexe RNP dans le cytoplasme de la cellule hôte. (ii) S'en suit les étapes de multiplication virale (transcription, traduction et réplication), relativement semblable chez les virus de l'ordre des *Mononegavirales*, et qui seront détaillées dans le *chapitre (4.2)*. Brièvement, chez les filovirus la transcription du génome viral est facilitée par les protéines du complexe RNP, NP, VP35, VP30 et la protéine L (100). La production des protéines virales, notamment la NP, conduit à la formation de corps d'inclusion qui sont les sites de réplication et transcription virales (101) ; (iii) Libération de nouvelles particules virales. La GP transite par le RE et l'appareil de Golgi où elle est maturée (clivage et glycosylations) puis adressée à la surface de la membrane de la cellule hôte. L'assemblage des nouveaux virions a lieu dans le cytosol. Les complexes RNP nouvellement formés se déplacent des corps d'inclusion vers les sites de bourgeonnement (102). VP40 permet la déformation de la membrane et joue un rôle central dans la libération de nouvelles particules virales (103,104).

3.5. Prise en charge et traitements

Le diagnostic initial est basé sur un examen clinique. Du fait des caractéristiques cliniques non spécifiques, la confirmation du diagnostic passe par une détection du génome viral par RT-qPCR (105–107) et/ou un test sérologique (détection des ACs) par ELISA (108,109). Le diagnostic en laboratoire est généralement effectué dans des centres de référence nationaux et internationaux sur des échantillons biologiques inactivés. Le virus est détectable dans le sang au début de la maladie et persiste à des concentrations élevées dans les cas mortels (110). Pour la détection des ACs (IgM et IgG), la technique couramment utilisée est le test ELISA. Les IgM sont des marqueurs de l'infection aiguë, et peuvent apparaître dès le lendemain des symptômes et disparaître entre 30 et 168 jours après l'infection (109,111,112). Les anticorps spécifiques (IgG) se développent 6 à 18 jours après l'apparition des symptômes et persistent

pendant de nombreuses années (113). Dans le cas des rémissions, le virus peut persister pendant des semaines voire des mois dans la chambre antérieure de l'œil ou les testicules, avec un risque potentiel évoqué de transmission sexuelle (48,108,114).

Au cours d'une épidémie, un système de santé publique coordonné impliquant une identification rapide et un isolement des cas (avérés et suspects) suivi d'une recherche des cas contacts, doit être rapidement mis en place. Le diagnostic est une étape cruciale dans le contrôle d'une épidémie. En effet, la mise en place de tests diagnostiques pour identifier les personnes infectées permet d'une part une prise en charge du patient mais également un suivi et éventuellement le traitement préventif des cas contacts présentant un risque de développer la maladie. Cependant, la détection de nouveaux cas est souvent tardive et l'arrivée d'un patient en centre de soin survient quand la maladie est déjà à un stade avancé. De plus, les centres de santé anti-Ebola déployés notamment en Afrique subsaharienne doivent faire face, en plus des difficultés logistiques, à des différences culturelles qui rendent la population méfiante, handicapent l'efficacité de la riposte et favorisent la propagation du virus.

Le risque de réémergence du virus Ebola est réel comme en témoigne l'épidémie récente qui toucha la RDC. Par conséquent, parallèlement aux mesures de santé publique mises en place, des stratégies de développement de molécules à potentiel antiviral et la recherche de nouveaux vaccins pour lutter contre cette maladie se sont intensifiées afin de répondre à cette urgence sanitaire.

Une prise en charge adaptée aux signes cliniques ainsi que des soins médicaux de soutien (réhydratation orale ou par intraveineuse) doivent être fournis aux patients infectés par le virus Ebola afin d'améliorer leur survie. Actuellement, il n'existe aucun traitement spécifique approuvé contre le virus Ebola. L'urgence sanitaire des infections à virus Ebola est un réel défi pour le développement de nouveaux vaccins et de molécules à potentiel antiviral pour le traitement des personnes infectées et des cas contacts (famille et personnel soignant). Actuellement, de nombreux traitements anti-Ebola sont entrés en phase d'essai clinique et ont montré leur efficacité dans le contrôle de la transmission, la réduction de la sévérité de la maladie et la durée d'hospitalisation dans les centres de soin anti-Ebola. En novembre 2018, un premier essai clinique de contrôle multi-médicaments a débuté en RDC au cours de la dernière épidémie afin de contrôler l'innocuité et l'efficacité des traitements anti-Ebola proposés (115). L'AC monoclonal MAb114 et 2 cocktails d'anticorps monoclonaux : REGN-EB3 et ZMapp sont au centre de cette étude. L'essai clinique mené en RDC a démontré que

les patients traités avec MAb114 et REGN-EB3 présentent un taux de mortalité plus faible que les patients traités avec ZMapp (115), ouvrant la voie vers leur homologation.

Les stratégies de développement de nouveaux vaccins, se sont orientées vers deux types de vaccin : les vaccins basés sur des vecteurs viraux non-réplicatifs et des vecteurs viraux réplicatifs. Les vecteurs viraux non-réplicatifs codent pour des antigènes viraux tels que la GP et sont dépourvus des gènes nécessaires à la réplication virale. Ces vaccins présentent une meilleure innocuité mais peuvent s'avérer moins efficace dans l'activation de la réponse immunitaire. Les vecteurs viraux réplicatifs ont l'avantage d'entraîner une activation de la réponse immunitaire plus efficace, mais peuvent présenter une pathogénicité résiduelle pouvant limiter leur utilisation. Le vaccin rVSV-EBOV produit à partir du virus de stomatite vésiculaire recombinant (rVSV) code pour le GP d'EBOV (souche Kikwit 1995). Des essais précliniques ont montré que ce vaccin protège 100% des souris contre une injection létale (116). Une protection totale a également été rapportée chez les primates non-humains (117). Un second vaccin développé par le groupe pharmaceutique Johnson & Johnson, a reçu le soutien des autorités de santé et a été déployé en RDC au cours de la dernière épidémie. Ce second vaccin (Ad26-ZEBOV) vient compléter la stratégie vaccinale déjà mise en place par l'administration du vaccin rVSV-ZEBOV-GP dans le but d'étendre la protection contre EBOV (118). En fin 2019, le vaccin rVSV-ZEBOV-GP a été le premier vaccin approuvé par la FDA pour prévenir les infections à EBOV (119,120).

4. Bases moléculaires de la réplication des *Mononegavirales*

4.1. Organisation génomique générale des mononégavirus

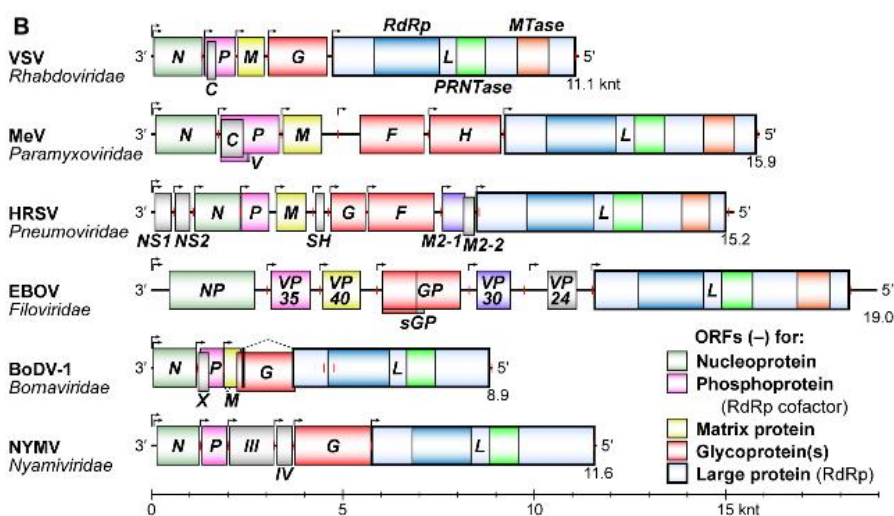


Figure 7 - Organisation génomique des *Mononegavirales*

Représentation de l'organisation génomique de virus appartenant à l'ordre des *Mononegavirales* : VSV (virus de la stomatite vésiculaire), MeV (virus de la rougeole), hRSV (virus respiratoire syncytial humain), EBOV (virus Ebola), BoDV-1 (virus de la maladie de Borna) et NYMV (virus Nyamanini). L'initiation et la terminaison de la transcription sont indiquées respectivement par des flèches noires au-dessus du génome et des barres rouges. Les gènes codant pour les différentes protéines sont indiqués par des boîtes de couleurs. Le gène de la protéine L code pour l'ARN polymérase ARN dépendante (RdRp en bleu), une polyribonucléotidyltransférase (PRNTase, en vert) et une méthyltransférase (MTase en orange).

Issue de Ogino, T et al., 2019

Au sein de l'ordre des *Mononegavirales*, les génomes viraux varient entre 8 et 19 kb. Les protéines communes aux virus à ARN négatif sont : la nucléoprotéine (NP), la phosphoprotéine (P), la protéine de matrice (M), la glycoprotéine (G ou GP) et la grande protéine L (L) (**Fig. 7**). Les fonctions de ces protéines sont généralement conservées chez les *Mononegavirales*. Les bornavirus et les nyamivirus codent en plus pour une protéine X impliquée dans la régulation de la synthèse des gènes et l'inhibition de la réponse immunitaire (121,122). Les paramyxovirus, les pneumovirus et les filovirus ont un génome plus complexe qui code pour des protéines supplémentaires. Chez les filovirus, VP35 et VP40 sont les homologues des protéines P et M (123–125). Les filovirus contiennent également les protéines VP24 et VP30. VP30 joue un rôle dans la régulation de la réplication et de la transcription (126,127). La protéine virale VP24 est présumée être une protéine de matrice impliquée dans l'assemblage des nucléocapsides et la formation de nouvelles particules virales (128–130). De plus, VP24 joue un rôle déterminant dans le tropisme cellulaire et est capable de contrecarrer la réponse aux interférons (IFN). De par ses propriétés, VP24 est considérée

comme un important facteur de virulence (131,132). Certains paramyxovirus codent pour une protéine de fusion (F) et une petite protéine hydrophobe (SH) qui s'associent à l'enveloppe virale pour faciliter l'entrée du virus. La protéine M2 présente chez hMPV et RSV participe à la régulation de la réplication/transcription (133). Enfin, les protéines non structurales NS1 et NS2 sont retrouvées uniquement chez les pneumovirus, et jouent un rôle important dans l'inhibition de la réponse immunitaire médiée par les INFs (134).

4.2. Cycle de réplication et transcription chez les *Mononegavirales*

Les virus à ARN négatif partagent la même stratégie de réplication/transcription. Le VSV, un virus de la famille des rhabdovirus, a servi de modèle pour élucider le mécanisme de réplication et transcription des virus à ARN négatif. Son génome est encapsidé par la NP formant la nucléocapside (135,136). Un complexe ARN polymérase ARN-dépendante (RdRp) composé de la P et de la protéine L est associé à la nucléocapside pour former le complexe RNP (100,137,138). Le génome du VSV commence et se termine par de courtes séquences : 3'-leader (Le) (139) et 5'-trailer (Tr), et de 5 gènes connectés entre eux par des régions intergéniques (**Fig. 8**).

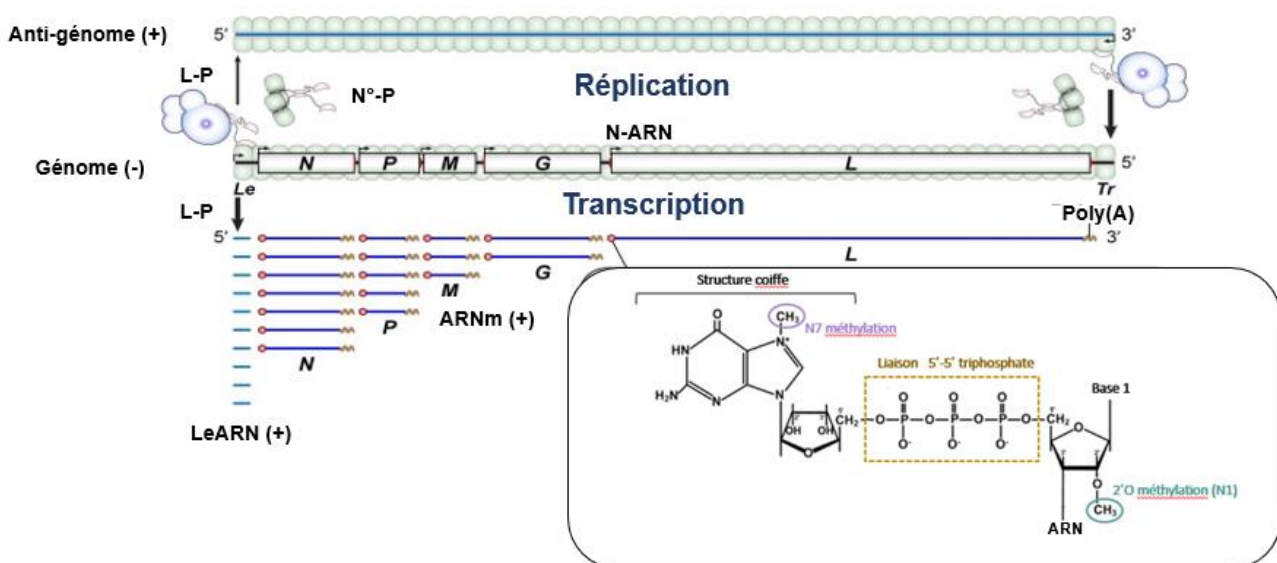


Figure 8 - Réplication et transcription des *Mononegavirales* (modèle VSV)

Le génome des *Mononegavirales* de polarité négative est encapsidé par la nucléoprotéine (N) formant le complexe (N-ARN) reconnu par la polymérase et son cofacteur (L-P). Le complexe N-ARN sert de matrice pour la réplication (synthèse de l'anti-génome) et transcription (synthèse des ARNm). « Le » désigne la région de début du génome et « Tr » la région de fin. La transcription (en bas) débute en 3' du génome. La polymérase reconnaît la séquence Le et transcrit séquentiellement les cinq gènes, selon un mode discontinu d'initiation/ré-initiation (« start/stop »). Tous les transcrits, sauf l'ARN leader (LeARN), sont coiffés (structure coiffe dans l'encadré, voir 6.1) et polyadénylés (poly(A)). La réplication (en haut) est un phénomène de synthèse continue d'ARN. La polymérase ignore les séquences de terminaison et synthétise l'anti-génome qui est encapsidé par la N. Les complexes L-P et N°-P (N°: N sans ARN) sont requis pour la réplication du génome.

Adaptée de Ogino T. et al., 2019

Le génome sert de matrice pour la synthèse de l'ARNm au cours de la transcription et de l'antigénome au cours de la réplication. Chacun des 5 gènes commence par une séquence « start », essentielle pour l'initiation de la transcription et la mise en place de la structure coiffe à l'extrémité 5' des transcrits viraux néosynthétisés (140–142), et se termine par une séquence « stop », jouant un rôle crucial pour la terminaison de la transcription et le mécanisme de polyadénylation (143). La polymérase virale transcrit les ARNs viraux dans le cytoplasme produisant ainsi les ARNm viraux qui serviront de matrice pour la traduction des protéines virales. La transcription chez les *Mononegavirales* est un phénomène discontinu, d'initiation/réinitiation de la transcription (144). Dans un premier temps, la polymérase virale se fixe au génome via l'extrémité 3'-leader et initie la synthèse d'un ARN leader (LeARN) de polarité positive d'environ 47 nucléotides. La transcription du premier ARNm débute à la séquence « start » et s'arrête à la séquence de terminaison, produisant un ARNm de la protéine N. Suite au relargage du premier transcrit, la polymérase initie et termine de façon similaire la transcription de chacun des gènes, synthétisant séquentiellement les ARNm des protéines P, M, G et L (145). La protéine L qui contient un domaine Cap (PRNTase) et méthyltransférase (MTase) ajoute en 5' de chaque transcrit néosynthétisé une structure coiffe et une queue poly-A à l'extrémité 3' (140–143). Au niveau de chaque région intergénique, la ré-initiation de la transcription perd en efficacité conduisant à un gradient de concentration des transcrits viraux et donc de protéines virales : [N]>[P]>[M]>[G]>[L] (144). Lorsque la concentration en protéines virales est suffisante dans le cytoplasme, la polymérase passe en mode réplcatif (146). Au cours de la réplication, la polymérase virale initie la polymérisation en 3' de l'ARN génomique et ignore les séquences de terminaison à la fin de chaque gène afin de générer un intermédiaire d'ARN de polarité positive, qui sera encapsidé par la N et qui servira de matrice pour la synthèse du génome viral (147–149). Les ARNm ne sont pas encapsidés par la nucléoprotéine, l'ajout de la structure coiffe et de la queue poly-A permet de les confondre au ARNm cellulaires permettant leur reconnaissance et leur prise en charge par la machinerie de traduction cellulaire.

4.2.1. Régulation de la réplication/transcription

Les mécanismes de réplication impliquent la polymérase virale, et nécessitent une régulation fine des différentes activités de l'enzyme. Chez les filovirus, la protéine VP30 joue un rôle clef dans la régulation de la transcription/réplication (100). Cette fonction est assurée par sa capacité à être phosphorylée, déphosphorylée et rephosphorylée. VP30 phosphorylé présente une affinité pour la nucléoprotéine et permet son transport vers les corps d'inclusion

où se déroule les événements de transcription et réplication (150). Pour initier la transcription virale, VP30 est déphosphorylée par des phosphatases cellulaires (151) et s'associe à la protéine L et VP35 au niveau du site d'initiation de la transcription, permettant la transcription du génome (124,152). Le phénomène de réplication est associé à VP30 phosphorylée. Ainsi la dynamique entre phosphorylation, déphosphorylation et rephosphorylation de VP30, permet le switch entre la transcription et la réplication virale (126). VP30 n'étant présente que chez les filovirus, ce mécanisme de régulation est spécifique à cette famille de virus. Chez les autres virus de l'ordre des *Mononegavirales*, la balance entre réplication et transcription fait intervenir le statut de phosphorylation de la protéine P (153–155). Une boucle de régulation de la phosphorylation, comparable à VP30, a été mise en évidence chez RSV qui présente une protéine additionnelle, M2-1, indispensable au mécanisme d'élongation de la transcription (133).

La protéine de matrice, codée par certains virus de l'ordre des *Mononegavirales*, participe également à la régulation du switch entre réplication et transcription, en inhibant la transcription virale et en stimulant la réplication (156,157). A un stade précoce du cycle viral, la protéine M présente une localisation nucléaire (158). Cette localisation participe au mécanisme de subversion des défenses de l'hôte (159) et permet la régulation du cycle transcription/réplication viral en permettant une transcription virale efficace dans le cytoplasme. Lorsque le niveau des protéines virales est suffisant, la protéine M quitte le noyau, favorisant le switch entre transcription et réplication virale (160–162).

4.2 Les protéines impliquées dans la réplication/transcription virale

4.2.1. La nucléoprotéine

Les mononégavirus ont la particularité d'avoir leur génome encapsidé par la nucléoprotéine virale, formant la nucléocapside (N-ARN). Ce phénomène d'encapsidation de l'ARN nécessite l'intervention de la phosphoprotéine (P) qui a pour rôle d'accompagner l'ARN, limitant ainsi l'association de la N à des ARNs non spécifiques (163). L'interaction ARN/N permet de protéger l'ARN viral de sa dégradation par des ribonucléases (RNases) cellulaires (164) et prévient sa détection par les senseurs intracellulaires de l'immunité innée. La nucléocapside sert de matrice pour la synthèse de l'ARN viral et forme le noyau structurel pour la formation de nouvelles particules virales (165). De plus, la N est un déterminant antigénique, pouvant être utilisé en diagnostic (166).

5.2.1.1. La nucléoprotéine du virus de la rage

La N du virus de la rage contient 450 acides aminés. Elle est composée d'un domaine N-terminal (NTD), impliqué dans l'interaction avec l'ARN, et un domaine C-terminal (CTD), indispensable à la fixation de la P (167,168). Le domaine C-terminal est une région désordonnée, stabilisée par la fixation de la protéine P. La N est retrouvée sous deux formes *in vitro* : soit associée à l'ARN formant la matrice utilisée pour la réplication et transcription, soit sous forme «soluble», en complexe avec la protéine P formant le complexe N°-P (169). L'étude du complexe N°-P par ultracentrifugation, spectrométrie de masse et microscopie électronique a montré qu'une N était associée à deux protéines P (170). La N du virus de la rage est phosphorylée sur la sérine en position 389 (SER389) lorsqu'elle est complexée avec l'ARN. La N engagée dans le complexe N°-P n'est pas phosphorylée sur la SER389. La N néosynthétisée s'associe dans un premier temps à la P. Le complexe N°-P ainsi formé participe à l'encapsidation de l'ARN. De plus, la liaison de la N à la protéine P permet à la N de lier spécifiquement l'ARN génomique viral. En effet, l'interaction avec la protéine P empêche la liaison de la N aux ARNs cellulaires (171,172). L'oligomérisation de la N autour de l'ARN entraîne un changement de conformation, exposant la SER389 permettant sa phosphorylation (173,174). De plus, cette phosphorylation joue un rôle dans le cycle viral. En effet, l'absence de phosphorylation de la N est associée à une diminution du taux de réplication et transcription (175–177). La phosphorylation de la N régulerait la réplication et transcription virale en modulant l'encapsidation de l'ARN. En effet, la N non phosphorylée se lie plus fortement à l'ARN (177). La phosphorylation de la N permettrait de stabiliser l'interaction entre la protéine P et le complexe N-ARN, jouant un rôle dans le recrutement de la protéine L (176). De plus, la phosphorylation de la N entraîne une diminution de l'affinité de N pour l'ARN permettant à la protéine L d'accéder au génome viral (175).

Lorsque la N est exprimée en cellules d'insecte, elle se lie à l'ARN cellulaire de façon non spécifique, conduisant à des structures de forme hélicoïdale (147,178). Les études de microscopie électronique ont montré que le complexe N-ARN adopte une forme en anneau pouvant contenir 9 à 13 sous-unités de N (N₉₋₁₃) (167). La résolution de la structure cristallographique du complexe N₁₁-ARN a permis sa caractérisation (179) (**Fig. 9**).

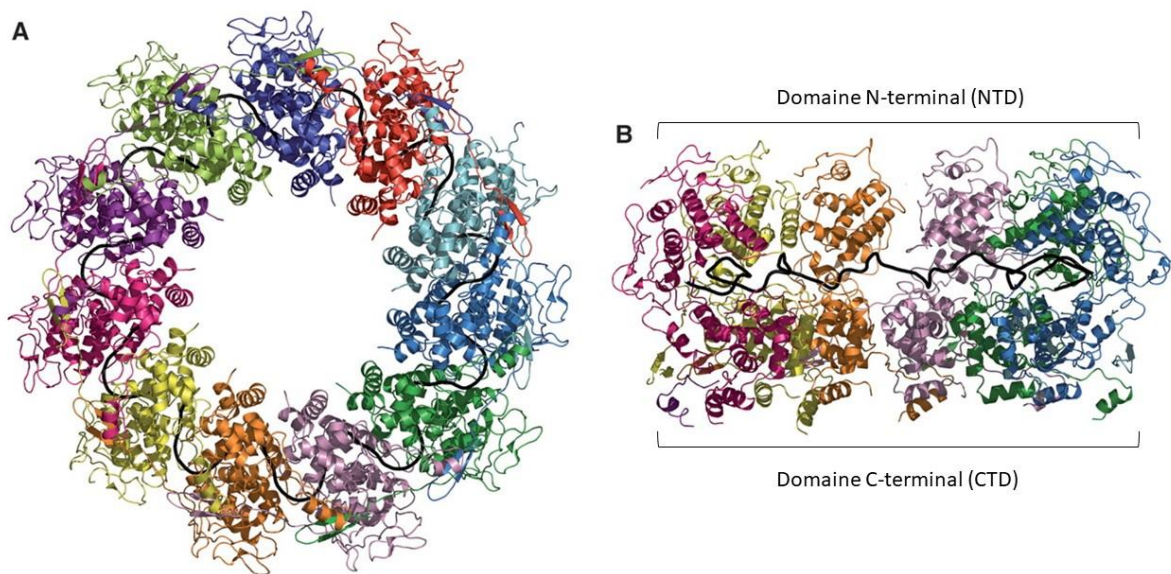


Figure 9 - La nucléoprotéine (N) du virus de la rage

(A) Structure d'un complexe de 11 Ns associées à l'ARN. Les protomères de N forme une structure en anneau, renfermant l'ARN au centre. Chaque protomère est identifié par une couleur distincte. (B) Un protomère de N est composé d'un domaine N-terminal (NTD) et d'un domaine C-terminal (CTD) formant un sillon accommodant l'ARN (en noir sur le schéma). L'ARN se retrouve masqué au centre de la structure hélicoïdale.

(Issue de *Albertini AA. et al. 2006*).

Le protomère de N est constitué de deux domaines : le domaine NTD présentant un repliement en forme hélicoïdal constitué de 6 hélices liées entre elles par de larges boucles, et le domaine CTD composé de 11 hélices reliées entre elles par des boucles plus étroites (179). L'interface entre ces deux domaines forme un sillon basique permettant d'accueillir l'ARN qui se retrouve ainsi protégé au centre de l'anneau (179) (**Fig. 9**). Cette séquestration par la N limite sa détection par le système immunitaire. En effet, elle empêche l'activation des récepteurs TLRs lors de son transit dans le compartiment endosomal (*voir chapitre immunité antivirale*) (180,181) et prévient la dégradation de l'ARN viral par les exonucléases cellulaires (182). Deux sous-domaines plus petits émergent des domaines NTD et CTD. Ils permettent d'établir la liaison entre les protomères (179). Les sous-domaines présentent une grande flexibilité et pourraient participer à l'ouverture de l'anneau, afin de faciliter l'accès de la protéine L au génome viral (183). L'interaction de la N avec la protéine P affecte les deux sous-domaines, permettant l'ouverture de la poche formée par NTD-CTD, rendant l'ARN accessible pour la polymérase virale (la protéine L) (179). Cette exposition partielle de l'ARN a été observée chez d'autres virus de l'ordre des *Mononegavirales* (184–186). Enfin, la N du

virus de la rage représente un déterminant antigénique. Il a été montré qu'elle présentait quatre sites antigéniques distincts (187–189).

4.2.1.2. La nucléoprotéine du virus Ebola

La NP, VP35, VP40 et la protéine L s'associe pour former le complexe RNP responsable de la réplication et transcription virales (190). L'ARN sb des virus à ARN de polarité négative est conditionné dans un complexe hélicoïdal qui comprend plusieurs copies de NP (**Fig. 10**).

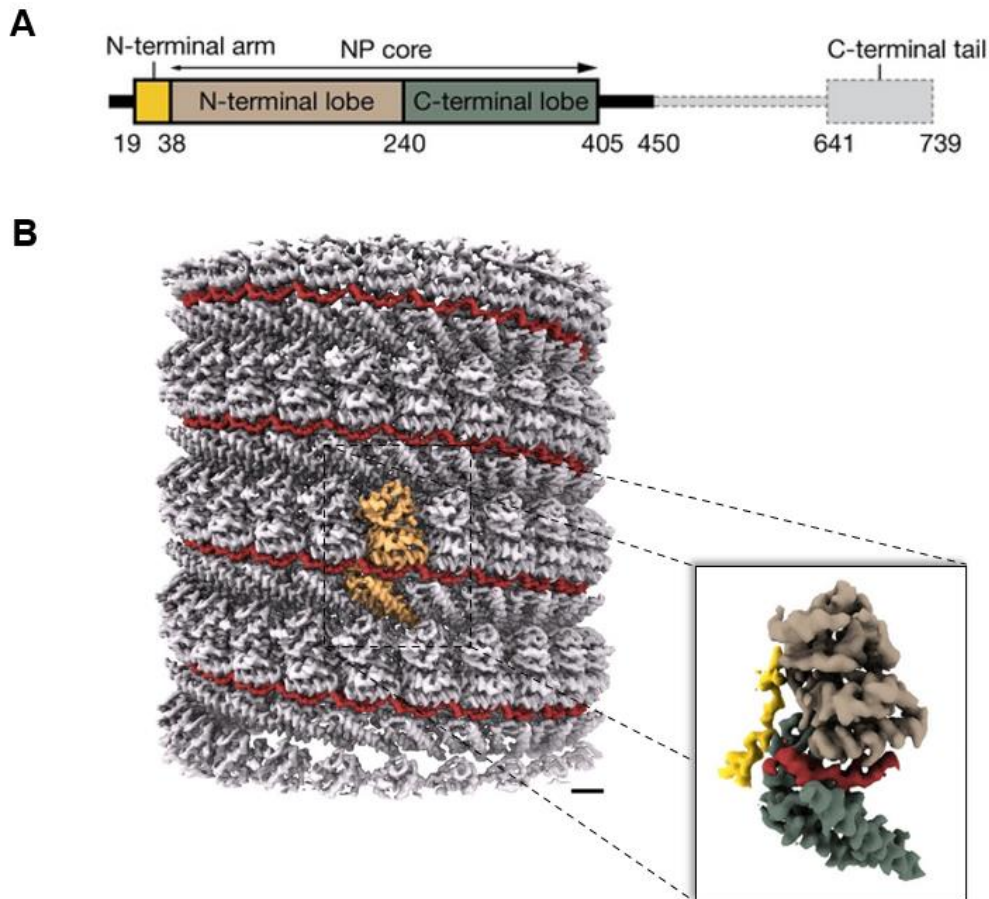


Figure 10 - La nucléoprotéine du virus Ebola.

(A) La nucléoprotéine (NP) contient 739 acides aminés et s'organise en deux modules. L'extrémité N-terminale composée d'une région nommée « arm » suivi du noyau de la protéine (« NP core ») formé d'un domaine N-terminal (« N-terminal lobe ») et d'un domaine C-terminal (« C-terminal lobe ») lié au module C-terminal par une région désordonnée se terminant par une queue C-terminale (« C-terminal tail »). (B) Représentation du complexe NP-ARN avec en rouge la molécule d'ARN. L'encadré noir représente un isomère de N (PDB : 5Z9W) avec l'ARN situé dans le sillon formé entre le « N-terminal lobe » (en beige) et le « C-terminal lobe (en vert) ».

Issue de Yukihiro Sugita, *et al.*, 2018

La NP présente un double rôle (structural et fonctionnel) dans le cycle de réplication viral. En effet, c'est un composant structural essentiel de la nucléocapside, jouant un rôle dans la morphogénèse virale. Par ailleurs elle joue un rôle fonctionnel, catalysant la réplication et la transcription virales (100,190,191). La NP des filovirus est constituée d'une chaîne polypeptidique plus longue (729 acides aminés) que celles des autres virus de l'ordre des *Mononegavirales*, avec une organisation en deux modules portant des fonctions distinctes (192,193). L'extrémité N-terminale se compose d'une petite région nommée « arm », suivi du noyau de la protéine (« core ») formé d'un domaine N-terminal (« N-lobe ») et C-terminal (« C-lobe ») lié à l'extrémité C-terminale par un domaine linker désordonné suivi d'une queue C-terminale (« C-terminal tail ») (93,194) (**Fig. 10A**). Plusieurs structures de la NP en complexe avec l'ARN ont été récemment résolues (93,195,196). L'extrémité N-terminale porte les fonctions d'oligomérisation (interaction NP-NP) et d'assemblage (192,193). Cette caractéristique semble être conservée chez d'autres virus de l'ordre des *Mononegavirales*. L'implication de ce domaine dans la fixation de l'ARN et son encapsidation a été confirmée par mutagenèse dirigée (193,197). De façon semblable à d'autres nucléoprotéines codées par des virus, les deux domaines de la NP d'Ebola forme un sillon chargé positivement pouvant accommoder l'ARN et permettre son encapsidation (100,192,195,198) (**Fig. 10B & 10C**). De plus, un autre domaine hydrophobe au niveau de l'extrémité N-terminale, conservé chez les filovirus, a été décrit comme pouvant former une poche de fixation de VP35 (93,195,199,200).

A la différence de l'extrémité N-terminale, l'extrémité C-terminale ne présente aucune homologie de séquence avec d'autres nucléoprotéines de mononégavirus. Elle est responsable de l'interaction protéine-protéine au sein de la nucléocapside (201). Ce module C-terminal hydrophobe est nécessaire pour la formation de la nucléocapside et intervient dans la réplication virale (193,202). De plus, ce domaine est responsable de l'interaction avec la protéine matricielle VP40 et de l'incorporation de la NP dans les virions (130,203). La structure de ce domaine a été déterminée par cristallographie aux rayons X et révèle un domaine globulaire organisé avec une alternance de feuillets β et d'hélices α (196,204). L'étude de la NP des filovirus a montré que le module C-terminal de la NP est un déterminant antigénique majeur, pouvant être utilisé en diagnostic (205). De plus, l'alignement des séquences des NPs des filovirus a mis en évidence, qu'en dépit de certaines similitude de structure, la séquence en acides aminés était faiblement conservée d'une espèce à une autre,

ouvrant la voie pour le développement de tests diagnostiques spécifiques de l'espèce virale (206–208).

La NP des filovirus subit également des modifications post-traductionnelles comme la O-glycosylation et la sialylation, nécessaires pour l'interaction avec la protéine VP35 suggérant un rôle essentiel dans le cycle de réplication virale (193,209). De plus, il a été montré que les NPs d'EBOV et de MARV pouvaient être phosphorylées (210). Les sites de phosphorylation sont principalement localisés au niveau du module C-terminal (211), et pourrait intervenir dans la régulation de l'affinité de la NP pour le génome viral (212). En effet, la phosphorylation de la NP entraîne une diminution de l'affinité NP/ARN, qui aurait pour conséquence de rendre le génome accessible à la polymérase pour la réplication et la transcription (212).

4.2.2. La phosphoprotéine

4.2.2.1. La phosphoprotéine du virus de la rage

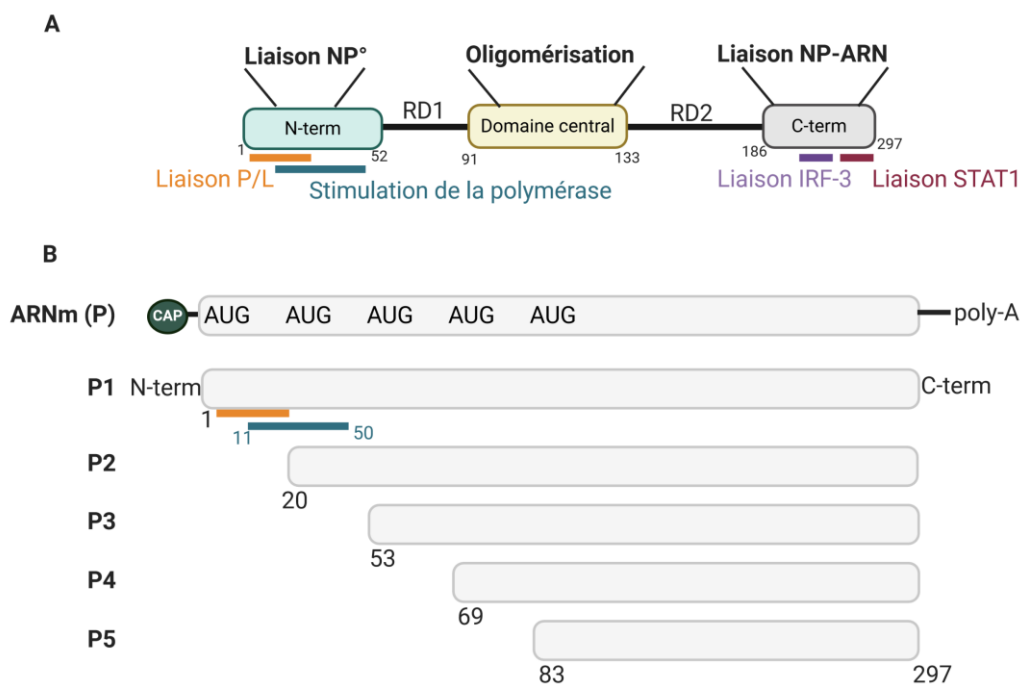


Figure 11 - La phosphoprotéine du virus de la rage

(A) Représentation de l'organisation de la phosphoprotéine (P) de RABV. La protéine P comprend trois domaines (N-terminal, domaine central et C-terminal) reliés entre eux par des régions désordonnées (RD). (B) Les différents isoformes de la protéine P. L'ARNm de la protéine P présente plusieurs codons de début de traduction (AUG) conduisant à la synthèse de protéine P complète (P1) et de protéines P tronquées (P2, P3, P4, P5).

La protéine P de la rage est une protéine multifonctionnelle qui joue un rôle dans la synthèse de l'ARN viral, participe à la régulation de la réplication et de la transcription et intervient également dans l'échappement de la réponse immunitaire innée. En effet, cette protéine représente le principal cofacteur de la polymérase et agit comme une chaperonne de la N. Elle intervient dans l'interaction de la N-ARN et de la protéine L formant le complexe ribonucléoprotéine (RNP). Néanmoins la protéine P seule n'a aucune activité enzymatique documentée (213). De plus, la protéine P interagit également avec le facteur de transcription IRF3, limitant la production d'IFNs, (214,215) et avec STAT1 et STAT2, limitant ainsi la production de protéines antivirales (*voir chapitre 6.3*) (216–218). La protéine P du virus de la rage (RABV) est retrouvée en solution sous forme d'un dimère (219,220). C'est un polypeptide de 297 acides aminés formé de trois domaines liés entre eux par deux régions désordonnées (RD) (**Fig. 11A**). Le domaine N-terminal (N-term) interagit d'une part avec la N et la maintient sous forme soluble, non liée à l'ARN (N°) (171), et d'autre part permet sa liaison avec la polymérase L (221). Le domaine central est responsable de l'oligomérisation de la protéine P (222–224). Enfin, le domaine C-terminal (C-term) se lie au complexe N-ARN (225–227). Les régions désordonnées reliant les différents domaines confèrent à la protéine P une grande flexibilité participant au fonctionnement de la machinerie répliquative. Les premières études *in vitro* de synthèse d'ARN chez le rhabdovirus VSV ont permis de mettre en évidence l'implication de la protéine P dans l'activité de la polymérase virale. La P, via son domaine N-terminal, interagit avec la protéine L. Cette interaction entraîne un changement de conformation de la L et stimule son activité polymérase (228,229). Cette stimulation de l'activité polymérase par la protéine P est indépendante de sa capacité de recrutement de la protéine L pour la formation du complexe RNP (230). Cette observation a également été rapportée pour la protéine P du RABV, suggérant un mécanisme de stimulation commun au sein des rhabdovirus. L'étude de la synthèse *in vitro* d'ARN chez RABV a mis en évidence la spécificité de l'interaction entre la protéine L et P dans la processivité de la polymérase. De plus, le domaine de liaison et le domaine de stimulation de la protéine L sont différents mais étroitement positionnés sur la protéine P. Les résidus 1-19 de la P participent à la liaison de la protéine L alors que les résidus 11 à 50 participent à la stimulation de l'activité polymérase (231) (**Fig. 11**).

In vitro, l'initiation de la traduction de la protéine P peut se faire au niveau plusieurs codons « start » (N=5), conduisant à la production de la protéine P complète (P1 de 297 acides aminés) et de protéines P tronquées (P2, P3, P4 et P5) (232) (**Fig. 11B**). Curieusement,

les protéines P tronquées ne sont pas retrouvées dans les virions matures. Il a été montré que le site d'interaction à la protéine L était situé au niveau des 19 premiers résidus de la protéine P (233), suggérant que les protéines tronquées n'interviennent pas dans sa fonction de cofacteur. De nombreuses études ont mis en évidence le rôle de ces protéines tronquées dans l'inhibition de la réponse immunitaire via notamment leur interaction avec le facteur de transcription STAT1 (217,218,234). Ces protéines participent à la pathogénicité du virus de la rage, supportant la réplication virale au niveau périphérique en inhibant la réponse antivirale médiée par l'IFN et facilitant ainsi l'évasion du virus vers le SNC (235,236).

La protéine P de nombreux mononégavirus subit des modifications post-transcriptionnelles. En effet, elle peut être phosphorylée sur plusieurs sites par la protéine kinase C et une autre protéine kinase cellulaire (237). Des études *in vitro* et *in vivo* prenant pour modèle la protéine P du VSV ont permis de caractériser le rôle de telles modifications. La protéine P non phosphorylée est biologiquement inactive, la phosphorylation par des kinases cellulaires permet son activation (238). De plus, la protéine P présente des sites de phosphorylation au niveau du domaine N-terminal et C-terminal, participant à la régulation de la réplication/transcription (239,240).

4.2.2.2. La phosphoprotéine du virus Ebola (VP35)

La protéine virale VP35 est l'équivalent des protéines P codées par les autres virus de l'ordre des *Mononegavirales*. Elle assure diverses fonctions, jouant un rôle important dans le cycle viral mais également dans l'inhibition de la réponse immunitaire antivirale. Il a été montré que VP35 joue un rôle de chaperonne pour la NP et représente le principal cofacteur de la polymérase (L). VP35 est constituée de deux domaines conservés chez les filovirus (241,242). Un domaine d'oligomérisation en N-terminal et un domaine d'inhibition de l'IFN (IID) en C-terminal (243) (**Fig. 12**).



Figure 12 - Organisation de VP35

VP35 est formée de deux domaines principaux. Un domaine d'oligomérisation en N-terminal permettant l'interaction protéine-protéine et un domaine d'inhibition de la réponse immunitaire (IID). Un domaine supplémentaire en amont du domaine d'oligomérisation joue un rôle dans la régulation de la réplication virale.

Le domaine d'oligomérisation participe à l'interaction protéine-protéine, incluant l'interaction de VP35 avec la polymérase L (244) et l'association de plusieurs VP35 entre

elles (245,246). De plus, l'étude des séquences des filovirus a mis en évidence un domaine additionnel conservé en amont du domaine d'oligomérisation nommé NPBP (« nucleoprotein binding domain ») (246) (**Fig. 12**). Cette région participe à la régulation de l'interaction NP-ARN. En effet, la liaison de VP35 via son domaine NPBP à la NP, limite l'oligomérisation de cette dernière et inhibe son association avec des ARNs non spécifiques, comme les ARNs cellulaires. Au cours de l'association de la NP à l'ARN viral, la NP est libérée de l'interaction avec la région NPBP de VP35 et interagit avec VP35 via un domaine distinct du domaine NPRB (200).

Le rôle de VP35 dans la limitation de la réponse immunitaire a également été montré *in vitro* chez EBOV et RESTV (247,248). En effet, VP35 inhibe la phosphorylation du facteur de transcription IRF3, empêchant sa translocation dans le noyau et inhibant ainsi la production d'IFN (*voir chapitre immunité antivirale*) (248). De plus, le domaine IID est capable de fixer l'ADN(db) et ainsi limite l'activation du senseur RIG-I (241).

4.2.3. La protéine L des mononégavirus

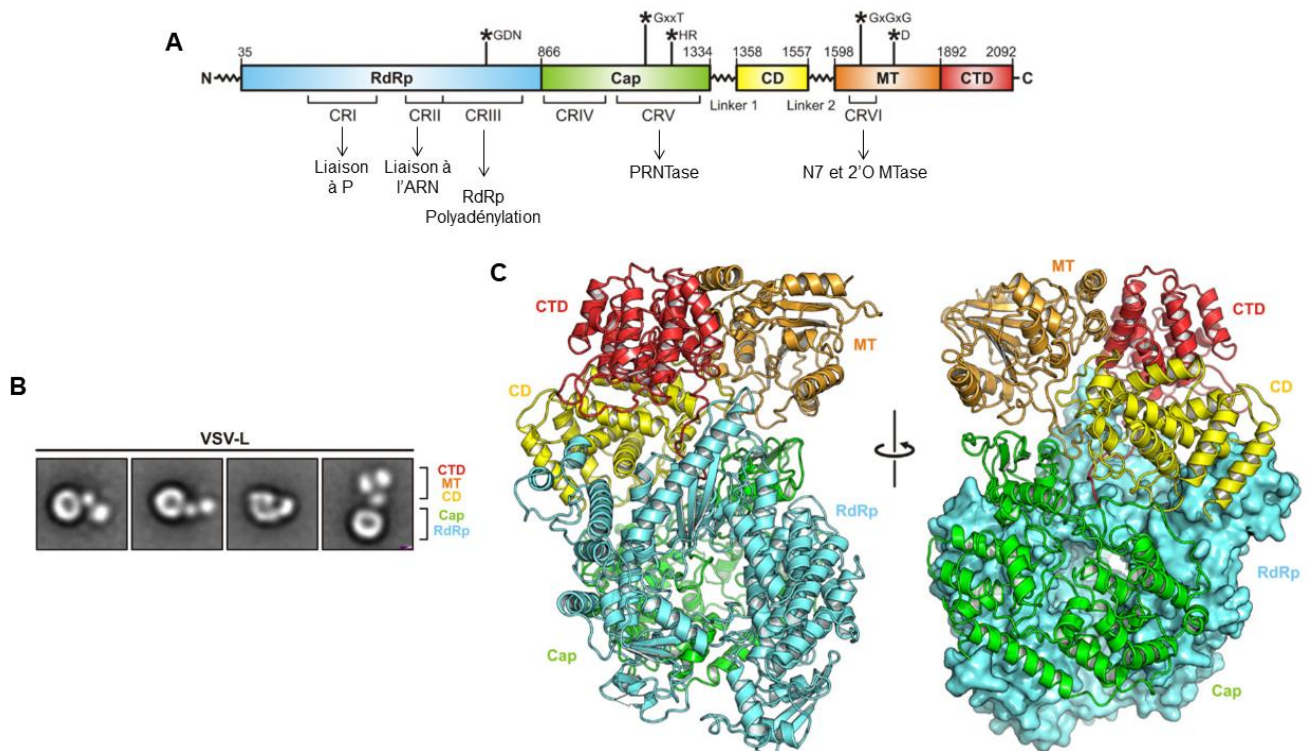


Figure 13- Structure de la protéine L des *Mononegavirales* (modèle VSV)

(A) La protéine L des *Mononegavirales* est organisée en plusieurs domaines. Les premières études bio-informatiques des séquences de protéine L ont mis en évidence 6 domaines conservés nommés CR-I à CR-VI dont certains motifs (indiqués par un astérisque) portent les activités enzymatiques de la L (indiquées en dessous par une flèche). La résolution de la structure de la L par cryo-EM a permis de définir l'organisation en 5 domaines fonctionnels : le domaine ARN polymérase ARN-dépendante (RdRp) (CR-I à CR-III) (colorée en bleu) suivi du domaine Cap (CR-IV et CR-V) (en vert) qui porte l'activité polyribonucléotidyltransférase (PRNTase) nécessaire à la synthèse de la structure coiffe, un domaine connecteur (CD) (en jaune), le domaine méthyltransférase (MTase ou MT) (CR-VI) (en orange) et le domaine C-terminal (CTD) (en rouge). (B) Clichés des premières images de microscopie électronique de la protéine L de VSV, révélant un anneau central (RdRp + PRNTase) et trois domaines globulaires (CD, MTase, CTD). (C) Structure de la protéine L de VSV résolue par cryo-ME à une résolution de 3.8 Å (PDB : 5A22).

Les protéines L sont les protéines les plus conservées au sein des *Mononegavirales*. Elle catalyse l'ensemble des étapes de la transcription et réplication virales. En plus de son activité ARN polymérase ARN-dépendante (RdRp), la protéine L est responsable de la mise en place de la structure coiffe et méthyle les extrémités 5' des transcrits (249,250). L'alignement des séquences des protéines L de différents mononégavirus a mis en évidence 6 régions conservées (notées CR-I à CR-VI), séparées par des régions de plus faibles homologies (**Fig. 13A**) (251,252). La première analyse en microscopie électronique (ME) de la protéine L de VSV révèle un anneau central et 3 domaines globulaires (**Fig. 13B**). Les premières études de l'activité des différents domaines ont conduit à l'hypothèse que l'anneau central était

responsable de la synthèse de l'ARN. Les domaines globulaires sont proposés comme étant responsable de la mise en place de la coiffe. Cette étude en ME a également permis de décrire l'interaction de la protéine L avec son cofacteur, la protéine P (253). Cinq ans après ces premières études, la structure de la protéine L du VSV a été déterminée par cryo-microscopie électronique (cryo-ME) à une résolution de 3.8 Å (254) (**Fig. 13C**). Cette structure révèle l'organisation de la protéine, formant deux blocs distincts : N-terminal et C-terminal, connectés via le « connecteur domaine » (CD) ne présentant pas de fonction enzymatique apparente. Elle contient 5 domaines dont l'ordre et la fonction sont conservés au sein des *Mononegavirales*: le domaine RdRp (CR-I à CR-III) suivi du domaine Cap (CR-IV et CR-V) qui porte l'activité polyribonucléotidyltransférase (PRNTase) nécessaire à la synthèse de la structure coiffe, le domaine connecteur (CD) (en jaune), le domaine méthyltransférase (MTase ou MT) (CR-VI) et le domaine C-terminal (CTD) (en rouge) (**Fig. 13A & 13C**).

4.2.3.1. L'extrémité N-terminale de la protéine L

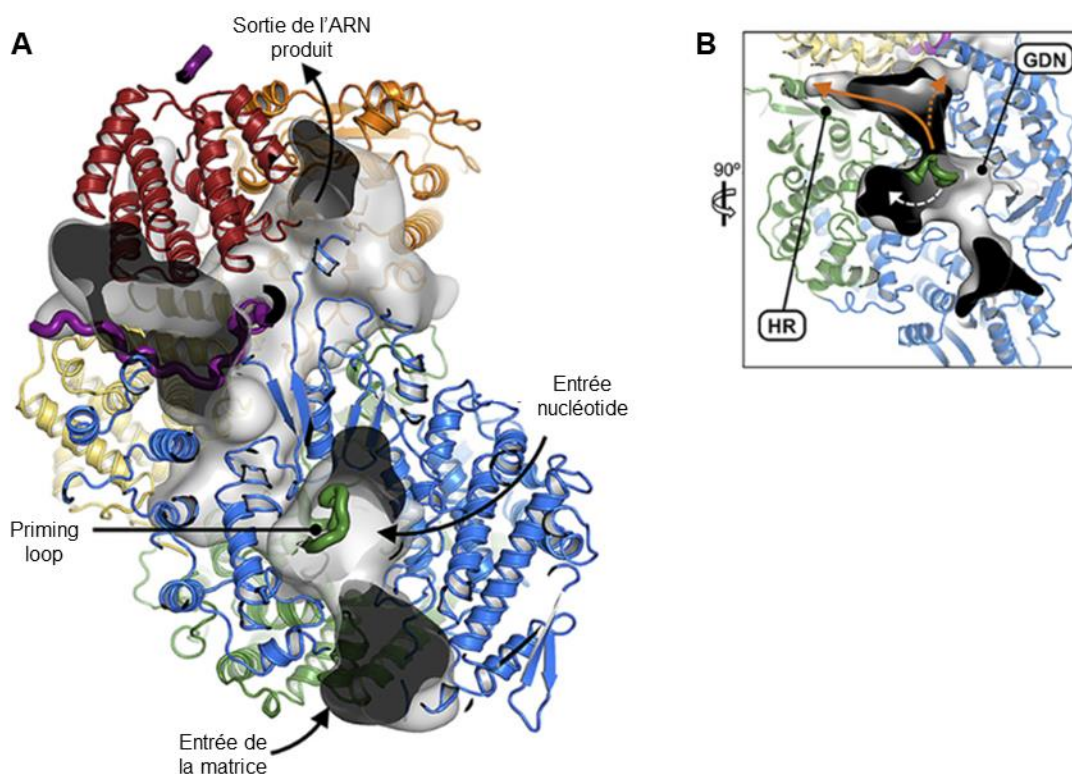


Figure 14- La protéine L du virus de la rage

(A) Modélisation du transit de l'ARN à travers la protéine L. La protéine L du virus de la rage a été résolue par cryo-microscopie électronique à une résolution de 3.3 Å (PDB : 6UEB). L'entrée et la sortie de l'ARN sont indiquées par des ombres noires. La « priming loop » (en vert) est en position d'initiation de la polymérisation. L'entrée des nucléotides est en regard. La cavité de transit de l'ARN est indiquée en gris. (B) Représentation de la « priming loop » en position initiation. Les pointillés blancs représentent le retrait de la boucle et le passage en mode élongation. Les deux flèches oranges représentent le transit de l'ARN vers le site catalytique (motif HR) de la polyribonucléotidyltransférase (PRNTase) pour la mise en place de la coiffe (flèche pleine) ou vers le domaine méthyltransférase (MTase) (flèche en pointillés).

L'extrémité N-terminale est constituée du domaine RdRp et du domaine Cap (**Fig. 14**). Le domaine RdRp présente un repliement commun à d'autres ARN et ADN polymérases, dit en main droite avec trois régions principales : les « doigts », la « paume », et le « pouce » (« finger-palm-thumb ») (254). Le site actif de la RdRp, correspondant à la région CR-III, présente un motif GDN conservé chez les virus de l'ordre des *Mononegavirales* (255) entouré de plusieurs régions formant un tunnel pouvant guider l'ARN vers le site actif. Le domaine RdRp possède également une activité de polyadénylation impliquant un mécanisme de glissement (« slippage ») au niveau d'une séquence spécifique en 3' des gènes (256). Le domaine « Cap » ou PRNTase est intimement lié au domaine polymérase et projette une boucle (« priming loop ») jouant un rôle clef dans l'initiation de la polymérisation près du site

catalytique de la RdRp (254). Le domaine PRNTase contient 2 motifs conservés, le motif HR et le motif GxxT qui participent à la formation de la structure coiffe et dont le mécanisme est décrit dans la partie (5.3.1.2) (254,257–259). La structure de la protéine L de VSV suggère que la RdRp et les activités de synthèse de la coiffe sont interconnectées, confirmant les études biochimiques (257,260). La fonction de la boucle flexible présente dans le domaine PRNTase a été étudiée par mutagenèse dirigée. Ces études ont confirmé le double rôle de cette boucle dans la transcription et la synthèse de la coiffe des transcrits. De plus, cette étude a mis en évidence l'importance du cofacteur (la protéine P) pour l'initiation de la transcription (261). Les structures cryo-EM récemment résolues de la protéine L des virus hMPV et RSV révèlent un repliement de la « priming loop » du domaine PRNTase différent de celui observé chez VSV et RABV, lié à la grande flexibilité de cette boucle. En effet, cette boucle est figée dans une conformation dite fermée mimant l'étape d'initiation de la polymérisation dans la structure de VSV et RABV (254,262,263). Les structures de RSV et hMPV révèlent une conformation dite ouverte, la boucle se retire de la cavité de la RdRp pour permettre l'élongation du transcrit vers le site catalytique de la PRNTase (264,265).

4.2.3.2. L'extrémité C-terminale de la protéine L

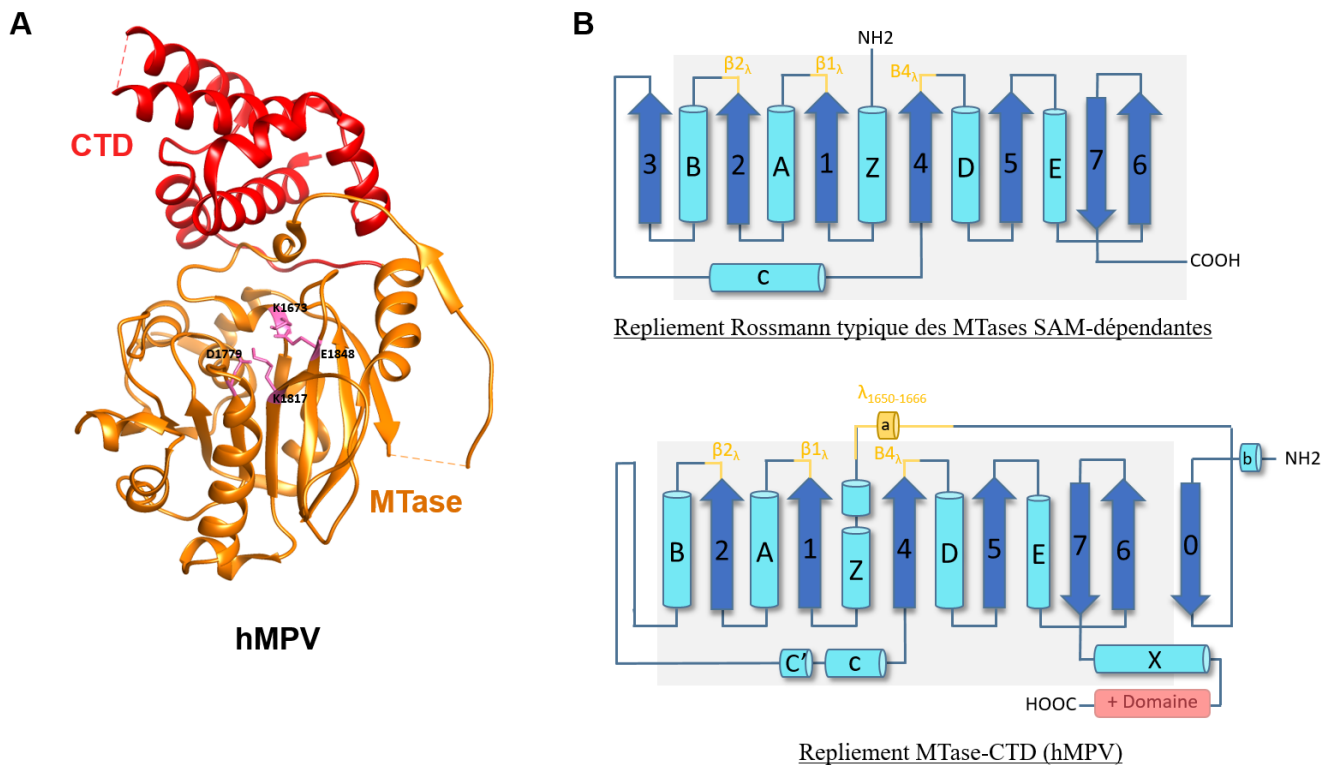


Figure 15 - Extrémité C-terminale de la protéine L de hMPV

(A) Le domaine méthyltransférase (MTase, en orange) et le domaine C-terminal (CTD, en rouge) ont été résolus à 2,2 Å par diffraction aux rayons X (PDB : 4UCI). Le site actif permettant la méthylation de la structure coiffe contient le motif catalytique **K**₁₆₇₃-**D**₁₇₇₉-**K**₁₈₁₇-**E**₁₈₄₈ (en rose). (B) Représentation schématique des structures secondaires du repliement Rossmann des MTases SAM-dépendantes (en haut) et de la MTase de hMPV (en bas). Les cylindres bleu clair représentent les hélices α et les flèches bleu foncé les feuillets β . Certaines boucles (en jaune) sont responsables du recrutement du SAM ($\beta_{1\lambda}$ et $\beta_{2\lambda}$) et de l'ARN substrat $\beta_{4\lambda}$. Le domaine MTase de hMPV présente un repliement semblable aux MTases SAM-dépendantes suggérant une conservation du mécanisme de méthylation des ARNm.

Adaptée de Paesen. G.C, Collet A. et al., 2015

L'extrémité C-terminale de la protéine L comprend les domaines CD, MTase et CTD. Les fonctions des domaines CD et CTD ne sont pas connues à ce jour. Le domaine MTase, quant à lui, catalyse la maturation des transcrits néosynthétisés en méthylant la structure coiffe. Le domaine MTase possède un repliement de type Rossmann (**Fig. 15**), caractéristique des MTases catalysant le transfert de groupement méthyle à partir de la S-adénosyl méthionine (SAM). Sa structure est constituée d'une alternance de feuillets β et d'hélices α . La résolution des structures de la protéine L de VSV, hMPV et RABV montrent que leurs domaines MTases sont formés d'une alternance de feuillets β et d'hélices α (254,262,263,265–267). Le domaine MTase contient un motif GxGxG, très conservé chez les virus de l'ordre des

Mononegavirales (et les MTases en général), qui participe au recrutement du substrat donneur de méthyl SAM pour la réaction de méthylation de la structure coiffe (249,267). En effet, le domaine MTase est responsable de la méthylation des structures coiffes sur le 2'OH du ribose du premier nucléotide puis en position N7 de la guanosine (*voir 5.3.1.2*) (268). Ce domaine contient également une tétrade catalytique K-D-K-E caractéristique des 2'O MTases (250) (**Fig. 15A**). Néanmoins, les études biochimiques menées sur le domaine MTase des *Mononegavirales* montrent que les deux activités 2'O et N7 MTases sont catalysées par le même site actif (250,269), impliquant des mécanismes de repositionnement de l'ARN pour expliquer ces deux activités (N7 et 2'O MTases). Une étude récente menée au sein de notre laboratoire a permis de caractériser l'activité MTase de SUDV (270). Il a été montré que la MTase de SUDV méthyle la structure coiffe des ARNs viraux (N7 et 2'O méthylation). De plus, la MTase de SUDV catalyse des modifications épi-transcriptomiques des ARNs viraux en méthylant des adénosines internes en position 2'O dans la séquence de l'ARN substrat. Récemment, la structure de la protéine L de RSV a été résolue. Cependant, bien que la protéine présente une activité MTase *in vitro*, les domaines CD, MTase et CTD ne sont pas visibles dans la structure suggérant une mobilité importante de ces domaines. Des observations similaires ont été réalisées chez VSV (253), confirmant la flexibilité de ces domaines liées aux domaines RdRp et Cap (264). Plus récemment, l'étude structurale de la protéine L d'un paramyxovirus a révélé une conformation unique de la protéine L où les domaines MTase et CTD présentent le site actif de la MTase au-dessus du domaine RdRp suggérant que l'ARN naissant est directement pris en charge par le domaine MTase afin d'assurer les méthylation de la structure coiffe (271).

4.2.3.3. L'interaction de la protéine L et son cofacteur P

La protéine L peut interagir avec son cofacteur, la protéine P, au niveau de plusieurs sites. Il a été notamment montré chez le VSV que l'extrémité N-terminale de la protéine L participe à cette interaction via la région conservée CR-I de son domaine RdRp (272–274). En effet, des mutations au niveau dans la région CR-I du virus Sendai, un mononégavirus de la famille des *Paramyxoviridae*, conduisent à une perte d'interaction entre la protéine L et son cofacteur (272). Des études d'immunoprécipitation ont montré que l'extrémité C-terminale de la protéine L présente également un site de liaison à la protéine P (233). Des études menées sur VSV ont mis en évidence que la région minimale de P nécessaire à l'activation de la L *in vitro* permet d'ancrer de manière stable les domaines C-terminaux (CD, MTase et CTD) contre les domaines N-terminaux (RdRp et PRNTase) de la polymérase (230,253). Une étude de

délétion de domaine de la protéine L de RABV a permis également d'identifier une séquence conservée au sein des lyssavirus (NPYNE) située en C-terminal (entre le domaine MTase et CTD) essentielle pour l'interaction avec la protéine P (275). Plus récemment, la structure du complexe polymérase (L/P) de RABV a été résolue par cryo-ME à une résolution de 3.3 Å confirmant l'implication du domaine C-terminal de la protéine L dans l'interaction avec son cofacteur (262). De plus, l'amélioration de la résolution de la structure L/P de VSV 3 Å a permis d'identifier 3 segments de P interagissant avec la L et mis en évidence leur implication dans la conformation de la protéine L (263) (**Fig. 16**).

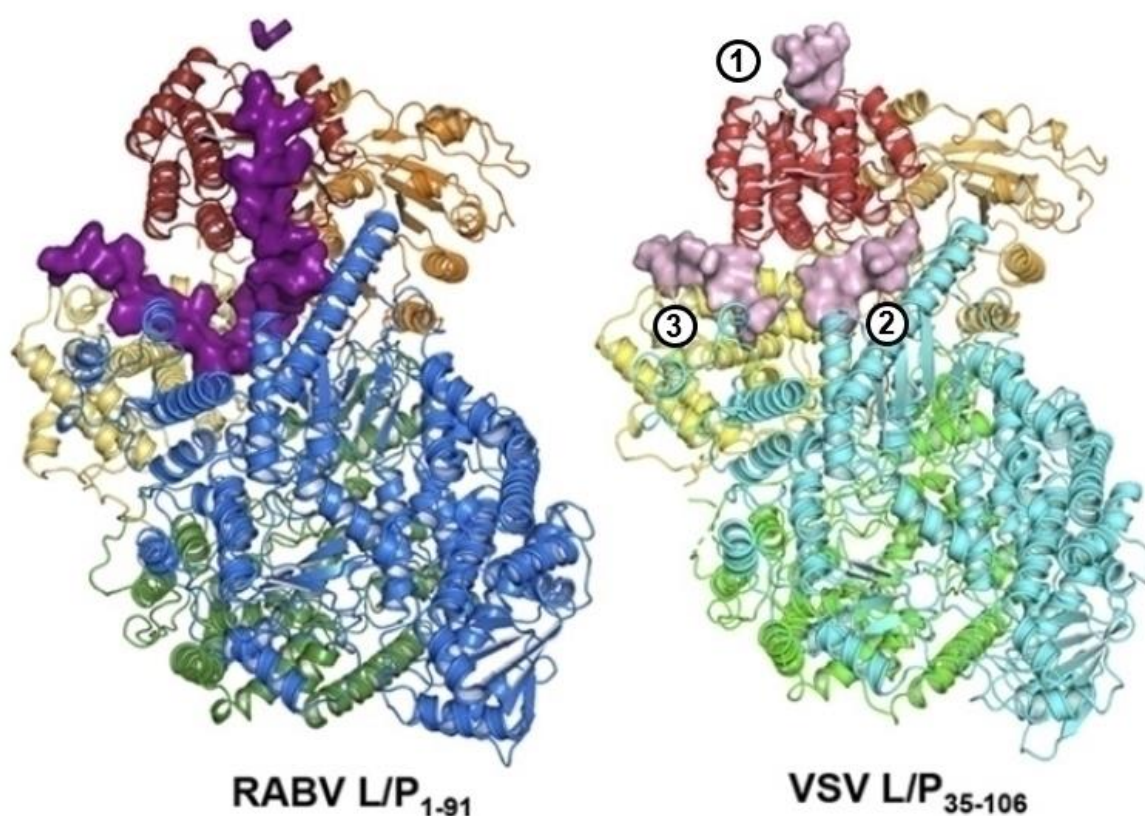


Figure 16 - Structure du complexe L/P chez les rhabdovirus

Structure du complexe L/P de RABV à gauche et complexe L/P de VSV à droite. La protéine P de RABV est indiquée en violet foncé, celle de VSV en violet clair. La protéine P de VSV présente trois segments d'interaction avec la protéine L numérotés de 1 à 3. Le segment 1 interagit avec le domaine CTD de la L (en rouge), le segment 2 se positionne entre le domaine CTD et RdRp (en bleu) de la L. Enfin le troisième segment de P se place dans le sillon formé par le domaine CD (en jaune) et CTD. L'interaction entre la L et la P de RABV est similaire à celle observée chez VSV. Cependant, la protéine P de RABV interagit également avec le domaine MTase (en orange).

Issue de J.A. Horwitz et al., 2020

Le premier segment interagit au niveau du domaine CTD de la protéine L et le second se lie entre le CTD et le domaine RdRp. Enfin, le troisième segment se situe dans une cavité entre le CTD et le CD. Le modèle indique que la protéine P induit une conformation de la protéine L fermée, compacte et compatible avec l'étape d'initiation (263). L'interaction entre L/P de RABV est similaire à celle observée pour VSV. La protéine P est par ailleurs en contact avec les domaines CD, MTase et CTD et maintient la protéine L dans une conformation dite fermée (262) (**Fig. 16**). Par opposition, les structures du complexe P/L chez hMPV et RSV, récemment résolues, indique que la P interagit uniquement avec les domaines RdRp et Cap, témoignant d'une flexibilité dans l'articulation des différents domaines de la polymérase (264,265).

En conclusion, la protéine L des *Mononegavirales* est un acteur clef dans le cycle de réplication de ces virus. En effet, en plus de l'activité RdRp, elle catalyse la synthèse et la méthylation de la structure coiffe. L'ajout de cette structure, dont le mécanisme est détaillé dans le chapitre suivant, permet de protéger les transcrits de la dégradation par des 5'-exonucléases cellulaires et contribue à l'initiation de la traduction des protéines virales. De part ces fonctions indispensable au cycle virale, la protéine L représente donc une cible attractive pour le développement de nouvelles molécules à potentiel antiviral. Les travaux exposés dans ce manuscrit s'intéressent plus particulièrement à l'activité MTase portée par la protéine L des *Mononegavirales*.

Le prochain chapitre s'attache à présenter les différents mécanismes de mise en place de la structure coiffe chez les eucaryotes et les virus.

5. Mécanisme de coiffage chez les eucaryotes et les virus

5.1. La structure coiffe

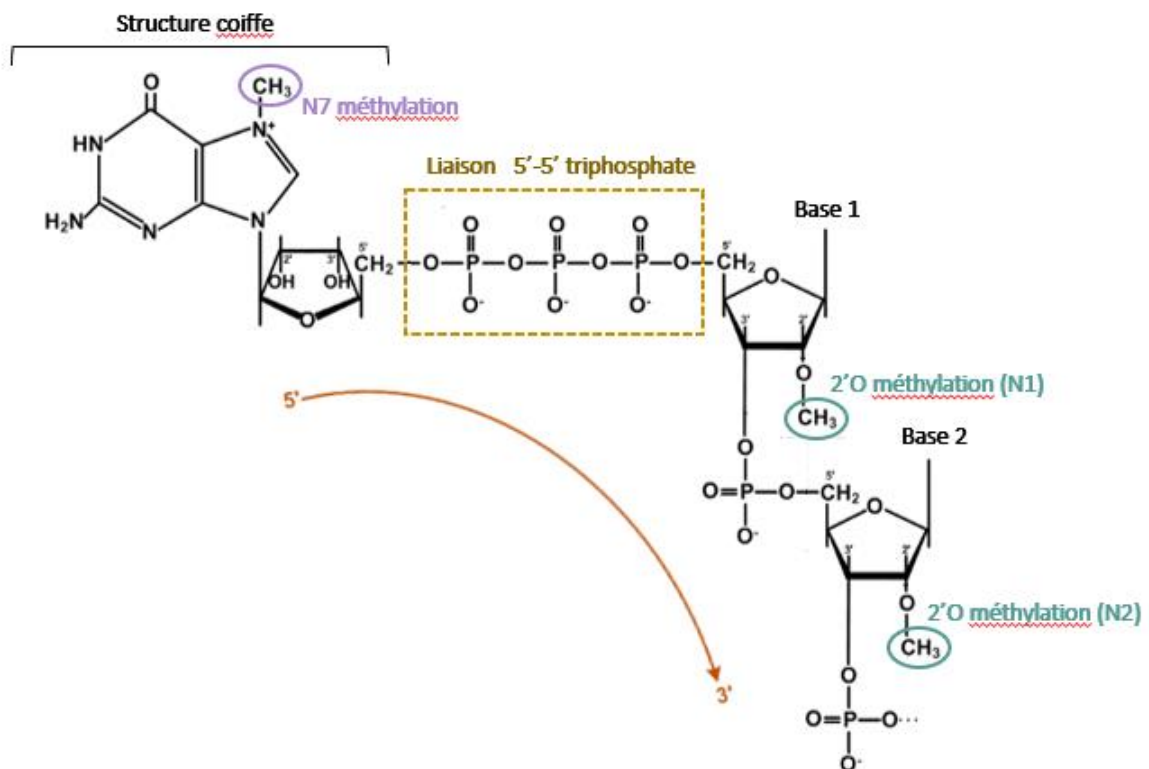


Figure 17 - La structure coiffe des ARNs messagers

La structure coiffe chez les eucaryotes est constituée d'une guanosine méthyliée en position N7 (entouré en violet) liée par une liaison 5'-5' triphosphate (encadré en doré) à l'extrémité 5' de l'ARN (représenté via la flèche 5'-3' en orange). Le premier et parfois le second nucléotide sont méthyliés sur l'hydroxyle en position 2' (entouré en bleu) formant, respectivement, la coiffe 1 et 2.

Adaptée de Ferron et al, 2012

La structure coiffe est une modification co-transcriptionnelle portée à l'extrémité 5' des ARNm. Cette structure chimique est constituée d'une guanosine liée au premier nucléotide de l'extrémité 5' de l'ARNm par une liaison 5'-5' triphosphate. Cette structure coiffe peut être méthyliée à diverses positions : en position N7 de la guanosine ($^m\text{GpppN}$ ou coiffe-0) et en position 2' hydroxyle du ribose du premier nucléotide ($^m\text{GpppN}_m$ ou coiffe-1). En fonction de la complexité des organismes eucaryotes, des méthylations supplémentaires peuvent être observées en position 2'O du second (N2) et du quatrième (N4) nucléotide de l'extrémité 5' de l'ARNm (appelées respectivement coiffe-2 et coiffe-4) (276,277) (**Fig. 17**).

5.2. La synthèse de la structure coiffe chez les eucaryotes

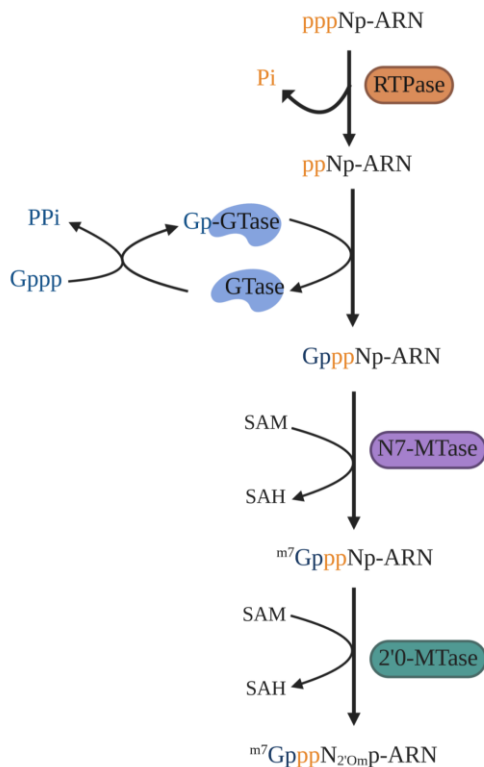


Figure 18 - Formation de la structure coiffe chez les eucaryotes

L'ARN triphosphatase (RTPase) (en orange) hydrolyse l'ARN triphosphate et libère l'ARN diphosphate (ppNp-ARN) et du Pi. Une guanylyltransférase (GTase) (en bleu) hydrolyse une molécule de GTP en GMP et libère du PPi. Le GMP est transféré sur ppNp-ARN pour former la coiffe (GpppNp-ARN) qui sera méthylée en position N7 par une N7 méthyltransférase (N7-MTase) (en violet). Le premier nucléotide (N) est méthylé en position 2'O par une 2'O MTase (en vert).

La voie de synthèse chez les eucaryotes est dite « conventionnelle » (**Fig. 18**). Dans le noyau des cellules eucaryotes, la synthèse de la structure coiffe se fait de façon séquentielle et co-transcriptionnelle. Elle fait intervenir différentes enzymes : une 5'-triphosphatase (RTPase), une guanylyltransférase (activité GTase) et des MTases. La RTPase hydrolyse l'ARN 5'-triphosphate (ppp-RNA) en ARN 5'-diphosphate (pp-RNA), entraînant la libération d'une molécule de phosphate inorganique (Pi) (278). La GTase hydrolyse une molécule de GTP (donneur de GMP) en GMP et libère une molécule de pyrophosphate (PPi). La GTase reste fixée de façon covalente au GMP et, de façon concomitante, transfère la molécule de GMP sur l'ARN diphosphate (accepteur de GMP) générant la structure coiffe à l'extrémité 5' de l'ARNm (GpppN1-ARN) (279). La guanosine est ensuite méthylée en position N7 de la coiffe par une N7 MTase cellulaire en présence du donneur de méthyle SAM (S-

adénylméthionine) conduisant à la formation de la coiffe-0 ($m^7\text{GpppGNp-ARN}$). Le premier comme le second nucléotide peuvent également être méthylés sur la position 2'O de leur résidu ribose par des 2'O MTases (2'O MTase), générant une coiffe-1 ($m^7\text{GpppGN}_{12'\text{Omp-ARN}}$) et une coiffe-2 ($m^7\text{GpppGN}_{12'\text{Omp-ARN}}$). La méthylation de ces coiffes 1 et 2 est assurée, respectivement, par CMTr1 (nucléaire) et CMTr2 (cytoplasmique). La synthèse des coiffe-0 est nucléaire tandis que la synthèse des coiffe-1 et coiffe-2 est cytoplasmique (280,281).

5.3. L'ajout de la structure coiffe chez les virus

Les infections virales entraînent le plus souvent le déclenchement d'une réponse antivirale dans la cellule hôte afin d'inhiber la réplication du virus. Les virus ont développé des mécanismes de subversion de la réponse engagée par l'hôte. Les transcrits viraux néosynthétisés subissent des modifications co- et post-transcriptionnelles comme l'ajout d'une structure coiffe à l'extrémité 5' des ARNm. Ces modifications sont essentielles à l'infection virale pour deux raisons majeures. L'ajout de cette structure permet au virus de détourner la machinerie traductionnelle cellulaire et d'assurer une traduction efficace des protéines virales. Elles permettent également de limiter la détection des ARNs viraux par les senseurs de l'immunité innée tels que RIG-I et MDA5 et de prévenir leur dégradation par des 5' exonucléases (*voir chapitre immunité antivirale*). Les trois différentes stratégies permettant au virus d'acquérir une structure coiffe au niveau l'extrémité 5' de leurs ARNm ont été décrits dans la revue suivante (282) (**Fig. 19A**).

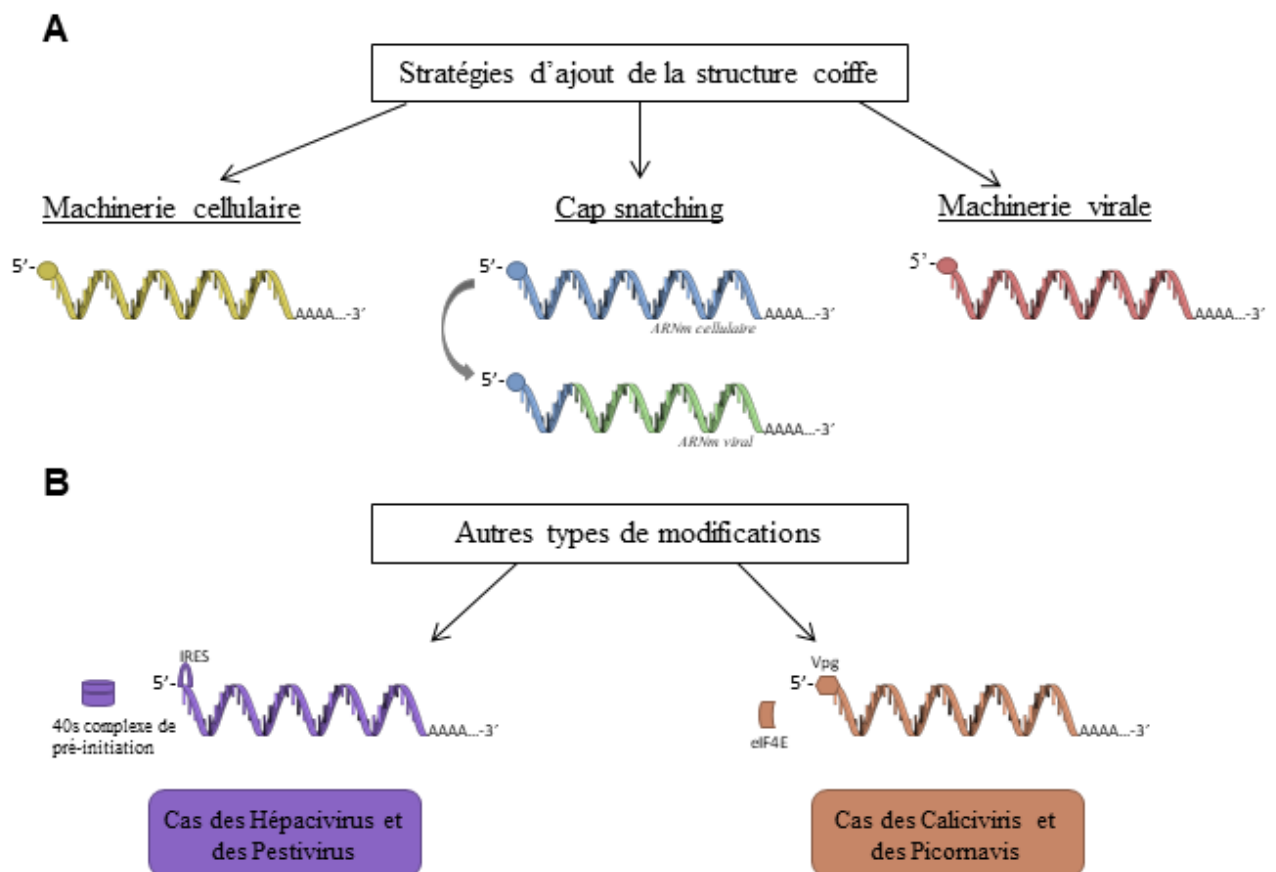


Figure 19 - Modifications 5' des ARNm viraux : les stratégies déployées par les virus

Les virus utilisent la machinerie de traduction cellulaire dépendante de la structure coiffe, nécessitant donc de synthétiser cette structure. (A) Les virus synthétisent leur structure coiffe en utilisant la machinerie cellulaire (en jaune), via le vol la structure coiffe des transcrits cellulaires (en bleu) ou en utilisant leurs propres enzymes virales (en rouge). (B) Certains virus possèdent un IRES (« internal ribosome entry site ») à leur extrémité 5' de l'ARNm, reconnu par le complexe de pré-initiation de traduction de la sous-unité 40S du ribosome (en violet). Pour d'autres virus, la traduction de leur ARNm est dépendante de la présence d'une protéine/peptide liée de façon covalente à l'extrémité 5' du génome viral (Vpg) (en orange).

Brièvement, pour les rétrovirus et la plupart des virus à ADN dont les ARNm sont synthétisés par l'ARN polymérase II, la synthèse de la structure coiffe est réalisée par les enzymes cellulaires décrites ci-dessus. Une seconde stratégie illustrée notamment pour les virus des familles des *Arenaviridae*, *Orthomyxoviridae* et des *Bunyaviridae* consiste à subtiliser une coiffe présente sur un ARNm de la cellule hôte via un mécanisme de “vol de coiffe” (« cap-snatching »). Ce mécanisme fait intervenir une endonucléase qui clive l'ARNm cellulaire une dizaine de nucléotides en aval de la structure coiffe. Les fragments d'ARN ainsi générés servent d'amorces pour la polymérase virale qui catalyse la synthèse des ARNm viraux, conduisant à la formation d'ARNs chimériques (une dizaine de nucléotides issus des ARNs cellulaires suivis par l'ARN viral). Enfin, le dernier mécanisme consiste à coder pour leur propre machinerie enzymatique nécessaire à la formation de leur structure coiffe. Cette

stratégie est notamment utilisée par les virus présentant un cycle répliatif cytoplasmique, comme les coronavirus, les flavivirus ou les *Mononegavirales*.

Toutefois, tous les virus ne synthétisent pas une structure coiffe décorant l'extrémité 5' de leurs ARNm. Ces virus utilisent des stratégies alternatives, dites « coiffe-indépendante », afin d'assurer la traduction des protéines virales et la protection de leurs transcrits. Par exemple, les virus de la famille de *Caliciviridae* masquent l'extrémité 5' de leurs ARNs en se liant de manière covalente à la protéine VPg (« viral protein genome-linked »), qui interagit directement avec la protéine eIF4E et initie ainsi la traduction des protéines virales (283). D'autres virus, comme les hépacivirus et les pestivirus, présentent à l'extrémité 5' de leurs transcrits un repliement spécifique appelé IRES (« internal ribosome entry site ») permettant d'initier la traduction des protéines virales sans passer par le facteur de traduction eIF-4E (284) (**Fig. 19B**).

5.3.1. Les voies de synthèse de la coiffe non-conventionnelles

5.3.1.1. Cas des *Alphavirus*

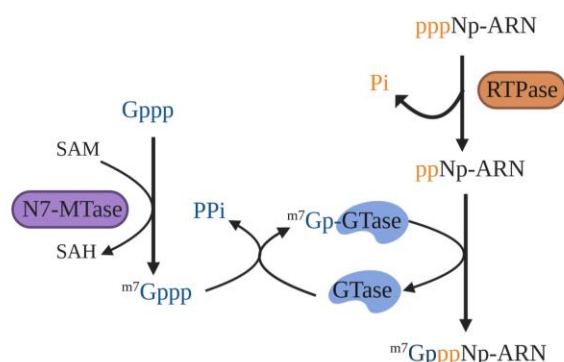


Figure 20 - Synthèse de la structure coiffe chez les virus du genre *Alphavirus*

L'ARN triphosphatase (RTPase) (en orange) hydrolyse le phosphate γ de l'ARN viral et libère ainsi l'ARN diphasphaté (ppNp-ARN) et le Pi. Une molécule de GTP est méthylée en position N7 par une N7 MTase SAM-dépendante. Ensuite, la GTase (en bleu) lie le mGTP d'une façon covalente et libère le Ppi. La GTase transfère la molécule de mGMP sur l'ARN 5' diphasphaté pour former mGpppNp -ARN.

Chez les virus du genre *Alphavirus*, virus à ARN de polarité positive comme le virus Chikungunya, la synthèse de la structure coiffe est dite « non conventionnelle » (**Fig. 20**). En effet, elle débute par la méthylation d'une molécule de GTP en position N7 par la N7 MTase virale en présence du SAM. La GTase hydrolyse la molécule de mGTP et forme un lien covalent avec la molécule de mGMP qui est transférée sur l'extrémité 5' diphasphaté de l'ARN viral générant ainsi une structure coiffe-0 (mGpppARN). Chez les alphavirus, la GTase et la N7 MTase sont portées par la protéine nsP1 (285), l'activité RTPase est portée par la protéine nsP2 (286). Les alphavirus sont dépourvus

d'activité 2'O MTase et ne synthétisent que des coiffes-0 à l'extrémité 5' de leurs ARNs, pouvant les rendre susceptibles à la reconnaissance par des senseurs de l'immunité innée tel

que IFIT. Néanmoins, il a été montré que la présence de structure tige-boucle à l'extrémité 5'UTR des ARNs des alphavirus limitent leur reconnaissance par les senseurs de l'immunité innée tels que RIG-I et les facteurs de restriction IFIT1/3 (287).

5.3.1.2. Cas des *Mononegavirales*

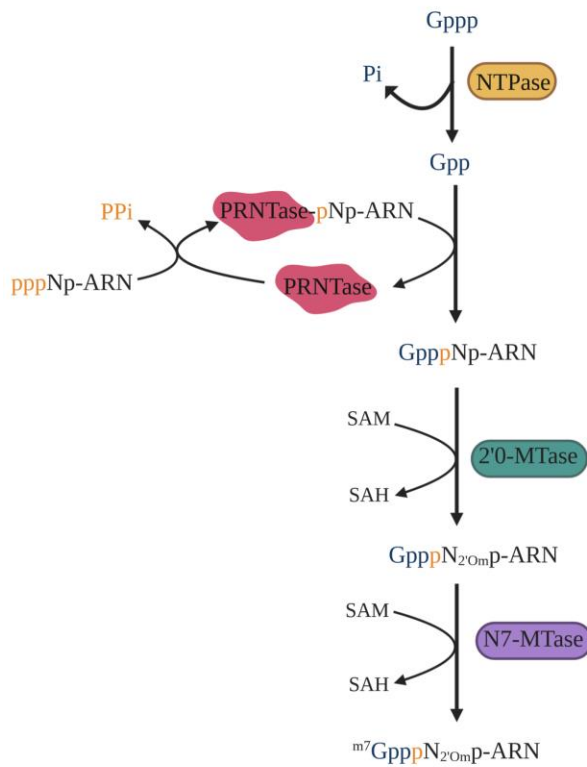


Figure 21 : Synthèse de la structure coiffe chez les *Mononegavirales*.

Le domaine PRNTase (en rose) hydrolyse l'ARN triphosphate et libère du PPI. Dans le même temps, une molécule de GTP est hydrolysée par une NTPase (en orange) et le Pi. Le GDP ainsi formé est transféré sur l'ARN monophosphate pour former la structure coiffe en 5' de l'ARN (GpppNp-ARN). Une 2'O MTase (en vert) méthyle le premier nucléotide (N) puis une N7-MTase (en violet) méthyle la guanosine de la coiffe en position N7.

La synthèse de la structure coiffe chez les virus de l'ordre des *Mononegavirales* est assurée par une protéine virale multifonctionnelle : la protéine L qui porte l'ensemble des activités enzymatiques nécessaire à la synthèse de leurs ARNs et à l'ajout de la coiffe en 5' des ARNm viraux (voir chapitre 5.2.3). L'ajout de la structure coiffe se fait de façon dite « non-conventionnelle » (288) (**Fig. 21**). En effet, dans un premier temps, la PRNTase, via l'histidine du motif HR, forme un lien covalent avec l'ARN viral 5'monophosphate, entraînant le relargage d'une molécule de PPI. De façon concomitante, une molécule de GTP est hydrolysée en GDP par une guanosine 5' triphosphatase (GTPase) (288). Le domaine de la protéine L portant cette activité GTase n'a, à ce jour, pas été clairement identifié bien que certaines études s'orientent vers le domaine MTase+CTD (266). La molécule de GDP ainsi formée se lie au motif GxxT de la

PRNTase. Cette interaction permet le transfert de l'ARN 5'monophosphate sur le GDP conduisant à la formation de la structure coiffe (GpppN-ARN, avec N correspondant au premier nucléotide de la séquence d'ARN). La méthylation de la structure coiffe chez les *Mononegavirales* suit également un ordre réactionnel non conventionnel (266,289). En effet, alors que chez les eucaryotes la N7 méthylation précède la 2'O méthylation (**Fig. 21**), chez les

Mononegavirales, la méthylation en position 2'O du premier nucléotide de l'extrémité 5' de l'ARN précède la méthylation en position N7 de la guanosine de la coiffe (266,289).

5.4. Rôles de la structure coiffe

La structure coiffe joue un rôle majeur dans plusieurs processus biologiques. Elle est nécessaire pour la maturation des ARNm, leur épissage alternatif, leur export nucléaire pour leur traduction et leur stabilité. De plus, elle représente un marqueur du « soi ».

Dans le cytoplasme, le recrutement des ribosomes et des facteurs de traduction permet une traduction efficace des protéines. Le facteur de traduction eIF-4E reconnaît spécifiquement la structure coiffe (290,291). Chez les virus qui utilisent la machinerie ribosomale cellulaire pour la traduction de leurs protéines, la structure coiffe en 5' de leur ARNm est reconnue par eIF-4E favorisant la traduction de leurs protéines virales. La structure coiffe permet également de stabiliser les ARNm en les protégeant de la dégradation des 5'-3' exonucléases, contribuant ainsi à augmenter leur durée de vie dans les cellules infectées (292). Dans le cytoplasme, les transcrits ont une demi-vie variable, quelques minutes à quelques jours, dépendante de l'efficacité de recrutement de la machinerie de dégradation. La dégradation des transcrits débute par le retrait de la queue poly-A en 3', puis soit le transcrit est décoiffé et dégradé par l'exonucléase 5'-3' Xrn1, soit via le recrutement de l'exosome qui dégrade les transcrits de 3' en 5' (293).

D'autre part, l'addition de la structures coiffe à l'extrémité 5' des ARNs viraux limite leur détection par les senseurs de l'immunité innée. De plus, la méthylation en position 2'O du premier nucléotide représente un marqueur du « soi ». Ainsi les ARNs coiffés dépourvus de cette modification seront reconnus par les senseurs de l'immunité innée de la famille des RLRs entraînant la production d'IFNs (*voir 6.2*) (294,295). La protéine IFIT1 produite suite à l'activation de la voie des IFNs se lie aux ARNs dépourvus de 2'O méthylation et limite leur traduction en séquestrant les ARNs (296,297). Par ailleurs, il a été montré que certaines 2'O MTases virales étaient capables de méthyler des positions internes à la séquence d'ARN (298,299). Le rôle de ces méthylations reste encore mal compris. Toutefois, il a récemment été montré que le virus de l'immunodéficience humaine (HIV) recrute une 2'O MTase cellulaire (FtsJ3) permettant la méthylation de son ARN sur 17 positions spécifiques. Ces méthylations internes spécifiques participent à la subversion des défenses de l'hôte en limitant la détection virale par le récepteur MDA5 et l'induction de la voie des INFs (300).

6. Immunité innée de l'hôte dans un contexte d'infection virale

Les virus sont des parasites obligatoires utilisant la machinerie des cellules hôtes pour se multiplier. L'émergence virale résulte souvent du franchissement de la barrière d'espèce à la suite de situations de stress écologique. Le virus doit s'adapter pour se propager dans une nouvelle espèce. Au cours de ma thèse, je me suis intéressée aux virus à ARN qui se répliquent via une ARN polymérase ARN-dépendante dépourvue d'activité « proof-reading » (correction d'erreur), ce qui favorise l'apparition de mutations à chaque cycle répliatif (301). En effet, le taux de mutation pour les virus à ADN est de 1 mutation pour 10^6 à 10^8 paires de base et par génération, alors que le taux de mutation pour les virus à ARN est de 1 mutation pour 10^4 à 10^6 paires de base (302). L'apparition de ces mutations participe à l'évolution des virus, jouant un rôle majeur dans l'adaptation des virus à un nouvel hôte et l'échappement à la réponse immunitaire. Par conséquent, les virus ont développé de nombreux mécanismes de subversion des défenses de l'hôte leur permettant de contrecarrer les réponses antivirales et ainsi de favoriser leur propagation. Ce chapitre introduit la réponse immunitaire innée engagée par l'hôte en réponse à une infection virale.

6.1. Les récepteurs de l'immunité innée

Les eucaryotes supérieurs ont développé divers mécanismes afin de lutter contre l'infection virale. Ces mécanismes de défense sont regroupés sous deux axes complémentaires : la réponse immunitaire innée, immédiate mais peu spécifique, et la réponse immunitaire adaptative, plus tardive mais qui confère une protection durable et spécifique contre l'agent infectieux. Cette réponse repose sur la capacité des cellules immunitaires à discriminer le « soi » du « non soi » via l'expression de récepteurs spécifiques (« pathogen recognition receptors » ou PRRs) capables de reconnaître des motifs moléculaires associés aux pathogènes (PAMPs : « pathogen-associated molecular patterns ») (303). Il existe trois grandes familles de PRR, les récepteurs membranaires de type Toll (« Toll-like receptors », TLRs), les récepteurs de type Nod (« Nod-like receptors », NLRs), et les récepteurs cytoplasmiques de type RIG-I (« RIG-I-like receptors », RLRs) (304). L'activation de ces récepteurs déclenche une cascade de signalisation qui entraîne la production d'interférons (IFNs) de type 1 (IFN α/β) et l'expression de gènes stimulés par l'IFN, les ISGs (« IFN-stimulated genes ») (305,306). Les NLRs reconnaissent des composants bactériens notamment le peptidoglycane et ne seront donc pas détaillés. Dans la cellule immunitaire, les TLRs et les RLRs représentent les principaux systèmes de détection des virus et seront détaillés dans les paragraphes suivants.

6.1.1. Les récepteurs de type Toll (TLRs)

Les TLRs, exprimés principalement par les cellules immunitaires telles que les monocytes, les macrophages et les cellules dendritiques (181), jouent un rôle essentiel dans leur activation. Les TLRs sont composés d'un domaine extracellulaire riche en leucine (LRRs), responsable de la reconnaissance des PAMPs, d'un domaine transmembranaire et d'un domaine cytoplasmique Toll/IL-1 (TIR) permettant la transduction du signal. Il existe différents types de TLRs, exprimés à la surface des cellules (TLR1, 2, 4, 6) ou dans les endosomes/lysosomes (TLR3, 7, 8, 9). La distribution cellulaire des TLRs est fonction du type de ligand qu'ils reconnaissent. TLR2 et TLR4, présents à la surface des cellules sont capables de détecter les protéines virales. Les TLRs 3, 8 et 9 présents à la surface des endosomes participent à la réponse antivirale en reconnaissant les ARNs viraux. En effet, TLR3 reconnaît l'ARN double brin viral alors que TLR8 reconnaît les ARNs viraux simples brins et TLR9 l'ARN double brin. L'activation des récepteurs TLRs déclenche la cascade de signalisation conduisant à la production d'interférons et de cytokines pro-inflammatoires (307).

6.1.2. Les récepteurs cytoplasmiques de type RIG-I (RLRs)

Les RLRs sont des récepteurs intracellulaires impliqués dans la reconnaissance des acides nucléiques issus d'une infection virale, le plus souvent par des virus à génome à ARN comme le virus de la rage, le virus Ebola ou encore le virus de la rougeole. La famille des RLRs comprend trois membres: RIG-I ("retinoic acid-inducible gene I"), MDA5 ("melanoma differentiation-associated gene 5") et LGP2 ("laboratory of genetics and physiology 2") (307,308). Ces trois RLRs sont composés de trois (MDA5, RIGI) ou deux (LGP2) domaines présentant des fonctions complémentaires (**Fig. 22A**). RIG-I et MDA5 sont des protéines solubles fonctionnelles dans le cytoplasme, présentant une organisation similaire. En effet, elles sont organisées en plusieurs domaines : deux domaines CARD (« caspase activation and recruitment domain ») en N-terminal, jouant un rôle dans le recrutement et l'activation de caspases, suivis d'un domaine hélicase à ARN présentant un site caractéristique DEX(D/H) permettant le recrutement des ARN et enfin, un domaine CTD (**Fig. 22A**).

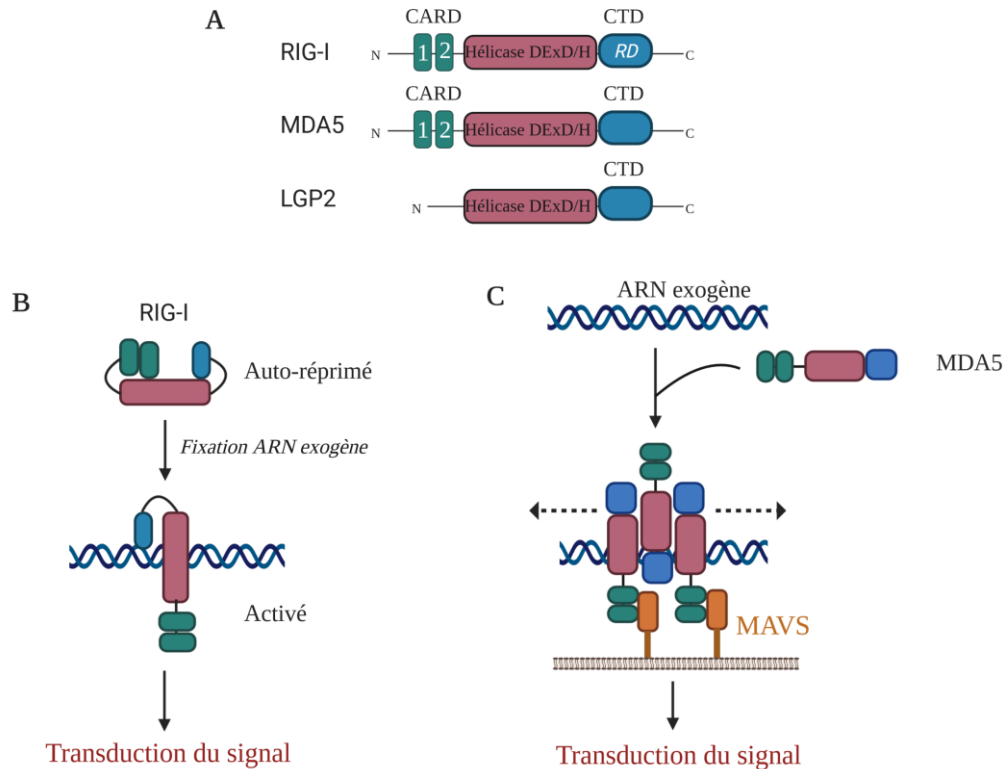


Figure 22 : Les récepteurs cytoplasmiques de type RIG-I (RLR)

A) Les famille des RLRs est composée de 3 types de récepteurs : RIG-I, MDA5 et LGP2. Alors que RIG-I et MDA5 présentent une organisation similaire, deux domaine CARD, un domaine hélicase central et un domaine C-terminal (CTD), LGP2 est dépourvu de domaine CARD. Cependant RIG-I et MDA5 présente un mécanisme d'activation différent. **B)** RIG-I dans sa forme inactive est autoréprimé par son domaine CTD. La reconnaissance d'un ARN exogène libère les domaines CARDS permettant ainsi le recrutement de partenaires protéiques nécessaire à la transduction du signal et la production d'IFN. **C)** MDA5 doit s'oligomériser pour être activé. La fixation d'ARN exogène permet la dimérisation puis l'oligomérisation de MDA5 autour de la molécule d'ARN. Les domaines CARD de MDA5 se fixe au domaine MAVS à la surface des mitochondries permet la transduction du signal.

Inspirée de Goubau D. *et al*, 2013

6.1.2.1. RIG-I

Bien que RIG-I et MDA5 présentent une organisation semblable, leur mode de fonctionnement présente des divergences. En effet, le domaine CTD de RIG-I comprend un domaine répresseur (RD) qui permet son auto-régulation (309). A l'état auto-réprimé, le domaine RD séquestre les domaines CARDS. Quand RIG-I détecte un ARN exogène, le domaine RD libère les domaines CARDS permettant le recrutement de partenaires protéiques nécessaires à la transduction du signal et la production d'IFN conduisant à l'inhibition de la réplication virale (306,310) (**Fig. 22B**). RIG-I a été initialement décrit reconnaissant l'extrémité triphosphate ou diphosphate des ARNs double brin (311) mais peut également reconnaître l'extrémité triphosphate des ARNs simple brin (312–315). Ainsi, les ARNs

portant une structure coiffe incomplète (ie : Cap-0). Le domaine hélicase est responsable de la reconnaissance de l'ARN et les domaines CARD activent les voies de signalisation en aval (316). Bien que RIG-I soit au contact d'ARNs cellulaires, ces derniers n'entraînent pas l'activation du senseur. En effet, les ARNs cellulaires présentent à leur extrémité 5' une structure coiffe méthylée en position 2'O du premier nucléotide (317). Plusieurs travaux récents ont permis de mettre en évidence l'impact de cette modification post-transcriptionnelle dans la discrimination du « soi » et du « non soi » par le senseur RIG-I. Ces études montrent notamment que les ARNs coiffés non méthylés en 2'O sont reconnus comme exogènes et entraînent l'activation de la réponse immunitaire médiée par RIG-I (294).

6.1.2.2. MDA5

Le senseur MDA5 est, quant à lui, impliqué dans la reconnaissance de long ARNdb viraux (318,319), cependant les mécanismes de reconnaissance de l'ARN sont différents de RIG-I. En effet, MDA5 n'est pas sensible aux extrémités 5' triphosphate des ARNs (320). Une étude structurale de l'interaction MDA5/ARN a permis de mieux comprendre son mécanisme d'action (321). En l'absence de ligand, MDA5 est retrouvé sous forme monomérique adoptant une conformation dite « ouverte ». Son activation nécessite son oligomérisation (322). En présence de son ligand ARN, MDA5 forme un dimère « tête bêche » via ses domaines hélicase et CTD puis s'oligomérisent en structure hélicoïdale autour de la molécule d'ARNdb (323). Cette organisation antiparallèle permet d'exposer les domaines CARDS permettant leur interaction avec les domaines CARDS des senseurs MAVS (« mitochondrial antiviral signaling ») à la surface des mitochondries. Cette interaction MDA5/MAVS permet la transduction du signal et la production d'IFNs (324) (**Fig. 22C**). MDA5 joue également un rôle dans la discrimination du « soi » et du « non-soi » (295). En effet, une étude a récemment montré que le génome du VIH était 2'O-méthylé sur des résidus spécifiques, et que ces modifications épitranscriptomiques permettaient au virus d'échapper à la reconnaissance par MDA5 (300).

6.1.2.3. LGP2

A la différence de MDA5 et RIG-I, le senseur LGP2 ne présente pas de domaines CARD (**Fig. 22**) (304). Il a été proposé que LGP2 modulait la réponse antivirale en jouant le rôle de régulateur positif et négatif (325,326). En effet, LGP2 régule de façon positive la réponse antivirale en favorisant l'interaction MDA5/ARN et son oligomérisation (327). Cependant, une concentration trop importante de LGP2 dans la cellule conduit à un effet antagoniste

(328,329). Une étude structure/fonction de LGP2 a montré que son domaine hélicase pouvait interagir avec des ARNs exogènes tout comme MDA5. De plus, son domaine CTD est capable de discriminer, tout comme RIG-I, les différentes extrémités de l'ARN. LGP2 pourrait compenser le manque de RIG-I, en interagissant avec les extrémités 5' des ARNs et entrainer le recrutement de MDA5 activant ainsi la cascade de signalisation et la production d'IFN. De façon surprenante, la régulation positive de MDA5 par LGP2 intervient à un stade précoce de l'infection favorisant la détection des ARNs viraux. Suite à l'augmentation de son expression et à la production d'IFN, LGP2 fonctionne comme un régulateur négatif, jouant un rôle dans le rétrocontrôle négatif de la réponse immunitaire médiée par MDA5 (330). Ce dernier récepteur cytosolique semble donc jouer un rôle dans la régulation de la réponse immunitaire en modulant la signalisation médiée par MDA5 à un stade précoce de l'infection (331).

6.2. La cascade de signalisation vers la synthèse d'interférons

La détection d'acides nucléiques exogènes par les récepteurs du système immunitaire tels que les RLRs et TLRs déclenche une cascade de signalisation conduisant à la production d'IFNs. L'activation des récepteurs cytosoliques RIG-I et MDA5 entraîne l'activation d'une cascade de signalisation déclenchée par une interaction CARD/CARD (**Fig. 23**).

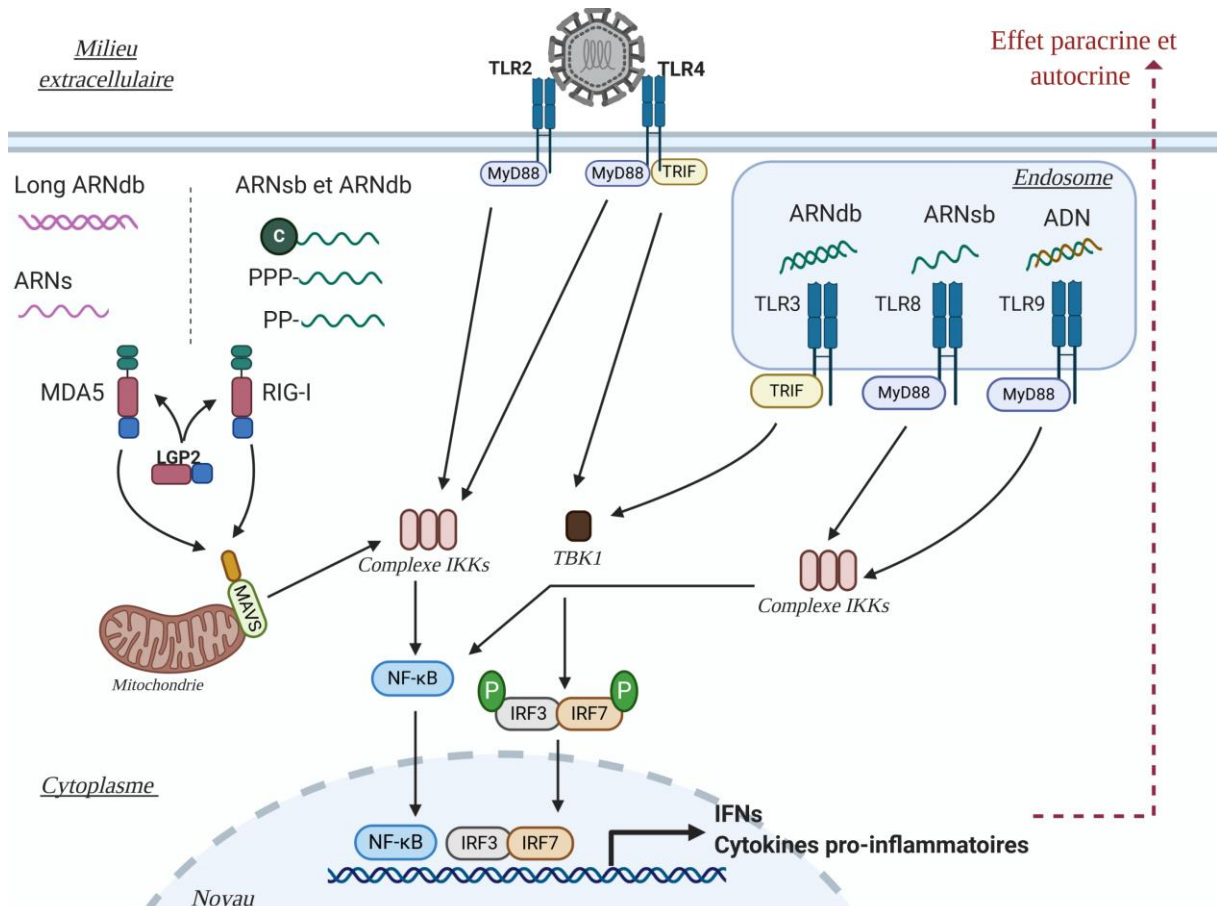


Figure 23 - Réponse immunitaire innée antivirale

L'activation des récepteurs de l'immunité innée tels que les « Toll-like receptors » (TLRs) et les « RIG-like receptors » (RLRs) entraîne l'activation d'une cascade de signalisation conduisant à la production de cytokines proinflammatoires et d'interférons (IFNs) participant à la réponse antivirale. L'activation des récepteurs cytosoliques RIG-I et MDA5 via la reconnaissance d'ARNs exogènes. Les ARNs reconnus par MDA5 sont indiqués en violet et les ARNs reconnus par RIG-I en vert. RIG-I est sensible aux extrémités 5' des ARNs (« ppp » pour triphosphate, « pp » pour diphosphate et « C » correspond à la structure coiffe). L'activation des senseurs entraîne leur interaction avec les domaines CARD (en orange) de MAVS à la surface des mitochondries. Cette interaction entraîne le recrutement du complexe IKKs (α, β, γ) qui entraîne l'activation du facteur de transcription NF- κ B qui transloque dans le noyau et stimulant la production de cytokines pro-inflammatoires. L'activation des TLRs à la surface de la cellule (TLR2 et TLR4) et ceux présent au niveau du compartiment endosomal (TLR 3, 8 et 9) activent la même voie de signalisation précédemment décrite exception faite des TLR4 et TLR3. TLR4 est capable de recruter les effecteurs cytosoliques MyD88 et TRIF. MyD88 active la voie de signalisation qui conduit à la translocation du facteur de transcription NF- κ B, précédemment décrite. L'adaptateur cytosolique TRIF (« TIR-domain-containing adapter-inducing interferon- β » entrainant l'activation de la kinase TBK-1 (« tankbinding kinase 1). TBK-1 phosphoryle IRF3 et IRF7 qui forment un dimère qui qui migrent vers le noyau pour interagir avec les promoteurs des gènes permettant la production d'IFNs. L'activation de TLR3 entraîne uniquement l'activation de la voie médiée par TRIF.

En effet, les domaines CARDS de RIG-I et MDA5 se lient au domaine CARD de la protéine MAVS à la surface externe des mitochondries (332–334). L'activation de MAVS permet le recrutement de différents facteurs notamment le facteur TRAF (« TNF receptor-associated factors ») entrainant l'activation de protéines kinases qui jouent un rôle dans

l'activation des gènes codant pour la production des IFNs (335–337). La kinase TBK-1 (« tankbinding kinase 1 ») est responsable de la phosphorylation des facteurs de transcription IRF-3 et IRF-7. La phosphorylation de ces facteurs entraîne la formation d'homodimères (IRF3/3 et IRF7/7) et/ou d'hétérodimères (IRF3/7) qui migrent vers le noyau pour interagir avec les promoteurs des gènes permettant la production d'IFNs. Un autre groupe d'IKKs (« inhibitor of NF- κ B (I κ B) kinase ϵ »), composé d'IKK- α , IKK- β et IKK- γ , permet la phosphorylation d'I κ B conduisant à sa dégradation et permettant la libération de NF- κ B et sa translocation vers le noyau stimulant la production d'IFNs et de cytokines pro-inflammatoires (336,338).

Tout comme pour les RLRs, l'activation des TLRs conduit à la production d'IFNs et de cytokines pro-inflammatoires. La transduction du signal fait intervenir des facteurs communs aux RLRs. La détection d'ARNs viraux entraîne l'activation des TLRs, qui vont s'homodimériser permettant le recrutement d'adaptateurs cytosoliques contenant un domaine TIR tel que MyD88 (« myeloid differentiation primary response 88 ») et TRIF (« TIR-domain-containing adapter-inducing interferon- β »). TLR-3 recrute TRIF, alors que TLR-2, TLR-8 et TLR-9 recrute MyD88. TLR-4 peut, quant à lui, interagir avec MyD88 et TRIF. La voie de signalisation MyD88 passe par l'activation du complexe de kinases IKKs (IKK- α , - β et - γ) et la libération du facteur de transcription NF- κ B. L'activation de l'adaptateur TRIF, active les kinases TBK-1/IKK ϵ et les facteurs de transcription IRF-3 et IRF-7 (339,340) (**Fig. 23**).

6.3. La réponse à l'IFN

La production d'IFNs, à la suite de l'activation des PRRs, représente la première ligne de défense face à une infection virale. En effet, ils sont sécrétés dans le milieu extracellulaire et présentent un mode d'action autocrine, mais également paracrine en signalant aux cellules non encore infectées la présence de l'agent pathogène (**Fig. 24**).

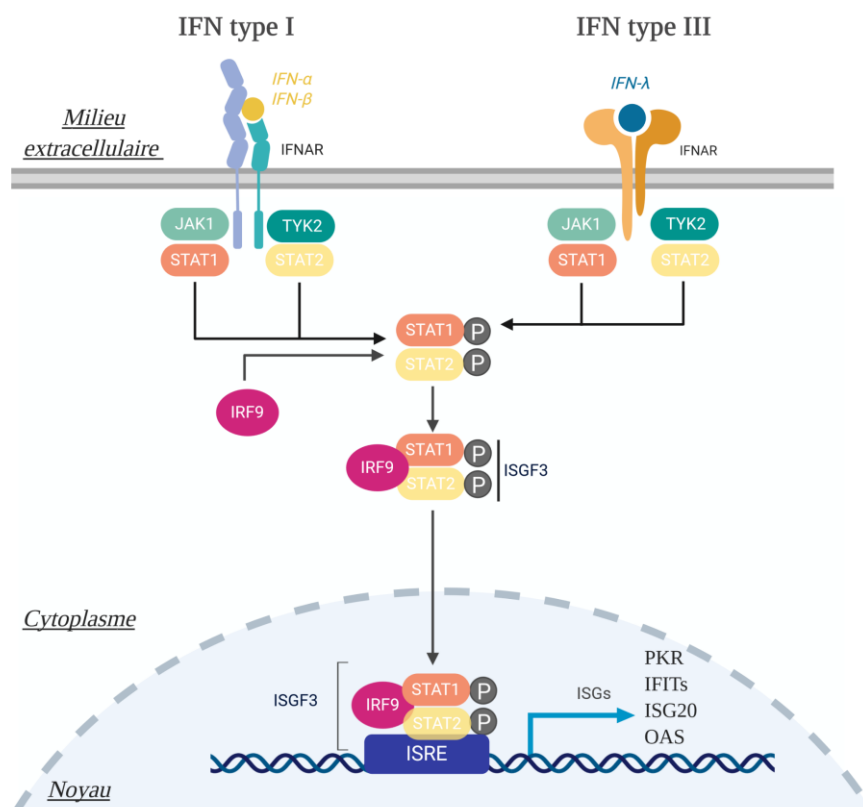


Figure 24 - La réponse aux IFNs

Les interférons (IFNs) produits en réponse à l'activation des récepteurs de l'immunité innée participent à la réponse antivirale. La fixation des IFNs à leurs récepteurs cellulaires permet le recrutement de protéines kinases (JAK1 et TYK2) qui entraîne la phosphorylation de STAT1 et STAT2. STAT1 et STAT2 phosphorylés se dimérisent et s'associent au facteur IRF9 conduisant à la formation du complexe ISGF3. Le complexe ISGF3 transloque dans le noyau et se fixe sur une séquence ISRE stimulant la transcription des gènes stimulés par l'IFN (ISGs) et la production de protéines antivirales telles que la protéine kinase R, ISG20, IFIT et la 2'-5' oligoadénylate synthétase (OAS).

Il existe trois types d'IFNs : les IFNs de type I (IFN α/β) et les IFNs de type III (IFN λ) qui jouent un rôle important dans la réponse antivirale et les IFNs de type II (IFN γ) qui interviennent dans la régulation de la réponse immunitaire, marquant la frontière entre le système immunitaire inné et adaptatif (340). La fixation des IFN α/β et IFN λ à leurs récepteurs permet le recrutement de tyrosines kinases JAK1 (« *Janus kinases* ») et TYK2 (« *Tyrosine kinase 2* ») entraînant la phosphorylation des facteurs de transcriptions STAT1 (« *Signal transducer and activator of transcription 1* ») et STAT2 (« *Signal transducer and activator of transcription 2* »). Ces facteurs ainsi phosphorylés se dimérisent et s'associent à un troisième facteur (IRF9) formant le complexe ISGF3 (« *interferon stimulated gene factor 3* »). Le complexe ISGF3 transloque dans le noyau et se fixe à une séquence ISRE (« *interferon-sensitive response element* »), stimulant la transcription de gènes stimulés par l'IFN, les ISGs (« *interferon stimulated genes* ») (340,341). Parmi les ISGs, la protéine kinase

R dépendante de l'ARN (PKR), la 2'-5' oligoadénylate synthétase (OAS), la RNase L et la protéine de résistance aux myxovirus sont essentielles pour permettre la clairance virale. De plus, d'autres protéines induites par les IFNs comme les IFITs, ISG15 et ISG20, participent à la réponse antivirale en ciblant diverses étapes du cycle viral. ISG20 est une exonucléase 3'-5' qui lie l'ARN exogène et permet sa dégradation (342,343). IFIT1 reconnaît l'extrémité 5' triphosphate des ARNs viraux et les coiffes incomplètes (non méthylées en 2'O) inhibant ainsi leur traduction en protéines (296,344) (**Fig. 24**).

En conclusion, la détection de l'infection virale et le déclenchement de la réponse aux IFNs permettent de limiter la propagation virale. Les virus ont développé divers mécanismes afin d'échapper à cette réponse antivirale. Le déclenchement de la réponse IFN faisant suite à une infection virale et son contrôle par le virus illustrent la relation établie entre le virus et son hôte. L'un des mécanismes est l'ajout d'une structure coiffe à l'extrémité 5' des transcrits viraux, identique à celle présente au niveau de l'extrémité 5' des ARNs cellulaires. Une étude récente a également mis en évidence le rôle de modifications épitranscriptomiques, telles que les méthylations internes du génome, dans l'échappement à la réponse immunitaire innée (300).

Ce projet de thèse s'est intéressé au virus de l'ordre des *Mononegavirales*, et aux mécanismes de capping qui jouent un rôle dans les subversions des défenses de l'hôte.

7. La subversion des défenses de l'hôte : vue d'ensemble

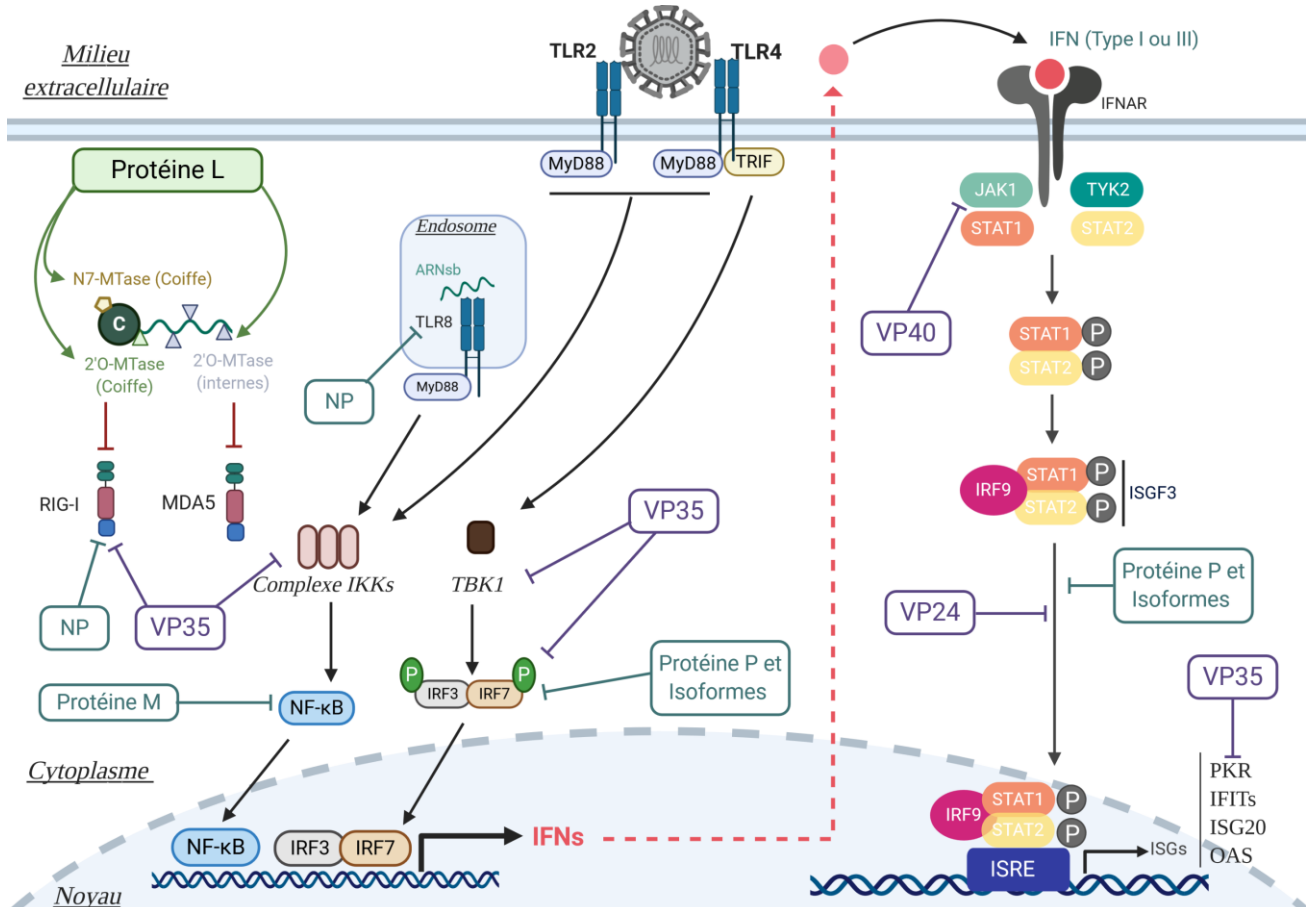


Figure 25 – Subversion des défenses antivirales de l'hôte

Les protéines virales participent à l'échappement de la réponse immunitaire à différents niveaux. Les protéines virales de RABV sont indiquées dans les cadres verts, les protéines d'EBOV participant à l'inhibition de la réponse immunitaire sont indiquées dans les cadres violets. La protéine L (indiquée en vert) joue un rôle dans la synthèse de l'ARN et la maturation des transcrits via l'ajout d'une structure coiffée en 5' (© vert) et d'une queue poly(A) (non représentée). La protéine L porte également les fonctions enzymatiques 2'O et N7 MTases. Les activités de la protéine L sont conservées au sein des *Mononegavirales*, exception faite de l'activité 2'O MTase interne rapportée pour la protéine L de EBOV (B. Martin *et al.*, 2018). L'ajout de la structure coiffée et des méthylation limite la détection par les senseurs de l'immunité innée tels que RIG-I et MDA5. La nucléoprotéine (N) de RABV et VP35 de EBOV participent à l'inhibition de RIG-I. La protéine M (RABV), P (RABV) et VP35 (EBOV) limitent la production d'IFNs et de cytokines pro-inflammatoires. VP40 (EBOV), VP24 (EBOV) et la protéine P (RABV) sont impliquées dans l'inhibition de la voie de signalisation activée par la production des IFNs. Enfin, VP35 (EBOV) intervient dans l'inhibition des protéines antivirales induites par les IFNs.

Au cours de leur évolution, les virus ont développé des mécanismes leur permettant d'échapper à la réponse immunitaire de l'hôte assurant ainsi leur multiplication. Le système immunitaire répond à une infection virale via la reconnaissance de PAMPs considérés comme des signatures du « non-soi ». De nombreux senseurs de l'immunité innée tels que les RLRs (RIG-I et MDA5) ont été montrés comme sensibles aux extrémités 5' des transcrits,

notamment à l'absence de structure coiffe ou de structures coiffes non 2'O méthylées. Le rôle de RIG-I dans la détection des ARNs viraux présentant un défaut de 2'O MTase a été mis en évidence chez les flavivirus. L'utilisation de système réplicon du virus de la fièvre jaune muté au niveau du site catalytique de la 2'O MTase a révélé que la réplication du virus est altérée dans les cellules compétentes pour l'IFN mais pas dans les cellules VERO, connues pour leur incapacité à produire de l'IFN (345). D'autres part, d'autres études menées en cellules infectées par SARS-CoV ont mis en évidence que l'introduction de mutations dans le site catalytique de la 2'O MTase conduit à des virus dont la capacité répliquative est altérée. Cependant, la virulence du virus muté dans des souris KO MDA5 ou IFIT, synthétisées en réponse à l'IFN, est similaire au virus sauvage (346). De plus, il a été montré que la 2'O MTase de résidus internes à la séquence d'ARN du HIV limite la détection par MDA5 (300). Ces études mettent en évidence le rôle des 2'O MTases comme marqueur du « soi ».

De nombreux virus se répliquant dans le cytoplasme, comme RABV et EBOV, ont évolué pour coder des 2'O MTases permettant de masquer leurs ARNs de la détection par le système immunitaire. La protéine L des *Mononegavirales* participe à l'activité de synthèse et de maturation des transcrits. Par conséquent, cibler cette protéine virale et notamment son activité 2'O MTase permettrait de booster la reconnaissance des ARNs viraux par le système immunitaire et d'autre part limiter la synthèse des protéines virales. De plus, le mécanisme de mise en place de la structure coiffe chez les virus implique une voie de synthèse non conventionnelle dont les mécanismes enzymatiques engagés diffèrent de ceux impliqués dans la synthèse de la structure coiffe chez les ARNm cellulaires (347,348). L'ensemble de ces observations font des MTases virales d'excellents candidats pour le développement de nouvelles stratégies antivirales.

En conclusion, la caractérisation des activités enzymatiques autour de la synthèse de la structure coiffe chez les virus constitue un enjeu important dans la compréhension des mécanismes répliquatifs viraux, en particulier leur échappement à l'immunité constituant une base intéressante pour le développement d'antiviraux.

OBJECTIFS DU PROJET DE THÈSE

L'objectif du projet de thèse présenté dans ce manuscrit est l'étude des domaines C-terminaux de la protéine L des *Mononegavirales*, un acteur central dans la transcription et la réplication virale. En effet, en plus de son activité ARN polymérase ARN-dépendante, la protéine L catalyse les étapes de maturation des transcrits en ajoutant une structure coiffe en 5' et une queue poly(A) en 3'. Elle contient 5 domaines dont l'ordre et la fonction sont conservés au sein de l'ordre des *Mononegavirales* : le domaine RdRp suivi du domaine Cap qui porte l'activité ribonucléotidyltransférase (PRNTase) nécessaire à la synthèse de la structure coiffe, le domaine connecteur (CD) de fonction inconnue, le domaine méthyltransférase (MTase) et le domaine C-terminal (CTD) au rôle non documenté également. Le domaine MTase catalyse la méthylation en 2'O du premier nucléotide de la séquence d'ARN et en position N7 de la structure coiffe. Ces méthylations jouent un rôle déterminant dans l'initiation de la traduction des protéines virales. De plus, elles limitent la détection des ARNm viraux par les senseurs de l'immunité innée. L'étude de ces activités enzymatiques doit permettre de comprendre le mécanisme de synthèse des structures coiffes des ARNs des *Mononegavirales* et d'apporter une preuve de concept que les MTases sont des cibles antivirales attractives. Ce projet d'étude prend pour modèle la protéine L du virus de la rage et du virus Ebola.

L'étude de la protéine L du virus de la rage s'imbrique dans un programme de recherche collaboratif basé sur la complémentarité de plusieurs domaines d'expertise (biochimique, structural et fonctionnel) visant à étudier le mécanisme de coiffage des *Mononegavirales*. Pour répondre à cet objectif, mon travail a consisté à :

- Identifier par une étude bio-informatique des bornes de constructions génétiques ciblant les domaines C-terminaux de la protéine L impliqués dans l'activité MTase ;
- Produire et purifier les protéines recombinantes ;
- Caractériser biochimiquement l'activité MTase de la protéine L du virus de la rage ;
- Étudier le rôle des domaines CD et CTD dans l'activité MTase ;
- Identifier des résidus clefs participant à l'activité MTase ;
- Mener une campagne de cristallogénèse ;
- Approfondir la relation structure/fonction du domaine MTase.

La première partie des résultats s'attachera à présenter les développements techniques mis en place pour l'expression de la protéine L du virus rabique afin de répondre aux objectifs. Cependant, en raison des difficultés rencontrées dans l'étape de production des

protéines recombinants ciblant les domaines C-terminaux de la protéine L du virus de la rage, et afin de mener à bien le projet, les objectifs ont été transférés sur un autre modèle d'étude : la protéine L du virus Ebola. Le choix du modèle reste tout à fait pertinent pour plusieurs raisons :

- Appartenant à l'ordre des *Mononegavirales*, ce virus, hautement pathogène pour l'Homme, représente un important problème de santé publique, nécessitant le développement de nouvelles stratégies antivirales ;
- Le mécanisme de synthèse de la structure coiffe chez les filovirus est encore peu documenté ;
- La région C-terminale de la protéine L de *Sudan ebolavirus* a été produite sous forme recombinante et purifiée au laboratoire ;
- Le domaine MTase du virus Ebola porte une activité originale : elle induit des méthylations internes ;
- Il n'y a pas de données structurales sur la protéine L des filovirus ;
- L'identification et l'exploitation de nouvelles cibles pour lutter contre le virus Ebola pourraient servir de modèle pour l'étude et la production d'antiviraux contre les virus de l'ordre des *Mononegavirales*.

RÉSULTATS

1. Etude des domaines C-terminaux de la protéine L du virus de la rage

1.1. Etude bio-informatique

Le but de cette étude étant d'exprimer la région C-terminale de la protéine L du virus de la rage. Nous avons dans un premier temps tenté d'identifier les bornes des domaines constitutifs de la protéine L de RABV par des analyses bio-informatiques. Notre objectif premier étant de produire le domaine de L portant l'activité MTase, les séquences en acides aminés des domaines C-terminaux des protéines L de RABV (ABZ81226.1) et de VSV, un rhabdovirus dont la structure a été résolue (PDB : 5A22) (253), ont été comparées. Dans un premier temps, un alignement a été généré par le serveur Clustal Omega (349) et retravaillé manuellement via le logiciel SeaView (350) et traité avec ESPript (351) (**Annexe 1 : alignement**). Les séquences C-terminales de la protéine L de RABV et de VSV s'alignent particulièrement bien dans la région MTase et les résidus de la tétrade catalytique K-D-K-E, caractéristique des 2'O MTases, ainsi que le motif conservé de fixation du SAM (GxGxG) sont facilement identifiés dans l'alignement. Les domaines C-terminaux de VSV étant bien définis dans les structures disponibles, nous avons défini les bornes des différentes constructions génétiques pour l'expression de protéines recombinantes ciblant les domaines C-terminaux de RABV. Les structures secondaires ont été insérées au-dessus de l'alignement grâce aux données structurales de VSV. Le domaine CD en amont du domaine MTase sert de connecteur entre le domaine PRNTase en N-terminal et le domaine MTase en C-terminal. La région séparant le domaine CD du domaine MTase ne semble pas structurée suggérant que cette région pourrait participer à la flexibilité entre les domaines N-terminaux (RdRp et PRNTase) et les domaines C-terminaux de la protéine L. La recherche de constructions permettant d'exprimer des protéines solubles nous a ainsi conduit à réaliser plusieurs constructions dont les bornes ont été définies dans les domaines désordonnés identifiées par le logiciel MeDor (352). Cette étude permet de délimiter des domaines désordonnés qui pourraient entraver la solubilité de la protéine. Les bornes de différentes constructions utilisées pour l'expression de protéines recombinantes ont été placées dans ces régions faiblement conservées. Les différentes bornes N- et C-terminales définies sont présentées dans la figure **Fig. 26**.

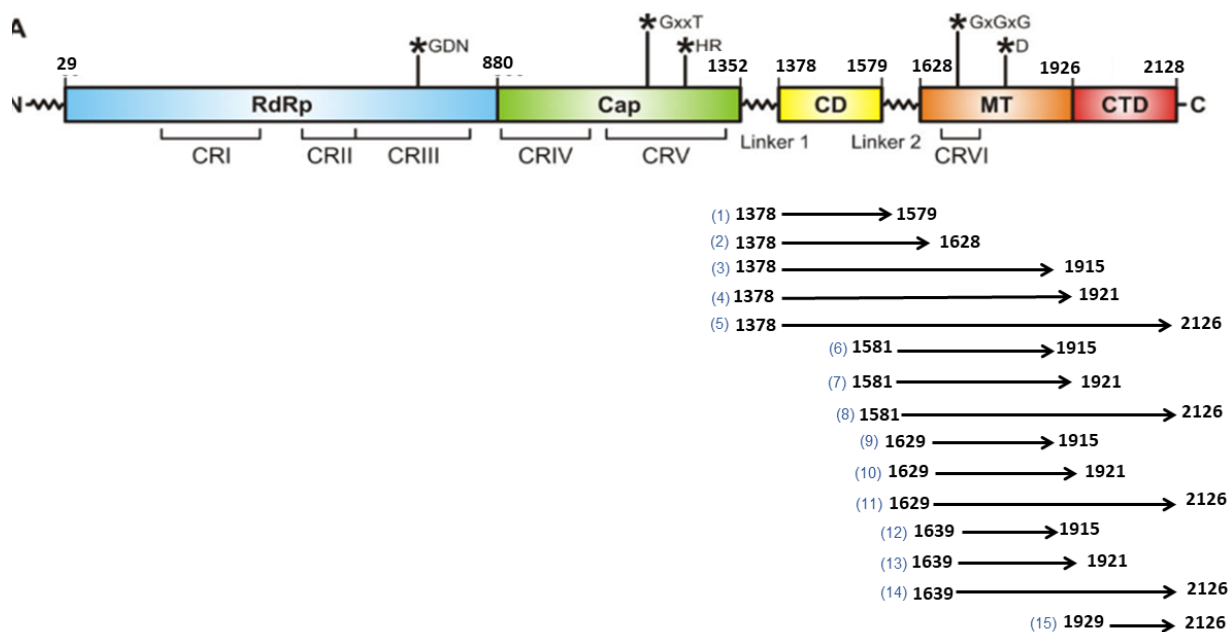


Figure 26 : Représentation des différentes constructions pour l’expression de protéines recombinantes.

Les études bio-informatiques des domaines C-terminaux de la protéine L de RABV ont permis de définir les bornes de différentes constructions génétiques, indiquées par des flèches noires. Afin de favoriser l’expression de domaines solubles, différentes bornes N- et C-terminales ont été définies pour un même domaine. Ces constructions permettront l’expression de protéines recombinantes exprimant les différents domaines (CD, MTase et CTD). Un total de 15 constructions différentes a été défini, numérotées de 1 à 15 en bleu.

Au total, 15 constructions ciblant les domaines C-terminaux de la protéine L (CD, MTase et CTD) ont été sélectionnées à la suite des études bio-informatiques. L’expression et la purification des différentes protéines recombinantes a pour but, d’une part de caractériser fonctionnellement l’activité MTase de la protéine L de RABV, et de s’interroger sur le rôle respectif des domaines CD et CTD dans l’activité méthyltransférase de la protéine L. Les protéines recombinantes ont été exprimées par ingénierie en système bactérien et nous avons ensuite tenté de les purifier et de les caractériser biochimiquement.

1.2. Stratégies d'expression en système bactérien

Quatre plasmides de type pDEST dont certains codent pour une protéine de fusion dont la séquence est insérée en amont de l'étiquette hexahistidine (6His) ont été utilisés pour le clonage des séquences d'intérêt : (i) pDEST 17 qui ne contient pas de domaine codant pour une protéine de fusion, (ii) pETG20A qui code pour la Thioredoxine (THX), (iii) pGST qui code pour la glutathion S-transférase (GST) et enfin (iv) pMBP qui code pour la Maltose Binding Protein (MBP). Les protéines de fusion ont été produites dans le but de favoriser l'expression et la solubilisation des protéines recombinantes (353). Le tag 6His permet de purifier la protéine d'intérêt par « *immobilized metal affinity chromatography* » (IMAC).

1.2.2. Criblage haut débit des conditions de culture

Il est important de noter que le facteur limitant de cette étude est l'obtention de protéines solubles et fonctionnellement actives. Afin de faire face à cette difficulté et mener à bien ce projet de recherche de nombreuses stratégies expérimentales ont été envisagées et testées.

Les conditions d'expression des protéines recombinantes sont des paramètres ajustables qui peuvent influencer l'expression et la solubilité des protéines. Parmi ces facteurs, la température, la souche bactérienne utilisée, la concentration en agent inducteur (IPTG), le milieu de culture sont tout autant de paramètres exploitables pour optimiser les conditions de production et/ou la solubilité de la protéine. Du fait du nombre important de constructions (60 constructions) et afin d'identifier les conditions d'expression favorables à la production de protéines solubles, des campagnes de criblage moyen/haut débit ont été menées selon le schéma de la **figure 27** via la plateforme haut débit d'expression et purification de protéine du laboratoire.

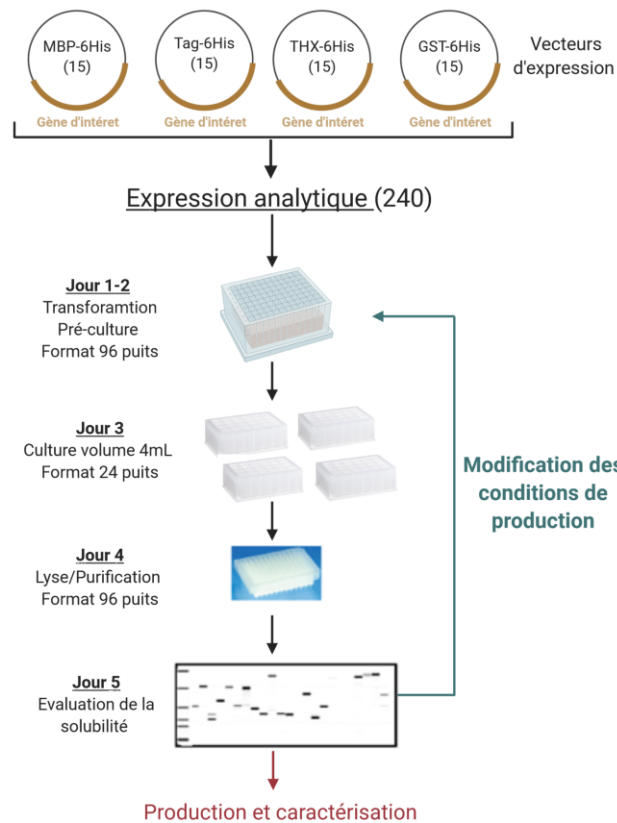


Figure 27 : Stratégie de criblage haut débit des conditions d'expression

Schéma du déroulement du criblage des conditions d'expression à partir des vecteurs d'expression jusqu'à l'analyse des résultats dans le but d'optimiser la solubilité des protéines. Les cultures bactériennes sont réalisées dans un format 24 puits permettant de gérer un grand nombre de culture, puis purifiées par chromatographie d'affinité automatisée sur un robot Tecan Freedom EVO200 (Switzerland) adapté au format 96 puits. L'analyse des résultats se fait par électrophorèse capillaire (Perkin Elmer GX II au format 96/384). Un total de 15 constructions génétiques clonées dans 4 vecteurs différents a ainsi été testé.

Le format haut débit nous a permis de tester un grand nombre de conditions de culture simultanément et de cribler des conditions d'expression de protéines solubles en parallèle. Dans un premier temps, des bactéries *E.coli* T7 sont transformées avec les plasmides d'expression codant pour les domaines présentés en figure 26, puis cultivées dans du milieu Terrific Broth (TB) supplémenté en ampicilline et chloramphénicol utilisés comme agents de sélection. Les cultures bactériennes sont cultivées à 37°C, correspondant à la température de croissance bactérienne optimale. Après induction l'isopropyl β -D-1-thiogalactopyranoside (IPTG), la température est abaissée afin de favoriser l'expression des protéines. Deux températures de culture sont testées : 25°C et 17°C. Après une incubation pendant la nuit, une étape de centrifugation permet de récupérer les bactéries qui sont re-suspendues dans un tampon de lyse (299). Les protéines sont purifiées par chromatographie d'affinité (NTa)

suivant les étapes décrites précédemment (354), automatisées sur un robot Tecan Freedom EVO200 (Switzerland) adapté au format 96 puits. Les résultats sont analysés par électrophorèse capillaire (Perkin Elmer GX II au format 96/384).

La campagne de criblage nous a permis de cribler 240 conditions cependant nous n'avons pas été en mesure d'identifier des conditions permettant l'expression de protéines solubles. Néanmoins, ces résultats ont permis d'adapter notre stratégie expérimentale pour une seconde campagne de criblage. En effet, cinq des quinze constructions, numérotées 5, 7, 8, 10 et 11 sur la figure 26, ciblant le domaine MTase seul ou en présence du domaine CD et/ou CTD ont été retenus pour poursuivre les essais d'optimisation des conditions d'expression de protéines solubles. Le choix a également été conditionnée afin d'adresser la question du rôle des domaines CD et CTD dans l'activité MTase de la protéine L du virus de la rage. Bien que le rendement de production reste faible, cette seconde campagne de criblage nous a permis de sélectionner une température de culture après induction (17°C) plus favorable à l'expression des protéines recombinante.

1.2.3. Stratégies d'expression en système bactérien

Ce chapitre a pour objectif de donner une vue de l'ensemble des tests engagés en s'attachant à la démarche scientifique. L'ensemble des facteurs qui ont été testés sont résumés dans la figure 28 et peuvent intervenir à différentes étapes du processus de production de protéine.

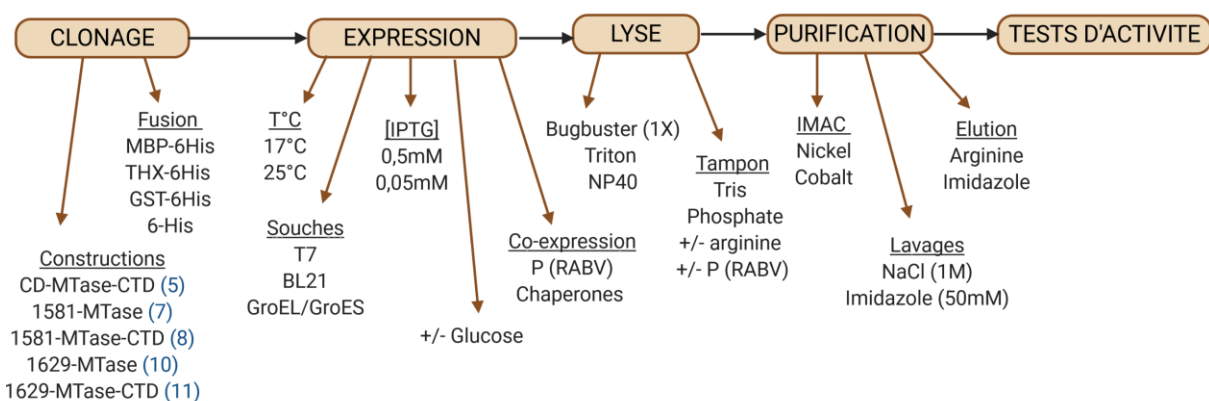


Figure 28 : Résumé des diverses conditions testées

Représentation du processus de production de protéines recombinantes. Les séquences codantes pour les domaines C-terminaux sont clonés dans divers vecteurs codant pour des protéines de fusion (MBP : Maltose Binding Protein, THX : thioredoxine, GST : glutathion S-transférase) afin de potentialiser l'expression de protéines solubles. Les conditions d'expression peuvent également affecter l'expression et la solubilité des protéines. Dans cette étude, plusieurs facteurs ont été modulés (T°C : températures, souches bactériennes, [IPTG] : concentration en agent inducteur, co-expression). Le mise au point d'un tampon de lyse optimal est une étape importante dans la purification des protéines. Plusieurs pistes ont été étudiées (détergents, composition du tampon). Les protéines ont été purifiées par chromatographie d'affinité sur colonnes de nickel ou cobalt et éluée avec un tampon hautement concentré en arginine ou imidazole.

Les conditions d'expression peuvent influencer le rendement de production et la solubilité des protéines. La concentration en agent inducteur est un paramètre exploitable, permettant de contrôler le niveau d'expression des gènes. Basé sur les travaux antérieurs menés au laboratoire ayant permis la caractérisation de l'activité de MTases virales (267,270,299), il a notamment été observé qu'une surexpression de la MTase virale du virus Ebola induisait son agrégation. Deux concentrations en IPTG ont donc été testées (0,05 mM et 0,5 mM) dans nos expériences. Après induction par l'IPTG (0,5 mM ou 0,05 mM), la température de culture est abaissée à 17°C. Après une nuit d'incubation, les bactéries sont récoltées par centrifugation et lysées selon le même protocole décrit précédemment (299). Les fractions solubles et insolubles sont séparées par centrifugation et analysées après séparation sur gel SDS-PAGE. L'analyse du gel a montré que l'induction à faible concentration en IPTG semblait être plus favorable à l'expression des MTases recombinantes. Cependant, bien que ces premiers criblages de conditions d'expression aient permis de définir une température et une concentration en agent inducteur, nous n'avons pas réussi à exprimer la MTase du virus de la Rage sous forme soluble en quantité suffisante. Suite à cette difficulté importante, nous avons concentré nos efforts sur les constructions contenant les séquences du domaine MTase-CTD et CD-MTase-CTD, clonées dans les vecteurs pMBP et pETG20A, présentant les résultats les plus prometteurs et essayer de trouver des conditions expérimentales permettant de solubiliser ces protéines.

Nous avons en particulier tenté d'optimiser les tampons de lyse utilisant des conditions expérimentales précédemment décrites pour la purification de la protéine L de VSV et RABV (230,252,258). De plus, nous avons testé le BugBuster (Merck Millipore), un réactif d'extraction de protéine, précédemment expérimenté sur la MTase du virus Ebola (269). Des essais de supplémentation des tampons de lyse en arginine ont également été conduits dans la mesure où l'arginine avait été utilisée pour la solubilisation du domaine MTase de hMPV (265). Cependant, les nombreux essais engagés n'ont pas conduit à la production de protéines solubles fonctionnellement actives.

Les protéines résistant aux différentes méthodes de stabilisation testées, des essais de renaturation (« refolding ») des agrégats de protéines insolubles présents dans les corps d'inclusion ont également été tentés. Dans ces expériences, après lyse de la culture bactérienne ayant exprimé le domaine CD-MTase-CTD ou Mtase-CTD de RABV (**Fig. 27 et 28**), les agrégats sont isolés par centrifugation. Les débris cellulaires sont ensuite resuspendus dans un tampon concentré en guanidine (6M), permettant la dénaturation des protéines. La

protéine est purifiée en condition dénaturante puis l'agent chaotropique est éliminé par dialyse progressive afin de favoriser la renaturation de la protéine. Cependant, cette technique n'a pas permis de récupérer des protéines solubles et actives.

Afin de favoriser la solubilisation des domaines de la L, nous avons également envisagé des essais de co-expression en présence de la protéine P qui est un cofacteur de la L connu pour favoriser l'expression soluble de la protéine L. De plus, il a été montré que le domaine C-terminal de la protéine L présentait un site d'interaction avec la P (233). Le gène de la protéine P de RABV est cloné dans le vecteur pMRBAD, dont l'expression dépend de l'induction par l'arabinose. Dans un premier temps, l'expression de la protéine P a été vérifiée expérimentalement (**Annexe 2A**). Les essais de co-expression ont été menés avec la construction du domaine MTase-CTD en fusion avec la THX. L'expression et la purification de ces domaines en co-expression avec la protéine P ont été analysées par western blot (**Figure 29A**). La protéine recombinante exprimée est majoritairement insoluble et nous n'avons pas détecté d'activité MTase.

D'autre part, des études précédentes avaient montré l'impact des protéines chaperonnes dans la solubilité des protéines recombinantes. Deux stratégies ont été abordées : la supplémentation en éthanol du milieu de culture permettant l'activation des chaperonnes (355) et l'utilisation de la souche bactérienne GroES/GroEL (356). Néanmoins, il est important de noter que dans le cas de protéine instable, l'induction de l'expression de chaperonnes peut entraîner un artéfact de solubilité de la protéine d'intérêt qui restera en interaction avec la chaperonne. L'étude par western blot a montré que la co-expression avec des chaperonnes favorisait la solubilité du domaine MTase-CTD de RABV, cependant la protéine purifiée se sont révélées inactives (**Figure 29B**).

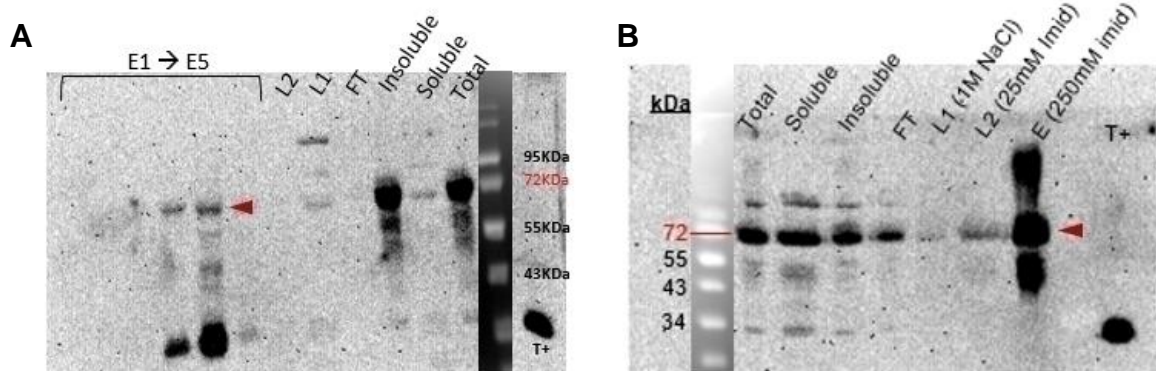


Figure 29 : Analyse expression et production du domaine MTase-CTD

Western blot d'un gel SDS-PAGE (12%) des différentes fractions de la purification par affinité Ni^{2+} -IMAC. La fraction « total » correspond au lysat bactérien. Les débris cellulaires (« insoluble ») sont éliminés par centrifugation. La fraction « soluble » contient la protéine d'intérêt, qui sera purifiée par chromatographie d'affinité. Le « flow-through » (FT) correspond à la fraction non retenue. L1 et L2 correspondent aux lavages 1 (1M NaCl, 10mM imidazole, 50mM Tris-HCL pH8) et 2 (300mM NaCl, 25mM imidazole, 50mM Tris-HCL pH8). La protéine est éluée dans un tampon contenant (300mM NaCl, 250mM imidazole, 50mM Tris-HCL pH8). Les fractions d'éluion sont notées E. (A) Analyse de la purification du domaine MTase-CTD en co-expression avec la P. La protéine d'intérêt est majoritairement insoluble. (B) Analyse de la purification du domaine MTase-CTD en co-expression avec les chaperonnes. Ces conditions d'expression favorisent la solubilité de la protéine.

En dépit des efforts engagés, l'expression des protéines en système bactérien n'a pas permis d'obtenir de protéines portant le domaine MTase du virus de la Rage solubles et actives. Dans la mesure où plusieurs études ont précédemment décrit l'expression la protéine L de VSV et RABV, ou du domaine MTase-CTD de hMPV, (265) en utilisant le système d'expression baculovirus (357) nous avons tenté de mettre en place ce système d'expression au laboratoire.

1.3. Stratégie d'expression en système baculovirus

Le système d'expression baculovirus est un outil d'expression génétique permettant la production de protéines recombinantes en cellules d'insecte. Brièvement, afin de produire les domaines C-terminaux de la L de RABV, des cellules d'insecte (Sf9) sont infectées par le baculovirus recombinant exprimant ces domaines. Quarante-huit heures post-infection (pi), les cellules sont récoltées par centrifugation puis re-suspendues dans un tampon de lyse. La protéine d'intérêt est purifiée par chromatographie d'affinité Ni^{2+} -IMAC. L'analyse du gel SDS-PAGE (12%) révèle une bande migrant à un poids moléculaire entre 66,2 et 116 kDa correspondant au domaine CD-MTase-CTD (85kDa) de RABV (Fig. 30) et l'activité enzymatique de la protéine purifiée est contrôlée.

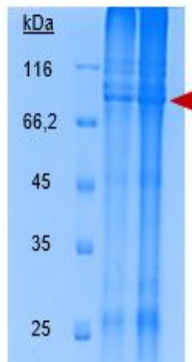


Figure 30 : Purification du domaine CD-MTase-CTD de RABV

Gel SDS-PAGE (12%) coloré au bleu de Coomassie de la fraction d'éluée de la purification par affinité Ni²⁺-IMAC du domaine CD-MTase-CTD de RABV exprimé en cellules d'insecte.

1.3.2. Caractérisation de l'activité MTase

Afin de caractériser l'activité enzymatique du domaine CD-MTase-CTD de RABV, des tests enzymatiques ont été conduits en utilisant différents ARNs synthétiques de 13 nucléotides mimant l'extrémité 5' de la séquence des transcrits de RABV (**Fig. 31A**).

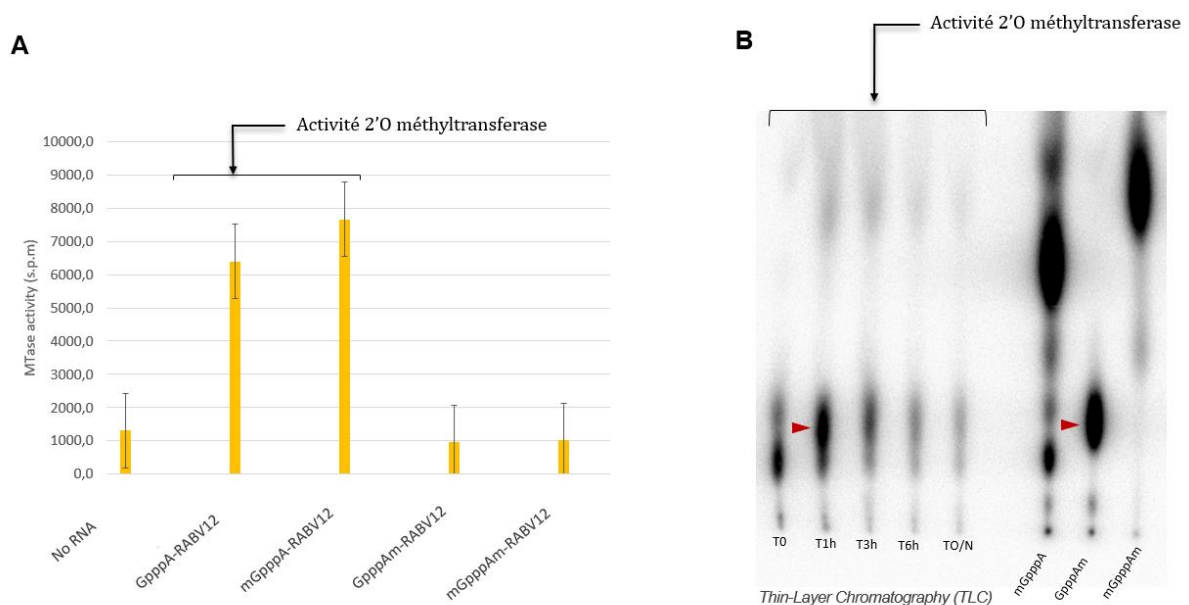


Figure 31 : Caractérisation de l'activité MTase du domaine CD-MTase-CTD

(A) Le domaine CD-MTase-CTD de RABV a été incubé avec des ARNs synthétiques mimant l'extrémité 5' des ARNm en présence du donneur de méthyl (SAM) marqué au tritium. Après un temps d'incubation (3h), les produits de réaction sont filtrés sur une membrane DEAE qui ne retient que les ARNs. La radioactivité détectée est fonction du transfert de groupement méthyl sur les ARNs et donc de l'activité MTase de la protéine étudiée. Les différents types de coiffe ont été utilisés (^mGpppA : N7 méthylée ; GpppA_m : 2'O méthylée ; ^mGpppA_m : 2'O et N7 méthylée) permettent de caractériser l'activité MTase du domaine CD-MTase-CTD (n=3). (B) Chromatographie sur couche mince des différentes coiffes générées au cours du temps (T0, 1h, 3h, 6h, sur la nuit (O/N)) par le domaine CD-MTase-CTD de RABV incubé avec l'ARN coiffé (GpppA₁₂-RABV) radiomarqué. Les pistes de droite correspondent aux contrôles, indiquant la migration des différents types de coiffe (^mGpppA : N7 méthylée ; GpppA_m : 2'O méthylée ; ^mGpppA_m : 2'O et N7 méthylées).

Les ARNs, synthétisés par l'équipe de Dr. Françoise Debart (IBMM), présentent une structure coiffée à leur extrémité 5', méthylée ou non en position N7 et 2'O du premier nucléotide. L'activité enzymatique du domaine CD-MTase-CTD a été évaluée en incubant les différents ARN substrats avec les protéines purifiées et le donneur de méthyle (SAM) radioactif (méthode *Filter Binding assay*, FBA). Les résultats sont présentés dans la figure 31A. Une activité MTase est détectée uniquement lorsque la protéine est incubée avec l'ARN coiffé non méthylé (GpppA-RABV₁₂) ou méthylé sur la position N7 (^mGpppA-RABV₁₂), suggérant que le domaine CD-MTase-CTD de RABV posséderait une activité 2'O-MTase. Cette technique sensible permet d'identifier des conditions expérimentales où la protéine est active, mais ne permet ni d'identifier définitivement les positions méthylées ni de décrire l'ordre réactionnel. Le test chromatographie sur couche mince (TLC) (**Fig. 31B**), permettant de caractériser le type de coiffe générée à la suite de l'incubation du domaine CD-MTase-CTD avec un ARN coiffé radiomarqué, confirme l'observation en suggérant que le domaine CD-MTase-CTD catalyse une activité 2'O MTase (**Fig. 31B**).

Cependant, de nombreuses difficultés d'expression et de purification du domaine CD-MTase-CTD de RABV ne nous ont pas permis d'optimiser suffisamment les conditions de purification de la protéine afin de confirmer ou d'infirmer ces résultats préliminaires.

A la suite de nombreux tests d'expression et de purification non concluants, nous avons débuté une collaboration avec la plateforme de production de protéines en système baculovirus de l'IGBMC à Strasbourg. En plus du domaine CD-MTase-CTD (1378-2126), la séquence du domaine MTase-CTD (1629-2126) a été clonée dans des vecteurs d'expression optimisés pour l'expression en cellules d'insecte. Les protéines ont été exprimées à l'IGBMC et les culots de cellules envoyés au laboratoire pour les tests de purification. Les premiers essais de purification basés sur le protocole de purification précédemment décrit ont été analysés sur gel SDS-PAGE (**Annexe 2B**). Les protéines sont récupérées grâce à une forte concentration en imidazole (250 mM) dans un tampon 50mM Tris-HCL [pH8], 300mM NaCl. Les fractions d'élution présentent des bandes migrant à un poids moléculaire entre 66,2 et 116 kDa, et entre 45 et 66,2 kDa correspondant respectivement au domaine CD-MTase-CTD et au domaine MTase-CTD. De façon surprenante, l'étude de l'activité MTase de ces domaines suggère que les deux protéines possèdent une activité N7-MTase, allant à l'encontre des résultats précédemment obtenus pour le domaine CD-MTase-CTD. Afin de confirmer que l'activité MTase était liée au domaine MTase de la protéine L de RABV, des mutations ont été introduites dans la tétrade catalytique K-D-K-E. Cependant, en raison de l'instabilité des

protéines et d'un très faible rendement de production, nous n'avons pas pu poursuivre cette étude.

Le principal facteur limitant de ce projet est l'obtention de protéines solubles et fonctionnellement actives. Afin de mener à bien ce projet de recherche, de nombreuses stratégies d'expression ont été testées, utilisant différents systèmes d'expression. En dépit des efforts fournis, l'expression des protéines en système bactérien n'a pas permis d'obtenir de protéines solubles actives. Bien qu'une seconde stratégie d'expression en système baculovirus présentait des résultats préliminaires prometteurs, nous avons rencontré d'importantes difficultés à purifier de façon reproductible les protéines. De plus, les rendements de production faibles ne permettaient pas de poursuivre l'étude vers la caractérisation biochimique robuste de l'activité MTase des domaines C-terminaux de la protéine L du virus de la rage.

1.4. Matériel et méthodes

1.4.1. Stratégie de clonage

La séquence codant pour les domaines C-terminaux de L RABV (CD-MTase-CTD) issu d'un gène synthétique optimisé pour l'expression en cellules d'insecte a été amplifiée par PCR utilisant divers jeux d'amorce permettant l'amplification des domaines définis par l'étude bio-informatique, qui sont clonés dans des vecteurs optimisés pour l'expression en système bactérien. La stratégie de clonage utilise le système de recombinaison Gateway (Invitrogen) (358). Cette technique permet d'obtenir un nombre important de vecteurs d'expression à partir d'un vecteur d'entrée. La première étape consiste à ajouter au gène d'intérêt les sites attB1 et attB2 par PCR d'assemblage (**Fig. 32**). Un site de clivage à la protéase TEV a également été intégré en amont du gène d'intérêt (**Annexe 3 : tableau amorces PCR**).

Les produits de PCR obtenus sont ensuite clonés dans le vecteur donneur pDONR201

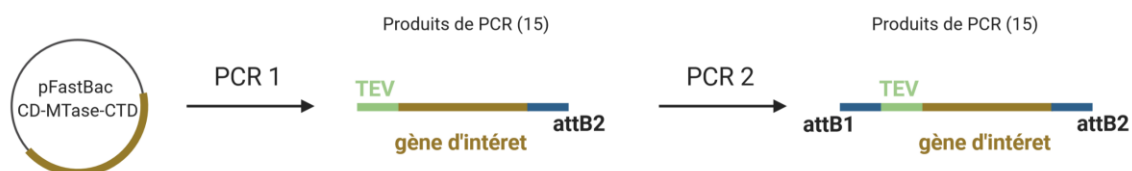


Figure 32 : PCR d'assemblage

La première PCR utilise comme matrice le plasmide pFastBac contenant la séquence des domaines CD-MTase-CTD. La combinaison de différentes amorces permet d'obtenir 15 produits de PCR contenant un site de coupure à la TEV (en vert), la séquence d'intérêt ciblant les domaines C-terminaux de la protéine L de RABV et la séquence de recombinaison attB2. Les produits issus de la PCR 1 serviront de matrice pour la PCR 2 qui produit les différents gènes flanqués des sites attB1 et attB2.

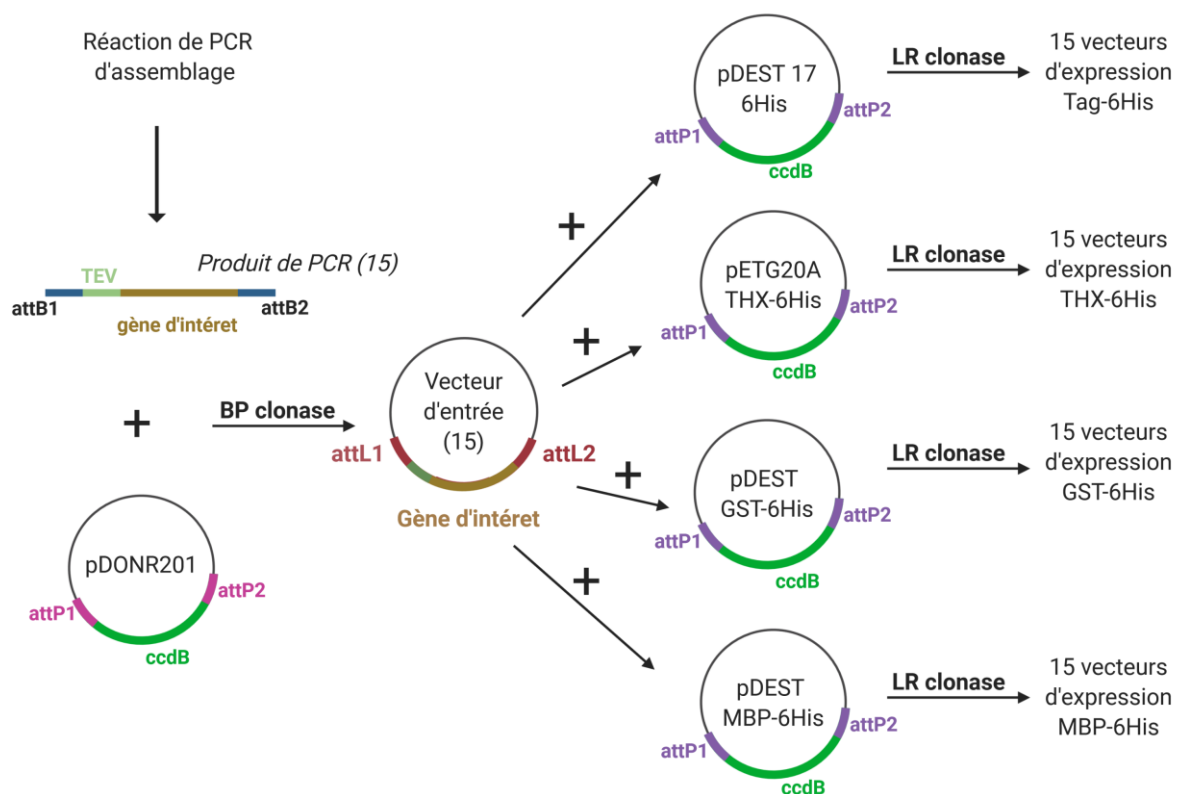


Figure 33 : Stratégie de clonage

L'ensemble des constructions définies ont été clonées dans 4 vecteurs d'expression différents, en utilisant la technique Gateway (Invitrogen), afin d'augmenter la probabilité d'exprimer des domaines solubles. Deux étapes de PCR ont été réalisées afin d'obtenir le produit de PCR qui sera inséré dans le plasmide donneur (pDONR201) formant le vecteur d'entrée à la suite d'une réaction de BP clonase. 15 vecteurs d'entrée correspondant au 15 constructions (**Fig. 25**) ont été produits. Une réaction de LR clonase permet le transfert du gène d'intérêt dans le vecteur d'expression (pDEST). Le clonage a permis d'obtenir 60 vecteurs d'expression différents.

(Invitrogen) par réaction de BP clonase (**Fig. 33**). Le vecteur d'entrée ainsi obtenu contient la séquence d'intérêt flanquée des sites de recombinaison homologue attL1 et attL2. Ce vecteur est utilisé pour l'obtention des vecteurs d'expression. Le gène d'intérêt sera transféré du vecteur d'entrée vers le vecteur d'expression (pDEST) en utilisant l'enzyme LR clonase (**Fig. 33**). Quatre vecteurs de type pDEST sont utilisés dans cette étude dont certains codent pour des protéines de fusion afin de favoriser la solubilité des protéines d'intérêt.

1.4.2. Expression et purification domaine CD-MTase-CTD en système baculovirus

Le système vecteur d'expression baculovirus est un outil d'expression génétique permettant la production de protéines recombinantes en cellules d'insecte. Le plasmide pFastBac contenant la séquence des domaines C-terminaux de la L de RABV (CD-MTase-CTD) précédée par la séquence du tag-6His a été synthétisé chez Twist. Des bactéries *E.coli* DH10Bac sont transformées avec ce plasmide permettant la transposition et la construction d'un bacmid recombinant. Le bacmid recombinant est transfecté dans des cellules d'insectes (sf9). Les stocks viraux récoltés à partir du surnageant des cellules transfectées peuvent ensuite être utilisés pour infecter des cultures de cellules sf9 fraîches pour l'expression de la protéine d'intérêt.

Pour produire les domaines C-terminaux de la L de RABV, des cellules d'insecte (sf9) sont infectées par le baculovirus recombinant exprimant ces domaines. Après 48h pi, les cellules sont récoltées par centrifugation puis re-suspendues dans un tampon de lyse : 50mM Tris-HCL [pH8], 300mM NaCl, 10mM imidazole, 10µg/mL DNase, 0,01% triton, 1mM de fluorure de phénylméthylsulfonyle (PMSF). Après 30 mins d'incubation, les débris cellulaires et les protéines insolubles sont éliminées par centrifugation avant la purification par chromatographie d'affinité sur une colonne His-Trap de 5mL (GE Healthcare). La protéine est éluée dans un tampon contenant 50mM Tris-HCL [pH8], 300mM NaCl, 250mM imidazole. L'imidazole est éliminé par une étape de dialyse contre 10mM Hepès [pH8], 300mM NaCl et 0,5mM de dithiothréitol (DTT), puis l'échantillon est concentré et dilué dans 50% de glycérol (v/v) et conservé à -20°C.

1.4.2. Test d'activité MTase

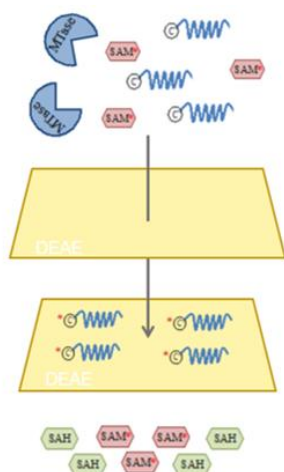


Figure 34 : Test d'activité MTase.

La MTase virale (en bleu) est incubée avec le SAM radioactif (en rouge) et l'ARN coiffé. Après un temps t d'incubation les produits de réaction sont transféré sur une membrane DEAE qui ne retient que les ARNs. La radioactivité mesurée est fonction de l'activité MTase de la protéine.

Le test d'activité MTase évalue le transfert du groupement méthyl marqué radioactivement (tritium) porté par le donneur de méthyl SAM vers un ARN synthétique (**Fig. 29**). La protéine virale purifiée ($4\mu\text{M}$) est incubée avec l'ARN substrat et le SAM radioactif dans des conditions de réaction optimisées. Après 3h d'incubation à 30°C , les produits de la réaction sont filtrés à travers une membrane diéthylaminoéthyle (DEAE) chargée positivement permettant de retenir les ARNs. La membrane est lavée pour éliminer la protéine, le SAM radioactif en excès et le produit de la réaction de méthylation, le SAH. La radioactivité de la membrane est mesurée par un compteur à scintillation et correspond au taux de méthylation de l'ARN et donc de l'activité

MTase de la protéine testée. Ce test permet d'évaluer les conditions expérimentales où la protéine est active, en ne donnant qu'une indication des positions méthylée (N7 ou 2'0) sans décrire l'ordre réactionnel.

1.4.3. Chromatographie sur couche mince (TLC)

Cette technique permet de caractériser le type de méthylation des structures coiffes (**Fig. 30**).

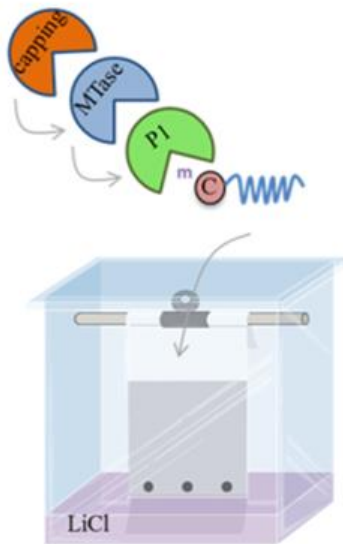


Figure 35 : Chromatographie sur couche mince (TLC)

L'ARN triphosphate est coiffé par du GTP radiomarqué via l'activité de l'enzyme de coiffage du virus de la vaccinec (en orange). Ces ARNs radiomarqués sont incubés avec la MTase virale étudiée (en bleu). La digestion par la nucléase P1 permet de libérer les structures coiffes qui sont déposés sur TLC et caractérisées en fonction de leur migration.

Dans un premier temps, l'ARN triphosphate est incubé 1h à 37°C avec de l' α [32 P]-GTP et l'enzyme de synthèse des structures coiffes du virus de la vaccine (Sigma). La réaction est stoppée par traitement à la chaleur 5mins à 70°C. L'enzyme de synthèse des structures coiffes est éliminée en utilisant des billes de Strataclean (Agilent) et l'ARN radiomarqué est purifié sur colonne de dessalage (G25, GE Healthcare). Les ARNs coiffés radiomarqué (Gp*ppN-ARN) sont incubés avec la protéine viral dont l'activité MTase doit être caractérisée dans les conditions de réaction du test de MTase. Après le temps d'incubation les ARNs sont digérés à la nucléase P1 (Sigma), une exonucléase 3'-5', qui digère l'ARN et relargue les structures coiffes. Les produits de la digestion sont déposés sur une plaque de silice et séparé par chromatographie sur

couche mince dont la phase mobile est une solution de lithium ou de sulfate d'ammonium. Après migration, les produits sont détectés par fluorographie au Phospho-Imageur. Les coiffes sont identifiées en fonction de leur distance de migration permettant leur caractérisation.

2. Etude des domaines C-terminaux de la protéine L du virus Soudan

La seconde partie des résultats porte sur la caractérisation des domaines C-terminaux de la protéine L du virus Soudan (SUDV). Dans un premier temps, les résultats de la caractérisation fonctionnelle du domaine MTase-CTD sont présentés. Nous nous sommes ensuite intéressés au rôle du domaine CTD dans l'activité MTase de la protéine L de SUDV et menée une étude de mutagenèse dirigée afin d'identifier des résidus clefs régulant l'activité MTase. Et enfin, nous avons pu caractériser structurellement le domaine MTase de la protéine L de SUDV.

2.1. Introduction vers une caractérisation fonctionnelle

Article 1 : Ebola virus L protein harbors a new enzymatic activity involved in the internal methylation of RNAs

Auteurs : Baptiste Martin*, **Coralie Valle***, Bruno Coutard, Bruno Canard, Françoise Debart, Etienne Decroly

**Contribution égale*

Références : Med Sci (Paris). 2018 Nov;34(11):919-921. doi: 10.1051/medsci/2018230. Epub 2018 Dec 10. French.

La protéine L de EBOV contient six régions conservées (CR) entrecoupées de régions variables. Son extrémité N-terminale (CRI, CRII et CRIII) contient le domaine ARN polymérase ARN-dépendante (RdRp) suivi du domaine Cap (CRIV et CRV) qui porte l'activité ribonucléotidyltransférase (PRNTase) nécessaire à la synthèse de la structure coiffe. La région C-terminale de la protéine L contient successivement les domaines « Connecteur » (CD), « méthytrasnférase » (MTase) et « C-terminal » (CTD). Une étude récente menée au sein du laboratoire a permis de caractériser la fonction MTase de la protéine L du virus Ebola (270). La région C-terminale (MTase-CTD) de la protéine L de *Sudan ebolavirus* (SUDV) a été produite sous forme recombinante en système d'expression bactérien et purifiée. En utilisant des ARNs substrats synthétiques mimant l'extrémité 5' des ARNm de SUDV, il a été montré que la MTase de SUDV méthyle successivement la position 2'O du premier nucléotide des ARNs et la position N7 du groupement guanosine pour former une structure Cap-1 qui permet l'initiation de la traduction des protéines virales et l'échappement aux senseurs de l'immunité innée tels que RIG-I. De façon surprenante, la MTase de EBOV catalyse également des modifications épi-transcriptomiques des ARNs viraux en méthylant la position 2'O des adénosines internes aux séquences ARN. Bien que de tels modifications aient été observées chez d'autres virus, leur rôle biologique reste peu documenté mais

quelques études suggèrent que ces modifications permettent d'échapper à la détection par le senseur MDA5. Cette revue présente les travaux menés au sein du laboratoire ayant permis de caractériser l'activité enzymatique des domaines C-terminaux de EBOV et mettre en évidence l'activité MTase non conventionnelle de du domaine MTase-CTD SUDV.



lyse de génomes actuels et de génomes anciens reconstitués à partir d'ossements retrouvés dans des sites archéologiques permettra de répondre à ces questions, et plus largement de reconstituer la chronologie des événements de sélection et les trajectoires évolutives des génomes lors de la domestication. Les travaux à venir devront également permettre de préciser, sur un plus large ensemble d'espèces, les cibles génétiques du syndrome de domestication. ♦

Homologous genomic signatures of domestication in goat and sheep

REMERCIEMENTS

Ce travail a été financé par l'Union européenne (projet NEXTGEN, FP7 n°244356), l'ANR (FACCE ERA-

NET Plus project CLIMGEN, ANR-14-JFAC-0002-01), le LabEx OSUG@2020 (Investissements d'avenir - ANR10LABX56), le Wellcome Trust (WT108749/Z/15/Z) et l'European molecular biology laboratory (EMBL).

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

1. Vigne JD. The origins of animal domestication and husbandry: a major change in the history of humanity and the biosphere. *C R Biol* 2011 ; 334 : 171-81.
2. Darwin C. *The variation of animals and plants under domestication*. London : John Murray, 1868 : vol 1 et 2.
3. Jordan B. Néanderthal et *Homo sapiens* : to meet or not to meet ? *Med Sci (Paris)* 2012 ; 29 : 1129-1132.
4. Axelsson E, Ratnakumar A, Arendt M-L et al. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* 2013 ; 495 : 360-4.
5. Frantz LAF, Schraiber JG, Madsen O, et al. Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nat Genet* 2015 ; 47 : 1141-8.
6. Carneiro M, Rubin C-J, Di Palma F, et al. Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. *Science* 2014 ; 345 : 1074-9.
7. Alberto FJ, Boyer F, Orozco-terWengel P, et al. Convergent genomic signatures of domestication in sheep and goats. *Nat Commun* 2018 ; 9 : 813.
8. An XP, Hou JX, Lei YN, et al. Two mutations in the 5'-flanking region of the KITLG gene are associated with litter size of dairy goats. *Anim Genet* 2015 ; 46 : 308-11.
9. Wang Z, Zhang H, Yang H, et al. Genome-wide association study for wool production traits in a Chinese Merino sheep population. *PLoS One* 2014 ; 9 : e107101.
10. Ottoni C, Flink LG, Evin A, et al. Pig domestication and human-mediated dispersal in western
11. Eurasia revealed through ancient DNA and geometric morphometrics. *Mol Biol Evol* 2013 ; 30 : 824-32.

NOUVELLE

La protéine L du virus Ébola porte une nouvelle activité enzymatique impliquée dans la méthylation interne des ARN

Baptiste Martin^{1,*}, Coralie Valle^{1,*}, Bruno Coutard¹, Bruno Canard¹, Françoise Debart², Étienne Decroly¹

¹AFMB, CNRS, Université Aix-Marseille, UMR 7257, Case 925, 163 Avenue de Luminy, 13288 Marseille cedex 09, France

²IBMM, Université de Montpellier, CNRS, ENSCM, Montpellier, France

*Contribution égale

etienne.decroly@afmb.univ-mrs.fr

► L'infection par le virus Ébola conduit à des fièvres hémorragiques souvent fatales. Depuis sa découverte en 1976 en République Démocratique du Congo, plus d'une vingtaine d'épidémies d'Ébola ont été recensées en Afrique. La plus importante a eu lieu en 2015-2016, entraînant le décès de plus de 11 000 personnes [1]. En raison du risque de réémergence de ce virus, et en l'absence de vaccins et d'agents thérapeutiques approuvés, la recherche de médicaments ciblant le virus Ébola est une priorité pour l'Organisation mondiale de la santé. Le virus Ébola est un virus enveloppé possédant un ARN génomique simple brin de polarité négative et appartient à l'ordre des *Mononegavirales*. Cet ordre est divisé en 5 familles de

virus contenant des pathogènes humains importants, comme les virus de la rage, de la rougeole, ou encore le virus respiratoire syncytial (VRS). Le virus Ébola fait partie de la famille des *Filovirus* (→) [2] et appartient au genre *Ebolavirus* qui contient 5 espèces : Zaïre, Soudan, Taï Forest, Bundibugyo et Reston. Le génome des filovirus code sept protéines virales produites à partir d'ARN messagers (ARNm) transcrits de manière séquentielle. Parmi ces protéines, la protéine « Large » (L) est un des acteurs majeurs de la propagation de l'infection puisqu'elle contient les activités enzymatiques nécessaires à la réplication/

(→) Voir la Synthèse de B. Martin et E. Decroly, *m/s* n° 8-9, août-septembre 2018 page 671

transcription et à la maturation des ARN viraux [3]. N'ayant pas d'homologues cellulaires, la protéine L constitue une cible antivirale attractive. Cette protéine contient six régions conservées (CR) entrecoupées de régions variables. La structure, récemment résolue par cryo-microscopie, de la protéine L du virus VSV (virus de la stomatite vésiculaire) indique que cette protéine est organisée en 4 domaines structuraux distincts (Figure 1) [4]. L'extrémité N-terminale de la protéine (régions CRI à CRIII) forme le domaine ARN polymérase ARN-dépendante (RdRp) intimement intriqué avec le domaine Cap (régions CRIV et CRV) qui porte l'activité polyribonucléotidyltransférase (PRNTase)

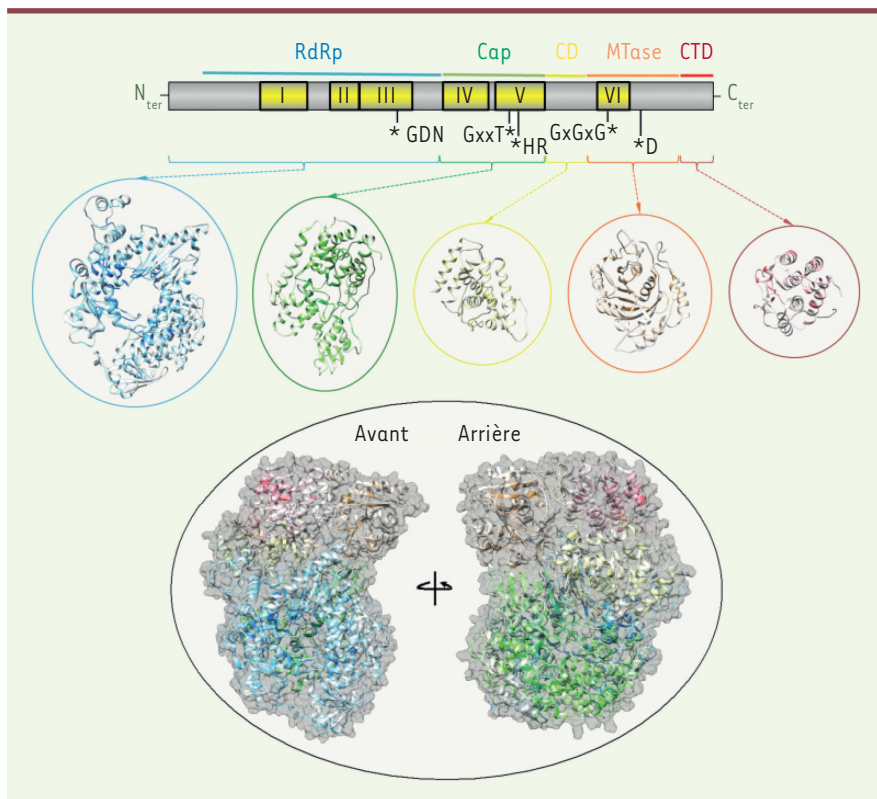


Figure 1. Domaines fonctionnels de la protéine L des Mononegavirales. Les régions conservées ont été identifiées après alignements multiples par bioinformatique. La structure présentée est celle de la protéine du virus VSV (virus de la stomatite vésiculaire) résolue par cryo-microscopie [3] qui révèle la présence de 5 domaines structuraux : le domaine ARN polymérase ARN-dépendante (RdRp) intimement intriqué avec le domaine Cap (régions CRIV et CRV) qui porte l'activité polyribonucléotidyltransférase (PRNTase) ; l'extrémité C-terminale de la protéine L contient successivement les domaines connecteur (CD), méthyltransférase (MTase) contenant la région CRVI et C-terminale (CTD).

nécessaire à la synthèse de la structure coiffe à l'extrémité 5' des ARNm. L'extrémité C-terminale de la protéine L contient successivement les domaines connecteur (CD), méthyltransférase (MTase) (CRVI) et C-terminal (CTD). La coiffe synthétisée par les mononégavirus est donc constituée d'une molécule de guanosine méthylée en position N7 reliée à l'extrémité 5' des ARNm par une liaison 5'-5' triphosphate. Cette coiffe est également méthylée en position 2'0 du ribose du premier nucléotide [5]. Par analogie avec d'autres *Mononegavirales*, le domaine MTase des filovirus a été proposé comme participant au processus de méthylation de la structure coiffe. Ce domaine contient un motif de reconnaissance d'une molécule donneuse de méthyle (S-adenosyl méthionine, SAM) à savoir Gly-X-Gly-X-Gly (X étant un acide aminé quelconque) et une tétrade catalytique caractéristique des 2'0-MTases (Lys-Asp-Lys-Glu). La présence de cette structure coiffe à l'extrémité des ARNm viraux est essentielle pour la propagation de l'infec-

tion virale. En effet, elle protège les ARNm viraux de la dégradation par des 5'-exonucléases cellulaires et assure le recrutement du facteur d'initiation de la traduction eIF4E, nécessaire à la traduction des ARNm en protéines [5]. Par ailleurs, il a été démontré que la 2'0-méthylation de cette structure coiffe était un marqueur du « soi » jouant un rôle dans le contrôle de la réponse antivirale [6]. Des études récentes ont démontré que la 2'0-méthylation des structures coiffes empêche la détection des ARN viraux par les senseurs cytoplasmiques de l'immunité innée tels que les protéines RIG-I (*retinoic-acid-inducible protein 1*) et/ou MDA5 (*melanoma-differentiation-associated gene 5*) et limite ainsi la production d'interféron [2]. De plus, il a été montré qu'il existe, parmi les ISG (*interferon-stimulated gene*), des facteurs de restriction comme les protéines IFIT (*IFN-induced proteins with tetratricopeptide repeats*), qui reconnaissent et séquestrent les ARN possédant des structures coiffées

incomplètement méthylées [7]. Les enzymes virales impliquées dans la synthèse des structures coiffes et leurs méthylation constituent donc des cibles thérapeutiques pertinentes. En effet, l'inhibition de ces enzymes devrait d'une part bloquer la réplication virale et, d'autre part, induire une détection précoce des agents infectieux par les mécanismes de l'immunité innée induisant une réponse antivirale protectrice.

La structure et l'activité MTase de la coiffe ont été récemment caractérisées chez le virus hMPV (*human metapneumovirus*, un autre virus de l'ordre des *Mononegavirales*) [8], mais nous manquons d'informations structurales et fonctionnelles concernant les activités MTases du virus Ébola. Nous avons donc cloné le domaine C-terminal de la protéine L du virus Ébola (MTase et CTD) pour le produire en système bactérien. Nous avons ensuite purifié l'enzyme afin de caractériser ses fonctions biologiques. En étudiant l'activité enzymatique de cette MTase du virus Ébola sur des ARN substrats synthétiques mimant l'extrémité 5' des ARN viraux, nous avons démontré son implication dans le processus de méthylation.

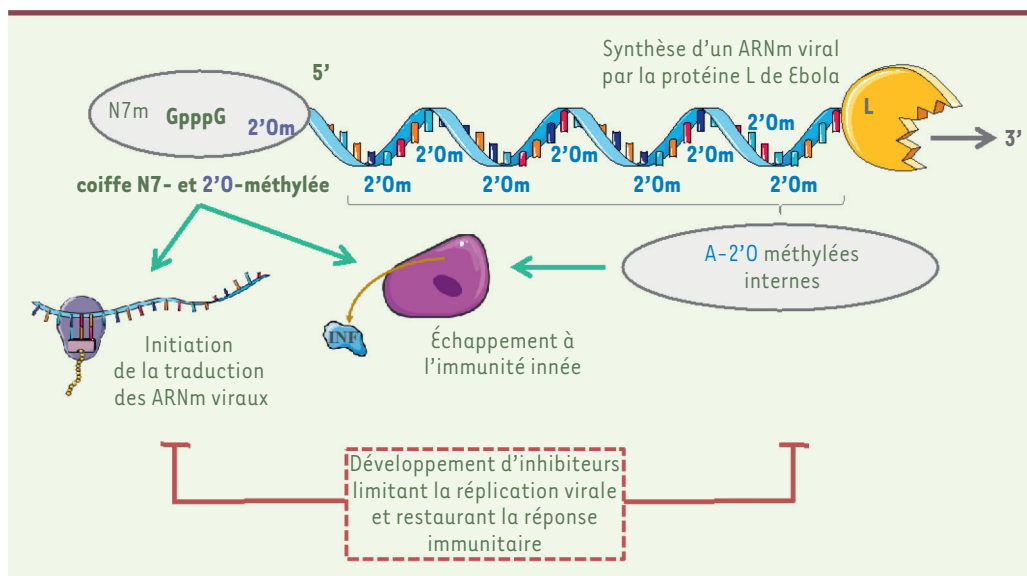


Figure 2. Modification des ARN viraux. Les différentes modifications épitranscriptomiques des ARN viraux induites par le domaine méthyltransférase (MTase) du virus Ebola sont représentées, à savoir des méthylations de la structure coiffe en position N7 et 2'OH du premier nucléotide ainsi que les méthylations internes des adénosines des ARN viraux. La figure indique également les rôles possibles de ces méthylations dans la

traduction des ARN messagers viraux et dans la régulation de la réponse antivirale. Des molécules thérapeutiques ciblant l'activité MTase du virus Ebola devraient bloquer la synthèse des protéines virales et restaurer l'immunité antivirale protectrice.

tion de la coiffe [9]. Nos observations indiquent que la méthylation de la coiffe, chez cette famille de virus, suit un ordre séquentiel obligatoire : la 2'O-méthylation du premier nucléotide en 5' de l'ARN précède la N7-méthylation du groupement guanosine de la coiffe. Nous avons de plus observé que la MTase du virus Ebola possède de façon surprenante, en plus de son activité de méthylation de la coiffe, une activité MTase cap-indépendante très significative : cette enzyme méthyle également des ARN non coiffés sur des résidus internes à leur séquence. En utilisant des ARN substrats homopolymériques (poly-A, poly-U, poly-A et poly-G), nous avons démontré que les méthylations concernent uniquement les résidus adénosines, et des analyses par chromatographie liquide de haute pression (HPLC) couplée à la spectrométrie de masse ont permis de démontrer que les méthylations modifient le groupement 2'OH du ribose des adénosines.

Ce travail relate, pour la première fois, l'existence de modifications épitranscriptomiques transmises par le domaine MTase de la protéine L d'un mononégavirus. Une activité de méthylation interne

similaire a récemment été caractérisée chez les *Flavivirus* comme le virus de la Dengue et le virus Zika [10]. De récentes études démontrent également l'existence de méthylations internes au sein de la séquence des ARN génomiques de différentes familles de virus [11]. Ces méthylations sont maintenant avérées, mais leur rôle dans la réplication virale et l'échappement au contrôle cellulaire via les mécanismes de l'immunité innée restent à déterminer. Ces modifications des ARN pourraient être impliquées dans la pathogenèse du virus Ebola liée, entre autres, à sa capacité d'échapper aux réponses antivirales précoces. Ces résultats ouvrent donc de nouvelles opportunités de recherche pour mieux comprendre les fonctions des modifications épitranscriptomiques des ARN viraux dans la relation hôte/pathogène et développer des stratégies antivirales efficaces contre le virus Ebola. ♦

Ebola virus L protein harbors a new enzymatic activity involved in the internal methylation of RNAs

LIENS D'INTÉRÊT

Les auteurs déclarent n'avoir aucun lien d'intérêt concernant les données publiées dans cet article.

RÉFÉRENCES

- Shultz JM, Espinola Z, Espinola M, et al. Distinguishing epidemiological features of the 2013–2016 West Africa Ebola virus disease outbreak. *Disaster Heal* 2016 ; 3 : 78–88.
- Martin B, Decroly E. Mécanisme d'échappement des filovirus à l'immunité innée. *Med Sci (Paris)* 2018 ; 34 : 671–7.
- Martin B, Canard B, Decroly E. Filovirus proteins for antiviral drug discovery: Structure/function bases of the replication cycle. *Antiviral Res* 2017 ; 141 : 48–61.
- Liang B, Li Z, Jenni S, et al. Structure of the L protein of vesicular stomatitis virus from electron cryomicroscopy. *Cell* 2015 ; 162 : 314–27.
- Decroly E, Ferron F, Lescar J et al. Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol* 2011 ; 10 : 51–65.
- Zust R, Cervantes-Barragan L, Habjan M, et al. Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat Immunol* 2011 ; 12 : 137–43.
- Daffis S, Szretter KJ, Schriewer J, et al. 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* 2010 ; 468 : 452–6.
- Paesen GC, Collet A, Sallamand C et al. X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat. Commun* 2015 ; 6 : 8749.
- Martin B, Coutard B, Guez T et al. The methyltransferase domain of the *Sudan ebolavirus* L protein specifically targets internal adenosines of RNA substrates, in addition to the cap structure. *Nucl Acids Res* 2018 ; <https://doi.org/10.1093/nar/gky637>.
- Coutard B, Barral K, Lichière J, et al. Zika virus methyltransferase: structure and functions for drug design perspectives. *J Virol* 2017 ; 91 : 5, e02202-16.
- McIntyre W, Netzband R, Bonenfant G, et al. Positive-sense RNA viruses reveal the complexity and dynamics of the cellular and viral epitranscriptomes during infection. *Nucleic Acids Res* 2018 ; 46 : 5776–91.

2.2. Caractérisation du domaine CTD de la protéine L du virus Soudan

Article 2: The C-terminal domain of the Sudan ebolavirus L protein is essential for RNA binding and methylation.

Auteurs : **Coralie Valle***, Baptiste Martin*, Françoise Debart, Jean-Jacques Vasseur, Isabelle Imbert, Bruno Canard, Bruno Coutard, Etienne Decroly

**Contribution égale*

Référence : J Virol. 2020 Apr 8. pii: JVI.00520-20. doi: 10.1128/JVI.00520-20.

Le domaine MTase, conservé au sein des *Mononegavirales*, est suivi du domaine CTD dont la fonction reste à élucider. L'alignement multiple des domaines C-terminaux des *Mononegavirales* a permis de définir des bornes des domaines MTase et CTD de SUDV (269). Du fait de la grande variabilité en séquence et en longueur, un alignement robuste des domaines CTD des *Mononegavirales* est difficile. Cependant, la comparaison des propriétés physico-chimiques des domaines CTD de différents mononégavirus a mis en évidence un enrichissement en acides aminés basiques. L'alignement des CTD des filovirus suggère que la présence de résidus basiques et aromatiques conservés pourrait être impliqués dans la fixation de l'ARN. De plus, l'étude des structures des domaines MTase-CTD de VSV et hMPV montre que les domaines MTase et CTD forment un sillon à proximité du site catalytique suggérant le rôle du domaine CTD dans le recrutement et positionnement de l'ARN dans le site catalytique. Dans un premier temps, afin de mettre en évidence le rôle du domaine CTD de la protéine L dans l'activité MTase, deux protéines recombinantes ont été produites et purifiées : le domaine MTase+CTD (noté MTase-CTD : résidus 1713 à 2211 de 58,4 kDa) et le domaine MTase délété du domaine CTD (noté MTase- Δ CTD : résidus 1713 à 2026 de 38,9 kDa). La comparaison des activités des protéines MTase- Δ CTD et MTase-CTD a montré que le domaine CTD joue un rôle crucial dans l'activité MTase, dans la mesure où la délétion du domaine entraîne une perte de l'activité MTase. Le second objectif de cette étude est l'identification de résidus clefs dans l'activité MTase. L'alignement des domaines CTD des filovirus nous a permis d'identifier des résidus probablement impliqués dans le recrutement du substrat. En mutant les résidus basiques et aromatiques conservés, nous avons ainsi mis en évidence leur rôle dans la reconnaissance des ARNs viraux et la régulation de l'activité MTase. L'ensemble de ces résultats contribue à une meilleure compréhension de l'activité MTase de la protéine L de EBOV, un acteur central dans la réplication virale et

l'échappement à la réponse immunitaire innée de l'hôte et ouvre la voie vers de nouvelles thérapies antivirales.

1 **Manuscript title: The C-terminal domain of the *Sudan ebolavirus* L protein**
2 **is essential for RNA binding and methylation.**

3 Coralie Valle^{1#}, Baptiste Martin^{1#}, Françoise Debart², Jean-Jacques Vasseur², Isabelle Imbert¹, Bruno
4 Canard¹, Bruno Coutard³ & Etienne Decroly^{1*}

5
6 ¹ AFMB, CNRS, Université Aix-Marseille, UMR 7257, Case 925, 163 Avenue de Luminy, 13288
7 Marseille Cedex 09, France

8 ² IBMM, UMR 5247 CNRS, Université de Montpellier, ENSCM, Montpellier, France

9 ³ Unité des Virus Emergents (UVE: Aix-Marseille Univ-IRD 190-Inserm, 1207-IHU Méditerranée
10 Infection) Marseille, France

11

12 # These authors contribute equally to this work. Author order was determined by mutual agreement.

13 * To whom correspondence should be addressed. Tel: +33 4 91 82 86 47; Fax: 04 91 82 86 38; Email:
14 etienne.decroly@afmb.univ-mrs.fr

15

16 **ABSTRACT**

17 The large (L) protein of Ebola virus is a key protein for virus replication. Its N-terminal region
18 harbors the RNA-dependent RNA polymerase activity, and its C-terminus contains a cap assembling
19 line composed of a capping domain and a methyltransferase domain (MTase) followed by a C-
20 terminal domain (CTD) of unknown function. The L protein MTase catalyzes methylation at the 2'-O
21 and N-7 positions of the cap structures. In addition, the MTase of Ebola virus can induce cap-
22 independent internal adenosine 2'-O-methylation. In this work, we investigated the CTD role in the
23 regulation of the cap-dependent and cap-independent MTase activities of the L protein. We found that
24 the CTD, which is enriched in basic amino acids, plays a key role in RNA binding and in turn
25 regulates the different MTase activities. We demonstrated that the mutation of CTD residues
26 modulate specifically the different MTase activities. Altogether, our results highlight the pivotal role
27 of the L protein CTD in the control of viral RNA methylation, which is critical for Ebola virus
28 replication and escape from the innate response in infected cells.

29

30 **IMPORTANCE**

31 Ebola virus infects human and non-human primates causing severe infections that are often fatal. The
32 epidemics, in West and Central Africa, emphasize the urgent need to develop antiviral therapies. The
33 Ebola virus large protein (L), which is the central protein for viral RNA replication/transcription,
34 harbors a methyltransferase domain followed by a C-terminal domain of unknown function. We show

35 that the C-terminal domain regulates the L protein methyltransferase activities and consequently
36 participates in viral replication and escape of the host innate immunity.

37

38 INTRODUCTION

39 Ebola virus is an emerging virus that causes severe epidemics, such as the unprecedented devastating
40 outbreak in West Africa of 2015 that caused more than 11,000 deaths (1). Although Ebola faded from
41 public attention, outbreaks continue to occur especially in the Democratic Republic of the Congo
42 (DRC) where the current epidemic, which was declared in August 2018, has already killed more than
43 2,000 people (2). International efforts deployed have led to control the epidemic (3). Ebola virus
44 infects human and non-human primates leading to hemorrhagic fever that is fatal in most cases.
45 Although vaccines are currently in development (4)(5) and some therapeutic antibodies have shown
46 antiviral activities (6), therapeutic compounds that limit viral replication are needed to treat infected
47 people or contact patients. However, the fine-tuning of the virus replication process remains poorly
48 understood.

49 *Ebolavirus* belongs to the *Mononegavirales* order that includes other important human pathogens,
50 such as measles, respiratory syncytial virus, human metapneumovirus (hMPV) and rabies virus. These
51 viruses have a similar genetic organization and common replication strategies. The *Ebolavirus* genus
52 contains six strains: *Zaire ebolavirus* (Ebola virus, EBOV), *Sudan ebolavirus* (Sudan virus, SUDV),
53 *Tai Forest ebolavirus* (Tai Forest virus, TAFV), *Bundibugyo ebolavirus* (Bundibugyo virus, BDBV),
54 *Reston ebolavirus* (Reston virus, RESTV), and *Bombali ebolavirus* (Bombali virus, BOMV). These
55 viruses, and *Marburg marburgvirus* (Marburg virus, MARV), which causes acute hemorrhagic fever,
56 *Lloviu cuevavirus* and *Mengla dianlovirus* form the *Filoviridae* family (7).

57 The negative-sense RNA genome of filoviruses is about 19 kb-long and encodes seven major
58 proteins: the nucleoprotein (NP), the viral proteins VP35 and VP40, the glycoprotein (GP), the viral
59 proteins VP30 and VP24, and the “large” protein L (8)(9). The viral life cycle is initiated by the
60 interaction of the GP envelope protein with cell surface determinants to allow the virus entry in the
61 cell by macro-pinocytosis. The NP-encapsidated viral genome is then released in the host cell
62 cytoplasm where the virus replication cycle is initiated. The L protein, which contains the RNA-
63 dependent RNA polymerase (RdRp) domain, as well as NP, VP35 and VP30 are required for
64 replication. First, viral mRNAs are transcribed using a discontinuous transcription mechanism that
65 generates seven monocistronic capped and poly-adenylated RNA species (10)(11). The polymerase
66 complex initiates mRNA synthesis at a conserved gene-start (GS) sequence and transcribes genes up
67 to the gene-stop sequence (GE) (12). The poly-A tail is then added by a polymerase-mediated
68 slippage mechanism at a conserved poly(U)₆ tract in the GE sequence (13). Then, the RdRp scans the
69 intergenic sequences and starts again RNA transcription at the next GS sequence. *Mononegavirales*
70 mRNAs are co-transcriptionally capped by an unusual sequence of capping reactions. This process,
71 which has been extrapolated from the vesicular stomatitis virus (VSV) model, takes place when the

72 nascent RNA chain length has reached a length of 31 nucleotides (14). The cap synthesis starts with
73 the formation of a covalent adduct between a conserved histidine residue of the polyribonucleotidyl
74 transferase (PRNTase) and the 5'-phosphate of the viral mRNA. In the presence of GDP, the
75 PRNTase transfers a GDP moiety to the 5' phosphate of the covalently bound RNA, leading to the
76 formation of capped RNA (Gppp-RNA) (15). The 5' cap structure is subsequently methylated at the
77 2'-O position of the first nucleotide (N1) and at the N-7 position of the guanosine, yielding cap-1
78 structures (^mGpppN_m) (16). This cap-1 structure is undistinguishable from cellular caps and plays a
79 critical role in virus replication because it protects viral mRNAs from the host cell 5' exonucleases
80 and allows the eukaryotic translation initiation factor 4 (eIF4e)-dependent initiation of mRNA
81 translation (revue decroly et al nature). Additionally, the cap-dependent 2'-O-methylation is a self-
82 marker that limits the detection of viral RNAs by the innate immunity sensors of the retinoic acid-
83 inducible gene-I (RIG-I) like family. Thus, viruses that express RNAs without cap 2'-O-methylation
84 (^mGpppN-RNA or cap-0) are detected early by RIG-I receptors in infected cells, leading to the
85 secretion type I interferon (IFN α/β) (17)(18). IFNs stimulate the expression of IFN-stimulated genes,
86 such as IFN Induced Protein with Tetratricopeptide Repeats (IFIT 1/3), that sequester miss-capped
87 RNAs (*i.e.* cap-0), leading to reduced virus propagation (19)(20). Therefore, capping enzymes, such
88 as cap-MTases, are considered potent antiviral targets. Indeed, N-7 methylation inhibition is supposed
89 to limit RNA translation into viral proteins, and 2'-O-MTase inhibitors should unmask viral RNA to
90 intracellular innate immunity sensors (21)(21).

91 Several cryo-electron microscopy (cryo-EM) structures of the full-length mononegavirus L protein
92 were recently described (23)(24)(25)(26). While L protein C-terminal domains were not visible in
93 most structures, the VSV L protein was fully unravelled (23). This study revealed an organization
94 with five different topological domains (Fig. 1A). The RdRp contains three conserved regions (CRI,
95 CRII, and CRIII) and is intimately associated with the PRNTase domain (CRIV), forming the “donut-
96 like” structure observed by negative staining (27). The RdRp domain is organized in a right-hand
97 “fingers-palm-thumb” structure, typical of polymerases. The cap domain (PRNTase), which has an
98 original fold, projects a loop near the catalytic site of the polymerase. This structural interplay
99 suggests that the PRNTase domain participates in polymerization initiation, similarly to polymerases
100 with a priming loop (28). The structure corroborates biochemical analysis results showing the
101 interplay between RdRp and capping activities (14). The “donut-like” cluster (RdRp and PRNTase
102 domains) is followed by three globular structures that corresponds to the connector domain (CD), the
103 MTase domain, and a small C-terminal domain (CTD). The CD might separates the PRNTase domain
104 from the MTase domain, which is composed of a bundle of eight helices (29). The MTase domain
105 contains a K-D-K-E catalytic tetrad, characteristic of 2'-O-MTases. The VSV MTase structure is
106 organized in a typical Rossman fold of S-adenosyl methionine (SAM)-dependent methyltransferases.
107 The X-ray structure of the hMPV MTase+CTD segment confirmed the presence of the

108 methyltransferase fold with a SAM-binding site (GxGxG) that positions the released methyl group in
109 front of the catalytic tetrad (30) (Fig. 1B-C). However, unlike other viral MTases, the hMPV
110 MTase+CTD lacks a canonical cap-binding site (31)(32)(33). Thus, it was proposed that the RNA
111 substrate is accommodated in unusual RNA-binding groove formed by the MTase domain overlaid by
112 the CTD enriched in basic and aromatic amino acids (in purple in Fig. 1B-C), known to interact with
113 the nucleotide bases by stacking interactions (30).

114 The MTase activities of the VSV L protein were first demonstrated *in vitro*. This protein
115 sequentially catalyzes the RNA cap structure methylation at the ribose 2'-O and guanosine N-7
116 positions (34). The MTase catalytic activity was confirmed using recombinant hMPV and SUDV
117 MTase+CTD proteins (30)(35). Indeed, it was reported that these MTase+CTD methylate the 2'-O
118 position of the first 5'-RNA nucleotide (N1) of capped and triphosphate RNAs. The subsequent
119 methylation of the N-7 guanosine of the cap structure was demonstrated by incubating the SUDV
120 MTase+CTD protein with longer RNA substrates. We also demonstrated that the SUDV MTase
121 catalyzes cap-independent methylation, leading to internal adenosine 2'-O-methylation (35), an
122 activity that is not obviously detected using the hMPV MTase (30). The role of these
123 epitranscriptomic RNA modifications is not yet fully understood. However, it was recently reported
124 that similar internal 2'-O-methylation on the HIV RNA genome protects viral RNAs from detection
125 by the RIG-like receptor Melanoma Differentiation-Associated protein 5 (MDA5) (36). Therefore, the
126 filovirus MTase might have evolved towards a dual activity, targeting the cap structure present at the
127 5' end of mRNA and adenosine residues inside the RNA.

128 In this work, we investigated the role of SUDV CTD domain of the L protein in the MTase
129 activities by mutational analysis. We found that the CTD plays a key role in RNA substrate binding
130 and thereby regulates the different MTase activities. In addition, the mutational analysis allowed
131 identifying key CTD residues that regulate the different MTase activities of the SUDV L protein.
132 These results suggest that SUDV CTD is involved in the RNA methylation process and in the control
133 of internal methylation *versus* cap methylation.

134

135 MATERIAL AND METHODS

136 Cloning and recombinant protein production

137 Codon-optimized SUDV MTase+CTD synthetic genes (Biomers) were cloned in the pET14b vector
138 for bacterial expression (37). The MTase- Δ CTD construct was derived from MTase+CTD by PCR
139 (Pfu Turbo DNA polymerase, Agilent technologies) using the following primers (Forward MTase-
140 Δ CTD:5'GGCTGAGCCACCTGACCAAGTATACCACCTGATGACTGCACAACAGCTATATCG
141 CGTTCGGCTTTCCG-3' and Reverse MTase- Δ CTD:
142 5'CGGAAAGCCGAACGCGATATAGCTGTTGTGTCAGTCATCAGGTGGTATACTTGGTCAGG
143 TGGCTCAGCC-3') that flank the sequence of interest, and was expressed in bacteria. Mutations
144 were introduced by PCR amplification of the wild type sequence using mutation-carrying primers

145 (designed using QuikChange Primer Design) and the Pfu Turbo DNA polymerase. After introduction
146 of the constructs in the *Escherichia coli* T7 strain (NEB) by transformation, bacteria were cultured at
147 30°C until the optical density at 600nm (OD_{600nm}) reached 0.6, after which the temperature was
148 reduced to 17°C and 20 μ M IPTG (Sigma) was added. After 16 hours of induction, bacteria were
149 centrifuged (8000 $\times g$ for 20 min at 4°C) in a Sorval Lynx 6000 centrifuge (Thermo), and pellets were
150 stored at -80°C.

151

152 **Recombinant protein purification**

153 All recombinant proteins were purified using a similar protocol, except for the elution step. Bacterial
154 pellets were thawed on ice, and lysed in lysis buffer (50 mM Tris pH 8, 150 mM NaCl, 5% glycerol,
155 30 mM imidazole, 1 mM PMSF, 100 μ g/mL lysozyme, 1 μ g/mL DNase, 0.1% Triton X100)
156 supplemented with the detergent mix BugBuster (Merck Millipore) only for MTase+CTD
157 purification. The volume of lysis buffer corresponded to 10 times the [final OD_{600nm} value \times culture
158 volume (mL)] value. After clarification (18,000 $\times g$, 30 min, 4 °C), lysates were incubated with the
159 CoNTA resin (Thermo Fisher; 0.5 mL/L culture) at 4°C for 1h, with gentle shaking. Beads were
160 transferred to 25-mL cartridges and washed with 2 \times 20 mL of 50 mM Tris pH 8, 1 M NaCl, 5%
161 glycerol, 30 mM imidazole and 2 \times 10 mL of buffer 50 mM Tris pH 8, 150 mM NaCl, 5% glycerol.
162 Proteins were eluted using two different buffers: 50 mM Tris pH 8, 150 mM NaCl, 5% glycerol 1 M
163 arginine for MTase+CTD, and 50 mM Tris pH 8, 150 mM NaCl, 5% glycerol, 150 mM imidazole for
164 MTase- Δ CTD. In this case, imidazole was removed by size exclusion chromatography using
165 Superdex S75 16/60 filtration columns (GE Healthcare). Finally, proteins were concentrated on
166 Amicon Ultrafilters (EMD Millipore), and the final concentration was determined by OD_{280nm} .
167 Proteins were stored in 50% glycerol at -20°C.

168

169 **Synthesis of RNA substrates**

170 RNA sequences were chemically synthesized on solid support using an ABI 394 oligonucleotide
171 synthesizer. After RNA elongation with 2'-O-pivaloyloxymethyl phosphoramidite ribonucleotides,
172 and 2'-O-methyl phosphoramidite ribonucleotides (Chemgenes, USA), the 5'-hydroxyl group was
173 phosphorylated and the resulting *H*-phosphonate derivative was oxidized and activated into a
174 phosphoroimidazolidate derivative to react with guanosine diphosphate (GpppRNA). After
175 deprotection and release from the solid support, GpppRNAs were purified by IEX-HPLC and their
176 purity (>95%) was confirmed by MALDI-TOF spectrometry. N-7 methylation of the purified
177 GpppRNAs was performed by incubation with N-7-hMTase (38, 39).

178

179 **MTase activity assay (filter binding assay)**

180 To evaluate methyltransferase activities, a radioactive test was set up by mixing 4 μ M of recombinant
181 proteins with 1 μ M of purified synthetic RNAs (RNA sequences in **Table 1**), 10 μ M SAM and 0.5 μ M

182 ³H-SAM (Perkin Elmer) in an optimized MTase assay buffer (50 mM Tris-HCl), supplemented with
183 10 mM arginine for MTase- Δ CTD, and optimum pH for the different MTase activities: pH 7 for the
184 N-7-cap MTase activity, pH 8 for the 2'-O-cap MTase, and pH 8.5 for internal methylation (35).
185 Reactions were stopped by 10-fold dilution in water after 3h at 30°C. Samples were transferred to
186 DEAE filtermats (Perkin Elmer) using a Filtermat Harvester (Packard Instruments). Methyl transfer
187 was then evaluated as described before (40). Briefly, the RNA-retaining filtermats were washed twice
188 with 10 mM ammonium formate pH 8, twice with water, and once with ethanol. Then, they were
189 soaked with liquid scintillation fluid to measure the ³H-methyl transfer to the RNA substrates using a
190 Wallac MicroBeta TriLux Liquid Scintillation Counter13. For comparison, it was assumed that the
191 different experimental groups were independent, and data followed a Gaussian distribution with the
192 same variance. Two-way ANOVA and multiple comparison Tukey tests were used (Prism) to
193 evaluate differences between groups. The level of significance for $\alpha=0.05$ is indicated as follows:
194 * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$.

195

196 **Fluorescence polarization (FP)**

197 Using T4 RNA ligase 1 (20 units; New England Biolabs), cyanine 5-cytidine-5-phosphate-3-(6-
198 aminohexyl)phosphate (12.5 μ M, Jena Bioscience) was ligated to the 3'ends of the RNA substrates
199 (10 μ M) in T4 RNA ligase 1 buffer (NEB), 1 mM ATP (16 °C, overnight). The ligase was removed
200 by RNA precipitation in 3 M sodium acetate supplemented with 1 μ g/ μ L of glycogen (Thermo
201 Scientific). Fluorescent RNA was incubated (at room temperature for 5 min) with increasing protein
202 concentrations in reaction buffer (20 mM Tris pH 8 or 8.5, 150 mM NaCl and 5% glycerol)
203 supplemented with 50 mM arginine for MTase- Δ CTD. This corresponded to the final concentration of
204 arginine in the fluorescence polarization (FP) assay. FP measurements were performed in a microplate
205 reader (PHERAstar FS; BMG Labtech) with an optical module equipped with polarizers and using
206 excitation and emission wavelengths of 590 and 675 nm, respectively. Dissociation constants (Kd)
207 were determined using Hill slope curve fitting (Prism).

208

209 **Thermal shift assay (TSA)**

210 A mix of 40 μ M recombinant proteins diluted (1/5) in 50 mM Tris pH 8, 150 mM NaCl, 5%
211 glycerol, and 0.02% (v/v) SYPRO Orange dye (Thermo Scientific) was prepared to a total volume of
212 25 μ L. To have the same arginine concentration as for MTase+CTD, the buffer reaction for MTase-
213 Δ CTD was supplement with 100 mM arginine, which corresponds to the final concentration of
214 arginine in the assay. Samples were placed in semi-skirted 96-well PCR plates (BioRad), sealed and
215 heated in a CFX96 Touch qPCR machine (BioRad) from 25 to 95°C at a rate of 1 °C.min⁻¹.
216 Fluorescence changes were monitored at 492 and 610 nm (excitation and emission wavelengths),

217 respectively. Melting temperatures (T_m) were calculated using a Boltzmann sigmoidal regression for
218 fitting (Prism).

219

220

221 RESULTS

222

223 **The CTD of the SUDV L protein contains basic and aromatic amino acids that could be**
224 **involved in RNA recognition.**

225 The C-terminus of the *Mononegavirales* L protein contains a conserved MTase domain followed
226 by a less conserved CTD domain with unknown function(s) (**Fig. 1**). Unfortunately, due to their
227 important differences in length (from ~120 residues in *Pneumoviridae* to ~240 residues in
228 *Rhabdoviridae*) and sequence, we could not align the CTD of *Mononegavirales* (**Table 2**).
229 Nevertheless, comparison of the isoelectric points (pI) of the different CTDs showed that they were
230 all quite basic and varied between 8.33 and 9.94, except for *Rubulavirus* (**Table 2**). To identify
231 conserved features among filovirus CTDs (residues 2037 to 2210, numbering corresponding to the
232 SUDV sequence), we aligned their sequences (**Fig. 1D**). The alignment revealed the presence of 10
233 conserved aromatic (n=5) and basic (n=5) amino acids (highlighted in magenta). In addition, we also
234 noted the presence of three positions with either arginine or lysine basic residues (highlighted in
235 orange) (**Fig. 1D**).

236

237 **The L protein CTD is essential for SUDV MTase activities.**

238 We produced the C-terminal region of the SUDV L protein (MTase+CTD) and the MTase domain
239 alone (MTase- Δ CTD) in bacteria. SDS-PAGE followed by Coomassie blue staining showed that
240 recombinant MTase+CTD and MTase- Δ CTD had the expected apparent molecular weight of 58 kDa
241 and 39 kDa, respectively (**Fig. 2A**). After confirming their identity by MALDI-TOF, we analyzed
242 their folding and stability by TSA (**Fig. 2B**). The analysis revealed the typical folded-to-unfolded
243 transition curves of globular proteins (41), with a T_m value of 54.3°C for MTase+CTD and 46.0°C for
244 MTase- Δ CTD. Although the T_m difference between MTase+CTD and MTase- Δ CTD was
245 significant (~8°C), the two proteins presented the classical native to unfolded phases, suggesting
246 that both adopt a folded conformation, and that the CTD contributes to MTase thermal resistance,
247 potentially indicative of a close structural relation. Altogether, these results confirmed that both
248 proteins are folded. We next investigated CTD involvement in MTase activity by comparing the
249 enzymatic activity of both proteins using a 13 nucleotide-long capped RNA (GpppG-SUDV₁₂)
250 substrate, which mimics the conserved 5' sequence of SUDV 5' transcripts, in the presence of a
251 radiolabeled methyl donor (³H-SAM). Measurement of the radioactivity transferred to the RNA
252 substrate (**Fig. 2C**) showed that MTase- Δ CTD was not active, whereas MTase+CTD induced the
253 methylation of the capped-RNA substrate. We previously confirmed the specificity of the observed

254 methylation by showing that mutation of each residue of the conserved K-D-K-E catalytic tetrad in
255 the MTase domain leads to a strong decrease of MTase+CTD MTase activities (37).

256 We also compared the RNA binding properties of MTase+CTD, MTase- Δ CTD, and the
257 MTase+CTD catalytic mutants by FP assay. To this aim, we labeled the capped RNA substrate at its
258 3' end by ligation of a fluorescent pCp-Cy5 residue (GpppG-SUDV₁₂-pCp-Cy5), incubated it with
259 increasing concentrations of MTase+CTD, MTase- Δ CTD, or catalytic mutants, and measured the
260 RNA-protein interaction by FP, except for the K1959A mutant that was not stable (**Fig. 2D & Table**
261 **3**). MTase- Δ CTD barely interacted with the capped RNA substrate. Conversely, MTase+CTD
262 induced an FP signal increase, confirming its interaction with the synthetic capped RNA substrate
263 (apparent K_d value of 740 nM). We also observed that the MTase+CTD catalytic mutants without
264 MTase activity (37) maintained most of their RNA binding properties. However, the K1813A
265 mutation led to a loss of the protein-RNA interaction ($K_d \approx 2.7 \mu\text{M}$) (**Table 3**). This indicates that the
266 catalytic site could also contribute to RNA interaction. Altogether, these results suggest that the
267 SUDV CTD plays a key role in RNA binding and promotes MTase activity.

268

269 **Effect on MTase activities of single mutations in the conserved residues of SUDV CTD.**

270 We previously demonstrated that in addition to cap N-7 and 2'-O-methylation, the SUDV
271 MTase+CTD induces internal adenosine 2'-O-methylation (37). To further characterize CTD role in
272 RNA recruitment and in the regulation of the different MTase activities catalyzed by MTase+CTD,
273 we mutated the conserved residues in CTD into alanine or aspartic acid (**Fig. 1C & Table 4**). Among
274 the 25 generated mutants, 13 led to the production of soluble proteins with a molecular weight similar
275 to that of wild type MTase+CTD (**Fig. 3A**) and with the typical denaturation curve of folded proteins
276 (T_m value from 50.6°C to 55.7°C) (**Table 4**). The T_m values of most mutants were close to that of wild
277 type MTase+CTD, except the T_m of the R2068A and R2172A mutants that were slightly lower (ΔT_m
278 = -3.7°C and ΔT_m = -3.1°C, respectively, compared with wild type MTase+CTD). This suggests that
279 these mutations slightly affected the protein stability, but to a lesser extent compared with CTD
280 deletion (ΔT_m = -8.3°C). We next tested the mutation effect on RNA binding by FP assay using
281 GpppG-RNA-Cy5 and ^mGpppG_m-RNA-Cy5 SUDV RNA as canonical substrates for all MTase
282 activities and for internal methylation, respectively (**Table 4**). All single mutations affected RNA
283 binding, with K_d values ranging from 1 to >9 μM ; however, the binding decrease was less important
284 than that observed with MTase- Δ CTD. Moreover, most mutations affected the binding of GpppG-
285 SUDV₁₂ and ^mGpppG_m-SUDV₁₂ in a similar manner (**Table 4**). Conversely, the K2043A and K2189A
286 mutations reduced the interaction with GpppG-SUDV₁₂ (by about 3 times), and particularly with
287 ^mGpppG_m-SUDV₁₂ ($K_d > 9 \mu\text{M}$ and $K_d = 8.6 \mu\text{M}$, respectively). Then, by incubating wild type and
288 mutated MTase+CTD proteins with GpppG_m(A_m)-SUDV₁₂, ^mGpppA(A_m)-SUDV₁₂ or ^mGpppG_m-
289 SUDV₁₂ in the presence of a radiolabeled methyl donor (³H-SAM), we determined the mutation
290 effects on the different MTase activities (i.e. cap N-7 MTase, cap 2'-O-MTase, and internal 2'-O-

291 MTase activities, respectively) (**Fig. 3B**). The results showed that some mutations affected almost all
292 MTase activities, whereas others impaired only some MTase activities. The mutations R2068A,
293 K2118A, K2189A/D, and R2193A led to a reduction by 90% or more of N-7- and 2'-O-MTase and
294 also internal 2'-O-MTase activities compared with wild type MTase+CTD. In these mutants, RNA
295 binding was also strongly reduced, except for the K2189A/D mutant that still partially recognized the
296 RNA substrate (**Table 4**). This suggested that this mutant protein accommodates the RNA in an
297 inappropriate way for the transfer of the methyl group to the RNA substrate. In addition, residues
298 K2189 and R2194 correspond to a conserved K-K-G motif in *Paramyxovirus* CTD that was
299 previously proposed to play a key role in RNA binding and N-7 and 2'-O-MTase activities (34). In
300 the second group of mutants, RNA binding and MTase activities were less affected. In this group, we
301 identified two mutations (K2043A, H2067D) that impaired more strongly the 2'-O-MTase activities
302 (cap 2'-O-MTase and internal 2'-O-A MTase) than the N-7 MTase activity. As expected, binding of
303 these two mutants to ³HmGpppG_m-SUDV₁₂ was more affected than that to GpppG-SUDV₁₂. In two other
304 mutants (H2112A and F2113A) of this second group, internal methylation was more strongly limited
305 than cap-dependent MTase activities. This suggests that these mutants can accommodate the RNA
306 substrate in a non-optimal manner for internal 2'-O-methylation. Altogether, these results confirmed
307 that the CTD plays an important role in the RNA substrate binding, and that some conserved basic
308 residues (R2068A, K2118A, K2189A, R2193A and K2189) are implicated in RNA binding and
309 MTase activities. We also identified basic and aromatic residues that impair specifically the 2'-O-
310 MTase activities more strongly than the N-7 MTase activity. These observations suggest that the CTD
311 of SUDV L protein selectively contributes to the right positioning of RNA in the MTase catalytic site,
312 thereby regulating the different MTase activities of the protein.

313

314 DISCUSSION

315

316 The L protein of mononegavirus plays a key role in replication/transcription and RNA capping.
317 Structural and biochemical studies have highlighted its organization in five main topological domains
318 (23) that include the RdRp, PRNTase and MTase activities. However, little is known about the roles
319 of the CD and CTD. Here, we show that in SUDV, the CTD regulates the MTase activities.

320 Here, we experimentally demonstrated that in SUDV, the CTD is essential for RNA binding
321 because binding of capped RNA is impaired in its absence (**Fig. 3D**), and single mutations of
322 conserved basic residues within the CTD affect RNA binding to different extents. We also evaluated
323 the effect of CTD deletion and of single mutations within the CTD on the different MTase activities
324 of the SUDV L protein. All MTase activities are abrogated by CTD deletion as well as RNA binding.
325 In contrast, single mutations in the CTD had different effects on the SUDV MTase activities.
326 Specifically, mutations that strongly impaired RNA binding also reduced all MTase activities
327 (R2193A, R2068A, and K2189A/D). Two other mutations (K2043A, H2067D) affected mainly the

328 2'-O-MTase activities, and three mutations (H2067A, H2112A and R2172A) strongly reduced
329 particularly the internal methylation. The different effect on MTase activities are consistent with the
330 RNA binding properties of the involved mutants, suggesting that the RNA substrate must be
331 differently accommodated in order to present the target nucleotide position to the catalytic site for the
332 specific MTase activity. In other words, in some mutants, RNA binding might occur in a position that
333 does not allow just a specific MTase activity (e.g. internal 2'O). The high mobility of CD, MTase and
334 CTD observed by cryo-EM analysis of the VSV L protein also suggests that different conformations
335 might correspond to different activation states of the MTase (23). Altogether, these results suggest
336 that in SUDV, the CTD regulates RNA substrate binding and in turn, participates in the fine-tuning of
337 the MTase activities.

338 Although CTD role in such regulation is unknown, it is noteworthy that some viral MTases
339 involved in cap-methylation require protein partners for their activities. For instance, interaction of the
340 N-7 MTase of vaccinia virus (VV), which is carried by the D1 subunit, with the stimulatory subunit
341 D12 promotes its MTase activity. In this system, D12 is an allosteric activator of the N-7 MTase.
342 Indeed, structural analysis of the VV D1/D12 complex revealed that the D12 protein is on the
343 opposite side of the MTase RNA binding site, and thus is not directly involved in RNA recognition.
344 Nevertheless, the D12 protein subunit stabilizes the MTase domain and increases substrate
345 recognition (42)(43)(44). More recently, it was reported that nsp16, the 2'-O-MTase of coronaviruses
346 (SARS & MERS-CoV), interacts with nsp10, a co-factor essential for its 2'-O-MTase enzymatic
347 activity. In this case, the mechanism of MTase activation is quite different. Indeed, in the absence of
348 nsp10, nsp16 recognizes the RNA substrate with low affinity (45)(46). In the presence of nsp10,
349 nsp16 is stabilized and its catalytic and RNA binding properties are increased. By interacting with
350 nsp16, nsp10 aligns one of the catalytic lysine (K46 in SARS-CoV nsp16) in helix αZ , and extends
351 the RNA binding groove, resulting in an overall stimulation of nsp16 MTase activity (45). The SUDV
352 MTase activation mechanism might be similar to that of CoV nsp16 because our results suggest a
353 direct participation of CTD in RNA binding. Moreover, our observation extends the list of viral
354 MTases that are regulated by an additional domain or protein, suggesting that such mechanism plays
355 an important role during the life cycle of several viruses.

356 Although the CTD domain of the mononegavirus L proteins shows a large degree of diversity in
357 size and sequence, its function as a regulator of the MTase activity might be conserved. Indeed, N-7
358 methylation is essential for RNA translation into viral proteins, and cap 2'-O-methylation avoids
359 detection by RIG-I sensors (18). Besides the canonical cap-dependent MTase activities, the SUDV
360 MTase shows also an unexpected high level of internal adenosine 2'-O-MTase activity *in vitro* (37).
361 Conversely, the hMPV MTase induces mainly N-7 and 2'-O-methylation of the cap structure (30).
362 These different MTase specificities might be related to the huge CTD sequence variability in
363 *Mononegavirales*.

364 We do not know whether SUDV L induces the methylation of adenosines on its own genomic
365 RNA and/or on viral or cellular mRNA in infected cells. The role of such methylation in the viral life
366 cycle is elusive and several hypotheses, not necessarily exclusive, can be proposed. It was recently
367 reported that other viruses, such as human immunodeficiency viruses (HIV), methylate their own
368 RNA. In this case, the virus recruits the host cell MTase FTSJ3 that catalyzes the genomic RNA
369 methylation to escape detection by MDA5 (36). It might be hypothesized that SUDV and filoviruses,
370 in general, have evolved to induce epitranscriptomic RNA modifications, instead of hijacking a
371 cellular enzyme. In both cases, the resulting methylation might limit detection of viral mRNA by
372 cellular sensors at the early stage of viral infection. Internal 2'-O-methylation of RNA might also
373 have pro-viral effects by inducing resistance toward cellular 3'-exonucleases involved in RNA decay.
374 Indeed in plants, it has been shown that microRNAs and small interfering RNAs are 2'-O-methylated
375 at their 3' end and such methylation induces resistance against 3'-5' exonucleases (47)(48) or
376 stabilizes RNAs, thus limiting their degradation (49). Such mechanisms might be used also by viruses
377 to limit exonuclease-mediated RNA turnover induced by IFN (50). Internal 2'-O-methylation of viral
378 RNA might also mediate an antiviral effect by interfering with the RNA replication process. Indeed,
379 the reverse transcriptase of retroviruses is impaired by 2'-O-methylation occurring inside the RNA
380 template at low dNTP concentration (51). Similarly, VSV and Dengue virus polymerases are sensitive
381 to 2'-O-methyl presence inside their RNA template (52)(53). Thus, the regulation of the MTase
382 activities of L is probably essential for the virus life cycle, and the CTD might be involved in such
383 process in order to control the temporality and the nature of RNA methylation (*i.e.* mRNA, poly-A tail,
384 RNA). Studies by cryo-EM of C-terminal domains of mononegavirus L proteins show that the MTase
385 and CTD domains (23)(26) harbor different positions, which might reflect the switching on/off of
386 specific MTase activities regulated by the CTD domain. Additional work is required to address these
387 hypotheses. Finally, 2'-O-methylation may regulate viral protein expression during the different
388 replication steps. For instance, it was reported that internal 2'-O-methylation in cellular mRNA
389 inhibits protein translation (54). The identification of mutations in the SUDV CTD that uncouple the
390 different MTase activities of the L protein might help to elucidate the role of epitranscriptomic RNA
391 modifications, such as internal methylation, in filovirus infection. In addition, this work suggests that
392 the regulation of viral MTase is probably a key factor for efficient virus replication. Designing
393 specific inhibitors of these activities might contribute to the development of combined antiviral
394 strategies that simultaneously lead to the activation of the innate immunity system, especially through
395 the RIG-I/MDA5 pathway, and to inhibition of viral replication.

396

397

398

399

400 **ACKNOWLEDGEMENT**

401 We thank Prof. Xavier De Lamballerie and Dr. Yannick Boehmann for their help in the initiation
402 of this study.

403

404 **FUNDING**

405 The program was initiated in a research program funded by the Délégation Générale pour
406 l'Armement (DGA) (grant 2009.34.0038)/ Aix-Marseille Université PhD fellowship for B.M. CV.
407 was funded by National Research Agency ANR under the program ANR Rab-cap (ANR-
408 16_CE11_0031_01).

409

410 **Author Contributions.** C.V, B.M., B.Co. and E.D. conceived and designed the study. C.V, B.M.,
411 B.Co. and E.D. performed experiments, using material prepared by F.D. and J-J.V. C.V, B.M., B.Co.
412 and E.D. analysed data and wrote the paper. B.Ca. contributed to the writing of the paper. I.I.
413 participated in the review of the paper.

414

415 **REFERENCES**

416

- 417 1. Shultz JM, Espinel Z, Espinola M, Rechkemmer A. 2016. Distinguishing epidemiological
418 features of the 2013-2016 West Africa Ebola virus disease outbreak. *Disaster Heal* 3:78–88.
- 419 2. 2019. Ebola | Ebola situation reports: Democratic Republic of the Congo. WHO.
- 420 3. Ebola DRC KIVU 2018 Dashboard redone.
- 421 4. Henao-Restrepo AM, Longini IM, Egger M, Dean NE, Edmunds WJ, Camacho A, Carroll
422 MW, Doumbia M, Draguez B, Duraffour S, Enwere G, Grais R, Gunther S, Hossmann S,
423 Kondé MK, Kone S, Kuisma E, Levine MM, Mandal S, Norheim G, Riveros X, Soumah A,
424 Trelle S, Vicari AS, Watson CH, Kéïta S, Kieny MP, Røttingen J-A. 2015. Efficacy and
425 effectiveness of an rVSV-vectored vaccine expressing Ebola surface glycoprotein: interim
426 results from the Guinea ring vaccination cluster-randomised trial. *Lancet* 386:857–866.
- 427 5. Wells CR, Pandey A, Parpia AS, Fitzpatrick MC, Meyers LA, Singer BH, Galvani AP. 2019.
428 Ebola vaccination in the Democratic Republic of the Congo. *Proc Natl Acad Sci*
429 116:201817329.
- 430 6. Saphire EO, Schendel SL, Gunn BM, Milligan JC, Alter G. 2018. Antibody-mediated
431 protection against Ebola virus. *Nat Immunol* 19:1169–1178.
- 432 7. Kuhn JH, Becker S, Ebihara H, Geisbert TW, Johnson KM, Kawaoka Y, Lipkin WI, Negrodo
433 AI, Netesov S V, Nichol ST, Palacios G, Peters CJ, Tenorio A, Volchkov VE, Jahrling PB.
434 2010. Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa
435 and viruses, and virus abbreviations. *Arch Virol* 155:2083–103.
- 436 8. Sanchez A, Kiley MP, Holloway BP, Auperin DD. 1993. Sequence analysis of the Ebola virus

- 437 genome: organization, genetic elements, and comparison with the genome of Marburg virus.
438 *Virus Res* 29:215–240.
- 439 9. Elliott LH, Sanchez A, Holloway BP, Kiley MP, McCormick JB. 1993. Ebola protein analyses
440 for the determination of genetic organization. *Arch Virol* 133:423–36.
- 441 10. Mühlberger E, Trommer S, Funke C, Volchkov V, Klenk HD, Becker S. 1996. Termini of all
442 mRNA species of Marburg virus: sequence and secondary structure. *Virology* 223:376–80.
- 443 11. Weik M, Modrof J, Klenk H-D, Becker S, Mühlberger E. 2002. Ebola virus VP30-mediated
444 transcription is regulated by RNA secondary structure formation. *J Virol* 76:8532–9.
- 445 12. Whelan SPJ, Barr JN, Wertz GW. 2004. Transcription and replication of nonsegmented
446 negative-strand RNA viruses. *Curr Top Microbiol Immunol* 283:61–119.
- 447 13. Jordan PC, Liu C, Raynaud P, Lo MK, Spiropoulou CF, Symons JA, Beigelman L, Deval J.
448 2018. Initiation, extension, and termination of RNA synthesis by a paramyxovirus polymerase.
449 *PLOS Pathog* 14:e1006889.
- 450 14. Tekes G, Rahmeh AA, Whelan SPJ. 2011. A freeze frame view of vesicular stomatitis virus
451 transcription defines a minimal length of RNA for 5' processing. *PLoS Pathog* 7:e1002073.
- 452 15. Ogino T, Banerjee AK. 2007. Unconventional mechanism of mRNA capping by the RNA-
453 dependent RNA polymerase of vesicular stomatitis virus. *Mol Cell* 25:85–97.
- 454 16. Shatkin AJ. 1976. Capping of eucaryotic mRNAs. *Cell* 9:645–53.
- 455 17. Schuberth-Wagner C, Ludwig J, Bruder AK, Herzner A-M, Zillinger T, Goldeck M, Schmidt
456 T, Schmid-Burgk JL, Kerber R, Wolter S, Stümpel J-P, Roth A, Bartok E, Drosten C, Coch C,
457 Hornung V, Barchet W, Kümmerer BM, Hartmann G, Schlee M. 2015. A Conserved Histidine
458 in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2'-O-Methylated Self RNA.
459 *Immunity* 43:41–51.
- 460 18. Devarkar SC, Wang C, Miller MT, Ramanathan A, Jiang F, Khan AG, Patel SS, Marcotrigiano
461 J. 2016. Structural basis for m7G recognition and 2'-O-methyl discrimination in capped RNAs
462 by the innate immune receptor RIG-I. *Proc Natl Acad Sci* 113:596–601.
- 463 19. Johnson B, VanBlargan LA, Xu W, White JP, Shan C, Shi P-Y, Zhang R, Adhikari J, Gross
464 ML, Leung DW, Diamond MS, Amarasinghe GK. 2018. Human IFIT3 Modulates IFIT1 RNA
465 Binding Specificity and Protein Stability. *Immunity* 48:487-499.e5.
- 466 20. Choi YJ, Bowman JW, Jung JU. 2018. A Talented Duo: IFIT1 and IFIT3 Patrol Viral RNA
467 Caps. *Immunity* 48:474–476.
- 468 21. Shuman S. 2002. What messenger RNA capping tells us about eukaryotic evolution. *Nat Rev*
469 *Mol Cell Biol* 3:619–625.
- 470 22. Züst R, Cervantes-Barragan L, Habjan M, Maier R, Neuman BW, Ziebuhr J, Szretter KJ,
471 Baker SC, Barchet W, Diamond MS, Siddell SG, Ludewig B, Thiel V. 2011. Ribose 2'-O-
472 methylation provides a molecular signature for the distinction of self and non-self mRNA
473 dependent on the RNA sensor Mda5. *Nat Immunol* 12:137–43.

- 474 23. Liang B, Li Z, Jenni S, Rahmeh AA, Morin BM, Grant T, Grigorieff N, Harrison SC, Whelan
475 SPJ. 2015. Structure of the L Protein of Vesicular Stomatitis Virus from Electron
476 Cryomicroscopy. *Cell* 162:314–327.
- 477 24. Gilman MSA, Liu C, Fung A, Behera I, Jordan P, Rigaux P, Ysebaert N, Tcherniuk S,
478 Sourimant J, Eléouët J-F, Sutto-Ortiz P, Decroly E, Roymans D, Jin Z, McLellan JS. 2019.
479 Structure of the Respiratory Syncytial Virus Polymerase Complex. *Cell*.
- 480 25. Pan J, Qian X, Lattmann S, El Sahili A, Yeo TH, Jia H, Cressey T, Ludeke B, Noton S,
481 Kalocsay M, Fearn R, Lescar J. 2020. Structure of the human metapneumovirus polymerase
482 phosphoprotein complex. *Nature* 577:275–279.
- 483 26. Abdella R, Aggarwal M, Okura T, Lamb RA, He Y. 2020. Structure of a paramyxovirus
484 polymerase complex reveals a unique methyltransferase-CTD conformation. *Proc Natl Acad Sci*
485 *Sci* 117:4931–4941.
- 486 27. Rahmeh AA, Schenk AD, Danek EI, Kranzusch PJ, Liang B, Walz T, Whelan SPJ. 2010.
487 Molecular architecture of the vesicular stomatitis virus RNA polymerase. *Proc Natl Acad Sci*
488 *U S A* 107:20075–80.
- 489 28. Selisko B, Potosopon S, Agred R, Priet S, Varlet I, Thillier Y, Sallamand C, Debart F, Vasseur
490 J-J, Canard B. 2012. Molecular Basis for Nucleotide Conservation at the Ends of the Dengue
491 Virus Genome. *PLoS Pathog* 8:e1002912.
- 492 29. Liang B, Li Z, Jenni S, Rahmeh AA, Morin BM, Grant T, Grigorieff N, Harrison SC, Whelan
493 SPJ. 2015. Structure of the L Protein of Vesicular Stomatitis Virus from Electron
494 Cryomicroscopy. *Cell* 162:314–327.
- 495 30. Paesen GC, Collet A, Sallamand C, Debart F, Vasseur J-J, Canard B, Decroly E, Grimes JM.
496 2015. X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat*
497 *Commun* 6:8749.
- 498 31. Egloff M-P, Benarroch D, Selisko B, Romette J-L, Canard B. 2002. An RNA cap (nucleoside-
499 2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and
500 functional characterization. *EMBO J* 21:2757–68.
- 501 32. Ray D, Shah A, Tilgner M, Guo Y, Zhao Y, Dong H, Deas TS, Zhou Y, Li H, Shi P-Y. 2006.
502 West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O
503 methylations by nonstructural protein 5. *J Virol* 80:8362–70.
- 504 33. Zhou Y, Ray D, Zhao Y, Dong H, Ren S, Li Z, Guo Y, Bernard KA, Shi P-Y, Li H. 2007.
505 Structure and Function of Flavivirus NS5 Methyltransferase. *J Virol* 81:3891–3903.
- 506 34. Rahmeh AA, Li J, Kranzusch PJ, Whelan SPJ. 2009. Ribose 2'-O Methylation of the Vesicular
507 Stomatitis Virus mRNA Cap Precedes and Facilitates Subsequent Guanine-N-7 Methylation
508 by the Large Polymerase Protein. *J Virol* 83:11043–11050.
- 509 35. Martin B, Canard B, Decroly E. 2017. Filovirus proteins for antiviral drug discovery:
510 Structure/function bases of the replication cycle. *Antiviral Res* 141:48–61.

- 511 36. Ringeard M, Marchand V, Decroly E, Motorin Y, Bennasser Y. 2019. FTSJ3 is an RNA 2'-O-
512 methyltransferase recruited by HIV to avoid innate immune sensing. *Nature* 565:500–504.
- 513 37. Martin B, Coutard B, Guez T, Paesen GC, Canard B, Debart F, Vasseur J-J, Grimes JM,
514 Decroly E. 2018. The methyltransferase domain of the Sudan ebolavirus L protein specifically
515 targets internal adenosines of RNA substrates, in addition to the cap structure. *Nucleic Acids*
516 *Res* 46:7902–7912.
- 517 38. Barral K, Sallamand C, Petzold C, Coutard B, Collet A, Thillier Y, Zimmermann J, Vasseur J-
518 J, Canard B, Rohayem J, Debart F, Decroly E. 2013. Development of specific dengue virus 2'-
519 O- and N7-methyltransferase assays for antiviral drug screening. *Antiviral Res* 99:292–300.
- 520 39. Thillier Y, Decroly E, Morvan F, Canard B, Vasseur J-J, Debart F. 2012. Synthesis of 5' cap-0
521 and cap-1 RNAs using solid-phase chemistry coupled with enzymatic methylation by human
522 (guanine-N⁷)-methyl transferase. *RNA* 18:856–68.
- 523 40. Bollati M, Milani M, Mastrangelo E, Ricagno S, Tedeschi G, Nonnis S, Decroly E, Selisko B,
524 de Lamballerie X, Coutard B, Canard B, Bolognesi M. 2009. Recognition of RNA cap in the
525 Wesselsbron virus NS5 methyltransferase domain: implications for RNA-capping mechanisms
526 in Flavivirus. *J Mol Biol* 385:140–52.
- 527 41. Pantoliano MW, Petrella EC, Kwasnoski JD, Lobanov VS, Myslik J, Graf E, Carver T, Asel E,
528 Springer BA, Lane P, Salemme FR. 2001. High-density miniaturized thermal shift assays as a
529 general strategy for drug discovery. *J Biomol Screen* 6:429–40.
- 530 42. Mao X, Shuman S. 1994. Intrinsic RNA (guanine-7) methyltransferase activity of the vaccinia
531 virus capping enzyme D1 subunit is stimulated by the D12 subunit. Identification of amino
532 acid residues in the D1 protein required for subunit association and methyl group transfer. *J*
533 *Biol Chem* 269:24472–9.
- 534 43. De la Peña M, Kyrieleis OJP, Cusack S. 2007. Structural insights into the mechanism and
535 evolution of the vaccinia virus mRNA cap N7 methyl-transferase. *EMBO J* 26:4913–25.
- 536 44. Higman MA, Christen LA, Niles EG. 1994. The mRNA (guanine-7)-methyltransferase domain
537 of the vaccinia virus mRNA capping enzyme. Expression in *Escherichia coli* and structural and
538 kinetic comparison to the intact capping enzyme. *J Biol Chem* 269:14974–81.
- 539 45. Chen Y, Su C, Ke M, Jin X, Xu L, Zhang Z, Wu A, Sun Y, Yang Z, Tien P, Ahola T, Liang Y,
540 Liu X, Guo D. 2011. Biochemical and Structural Insights into the Mechanisms of SARS
541 Coronavirus RNA Ribose 2'-O-Methylation by nsp16/nsp10 Protein Complex. *PLoS Pathog*
542 7:e1002294.
- 543 46. Aouadi W, Blanjoie A, Vasseur J-J, Debart F, Canard B, Decroly E. 2017. Binding of the
544 Methyl Donor S-Adenosyl-L-Methionine to Middle East Respiratory Syndrome Coronavirus
545 2'-O-Methyltransferase nsp16 Promotes Recruitment of the Allosteric Activator nsp10. *J Virol*
546 91.
- 547 47. Ramachandran V, Chen X. 2008. Degradation of microRNAs by a Family of Exoribonucleases

- 548 in *Arabidopsis*. *Science* (80-) 321:1490–1492.
- 549 48. Kirino Y, Mourelatos Z. 2007. Mouse Piwi-interacting RNAs are 2'-O-methylated at their 3'
- 550 termini. *Nat Struct Mol Biol* 14:347–348.
- 551 49. Prusiner P, Yathindra N, Sundaralingam M. 1974. Effect of ribose O(2')-methylation on the
- 552 conformation of nucleosides and nucleotides. *Biochim Biophys Acta - Nucleic Acids Protein*
- 553 *Synth* 366:115–123.
- 554 50. Cooper DA, Jha BK, Silverman RH, Hesselberth JR, Barton DJ. 2014. Ribonuclease L and
- 555 metal-ion-independent endoribonuclease cleavage sites in host and viral RNAs. *Nucleic Acids*
- 556 *Res* 42:5202–16.
- 557 51. Motorin Y, Muller S, Behm-Ansmant I, Branlant C. 2007. Identification of Modified Residues
- 558 in RNAs by Reverse Transcription-Based Methods. *Methods Enzymol* 425:21–53.
- 559 52. Morin B, Whelan SPJ. 2014. Sensitivity of the polymerase of vesicular stomatitis virus to 2'
- 560 substitutions in the template and nucleotide triphosphate during initiation and elongation. *J*
- 561 *Biol Chem* 289:9961–9.
- 562 53. Dong H, Chang DC, Hua MHC, Lim SP, Chionh YH, Hia F, Lee YH, Kukkaro P, Lok S-M,
- 563 Dedon PC, Shi P-Y. 2012. 2'-O methylation of internal adenosine by flavivirus NS5
- 564 methyltransferase. *PLoS Pathog* 8:e1002642.
- 565 54. Elliott BA, Ho H-T, Ranganathan S V., Vangaveti S, Ilkayeva O, Abou Assi H, Choi AK,
- 566 Agris PF, Holley CL. 2019. Modification of messenger RNA by 2'-O-methylation regulates
- 567 gene expression in vivo. *Nat Commun* 10:3401.
- 568
- 569

570 **Figure 1 – Bioinformatic analysis of the C-terminal domains (MTase+CTD) of mononegavirus L**

571 **protein based on hMPV and VSV structures and alignment of the filovirus CTD sequences**

572 (A) Domain organization of the mononegavirus L protein with the six conserved regions (CRI to CRV,

573 black boxes and CRVI, grey box) that contain motifs important for the different activities of the L

574 protein (motifs mapped with asterisks). The *Sudan ebolavirus* (SUDV) methyltransferase domain

575 (MTase) encompasses amino acids 1693-2036, based on the alignment with the vesicular stomatitis

576 virus (VSV) L protein. The SAM-binding site (GxGxG) and the 2'-O catalytic tetrad K-D-K-E are

577 also shown (asterisks). The C-terminal domain (CTD) of SUDV L follows the MTase domain (amino

578 acids 2037–2210). (B) X-ray structure of human metapneumovirus (hMPV) MTase (grey) and CTD

579 (blue) domains at a resolution of 2.2 Å (27) (PDB: 4UCZ). The residues of the catalytic tetrad and of

580 the SAM-binding site are in red and yellow, respectively. Aromatic and basic residues are in purple.

581 WebLogo (52) was used to illustrate the degree of amino acid conservation based on previously

582 published alignments (36). (C) Structure of VSV MTase (grey) (PDB: 5A22) and CTD (blue) (22).

583 The residues involved in the catalytic tetrad and SAM-binding site are in red and yellow, respectively.

584 Aromatic and basic residues are in purple. (D) CTD domains in the following L protein sequences

585 selected from the NCBI Protein Data Bank: EBOV (*Zaire ebolavirus*, AAG40171.1), SUDV (*Sudan*
586 *ebolavirus*, YP_138527.1), TAFV (*Tai Forest ebolavirus*, ALT19766.1), BDBV (*Bundibugyo*
587 *ebolavirus*, AKB09568.1), RESTV (*Reston ebolavirus*, APA16576.1), MARV (*Marburg virus*,
588 CAA82542.1), LLOV (*Lloviu cuevavirus*, YP_004928143.1), and MLAV (*Mengla dianlovirus*,
589 AZL87829.1). * indicates the strain under study. The alignment was generated with Seaview and
590 analyzed with ESPript. The numbers on top of the alignment indicate the amino acid positions in the
591 SUDV sequence. Strictly conserved aromatic and basic amino acids are highlighted in purple, and
592 arginine or lysine basic residues are highlighted in orange. Mutations leading to soluble proteins are
593 indicated by ▲, and mutations leading to insoluble proteins are indicated by ▲. Spirals just above the
594 alignment indicate the position of α helices, based on the secondary structure prediction (*NPS@: PHD*
595 *secondary structure prediction*).

596

597 **Figure 2 – Production and purification of recombinant SUDV MTase+CTD and MTase- Δ CTD,**
598 **and CTD effect on MTase activity and RNA binding**

599 (A) Recombinant Sudan ebolavirus (SUDV) MTase+CTD (58 kDa) and MTase- Δ CTD (39 kDa) were
600 purified by affinity chromatography, then separated by SDS-PAGE before Coomassie blue staining.
601 (B) Thermal shift assay of recombinant SUDV MTase+CTD (dark grey) and MTase- Δ CTD (light
602 grey). The melting temperatures (T_m) were calculated using the Boltzmann sigmoidal regression and
603 were evaluated at 54.6°C and 46°C, respectively. Raw data are shown. (C) MTase activity evaluation
604 of SUDV MTase+CTD (grey) and MTase- Δ CTD (dark) incubated with a 13-mer capped RNA
605 substrate that mimics the capped 5' extremity of viral mRNA (GpppG-SUDV12). MTase activity was
606 determined by filter binding assay. Data are the mean \pm standard deviation (n=6). (D) RNA-binding
607 analysis by fluorescence polarization using a 13-mer capped SUDV RNA substrate labeled at the
608 3' end pCp-Cy5 and SUDV MTase- Δ CTD (light grey) and MTase+CTD (dark grey). The dissociation
609 coefficient (Kd) was calculated by site-specific binding regression analysis with Hill slope.
610 MTase+CTD affinity for GpppG-SUDV12 was 740 ± 108 nM. Data are the mean \pm standard
611 deviation (n=3).

612

613 **Figure 3 - Effect of single mutations of CTD conserved residues on MTase activities of Sudan**
614 **ebolavirus MTase+CTD**

615 (A) Wild type (WT) and mutated SUDV MTase+CTD (58 kDa) were purified by affinity
616 chromatography, then separated by SDS-PAGE analysis before Coomassie blue staining. (B) MTase
617 activities of SUDV WT and mutated MTase+CTD incubated with GpppGm(Am)-SUDV12,
618 mGpppA(Am)SUDV12, or mGpppGm-SUDV12 and a radiolabeled methyl donor (3H-SAM), to
619 assess the cap N-7 MTase (black), cap 2'-O-MTase (light grey), and internal 2'-O-A MTase (dark
620 grey) activities, respectively, by filter binding assay. Data are the mean \pm standard deviation (n=3);

621 * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (two-way ANOVA and multiple comparison Tukey test, WT vs.
622 mutation).

623

624 **Table 1 – List of synthetic RNAs**

625 m^7X : nucleotide methylated on position N7 of the base

626 X_m : nucleotide methylated on position 2' O of the ribose

627 **Table 2 – Isoelectric point analysis of mononegavirus MTase and CTD domains**

628 Based on the alignment of different viruses belonging to the Mononegavirales order, the residue
629 content of the methyltransferase (MTase) and C-terminal (CTD) domains of their L protein was
630 analyzed. For each virus, the MTase, CTD and MTase+CTD total number of residues was determined.
631 The isoelectric points (pI) were estimated with ProtParam. The difference between the MTase and
632 CTD pIs were calculated and modeled by a blue gauge normalized to the greatest difference

633

634 **Table 3 - RNA binding assay of catalytic mutants**

635 RNA-binding assays were performed by fluorescence polarization using a 13-mer SUDV-specific
636 capped RNA (GpppG-SUDV₁₂) labeled at the 3' end with pCp-Cy5 and wild type (WT) SUDV
637 MTase+CTD and the indicated catalytic mutant MTase+CTD proteins. The dissociation coefficient
638 (K_d) was calculated by site-specific binding regression with Hill slope. Data are the mean (\pm standard
639 deviation) (n=3). ND: not determined.

640

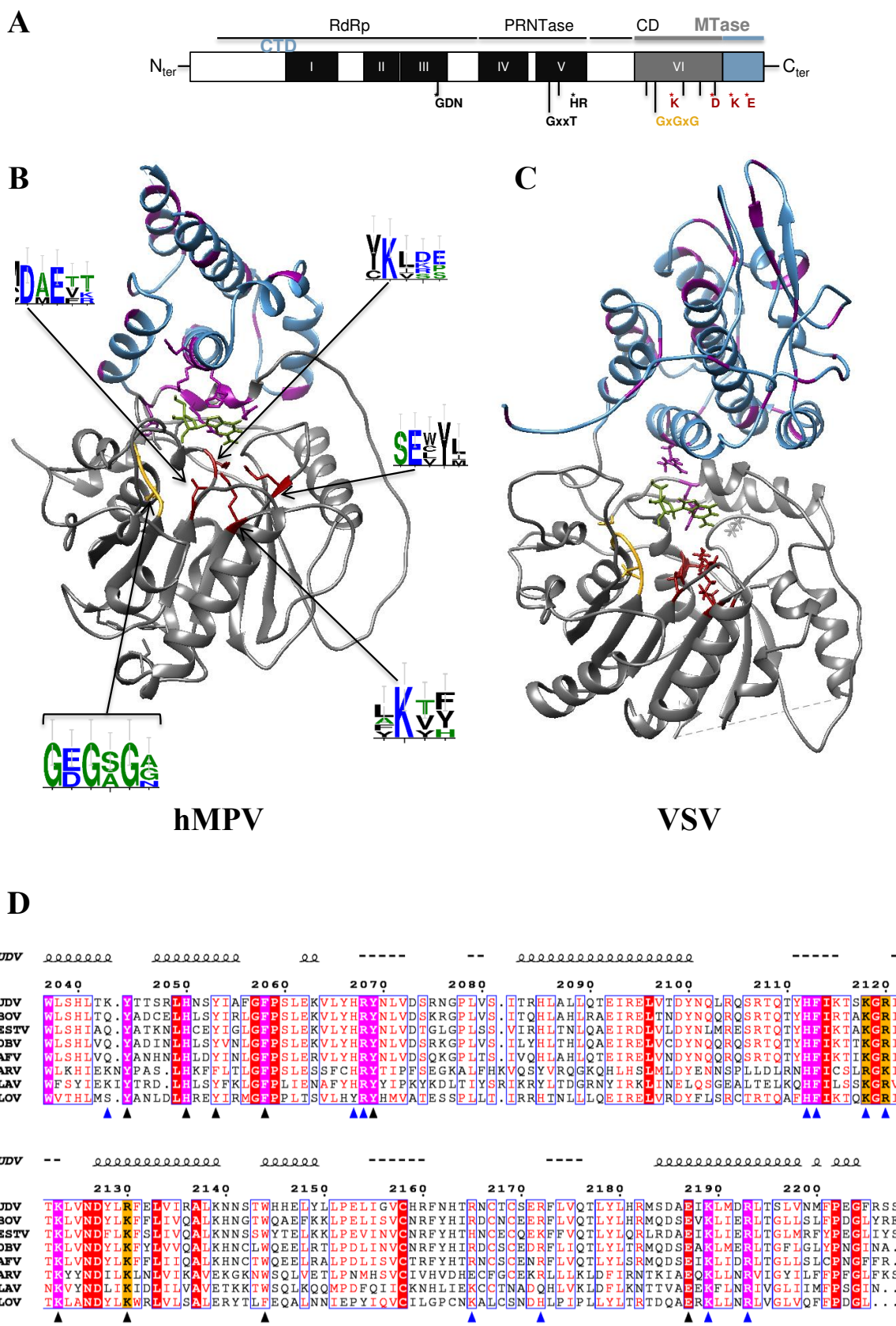
641 **Table 4 – Effects of single mutations of CTD conserved residues on SUDV MTase+CTD stability
642 and RNA binding**

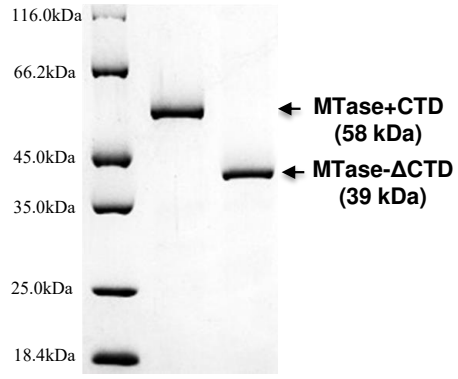
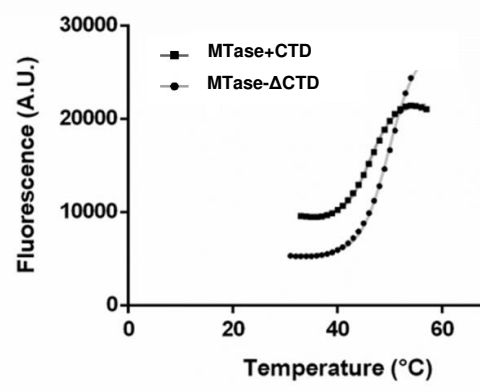
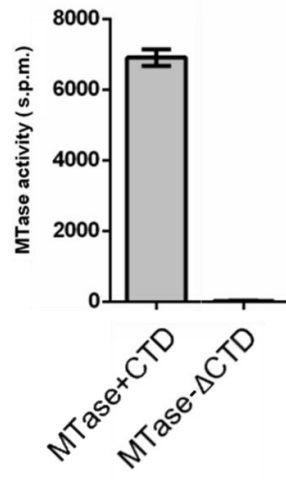
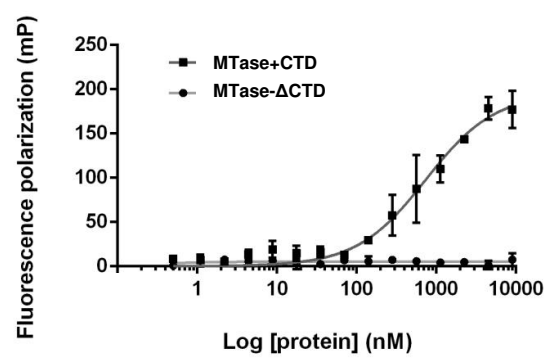
643 Thermostability analysis by thermal shift assay of wild type (WT) *Sudan ebolavirus* (SUDV)
644 MTase+CTD and the indicated mutated proteins. Melting temperatures (T_m) were calculated by
645 Boltzmann sigmoidal regression. T_m data are the mean (\pm standard deviation) of n=2. ΔT_m
646 corresponds to the difference between the T_m of the mutant and WT proteins. RNA-binding was tested
647 by fluorescence polarization using 13-mer SUDV-specific capped RNAs (GpppG-SUDV₁₂ and
648 m^7GpppG_m -SUDV₁₂) labeled at the 3' end with pCp-Cy5 and SUDV MTase+CTD and mutated
649 MTase+CTD proteins. The dissociation coefficient (K_d) was calculated by site-specific binding
650 regression with Hill slope. Data are the mean (\pm standard deviation) of n=3.

651

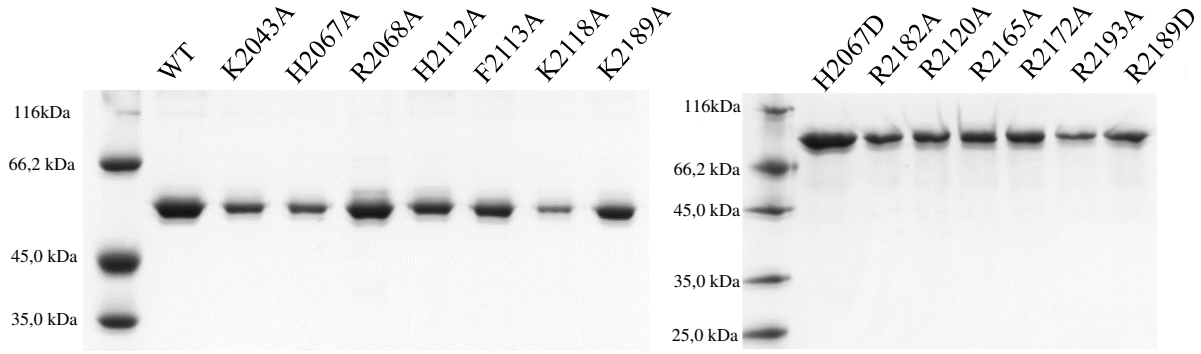
652

653



A**B****C****D**

A



B

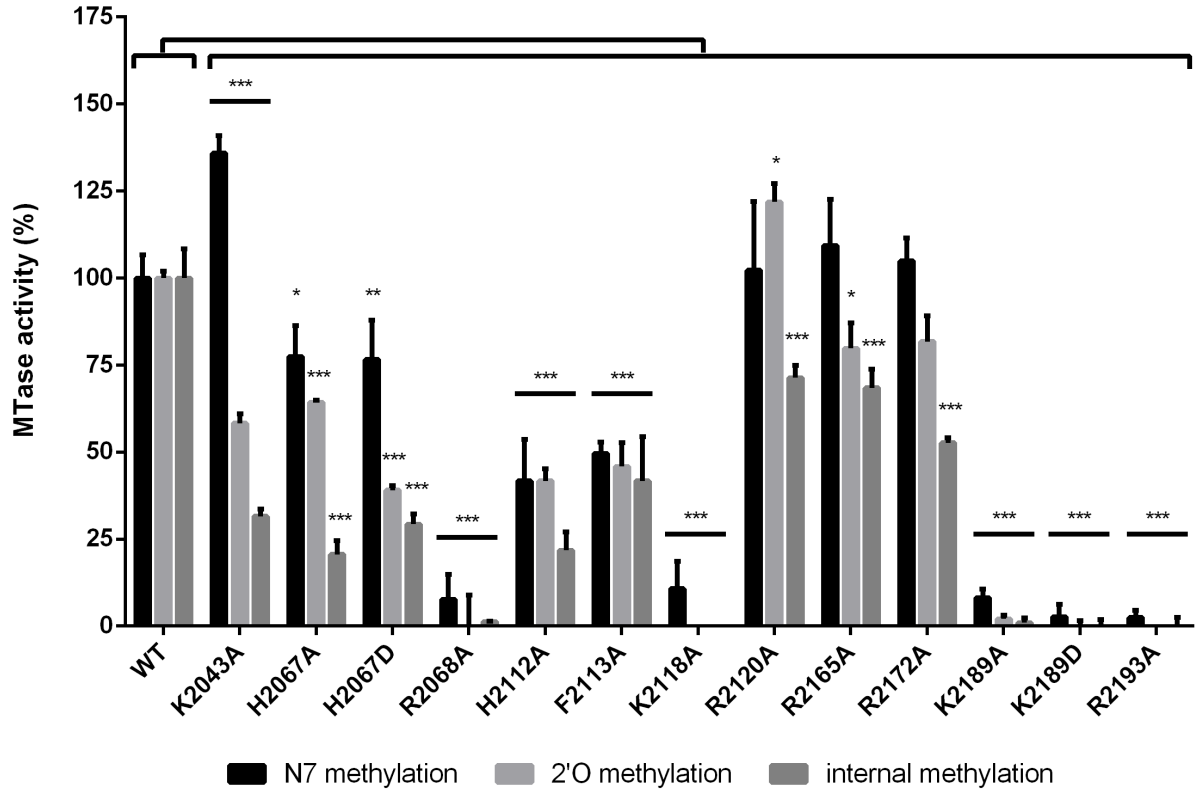


Table 1 – List of synthetic RNAs

RNA	Sequence	Origin
GpppG-SUDV ₁₂	GpppGAUGAAGAUUAAG	chemical synthesis
GpppG _m (A _m)-SUDV ₁₂	GpppG _m A _m UGA _m A _m GA _m UUA _m A _m G	chemical synthesis
^m GpppG _m -SUDV ₁₂	^m GpppG _m AUGAAGAUUAAG	chemical synthesis
^m GpppA(A _m)-SUDV ₁₂	^m GpppAUGA _m UGA _m A _m GA _m UUA _m	chemical synthesis

^mX : nucleotide methylated on position N7 of the base

X_m : nucleotide methylated on position 2'O of the ribose

Order	Family	Genus	Species	Domain	Total residue number	pI	Δ pI
Mononegavirales	Filoviridae	Ebolavirus	EBOV (AAG40171.1)	MTase	363	7,73	1,46
				CTD	171	9,19	
				MTase-CTD	534	8,81	
			SUDV (YP_138527.1)	MTase	346	6,49	3,22
				CTD	172	9,71	
				MTase-CTD	518	8,68	
			TAFV (ALT19766.1)	MTase	354	8,50	0,66
				CTD	169	9,16	
				MTase-CTD	523	8,93	
			BDBV (AKB09568.1)	MTase	354	7,66	1,12
				CTD	169	8,78	
				MTase-CTD	523	8,53	
			RESTV (APA16576.1)	MTase	351	7,20	1,94
				CTD	173	9,14	
				MTase-CTD	524	8,65	
	Marburgvirus	MARV (CAA82542.1)	MTase	445	8,48	0,88	
			CTD	176	9,36		
			MTase-CTD	621	8,95		
	Cuevavirus	LLOV (YP_004928143.1)	MTase	339	8,55	0,86	
			CTD	167	9,41		
			MTase-CTD	506	8,99		
	Paramyxoviridae	Respirovirus	SeV (AAB06283.1)	MTase	343	5,69	3,85
				CTD	205	9,54	
				MTase-CTD	548	7,26	
		Morbillivirus	MeV (BAB60955.1)	MTase	338	7,78	2,16
				CTD	165	9,94	
				MTase-CTD	503	9,32	
Rubulavirus		MuV (BAA01432.1)	MTase	340	4,95	0,49	
			CTD	227	5,44		
			MTase-CTD	567	5,13		
Henipavirus		HeV (O89344.3)	MTase	342	6,02	3,61	
			CTD	165	9,63		
			MTase-CTD	503	8,62		
Pneumoviridae	Metapneumovirus	hMPV (Q91L20)	MTase	285	8,57	0,92	
			CTD	124	9,49		
			MTase-CTD	409	8,97		
Orthopneumovirus	hRSV (AAx23996.1)	MTase	323	8,94	0,48		
		CTD	127	9,42			
		MTase-CTD	450	9,14			
Rhabdoviridae	Vesiculovirus	VSV (5A22_A)	MTase	317	5,81	3,54	
			CTD	197	9,35		
			MTase-CTD	514	7,58		
	Lyssavirus	RABV (ABZ81226.1)	MTase	285	6,98	1,36	
CTD			217	8,34			
			MTase-CTD	502	8,02		

Table 2 – Isoelectric point analysis of mononegavirus MTase and CTD domains

Based on the alignment of different viruses belonging to the *Mononegavirales* order, the residue content of the methyltransferase (MTase) and C-terminal (CTD) domains of their L protein was analyzed. For each virus, the MTase, CTD and MTase+CTD total number of residues was determined. The isoelectric points (pI) were estimated with ProtParam. The difference between the MTase and CTD pIs were calculated and modeled by a blue gauge normalized to the greatest difference.

	Fluorescence polarization
	GpppG-SUDV ₁₂
	K _d (μM)
WT	0.740 (± 0.108)
K1813A	4.6 (± 0.943)
D1924A	0.581 (± 0.05)
K1959A	ND
E1994A	1.3 (± 0.206)

Table 3 - RNA binding assay of catalytic mutants

RNA-binding assays were performed by fluorescence polarization using a 13-mer SUDV-specific capped RNA (GpppG-SUDV₁₂) labeled at the 3'end with pCp-Cy5 and wild type (WT) SUDV MTase+CTD and the indicated catalytic mutant MTase+CTD proteins. The dissociation coefficient (K_d) was calculated by site-specific binding regression with Hill slope. Data are the mean (± standard deviation) (n=3). ND: not determined.

List of mutants*	Thermal Shift Assay		Fluorescence polarization	
	T _m (°C)	ΔT _m	GpppG-SUDV ₁₂	^m GpppG _m -SUDV ₁₂
			K _d (μM)	K _d (μM)
WT	54.3 (± 0.13)	-	0.740 (± 0.108)	0.570 (± 0.110)
K2043A	55.7 (± 0.03)	1.4	2 (± 0.280)	> 9
H2067A	54.7 (± 0.58)	0.4	5 (± 1.5)	4.4 (± 1.5)
H2067D	53.1 (± 0.70)	-0.7	3.5 (± 1)	6.5 (± 2.3)
R2068A	50.6 (± 0.01)	-3.7	> 9	> 9
H2112A	54.0 (± 0.39)	-0.3	4.5 (± 1.1)	5.4 (± 2.1)
F2113A	53.5 (± 0.33)	-0.8	6.6 (± 1.6)	3.1 (± 0.567)
K2118A	53.7 (± 0.24)	-0.6	> 9	> 9
R2120A	51.3 (± 0.16)	-3	2.8 (± 0.649)	2.2 (± 0.705)
R2165A	53.05 (± 0.63)	-1.2	1 (± 0.238)	1.4 (± 0.254)
R2172A	51.2 (± 0.41)	-3.1	3.1 (± 0.988)	1.7 (± 0.835)
K2189A	55.0 (± 0.83)	0.7	1.9 (± 0.507)	8.6 (± 2.8)
K2189D	53.85 (± 0.28)	-0.45	2.4 (± 0.509)	1.3 (± 0.478)
R2193A	51.9 (± 0.70)	-2.4	> 9	7.7 (± 2.7)

* Mutations that led to insoluble proteins: F2113D, K2043D, W2144A, R2130A, Y2053A, Y2044A, Δ-CTD(2105), H2050A, F2058A, Y2069A, K2123A, and E2187A

Table 4 – Effects of single mutations of CTD conserved residues on SUDV MTase+CTD stability and RNA binding

Thermostability analysis by thermal shift assay of wild type (WT) *Sudan ebolavirus* (SUDV) MTase+CTD) and the indicated mutated proteins. Melting temperatures (T_m) were calculated by Boltzmann sigmoidal regression. T_m data are the mean (± standard deviation) of n=2. ΔT_m corresponds to the difference between the T_m of the mutant and WT proteins. RNA-binding was tested by fluorescence polarization using 13-mer SUDV-specific capped RNAs (GpppG-SUDV₁₂ and ^mGpppG_m-SUDV₁₂) labeled at the 3' end with pCp-Cy5 and SUDV MTase+CTD and mutated MTase+CTD proteins. The dissociation coefficient (K_d) was calculated by site-specific binding regression with Hill slope. Data are the mean (± standard deviation) of n=3.

2.3. Etude structure/fonction du domaine MTase de la protéine L du virus Soudan

Article 3: New insights into structural features of Ebolavirus methyltransferase activities (Draft).

Auteurs : Coralie Valle*, Baptiste Martin*, François Ferron, Véronique Zamboni, Aline Desmyter, Françoise Debart, Jean-Jacques Vasseur, Bruno Canard, Bruno Coutard, Etienne Decroly

**Contribution égale*

En cours de préparation.

Au cours de ces cinq dernières années, de nombreuses structures de la protéine L de virus appartenant à l'ordre des *Mononegavirales*, tels que VSV, hMPV, RSV, RABV, ont pu être résolues, ouvrant la voie vers des études structure/fonction pertinentes. Cependant aucune donnée fonctionnelle ni structurale ne sont encore documentées chez les filovirus. Les études précédentes ont permis de caractériser l'activité enzymatique du domaine MTase-CTD de la protéine L de SUDV révélant, pour la première fois chez un virus de l'ordre des *Mononegavirales*, une activité MTase capable d'induire des méthylations internes des ARNs. Une étude de délétion de domaine a permis de caractériser le rôle du domaine CTD dans le recrutement de l'ARN et l'activité MTase de SUDV. De plus, une étude de mutagenèse dirigée nous a permis d'identifier des résidus clefs dans le domaine CTD régulant les différentes activités MTases portées par l'extrémité C-terminale de la protéine L de SUDV. Afin de compléter les études fonctionnelles, une campagne de cristallogénèse a été conduite. Dans cette étude, la structure du domaine MTase de SUDV en complexe avec une VHH a été résolue par cristallographie aux rayons X. La structure révèle que le domaine MTase SUDV adopte un repliement de type "Rossmann", typique des MTases catalysant le transfert de groupement méthyl à partir du SAM. De plus, les résidus catalytiques sont exposés à la surface de la structure confirmant l'implication du domaine CTD dans le recrutement de l'ARN. De façon surprenante, le domaine MTase de SUDV présente une hélice supplémentaire absente des structures MTases des virus VSV, hMPV et RABV. Enfin, l'analyse de la structure nous a permis d'identifier des résidus pouvant être impliqués dans le recrutement de l'ARN ou du SAM. Une étude de mutagenèse dirigée nous a permis de mettre en évidence le rôle de certains résidus dans la régulation des différentes activités MTases de SUDV. Cette étude nous a permis de compléter la caractérisation fonctionnelle des domaines C-terminaux de la protéine L de SUDV en apportant des éléments structuraux permettant

d'établir des relations structure/fonction pertinentes. De plus, la résolution de la structure du domaine MTase SUDV permet d'ouvrir la voie vers de nouvelles perspectives de criblage de molécules à potentiel antiviral.

New insights into structural features of Ebolavirus methyltransferase activities

Coralie Valle^{1#}, Baptiste Martin^{1#}, François Ferron¹, Véronique Roig-Zamboni¹, Aline Desmyter¹, Françoise Debart², Jean-Jacques Vasseur², Bruno Canard¹, Bruno Coutard³ & Etienne Decroly^{1*}

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

² IBMM, UMR 5247 CNRS, Université de Montpellier, ENSCM, Montpellier, France

³ Unité des Virus Émergents (UVE: Aix-Marseille Univ-IRD 190-Inserm, 1207-IHU Méditerranée Infection) Marseille, France

These authors contributed equally to this work.

* To whom correspondence should be addressed. Tel: +33 4 91 82 86 47; Fax: 04 91 82 86 38; E-mail: etienne.decroly@afmb.univ-mrs.fr

ABSTRACT

Since its discovery in the 70's, Ebola virus has been responsible for several epidemics in West and Central Africa. Ebola virus belongs to the *Mononegavirales* order and its genome codes for the “large” protein L, embedding polymerase, capping and methyltransferase (MTase) activities, essential for virus replication. The MTase induces a set of RNA post-transcriptional modifications as it methylates the N7 position of the cap structure and the 2'O position of the first nucleotide present at the RNA 5'end. In contrast to other mononegaviruses, Ebola virus MTase also catalyses 2'O methylation of adenosines located within the RNA sequence. Here, we described the 1.8 Å structure of the Ebola virus MTase domain crystallised in complex with a VHH. The MTase domain adopts the canonical Rossmann fold organization with a conserved S-adenosylmethionine (SAM) binding pocket close to the K-D-K-E 2'O MTase catalytic tetrad. Additionally, structural features with hMPV and VSV MTase domains are revealed as the 7 central β-strands form a planar structure. SUDV MTase catalytic residues are exposed at the surface of the structure lacking an obvious cap/RNA-binding site. Mutagenesis study confirms the role of conserved residues for the enzymatic activities and identifies key residues in the recently characterised 2'O internal methylation. Altogether, these results provide the first high resolution of a domain the Ebola virus L protein, contributing to the better understanding of the capping of Ebola virus RNA and providing a framework to design new effective Ebola virus inhibitors.

INTRODUCTION

Mononegavirales order includes viruses with linear, non-segmented, single-stranded, negative-sense RNA genomes. Despite their genetic organization and replication similarities, these viruses induce a wide range of symptoms associated with variable pathogenicity. Ebola is responsible for important epidemics as evidenced by the latest outbreaks in West Africa (2013-2016) and in the Democratic Republic of the Congo (2018-2020), which have respectively caused at least 11,000 and more than 2,200 deaths (1). Fruit bats are considered as the main reservoir of this zoonotic virus. In rare circumstances, Ebola virus can be transmitted to non-human primates and humans. The inter-human contaminations mainly rely on direct contacts with biological fluids of infected patients, which leads to dissemination of the virus in the population (2). After 2 to 10 days of incubation, Ebola infection can cause a haemorrhagic fever fatal in 50-60% of cases. To date, treatments against Ebola and main treatments are supportive. Indeed, while rVSV-ZEBOV-GP and Ad26-ZEBOV vaccines have recently shown a good efficacy to limit the last Ebola outbreak (3), absence of effective antiviral therapeutics highlights the need to develop new antiviral strategies.

Within the *Mononegavirales* order, *Ebolavirus* genus, which contains six species (*Zaire ebolavirus*, EBOV; *Sudan ebolavirus*, SUDV; *Tai Forest ebolavirus*, TAFV; *Bundibugyo ebolavirus*, BDBV; *Reston ebolavirus*, RESTV; *Bombali ebolavirus*, BOMV), belongs to the *Filoviridae* family like the highly pathogenic Marburg virus (MARV) and other recently identified viruses (*Lloviu cuevavirus*, LLOV; *Mengla dianlovirus*, MLAV; *Xilang striavirus*, XILV; *Huangjiao thamnivirus*, HUV) (4). The viral genome of about 19 kb encodes seven proteins: the nucleoprotein (NP), VP35, VP40, the glycoprotein (GP), VP30, VP24 and the “large” protein L (5, 6). The L protein drives virus replication as it embeds all enzymatic activities required for genome replication, transcription and mRNA capping and polyadenylation. Mononegavirus mRNAs are co-transcriptionally capped by an unconventional capping reaction (7). Indeed, the nascent viral mRNA was shown to make a covalent bond with a conserved catalytic histidine residue of the polyribonucleotidyltransferase (PRNTase) present in the cap domain of the L protein. The PRNTase next transfers a GDP molecule to the 5' phosphate of the covalently bound RNA forming the cap structure (GpppN1). The cap is subsequently methylated by the MTase domain at the 2'OH position of the first nucleotide (N1) ribose and at the N7 position of the cap guanosine (cap-1, ^mGpppN_m) (7–9). The N7 methylation of the cap structure is required for viral mRNA translation into protein as it allows the recognition of mRNA by the translation initiation factor eIF4E (10). The 2'O methylation masks the viral mRNA from the detection by cytoplasmic sensors belonging to retinoic acid-inducible gene-I (RIG-I)-like receptor family (11). In addition to its role during RNA transcription, the L protein also ensures genome replication when the NP protein concentration is increased. The pleiotropic activities carried by the L protein suggest that the different enzymatic activities are timely regulated, in order to ensure the different specific functions required for virus replication.

The multi-alignment analysis reveals that L protein of *Mononegavirales* contains six conserved regions (CRI to CRVI) located between the RNA-dependent RNA polymerase (RdRp) domain (CRI to CRIII), the Cap or PRNTase domain (CRIV & V), and the methyltransferase (MTase) domain (CRVI) (**Fig. 1A**). On the topological point of view, electron microscopy with negative staining experiments originally performed on the vesicular stomatitis virus (VSV) L protein reveal that the RdRp together with the Cap domain form a “donut-like” structure, followed by three flexible globular domains corresponding to the connector domain (CD), the MTase domain and a small C-terminal domain (CTD) (12). The recent technological advances of cryo-EM have allowed the resolution of several mononegavirus L protein structures in which the C-terminal domain (CD+MTase+CTD) is not always clearly defined suggesting an important mobility (13–16). The VSV structure reveals that the RdRp domain forms a typical right-hand polymerase structure with “fingers-palm-thumb” subdomains (13). The subsequent PRNTase domain, organized in an original fold, is intimately intricated with the polymerase domain and projects a loop near the catalytic site of the polymerase.

The C-terminus of the L protein contains the conserved MTase domain followed by the CTD domain, which is barely conserved between viruses of the *Mononegavirales* order. The MTase domain contains a typical 2’O MTase catalytic tetrad (K-D-K-E) and a GxGxG motif characteristic of the S-adenosylmethionine (SAM)-binding site whereas the CTD domain shows no conserved signature (17). The VSV L cryo-EM structure has revealed a typical Rossman fold organization of the MTase domain (13). Remarkably, the cryo-EM structure of RSV reveals that, although the C-terminal domain of L were present in the constructs, the CD, MTase and CTD domains were not visible in the density map suggesting an important flexibility of these regions (18). This flexibility was also observed in the recent resolved structure of parainfluenza virus 5 (PIV5) L protein, showing a rearrangement of the MTase and CTD domains in the vicinity of the RdRp domain. *Mononegavirales* MTase domains are organised in Rossman fold with a canonical SAM binding site (8). However, these MTase domain apparently lack cap-binding site observed in most viral MTases (8, 13, 16). Rather, the MTase and CTD domains are associated and form a narrow RNA-binding groove enriched in basic amino acids close to the catalytic site (8). The role of the CTD of L protein was next demonstrated by biochemical studies showing that RNA-binding properties and the MTase activities of SUDV MTase depend on the presence of the CTD domain (19).

Biochemical studies reported that cap methylations driven by VSV MTase follow an unconventional pathway in which the ribose 2’O methylation of the first nucleotide precedes the N7 methylation of the cap structure (9). The catalytic activity of the MTase has been next confirmed for other mononegaviruses (hMPV & SUDV) (8, 20). Beside this shared feature, hMPV MTase is able to methylate uncapped RNAs on the 2’OH of the first nucleotide (8) and SUDV MTase carries an additional activity inducing the 2’O methylation of adenosine residues inside the RNA sequence (20). The role of such post-transcriptional RNA modifications is not yet understood in the context of Ebola virus infection. However, similar epitranscriptomic RNA modifications were evidenced in other viral

RNA, such as Zika virus (ZIKV) and Dengue virus (DENV) and HIV (21–23). Recent studies have shown their involvement in the regulation of host-pathogen interaction. This is illustrated by the HIV which recruits the cellular MTase FtsJ3 to induce internal 2'O methylations on the viral RNA genome, participating in the subversion of host defences by limiting viral detection by the RIG-like receptor Melanoma Differentiation-Associated protein 5 (MDA5) (21). Thus, increasing evidence show pleiotropic functions of L-associated MTase activities that catalyse different post-transcriptional modifications of viral RNA regulating the virus life cycle and the early antiviral response.

In this work, we present the structure of the SUDV MTase used to perform a structure guided mutagenesis study in order to elucidate how the viral MTase activities and RNA binding are regulated.

MATERIALS & METHODS

Cloning and expression

Codon-optimized SUDV MTase+CTD synthetic genes (Biomers) were cloned into a pET14b vector for expression in bacteria. The construct MTase without the CTD domain (MTase) was synthesized by cloning the sequence of interest from MTase+CTD into a pET14b plasmid with Gateway system (ThermoFischer). Mutations were introduced by PCR-amplifying the SUDV MTase+CTD sequence using primers carrying mutations, and Turbo DNase (Ambion). PCR-products were purified using the Wizard SV PCR Clean-Up System (Promega). Transformed *E. coli* bacteria (NEB) were cultured at 30°C until OD_{600nm} of 0.6 was reached. Then, temperature was shifted to 17°C and IPTG (Sigma-Aldrich) was added to 20 µM. The next day, bacteria were spun down (8000 × g for 20 min at 4°C) using a Sorval Lynx 6000 centrifuge (Thermo) before storage at -80°C.

Recombinant protein purification

Wt and mutant MTase+CTD proteins were purified as previously described (19). For the MTase, pellets were thawed at room temperature (RT), and 1 mM PMSF, 10 µg/mL DNase, 20 mM MgSO₄. After 30 min of incubation at 4°C, cells were sonicated and clarified by centrifugation (18,000 xg, 30 min, 4°C) and a tablet of EDTA-free antiprotease cocktail (Roche) was added. The Proteins was then purified by immobilized metal affinity chromatography (IMAC) purification on 5mL his-trap columns (GE Healthcare). Protein was eluted with 300 mM NaCl, 50 mM Tris (pH 8.0), 250 mM imidazole. Eluted proteins were next loaded on Superdex S75 16/60 (GE Healthcare) equilibrated with 10 mM Tris (pH 8.0) and 150 mM NaCl.

Generation of VHH against MTase domain of SUDV

A healthy llama (*Llama glama* from Ardèche lamas, France) was immunized with 0.8 mg purified Ebola MTase, produced as previously described (20) and stored in 50 mM Tris (pH 8.0), 150 mM NaCl, 5% glycerol. During immunization, the protein was injected five times, at one-week intervals. Lymphocytes

were isolated from blood samples obtained five days after the last immunization. The cDNA was synthesized from purified total RNA by reverse transcription and was used as a template for PCR to amplify the sequences corresponding to the variable domains of the heavy-chain antibodies. PCR fragments were then cloned into the phagemid vector pHEN4 (24) to create a nanobody phage display library. Selection and screening of VHH were performed as described previously (25).

Expression and purification of VHH

Selected nanobodies were cloned in a pHEN6 plasmid encoding a N-terminal pelB periplasmic signal sequence in frame with a VHH expression cassette and a C-terminal 6His tag for detection and purification. Protein expression was achieved using WK6 bacteria cultivated in Terrific Broth medium (AthenaES) supplemented with 100 µg/ml ampicillin and 0.1% glucose at 37° until reaching an OD_{600nm} 0.5 –0.8. Expression was then induced by the addition of 1 mM IPTG and growth was continued overnight at 28°C.

The periplasmic proteins were extracted according to Skerra *et al.* (26). Bacteria were pelleted by centrifugation at 3500 g for 15 min at 4°C. The pellet was resuspended in 9 mL cold TES buffer (0.2 M Tris-HCl (pH 8.0), 0.5 mM EDTA, 0.5 M fsucrose) per liter of culture and kept on ice for 1h. The periplasmic proteins were removed by osmotic shock by addition of 13.5 mL of cold TES diluted four times with water. After 1-2h on ice, the suspension was centrifuged at 21,700 g for 30 min at 4°C. Vhh were purified by immobilized metal affinity chromatography (IMAC) purification using Ni-NTA resin (Thermo Fisher Scientific). After loading samples for 1h at 4°C, resin was cleared of contaminants with wash buffer (50 mM phosphate buffer (pH 8.0), 300 mM NaCl, 10% glycerol) then eluted by step gradient using dilutions from 1:10 to 1:2 of elution buffer (50 mM phosphate buffer (pH 8.0), 300 mM NaCl, 10% glycerol, 0.5 M imidazole) in wash buffer. Fractions containing the purified nanobody were concentrated on Amicon Ultra-MW10000 (Millipore) to 1 mL. Finally, SEC of VHH fragments was performed using a Superdex 75 16/60 column equilibrated with 10 mM Tris (pH 8.0) and 150 mM NaCl.

Purification of SUDV MTase / VHH heterodimer

Monomeric fractions of MTase were pooled and complexed to purified VHH (1:1.5). After 1h incubation at 4°C, SEC was applied to separate the complex (MTase and VHH) and the free VHH using Superdex S75 16/60 (GE Healthcare) (**Fig. 1C**) equilibrated with 10 mM Tris (pH 8.0) and 150 mM NaCl. Purified complex (MTase and VHH) was concentrated to 7 mg/mL and stored at 4°C.

Crystallogensis and X-ray crystallography

Concentrated complex (7 mg/mL) was crystallized by vapor diffusion at 20°C using a 96-well sitting drop plate (SWISSCI 3 Lens Crystallisation Microplate). Crystals grew spontaneously within 48 h by equilibrating 300 nL of protein with 100 nL of 0.1 M Tris (pH 8.0), 0.1 M NaCl, 8% (w/v) of PEG 8,000 as F5 solution of the ProPlex HT-96 screen (Molecular Dimensions).

Crystals were cryo-protected with reservoir solution with 20% PEG 200, prior being flash cooled in liquid nitrogen. X-ray diffraction was performed on beamline Proxima1 at Soleil synchrotron. Data from crystals were collected as SAD data at peak ($\lambda = 1.28242 \text{ \AA}$) of zinc. Data sets were processed individually and analysed with the autoPROC toolbox. Anomalous signal was detected for 4 crystals that presented weak anomalous signal around 4 \AA . Merging data sets with AIMLESS allowed to enrich the anomalous signal at low resolution corresponding to S scatter. Structure was solved by a combination of molecular replacement and anomalous signal method, using PHASER. Original placement of VHH allowed a partial map and the anomalous difference map helped to calculate an initial electron density map and align the sequence identifying the MTase. Density modification was performed with PARROT and auto-building with BUCCANNEER allowed to extend the model, which was further manually built with Coot and refined with BUSTER. The final model has a $R_{\text{work}} = 19.6\%$ and a $R_{\text{free}} = 23.4\%$ and was confirmed to have good stereochemistry according to MOLPROBITY. Using that model we did a molecular replacement on a single data set, and refined with BUSTER the structure up to 1.8 \AA . Data collection and refinement statistics are listed in **Table 1**.

Comparison of domains, modelling of missing loops and structural analysis

Homologous structures were searched with DALI and 3D-fold search, retrieving MTase structures. Comparison with MTase structures of VSV and hMPV (PDB: 4UCZ and 5A22, respectively) allowed us to confirm the coiled structure corresponding to the missing loop. To ensure the complete continuity of the main chain and ensure a proper surface analysis, we decided to model the two missing loops (regions 1-1 and 1-1). Structural analysis and reference files for modelling were prepared with CHIMERA and missing loops were then further modelled using MODELLER 9.23. Surface electrostatics was calculated with APBS. Sequences and interface were analysed using ENDSCRIPT (27).

Synthesis of RNA substrates

RNA sequences were chemically synthesized as previously described (19). RNA sequences were chemically synthesized on solid support using an ABI 394 oligonucleotide 171 synthesizer. RNA elongation was performed with 2'-O-pivaloyloxymethyl phosphoramidite ribonucleotides and 2'-O-methyl phosphoramidite ribonucleotides (Chemgenes). Then, the 5'-hydroxyl group was phosphorylated and the resulting H-phosphonate derivative was oxidized and activated into a phosphoroimidazolidate derivative to react with guanosine diphosphate (GpppRNA). After deprotection and release from the solid support, GpppRNAs were purified by IEX-HPLC and their purity (>95%) was confirmed by MALDI-TOF spectrometry. N7 methylation of the purified GpppRNAs was performed by incubation with human N7 MTase (38, 39).

MTase activity assay

Methyltransferase activities were assessed using a filter-binding assay as previously reported (19). Briefly, this radioactive assay was set up by mixing 4 μM of protein with 1 μM of purified synthetic RNAs, 10 μM of SAM and 0.5 μM of ^3H -SAM (Perkin Elmer) in 50 mM Tris-HCl at different pH to evaluate the different MTase activities (pH 7.0 for the cap-N7 MTase activity, pH 8.0 for the cap-2'O MTase and pH 8.5 for internal methylation). After 3 h at 30°C, reactions were stopped by 10-fold dilution in water and mixes were loaded onto DEAE filtermats (Perkin Elmer) using a Filtermat Harvester (Packard Instruments). Methyl transfer was then evaluated after two washes with 10 mM ammonium formate pH 8.0, two washes with water, and a last wash with ethanol. Then, filters were soaked with liquid scintillation fluid to measure the ^3H -methyl transfer to the RNA substrates using a Wallac MicroBeta TriLux Liquid Scintillation Counter13. For statistical analysis, it was assumed that the different experimental groups were independent, and data followed a Gaussian distribution with the same variance. Two-way ANOVA and multiple comparison Dunnett test were performed (Prism) to evaluate differences between groups. The level of significance for $\alpha=0.05$ is indicated as follows: * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Fluorescence polarization (FP)

The ligation of a fluorochrome at the RNA 3' ends was performed by mixing T4 RNA ligase 1 (20 units; New England Biolabs), cyanine 5-cytidine-5-phosphate-3-(6-aminoethyl)phosphate (12.5 μM , Jena Bioscience) and RNA substrates of interest (10 μM) with T4 RNA ligase 1 buffer (NEB) and 1 mM ATP (16 °C, overnight). After ligase removal by RNA precipitation in 3 M sodium acetate supplemented with 1 $\mu\text{g}/\mu\text{L}$ of glycogen (Thermo Scientific), fluorescent RNA was incubated (at room temperature for 5 min) with increasing protein concentrations in reaction buffer (20 mM Tris pH 8 or 8.5, 150 mM NaCl and 5% glycerol). FP levels were evaluated in a microplate reader (PHERAstar FS; BMG Labtech) using excitation and emission wavelengths of 590 and 675 nm, respectively. Dissociation constants (K_d) were determined using Hill slope curve fitting (Prism).

RESULTS

Production, purification and crystallization of SUDV MTase in complex with VHH

Crystallization trials using SUDV MTase+CTD constructs did not lead crystals and the protein was submitted to mild proteolysis analysis. We identified a protease resistant MTase domain corresponding to residues 1744 to 2046 of the L protein (35.5 kDa). The MTase domain was cloned in a bacterial expression system and expressed and purified by IMAC purification followed by size exclusion chromatography (SEC). The soluble MTase domain elute in a monomeric form (**Fig. 1B**), assessed by Size Exclusion Chromatography. In parallel, we purified a llama nanobody (VHH) (**Fig. 1B**) targeting the SUDV MTase domain. The purified SUDV MTase was incubated together with the purified VHH at a ratio 1:1.5 and the complex was recovered after SEC (**Fig. 1C**). SDS-PAGE analysis shows that

purified heterodimer is detected as two bands corresponding respectively to MTase and VHH proteins migrating at their expected molecular weights on SDS-PAGE gel (**Fig. 1B**). The purified complex was concentrated up to 7 mg/ml and used for crystallogenes. Crystals were obtained upon vapor diffusion in 0.1 M Tris (pH 8.0), 0.1 M NaCl, 8% (w/v) of PEG 8,000 (**Fig. 1D**) and the structure was obtained by X-ray crystallography.

Structural overview of SUDV MTase domain in complex with a VHH

The crystal structure of the SUDV MTase domain/VHH heterodimer (**Fig. 2A**) was determined by molecular replacement coupled with single wavelength anomalous diffraction (MR-SAD) using sulfur as the anomalous scatterer. The crystal belongs to the space group P6222, with cell dimensions as follows: $a = b = 153.76 \text{ \AA}$, $c = 105.35 \text{ \AA}$ and $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$, containing twelve heterodimers in the unit cell. The higher order of the crystal shows that the structural assembly forms a compact mesh crossed by large hexameric solvent channels for which dimensions of a side is about 60 \AA long and diagonal $\sim 107 \text{ \AA}$ (**Fig. S1**).

The crystal structure of the SUDV MTase domain in complex with this nanobody was solved with a resolution of 1.8 \AA . The MTase domain presents a Rossmann fold (**Fig. 2B**), typical of most MTases catalyzing the methyl transfer from S-adenosylmethionine (SAM). The domain contains 8 β -strands of which 2 to 8 are part of the core of the Rossmann fold adopting the classical parallel strands organisation except for the last one being antiparallel and sandwiched between the antepenultimate and penultimate strands (**Fig. 2C**). This central β -sheet is surrounded by 6 α -helices. Additional to the Rossmann core, the additional first β -strand is antiparallel to the penultimate strand ($\beta 7$), and two long α -helices ($\alpha 1$ and $\alpha 8$) interact with each other and the first strand (**Fig. 2B&2C**). The linker regions between $\beta 1$ and $\alpha 1$ (residues 1764 to 1775) and $\alpha 1$ and $\alpha 2$ (residues 1795 to 1808) were not built due to lack of density and probable high flexibility (**Fig. 2A**). The equivalent $\alpha 1 - \alpha 2$ segment is in the 3 other structures stabilized by the CTD and present the same structure: a long loop punctuated by a short helix. At a sequence level, this short helix is signalled by a hydrophobic patch. As a similar hydrophobic signature exists in SUDV MTase and because the sequence length is compatible with the structure, we decided to proceed to the modelling of the missing part (highlighted in green, **Fig. 2B&2C**). The 2' O MTase catalytic tetrad (K1813/D1924/K1959/E1996) highlighted in yellow (**Fig. 2D**) as well as the SAM-binding motif (G1833-X-G1835-X-G1837, shown below in **Fig. 3**) are localized in the Rossmann fold secondary structure but no cap-binding site has been identified.

The VHH is a β - sandwich composed of 8 strands connected by flexible loops (**Fig. 2A**). The VHH antigen interface binds to the MTase on a single epitope formed by the bottom part of three helices ($\alpha 4$, $\alpha 5$, $\alpha 6$) and on the opposite side of its catalytic site. This observation is consistent with the MTase activity test performed in presence of different concentrations of nanobody. Indeed, while the nanobody recognizes both MTase and previously described MTase+CTD domains of SUDV L protein, it does not inhibit the MTase activities of SUDV MTase+CTD (**Fig. S2A&S2B**) suggesting that the VHH does not

alter the MTase folding. However, the VHH binds to the MTase through an extended surface composed of several hydrophobic residues located in two side loops (residues 55-61 and residues 100-109) and a turn (residues 27-31) closed to the N-terminus. The overall binding interface area is about 833.6 Å² (**Fig.S3**).

Structural comparison with other viral MTase domains

From a topological point of view, *Mononegavirales* MTases present a structural signature to the Rossman fold (PMID: 26549102) where the first amino terminal β strand is anti-parallel to the β -sheet and interacting with the C-terminal α -helix. Superimposition of SUDV MTase structure with hMPV and DENV2 MTase domains, as well as the cellular 2' O MTase CMTr1, reveals that the central β -sheet is strictly superposed, revealing a high structural conservation despite the low sequence identity between the different proteins ($\leq 10\%$) (**Fig. 3A**). Interestingly, most SUDV helices are structurally shifted as compared to other MTases. Moreover, there is no equivalent to $\alpha 1$ in hMPV MTase nor in other mononegavirus MTases (**Fig. S4A**), for which it is replaced by a long flexible loop. Conversely, DENV2 and CMTr1 MTases (**Fig. 3A**), as well as ZIKV MTase (**Fig. S4A**), present a α -helix that can be superimposed to $\alpha 1$ of SUDV MTase. This supplementary α -helix is largely involved in a positively charged patch, forming a large positive groove right next to the active site (**Fig.S4B**).

Unlike in other viral MTase structures, the structure of the SUDV MTase is lacking methyl donor or acceptor. A SAM molecule was thus modelled in the SUDV ^{SAMP} using hMPV MTase domain superimposition to spot residues implicated in SAM recruitment (**Fig.3C**). As it was previously observed for hMPV (residues 1718-1729) and flaviviruses, SUDV MTase presents a long flexible loop that participates in the ^{SAMP} (residues 1854-1875). However, this loop adopts an “open” conformation in SUDV structure where no SAM molecule is co-crystallized while that of hMPV shows a “close” conformation, clamping the SAM substrate in the ^{SAMP}. Similarly, a GTP molecule was modelled in the SUDV MTase structure to identify residues of a putative RNA and/or cap substrate pocket (^{SUBP}) (**Fig.3D**). Two loops were identified to participate in the ^{SUBP} of SUDV MTase. The first one, from residue 1803 to residue 1811, is quite variable within mononegaviruses. Indeed, while this loop contains charged residues, its equivalent loop in hMPV MTase (residues 1663-1670) is more hydrophobic. The second loop, between residues 1988 and 1995 of SUDV MTase, immediately followed by the catalytic residue E1996, is rather conserved with hMPV MTase sequence (residues 1940-1947). Lastly, hMPV structure shows a deep hydrophobic cavity that can bind adenosine moiety of SAM or ATP (^{NSP}). This cavity is associated with a flexible loop in SUDV MTase, which is assumed to adopt a “close” position. The length of this loop is a little longer than those of hMPV or VSV, with an enrichment in charged and polar residues.

Effects of single mutations in the SUDV MTase domain on MTase activities

As the MTase domain produced in absence of the CTD is not able to recruit RNA, the functional MTase+CTD (from amino acid 1750 to amino acid 2126, according to SUDV L protein numbering) was used to perform structure-guided functional studies focusing on residues of the SAM-binding site and the RNA-binding groove (**Fig. S4B**), highlighted in orange and pink, respectively (**Fig.4A&4B**). In order to specifically evaluate cap-N7, cap-2'O and internal A-2'O methylations, we used GpppG_m(A_m)-SUDV₁₂, ^mGpppA(A_m)-SUDV₁₂, ^mGpppG_m-SUDV₁₂, respectively (**Fig.4C&4D**). For residues of the putative SAM-binding site and those potentially involved in RNA recruitment, we observe that some of the mutations lead to a drastic loss of all MTase activities (cap-N7, cap-2'O and internal A-2'O methylations) whereas some affect differentially the MTase activities. Indeed, mutations G1835S, G1837S, T1854A and L1855A in the SAM-binding site and mutations Y1800A and S1809A in the putative RNA-binding groove lead to a complete loss of overall activities. Other residues like L1855A of the SAM-binding pocket and I1806A, V1807A, T1927A and S1990A possibly involved in RNA binding show a significant reduction of all MTase activities. Interestingly, several mutations in the RNA-binding groove result in an uncoupling of the different MTase activities. S1991A and K1993A mutants show a drastic reduction (approximately 50%) and almost a complete loss of 2'O MTase activities (cap and internal A-2'O MTase activities), respectively, but keep the N7 MTase activity. Similarly, S1808A and R1792A mutants impair at the same respective levels only internal A-2'O activity without altering MTase activities associated to the cap synthesis. Finally, E1834A mutation in the SAM-binding pocket shows promotes the three methylations, with a preference for the A-2'O MTase activities.

DISCUSSION

In this study, the structure of the SUDV MTase domain in complex with a nanobody was determined by X-ray crystallography at a resolution of 1.8 Å. The L protein being the most conserved protein of viruses belonging to the *Mononegavirales* order (28), we compared structures by superimposing the SUDV MTase onto other mononegavirus MTase domains, such as hMPV and VSV (**Table S1**). This observation supports that the SUDV MTase domain is correctly folded and its structure is independent from other L protein domains, as suggested for VSV (13). Nonetheless, previously described pockets ^{SAMP}P, ^{SUBP}P and ^{NSP}P responsible for SAM, RNA and nucleoside binding (8), respectively, are structurally and functionally preserved. Furthermore, some original features in the SUDV MTase structure were described and seem to be involved in MTase activities regulation.

First, as ^{SAMP}P is largely conserved (GxGxG motif) among mononegaviruses, residues were mapped based on SAM modelization and several residues were mutated (E1834A, G1835S, G1837S). Unsurprisingly, mutations of glycines led to a complete loss of MTase activities due to a default of SAM recruitment as it has already been described in many studies. However, the residue E1834, which

is conserved in both hMPV and SUDV MTases, shows a loss of catalysis when mutated in hMPV (8) but greater MTase activities for SUDV. Structure analysis of SUDV MTase also reveals a flexible loop close to ^{SAMP} (residues 1854-1875). In hMPV, this loop (residues 1718-1729) adopts an “open” or “close” conformation to bind the SAM molecule (PDB: 4UCI vs. 4UCK), suggesting that it may participate in the turn-over of the MTase reaction. Like mutations in the SAM-binding motif, T1854A and L1855A mutants lead to a complete loss of MTase activities, indicating their importance for MTase activities. Since mutated residues T1854 and L1855 are localised in a flexible loop, these results suggest that these residues would alter its flexibility, thus impairing SAM recruitment.

Then, a similar approach based on SUDV MTase model with GTP was used to define ^{SUBP}. This pocket likely forms a groove defined by two loops (residues 1803-1811 and 1988-1995) close to the catalytic residues. Although the loop 1803-1811 is variable, the loop 1988-1995 is conserved among viruses of the *Mononegavirales* order. Conversely to ^{SAMP}, mutations in ^{SUBP} result in variable MTase activity patterns. Indeed, phenotypes can be divided into two categories: mutations inducing a complete or major alteration of overall MTase activities and mutations uncoupling MTase activities. In the loop 1803-1811, most mutations (I1806A, V1807A and S1809A) cause an extensive loss of function. Similarly to most mutations in SUDV MTase, mutations in this loop in hMPV MTase lead to strong impairing of MTase activities (8). The S1808A mutation however results in a specific decrease of internal A-2'O MTase activity and keeps both cap-N7 and cap-2'O MTase activities. These observations suggest that this loop plays a key role for the different MTase activities but the central position of the serine in SUDV might participate in the internal 2'O MTase activity. On the other hand, in the conserved loop 1988-1995, mutation S1990A leads to a broad loss of function whereas mutants S1991A and K1993A only alter cap-2'O and internal A-2'O MTase activities. Interestingly, mutation of the corresponding lysine to alanine in hMPV leads to a slight decrease of 2'O MTase activity while N7 MTase activity is less impacted (8), suggesting that this conserved loop can be crucial for the RNA positioning supporting 2'O MTase activity regulation.

The ^{NSP} pocket sustaining the catalytic residue D1924 is constituted by the loop 1924-1933 in SUDV MTase that is conformed differently from hMPV structure suggesting some flexibility. Mutations of the conserved E1926 and T1927 into alanine lead to a decrease of overall SUDV MTase activities, T1927 slightly uncoupling internal 2'O methylation from cap-N7 methylation and, to a less extent, cap-2'O methylation from cap-N7 methylation. These results would indicate that E1926 may be involved in the stability of the catalytic pocket and that T1927 might participate in the stability of the loop but also in the groove for RNA positioning as proposed for ^{SUBP}. Interestingly, similar mutation of E1981 in hMPV results only in a modest decrease of MTase activities indicating that length and sequence of this loop may involve distinct structural roles for MTase regulation in these viruses.

Furthermore, SUDV MTase also contains an additional α -helix (α 1), close to its N-terminus. This α -helix is not present in hMPV, VSV and PIV5 MTases. However, a homologous helix is found in the NS5 MTase domain of flaviviruses (DENV2 and ZIKV). Flavivirus NS5 is known to catalyse internal

A-2'O methylations similarly, but to a lower extent compared to SUDV MTase (21–23). Similarly, the cellular 2'O MTase CMTr1 exhibits the same structural features, including the additional α -helix near the N-terminus (29). Interestingly, recent studies suggest that this protein could also methylate internal residues at the 2'O position (30, 31). Because of the presence of positively charged and polar residues in this α -helix, it has thus been hypothesized that this structure might be involved in the regulation of this internal A-2'O MTase activity. To evaluate this assumption, R1792 was mutated to alanine and this mutant showed a decoupling of cap methylations from internal methylation activity. These observations confirm that this α -helix participates in the 2'O internal MTase activity by bringing residues necessary to accommodate the RNA substrate properly to allow these methylations.

Altogether, these results contribute to drawing a comprehensive model of Ebola MTase activity regulation through RNA accommodation, supporting the absence of a canonical cap-binding site. Indeed, when structure and function are analysed together (**Fig. 5**), two patches can be depicted. The first patch comprises residues around the catalytic tetrad, including ^{SAMP}, ^{NSP} and the loop 1803-1811 of ^{SUBP}, which are all critical for the different MTase activities. The second patch, which is supported by the loop 1987-1995 and the α 1-helix, creates a groove, specifically recruiting the RNA for 2'O methylations. It is thus possible to postulate a model where the cap-N7 methylation occurs when the RNA is accommodated by the first patch, positioning the cap directly into the active site while the cap-2'O and the internal 2'O methylations are catalysed when the RNA is recruited through the second patch, projecting 2'OH of nucleotide ribose into the catalytic site. Interestingly, the structures of VSV and PIV5 L full length revealed that MTase+CTD domain adopt different positions regarding RdRp and PRNTase domains (13, 16, 32). The dynamic of this flexibility suggests different conformations of the MTase+CTD domain causing preference for transcription over replication (and vice versa). This flexibility of the C-terminal region around the N-terminal domains has been previously observed (32). In PIV5 cryo-EM structure (16), (MTase+CTD position suggests a transcriptionally conformation whereas MTase+CTD position in VSV leads to a preference for replication. In the model proposed by *Abdella et al.*, the nascent capped mRNA transits through the groove between the MTase and CTD domains which are closed to the PRNTase domain whereas replicative conformation positions MTase+CTD domains away from the PRNTase domain. In this conformation, the RNA did not transit through the MTase domain suggesting an absence of methyl transfer insured by MTase domain. In addition, a previous study highlighted the role of internal methylations in the subversion of host defences by limiting viral detection (21). In conclusion, it is possible that transcription conformation opens the way to effective methylation activities on nascent mRNA while replication events did not and led to encapsidation of nascent genomic or antigenomic RNA, highlighting the role of such epitranscriptomic modifications in the context of Ebola virus infection.

Accession numbers

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data bank under accession number 6YU8.

Supplementary Information is available in the online version of the paper.

Acknowledgements

The program was initiated in a research program funded by the Délégation Générale pour l'Armement (DGA) (grant 2009.34.0038)/ Aix-Marseille Université PhD fellowship for B.M. CV. was funded by National Research Agency ANR under the program ANR Rab-cap (ANR-16_CE11_0031_01). The French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INSB-05-01. The authors acknowledge SOLEIL for provision of synchrotron radiation facilities, and we would like to thank the beamline team of Proxima-1 for their assistance during data collection. We thank Sarah Attoumani and Anais Gaubert for technical support.

C.V., B.M., B.Co. and E.D. designed and performed experiments, using material prepared by F.D. and J-J.V.. F.F. resolved the structure. V.Z. participated in protein purification and crystallogensis. A.D. identified nanobodies against SUDV MTase. C.V., B.M., F.F., B.Co. and E.D. analysed data and wrote the paper. B.Ca. contributed to the writing and participated in the review of the paper.

Conflict of interest

Authors declare no conflict of interest.

REFERENCES

1. Ebola Virus Disease | WHO | Regional Office for Africa.
2. Lawrence,P., Danet,N., Reynard,O., Volchkova,V. and Volchkov,V. (2017) Human transmission of Ebola virus. *Curr. Opin. Virol.*, **22**, 51–58.
3. Ollmann Saphire,E. (2020) A Vaccine against Ebola Virus. *Cell*, **181**, 6.
4. Kuhn,J.H., Adachi,T., Adhikari,N.K.J., Arribas,J.R., Bah,I.E., Bausch,D.G., Bhadelia,N., Borchert,M., Brantsæter,A.B., Brett-Major,D.M., *et al.* (2019) New filovirus disease classification and nomenclature. *Nat. Rev. Microbiol.*, **17**, 261–263.
5. Elliott,L.H., Sanchez,A., Holloway,B.P., Kiley,M.P. and McCormick,J.B. (1993) Ebola protein analyses for the determination of genetic organization. *Arch. Virol.*, **133**, 423–36.
6. Sanchez,A., Kiley,M.P., Holloway,B.P. and Auperin,D.D. (1993) Sequence analysis of the Ebola virus genome: organization, genetic elements, and comparison with the genome of Marburg virus. *Virus Res.*, **29**, 215–240.

7. Ogino, T. and Banerjee, A.K. (2007) Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Mol. Cell*, **25**, 85–97.
8. Paesen, G.C., Collet, A., Sallamand, C., Debart, F., Vasseur, J.-J., Canard, B., Decroly, E. and Grimes, J.M. (2015) X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat. Commun.*, **6**, 8749.
9. Rahmeh, A.A., Li, J., Kranzusch, P.J. and Whelan, S.P.J. (2009) Ribose 2'-O methylation of the vesicular stomatitis virus mRNA cap precedes and facilitates subsequent guanine-N-7 methylation by the large polymerase protein. *J. Virol.*, **83**, 11043–50.
10. Shatkin, A.J. (1976) Capping of eucaryotic mRNAs. *Cell*, **9**, 645–53.
11. Devarkar, S.C., Wang, C., Miller, M.T., Ramanathan, A., Jiang, F., Khan, A.G., Patel, S.S. and Marcotrigiano, J. (2016) Structural basis for m7G recognition and 2'-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proc. Natl. Acad. Sci.*, **113**, 596–601.
12. Rahmeh, A.A., Schenk, A.D., Danek, E.I., Kranzusch, P.J., Liang, B., Walz, T. and Whelan, S.P.J. (2010) Molecular architecture of the vesicular stomatitis virus RNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 20075–80.
13. Liang, B., Li, Z., Jenni, S., Rahmeh, A.A., Morin, B.M., Grant, T., Grigorieff, N., Harrison, S.C. and Whelan, S.P.J. (2015) Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy. *Cell*, **162**, 314–327.
14. Horwitz, J.A., Jenni, S., Harrison, S.C. and Whelan, S.P.J. (2020) Structure of a rabies virus polymerase complex from electron cryo-microscopy. *Proc. Natl. Acad. Sci. U. S. A.*, **117**, 2099–2107.
15. Pan, J., Qian, X., Lattmann, S., El Sahili, A., Yeo, T.H., Jia, H., Cressey, T., Ludeke, B., Noton, S., Kalocsay, M., *et al.* (2020) Structure of the human metapneumovirus polymerase phosphoprotein complex. *Nature*, **577**, 275–279.
16. Abdella, R., Aggarwal, M., Okura, T., Lamb, R.A. and He, Y. (2020) Structure of a paramyxovirus polymerase complex reveals a unique methyltransferase-CTD conformation. *Proc. Natl. Acad. Sci.*, **117**, 4931–4941.
17. Ferron, F., Longhi, S., Henrissat, B. and Canard, B. (2002) Viral RNA-polymerases -- a predicted 2'-O-ribose methyltransferase domain shared by all Mononegavirales. *Trends Biochem. Sci.*, **27**, 222–224.
18. Gilman, M.S.A., Liu, C., Fung, A., Behera, I., Jordan, P., Rigaux, P., Ysebaert, N., Tcherniuk, S., Sourimant, J., Eléouët, J.-F., *et al.* (2019) Structure of the Respiratory Syncytial Virus Polymerase Complex. *Cell*, 10.1016/j.cell.2019.08.014.
19. Valle, C., Martin, B., Debart, F., Vasseur, J.-J., Imbert, I., Canard, B., Coutard, B. and Decroly, E. (2020) The C-terminal domain of the *Sudan ebolavirus* L protein is essential for RNA binding and methylation. *J. Virol.*, 10.1128/JVI.00520-20.
20. Martin, B., Coutard, B., Guez, T., Paesen, G.C., Canard, B., Debart, F., Vasseur, J.-J., Grimes, J.M. and Decroly, E. (2018) The methyltransferase domain of the *Sudan ebolavirus* L protein specifically targets internal adenosines of RNA substrates, in addition to the cap structure. *Nucleic Acids Res.*, **46**, 7902–7912.
21. Ringeard, M., Marchand, V., Decroly, E., Motorin, Y. and Bennasser, Y. (2019) FTSJ3 is an RNA

- 2'-O-methyltransferase recruited by HIV to avoid innate immune sensing. *Nature*, **565**, 500–504.
22. Dong,H., Chang,D.C., Hua,M.H.C., Lim,S.P., Chionh,Y.H., Hia,F., Lee,Y.H., Kukkaro,P., Lok,S.-M., Dedon,P.C., *et al.* (2012) 2'-O methylation of internal adenosine by flavivirus NS5 methyltransferase. *PLoS Pathog.*, **8**, e1002642.
 23. Coutard,B., Barral,K., Lichièrre,J., Selisko,B., Martin,B., Aouadi,W., Lombardia,M.O., Debart,F., Vasseur,J.-J., Guillemot,J.C., *et al.* (2017) Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. *J. Virol.*, **91**, e02202-16.
 24. Arbabi Ghahroudi,M., Desmyter,A., Wyns,L., Hamers,R. and Muyldermans,S. (1997) Selection and identification of single domain antibody fragments from camel heavy-chain antibodies. *FEBS Lett.*, **414**, 521–6.
 25. Desmyter,A., Farenc,C., Mahony,J., Spinelli,S., Bebeacua,C., Blangy,S., Veessler,D., van Sinderen,D. and Cambillau,C. (2013) Viral infection modulation and neutralization by camelid nanobodies. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, E1371-9.
 26. Skerra,A. and Pluckthun,A. (1988) Assembly of a functional immunoglobulin Fv fragment in *Escherichia coli*. *Science (80-)*, **240**, 1038–1041.
 27. Robert,X. and Gouet,P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.*, **42**, W320-4.
 28. Poch,O., Blumberg,B.M., Bougueleret,L. and Tordo,N. (1990) Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *J. Gen. Virol.*, **71 (Pt 5)**, 1153–1162.
 29. Smietanski,M., Werner,M., Purta,E., Kaminska,K.H., Stepinski,J., Darzynkiewicz,E., Nowotny,M. and Bujnicki,J.M. (2014) Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nat. Commun.*, **5**, 3004.
 30. Inesta-Vaquera,F., Chaugule,V.K., Galloway,A., Chandler,L., Rojas-Fernandez,A., Weidlich,S., Peggie,M. and Cowling,V.H. (2018) DHX15 regulates CMTR1-dependent gene expression and cell proliferation. *Life Sci. Alliance*, **1**, e201800092.
 31. Toczydlowska-Socha,D., Zielinska,M.M., Kurkowska,M., Astha, Almeida,C.F., Stefaniak,F., Purta,E. and Bujnicki,J.M. (2018) Human RNA cap1 methyltransferase CMTr1 cooperates with RNA helicase DHX15 to modify RNAs with highly structured 5' termini. *Philos. Trans. R. Soc. B Biol. Sci.*, **373**, 20180161.
 32. Jenni,S., Bloyet,L.-M., Diaz-Avalos,R., Liang,B., Whelan,S.P.J., Grigorieff,N. and Harrison,S.C. (2020) Structure of the Vesicular Stomatitis Virus L Protein in Complex with Its Phosphoprotein Cofactor. *Cell Rep.*, **30**, 53-60.e5.

FIGURE LEGENDS

Figure 1 - Crystallogenesis of SUDV MTase domain in complex with a nanobody

(A) Functional domains of the mononegavirus L protein with the six conserved regions (CRI to CRVI) containing motifs important for L protein activities (mapped with asterisks). Among these motifs, mononegavirus methyltransferase domain (MTase) encompasses a SAM-binding site (GxGxG) and the 2'-O catalytic tetrad K-D-K-E. (B) Recombinant *Sudan ebolavirus* (SUDV) methyltransferase domain (MTase, 35.5 kDa) (left), a specific nanobody (VHH, 15 kDa) (middle) and the complex of both proteins (right) were purified, then analyzed by SDS-PAGE and Coomassie blue staining. (C) After affinity chromatography, the SUDV MTase and nanobody were mixed at the ratio 1:1.5. Complexes were purified by size exclusion chromatography using a Superdex S75 16/60. Chromatogram was plotted according to fraction absorbance at 260 nm. (D) Protein complex crystals were obtained by vapor diffusion at 20°C using crystallisation buffer (0.1 M Tris pH 8.0 with 8% (w/v) polyethylene glycol 8000) and a suspended drop of 300 nL of 7 mg/mL of purified protein complex mixed with 100 nL of crystallisation buffer. Images were taken using a Nikon E600 microscope.

Figure 2 - SUDV MTase structure resolved by X-ray crystallography

(A) Crystal structure of *Sudan ebolavirus* (SUDV) methyltransferase (MTase) domain (blue) in complex with a nanobody (orange). N- and C-terminal extremities are indicated by N and C letters, respectively, for both proteins. The SUDV MTase structure shows an alternance of 8 β -sheets and 8 α -helices and 2 undefined loops (residues K1764 to P1775 and residues V1795 to S1808, dash lines). A Mg^{2+} ion (light green) is found in the structure, linked to Y1845 (sticks). The VHH reveals a classic β -sandwich fold composed of 8 strands connected by flexible loops. Two loops (residues G55 to A61 and residues A100 to Y109) and a turn (residues R27 to R31) are involved in the interaction with the SUDV MTase domain, at the opposite side of the catalytic pocket. (B) Numbering of secondary structures and modelization of unresolved segments (residues V1795 to S1808, green) using their equivalence with 3 other virus MTase structures (PDB: 4UCI, 5A22 and 6UEB) stabilized by the CTD and presenting the same long loop punctuated by a short α -helix (α'). (C) Topological organization of SUDV MTase domain revealing a conserved Rossmann fold (grey box) of prototypical SAM-dependent MTase. Helices are depicted as cyan barrels, β -strands as dodger blue arrows and coils as sky blue lines. N- and C-terminal extremities are indicated by N and C letters, respectively. This representation reveals also additional features at the N-terminus (1 β -strand and 1 α -helix) and at the C-terminus (1 α -helix). (D) Surface representation of the SUDV MTase domain with a close-up on the typical K-D-K-E catalytic tetrad (yellow) of 2'-O MTases. Yellow sticks show the K1813-D1924-K1959-E1996 motif with side-chains exposed at the surface of the structure.

Figure 3 - Structural and functional comparison with other MTase domains

(A) Superimposition of *Sudan ebolavirus* (SUDV) methyltransferase domain (MTase, blue) with human metapneumovirus (hMPV) L MTase domain (PDB: 4UCI, top of the panel, red), dengue 2 virus (DENV2) non-structural protein 5 (NS5) MTase domain (PDB: 5ZQK, middle of the panel, green) and human 2'-O MTase (CMTr1, PDB: 4N49, bottom of the panel, orange) MTase domains. Central β -sheet of all structures are represented in dark blue, dark red, dark green and dark orange for SUDV, hMPV, DENV3 and CMTr1 MTase domains, respectively. Compared to other mononegaviruses, a supplementary N-terminal α -helix in SUDV MTase is also found in both DENV2 and CMTr1 MTase structures (red). (B) Close-up on the superimposition of the typical K-D-K-E catalytic tetrad of SUDV (blue) and hMPV (red) MTase domains. Residues K1813, D1924, K1959 and E1996 of SUDV MTase and residues K1673, D1779, K1817, E1848 of hMPV MTase are represented in yellow and red sticks, respectively. (C) Modelization of the S-adenosylmethionine (SAM) molecule (red) within the SUDV MTase structure based on hMPV MTase structure. Residues involved in SAM-binding pocket (^{SAM}P) are represented in orange and catalytic residues are in yellow. (D) Modelization of a GTP molecule (red) within the SUDV MTase structure based on hMPV MTase structure. Residues involved in the substrate-binding pocket (^{SUB}P) are represented in pink and catalytic residues are in yellow.

Figure 4 - Single mutation analysis of SAM-binding and RNA-binding pockets

(A) Mapping of mutated residues potentially involved in the S-adenosylmethionine (SAM)-binding pocket (^{SAMP}) within *Sudan ebolavirus* (SUDV) methyltransferase domain (MTase). (B) MTase activity of SUDV WT and mutated MTase+CTD in the ^{SAMP}. Proteins were incubated with GpppG_m(A_m)-SUDV₁₂, ^mGpppA(A_m)-SUDV₁₂, or ^mGpppG_m-SUDV₁₂ and a radiolabeled methyl donor (³H-SAM) to assess the cap N7 MTase, cap-2'O MTase, and internal A-2'O MTase activities, respectively, by filter binding assay. Data represent mean ± standard deviation (n=3); *p<0.05, **p<0.01, ***p<0.001 (two-way ANOVA and multiple comparison Dunnett test, WT vs. mutation). (C) Mapping of mutated residues potentially involved in the RNA-binding groove (^{SUBP}) within SUDV MTase. (D) MTase activity of SUDV WT and mutated MTase+CTD in the RNA-binding groove. Proteins were incubated with GpppG_m(A_m)-SUDV₁₂, ^mGpppA(A_m)-SUDV₁₂, or ^mGpppG_m-SUDV₁₂ and a radiolabeled methyl donor (³H-SAM) to assess cap-N7 MTase, cap-2'O MTase, and internal A-2'O MTase activities, respectively, by filter binding assay. Data represent mean ± standard deviation (n=3); *p<0.05, **p<0.01, ***p<0.001 (two-way ANOVA and multiple comparison Dunnett test, WT vs. mutation). (E) RNA-binding assay by fluorescence polarization using 13-mers SUDV-specific capped RNAs (GpppG-SUDV₁₂ and ^mGpppG_m-SUDV₁₂) labeled at the 3' end with pCp-Cy5 and SUDV MTase+CTD and ^{SUBP} mutants. The dissociation coefficient (K_d) was calculated by site-specific binding regression with Hill slope. Data represent means (± standard deviation, n=3).

Figure 5 - Proposition for structural bases of SUDV cap and internal methylations

Mapping of single mutations onto *Sudan ebolavirus* (SUDV) methyltransferase domain (MTase, light blue), colored according to phenotypes: loss of all MTase activities (dark orange), decrease of all MTase activities (light orange), stimulation of all MTase activities (green), decoupling 2'O MTase activities (purple), decoupling internal 2'O activity (dark blue). A capped RNA (grey) has been modelled according to its putative position for cap-N7 MTase (left), cap-2'O MTase (middle) and internal adenosine 2'O MTase reactions (right).

Table 1 - Crystallographic data collection and refinement statistics

* Referring to anomalous statistics

Statistics for the highest-resolution shell are shown in parentheses.

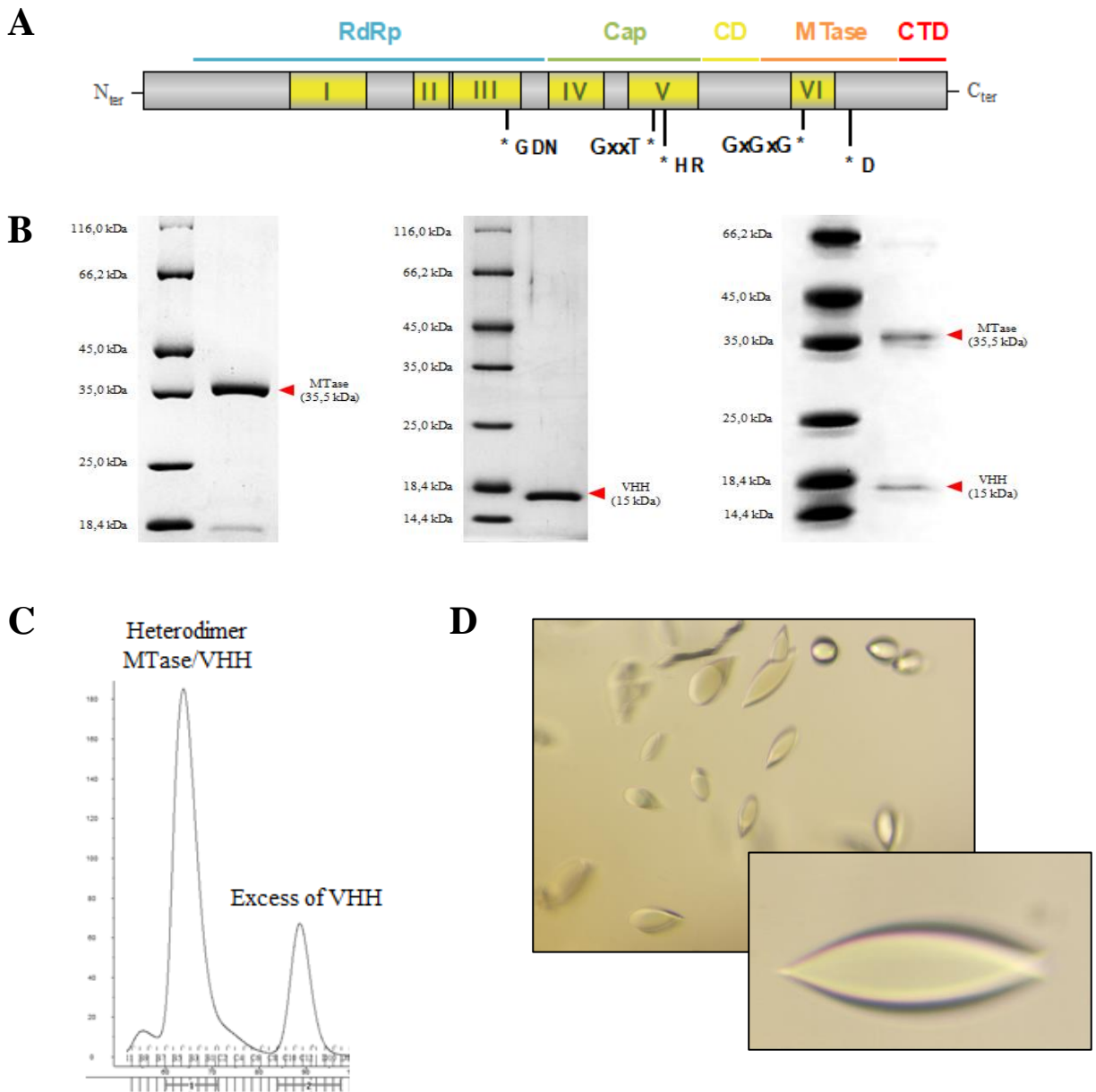


Figure 1 - Crystallogenesis of SUDV MTase domain in complex with a nanobody

(A) Functional domains of the mononegavirus L protein with the six conserved regions (CRI to CRVI) containing motifs important for L protein activities (mapped with asterisks). Among these motifs, mononegavirus methyltransferase domain (MTase) encompasses a SAM-binding site (GxGxG) and the 2'-O catalytic tetrad K-D-K-E. (B) Recombinant *Sudan ebolavirus* (SUDV) methyltransferase domain (MTase, 35.5 kDa) (left), a specific nanobody (VHH, 15 kDa) (middle) and the complex of both proteins (right) were purified, then analyzed by SDS-PAGE and Coomassie blue staining. (C) After affinity chromatography, the SUDV MTase and nanobody were mixed at the ratio 1:1.5. Complexes were purified by size exclusion chromatography using a Superdex S75 16/60. Chromatogram was plotted according to fraction absorbance at 260 nm. (D) Protein complex crystals were obtained by vapor diffusion at 20°C using crystallisation buffer (0.1 M Tris pH 8.0 with 8% (w/v) polyethylene glycol 8000) and a suspended drop of 300 nL of 7 mg/mL of purified protein complex mixed with 100 nL of crystallisation buffer. Images were taken using a Nikon E600 microscope .

Table 1 - Crystallographic data collection and refinement statistics

Data processing	MTase Ebola SUDV (S)	MTase Ebola SUDV (native)
Wavelength (Å)	1.282	1.282
Space group	P62 2 2	P 62 2 2
- a, b, c (Å)	153.98, 153.98, 105.41	153.76, 153.76, 105.35
- α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Resolution range (Å)	76.87- 2.00 (2.05 -2.00)	76.88 - 1.84 (1.907 - 1.84)
Total No. of reflections	8916413 (355358)	2566705 (257401)
No. of unique reflections	49877 (3627)	63699 (6280)
Completeness (%)	100 (99.7)*	99.98 (100)
Multiplicity	94.1 (50.5)*	40.3 (41)
I/σ(I)	30.5 (3.5)	13.74.54 (2.14)
R_{meas}		0.2135 (1.601)
CC1/2	--	0.99 (0.83)
Wilson B-factor (Å²)	-	32.08
Structure solution & refinement		
No. of reflections, working set	-	63691 (6280)
No. of reflections, test set (%)	-	3164 (311), 5 %
R_{cryst}	-	0.1782 (0.2455)
R_{free}	-	0.2009 (0.2663)
No. of non-H atoms	-	3571
- Protein	-	3070
- Ligand	-	73
- Water	-	428
R.m.s. deviations		
- Bonds (Å)	-	0.01
- Angles (°)	-	1.51
Average B-factors (Å²)	-	39.00
- Protein	-	36.5.
- Ligand	-	56.05
- Water	-	54.12
Ramachandran Plot		
- Favoured (%)	-	97.92
- Allowed (%)	-	1.82
- Outliers (%)	-	0.26
Clashcore	-	5.15
PDB code	-	6YU8

* Referring to anomalous statistics

Statistics for the highest-resolution shell are shown in parentheses.

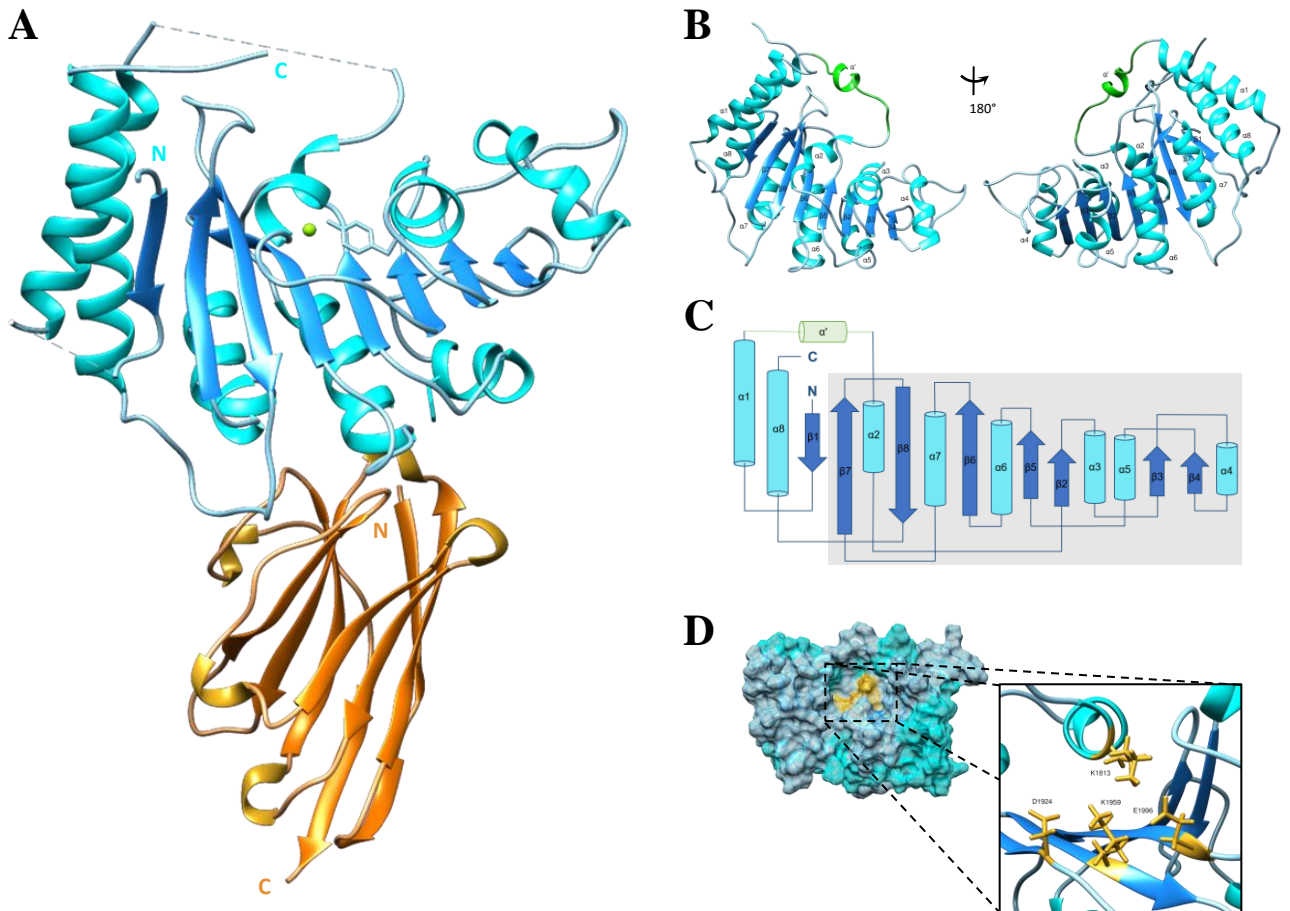


Figure 2 - SUDV MTase structure resolved by X-ray crystallography

(A) Crystal structure of *Sudan ebolavirus* (SUDV) methyltransferase (MTase) domain (blue) in complex with a nanobody (orange). N- and C-terminal extremities are indicated by N and C letters, respectively, for both proteins. The SUDV MTase structure shows an alternance of 8 β -sheets and 8 α -helices and 2 undefined loops (residues K1764 to P1775 and residues V1795 to S1808, dash lines). A Mg^{2+} ion (light green) is found in the structure, linked to Y1845 (sticks). The VHH reveals a classic β -sandwich fold composed of 8 strands connected by flexible loops. Two loops (residues G55 to A61 and residues A100 to Y109) and a turn (residues R27 to R31) are involved in the interaction with the SUDV MTase domain, at the opposite side of the catalytic pocket. (B) Numbering of secondary structures and modelization of unresolved segments (residues V1795 to S1808, green) using their equivalence with 3 other virus MTase structures (PDB: 4UCI, 5A22 and 6UEB) stabilized by the CTD and presenting the same long loop punctuated by a short α -helix (α'). (C) Topological organization of SUDV MTase domain revealing a conserved Rossmann fold (grey box) of prototypical SAM-dependent MTase. Helices are depicted as cyan barrels, β -strands as dodger blue arrows and coils as sky blue lines. N- and C-terminal extremities are indicated by N and C letters, respectively. This representation reveals also additional features at the N-terminus (1 β -strand and 1 α -helix) and at the C-terminus (1 α -helix). (D) Surface representation of the SUDV MTase domain with a close-up on the typical K-D-K-E catalytic tetrad (yellow). Yellow sticks show the K1813-D1924-K1959-E1996 motif with side-chains exposed at the surface of the structure.

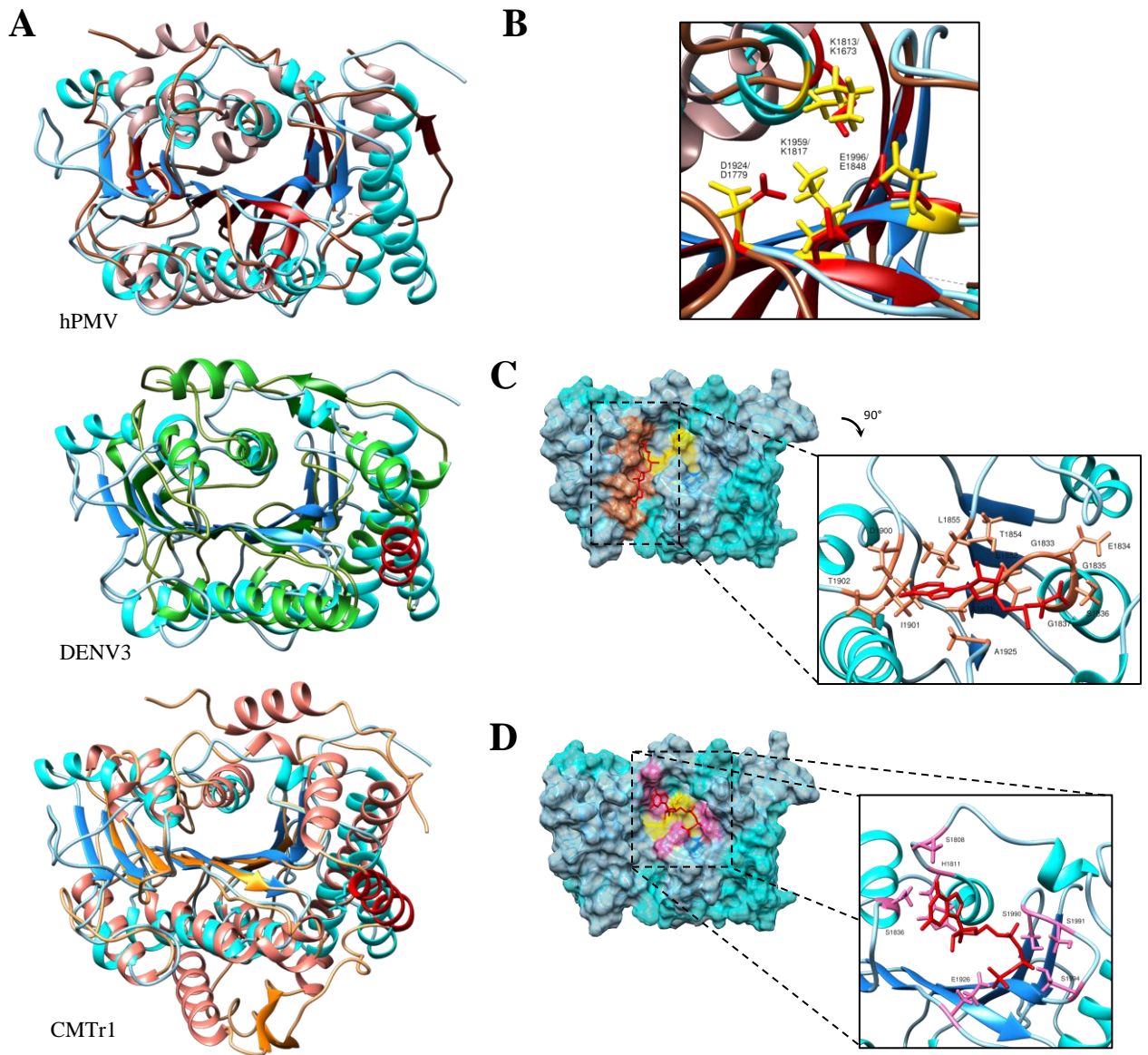


Figure 3 - Structural and functional comparison with other MTase domains

(A) Superimposition of *Sudan ebolavirus* (SUDV) methyltransferase domain (MTase, blue) with human metapneumovirus (hMPV) L MTase domain (PDB: 4UCI, top of the panel, red), dengue 2 virus (DENV2) non structural protein 5 (NS5) MTase domain (PDB: 5ZQK, middle of the panel, green) and human 2' O MTase (CMTr1, PDB: 4N49, bottom of the panel, orange) MTase domains. Central β -sheet of all structures are represented in dark blue, dark red, dark green and dark orange for SUDV, hMPV, DENV3 and CMTr1 MTase domains, respectively. Compared to other mononegaviruses, a supplementary N-terminal α -helix in SUDV MTase is also found in both DENV2 and CMTr1 MTase structures (red). (B) Close-up on the superimposition of the typical K-D-K-E catalytic tetrad of SUDV (blue) and hMPV (red) MTase domains. Residues K1813, D1924, K1959 and E1996 of SUDV MTase and residues K1673, D1779, K1817, E1848 of hMPV MTase are represented in yellow and red sticks, respectively. (C) Modelization of the S-adenosylmethionine (SAM) molecule (red) within the SUDV MTase structure based on hMPV MTase structure. Residues involved in SAM-binding pocket ($SAMP$) are represented in orange and catalytic residues are in yellow. (D) Modelization of a GTP molecule (red) within the SUDV MTase structure based on hMPV MTase structure. Residues involved in the substrate-binding pocket ($SUBP$) are represented in pink and catalytic residues are in yellow.

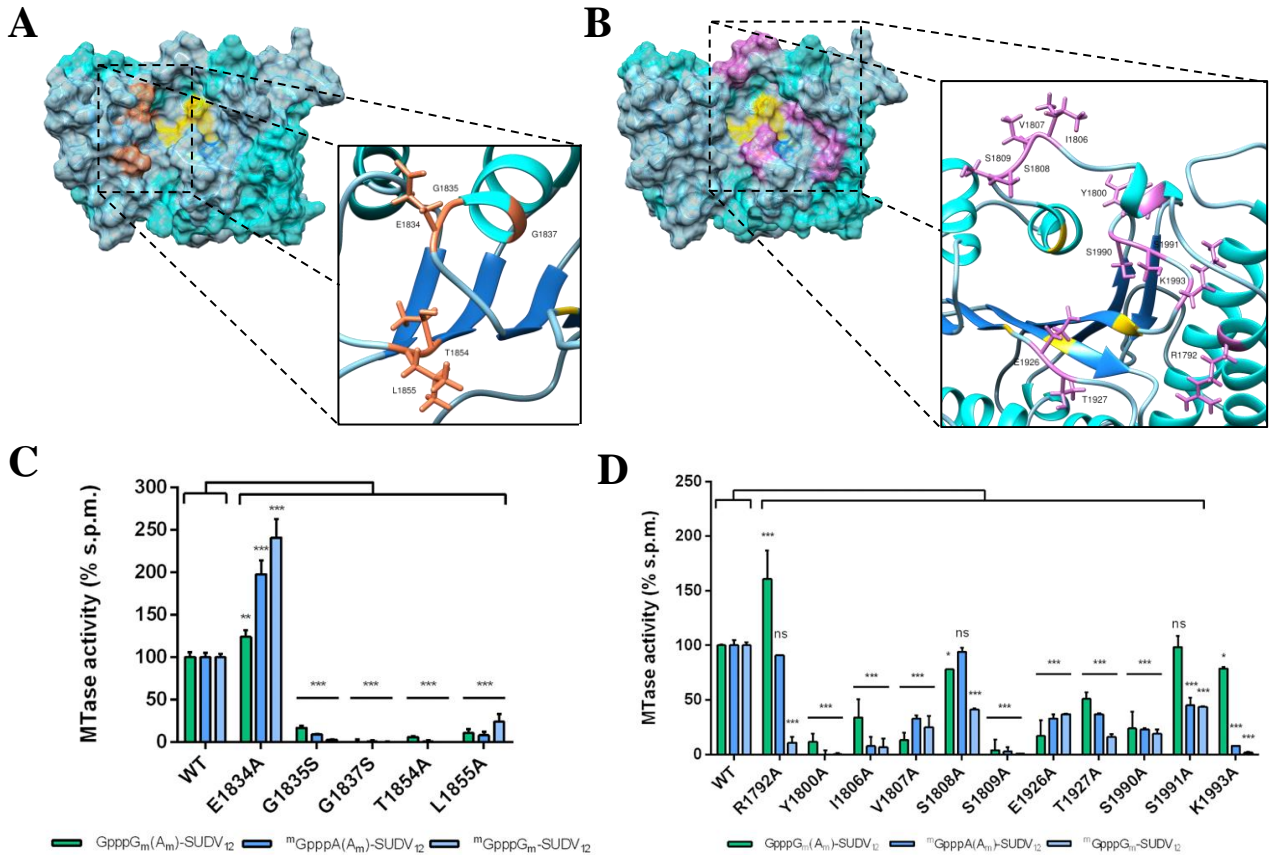


Figure 4 - Single mutation analysis of SAM-binding and RNA-binding pockets

(A) Mapping of mutated residues potentially involved in the S-adenosylmethionine (SAM)-binding pocket (^{SAMP}) within *Sudan ebolavirus* (SUDV) methyltransferase domain (MTase). (B) MTase activity of SUDV WT and mutated MTase+CTD in the ^{SAMP}. Proteins were incubated with GpppG_m(A_m)-SUDV₁₂, ^mGpppA(A_m)-SUDV₁₂, or ^mGpppG_m-SUDV₁₂ and a radiolabeled methyl donor (³H-SAM) to assess the cap N7 MTase, cap-2' O MTase, and internal A-2' O MTase activities, respectively, by filter binding assay. Data represent mean ± standard deviation (n=3); *p<0.05, **p<0.01, ***p<0.001 (two-way ANOVA and multiple comparison Dunnett test, WT vs. mutation). (C) Mapping of mutated residues potentially involved in the RNA-binding groove (^{SUBP}) within SUDV MTase. (D) MTase activity of SUDV WT and mutated MTase+CTD in the RNA-binding groove. Proteins were incubated with GpppG_m(A_m)-SUDV₁₂, ^mGpppA(A_m)-SUDV₁₂, or ^mGpppG_m-SUDV₁₂ and a radiolabeled methyl donor (³H-SAM) to assess cap-N7 MTase, cap-2' O MTase, and internal A-2' O MTase activities, respectively, by filter binding assay. Data represent mean ± standard deviation (n=3); *p<0.05, **p<0.01, ***p<0.001 (two-way ANOVA and multiple comparison Dunnett test, WT vs. mutation).

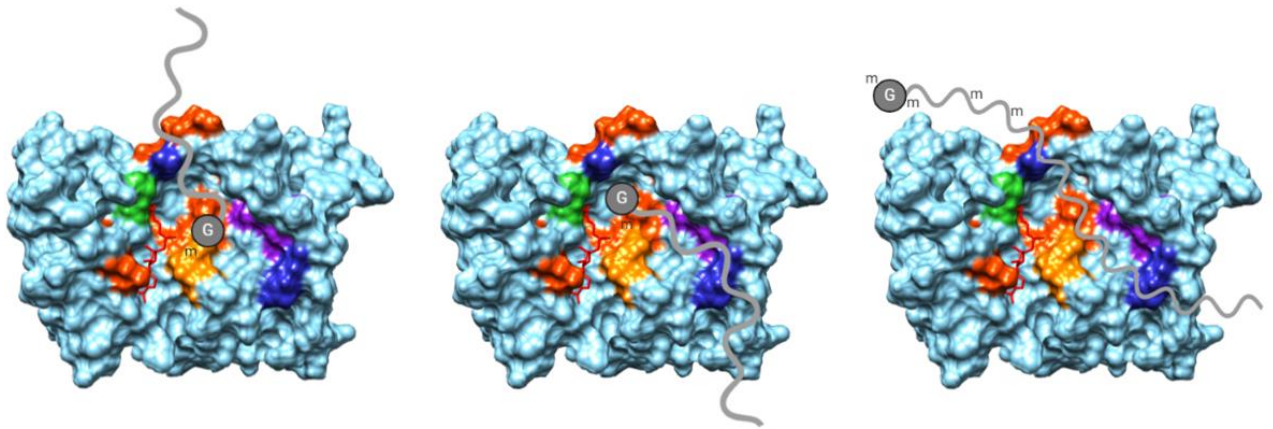


Figure 5 - Proposition for structural bases of SUDV cap and internal methylations

Mapping of single mutations onto *Sudan ebolavirus* (SUDV) methyltransferase domain (MTase, light blue), colored according to phenotypes: loss of all MTase activities (dark orange), decrease of all MTase activities (light orange), stimulation of all MTase activities (green), decoupling 2'O MTase activities (purple), decoupling internal 2'O activity (dark blue). A capped RNA (grey) has been modelled according to its putative position for cap-N7 MTase (left), cap-2'O MTase (middle) and internal adenosine 2'O MTase reactions (right).

SUPPLEMENTARY INFORMATION

Supplementary Methods

Enzyme-Linked Immunosorbent Assay (ELISA)

Purified protein was diluted in PBS (20 µg/ml) and used to coat 96-well immunoplates (Maxisorb Nuc[®]) at 4°C overnight. Coated plates were then blocked by adding PBS with 2% of non-fat dry milk at room temperature for 2h. Purified VHH (2 µg/mL) was added to the wells at room temperature for 2h. Plates were rinsed three times with PBS, 0.1% Tween 20 (PBST) and three times with PBS. For detection, 100 µL of primary mouse anti-HA antibody (1:1500, Sigma) was added at 37°C for 1h. After three washes with PBST and then PBS, the secondary HRP-coupled anti-mouse antibody (1:1000, Agilent) was added at 37°C for 1h. Then, tetramethylbenzidine and H₂O₂ (Sigma) were added and the reaction was stopped after 10 min with 1 M H₂SO₄. The optical density at 450nm was determined using a microplate reader (Safire2, Tecan). BSA was used as negative control. Each assay was carried out in triplicate.

Supplementary Figures and Tables

Figure S1 - Analysis of the crystal packing

Ribbon representation of the heterodimer MTase-VHH colored in purple and in green, respectively. (A) Orientation along the P6 axis marked by a red hexagone. Right close caption presents the hexameric shape solvent channel of large dimension with a side ~60 Å. (B) Left and right panels show 90° rotation of panel (A) following different axes, both axes presenting a tight packing.

Figure S2 - Characterization of nanobody interaction with SUDV MTase+CTD domain

(A) Evaluation of *Sudan ebolavirus* (SUDV) methyltransferase alone (MTase) or along with the L C-terminal domain (MTase+CTD) with the nanobody by enzyme-linked immunosorbent assay (ELISA). Data represent mean ± standard deviation (n=5). (B) MTase activity of SUDV MTase+CTD incubated with GpppG-SUDV₁₂, a radiolabeled methyl donor (³H-SAM) and different concentration ratios of nanobody (VHH). Data represent mean ± standard deviation (n=3).

Figure S3 - Surface interaction between SUDV MTase domain and the nanobody (VHH)

(A) Surface (left panel) and ribbon (right panel) representations of the *Sudan ebolavirus* (SUDV) methyltransferase (MTase) domain in complex with a nanobody (VHH). MTase is depicted in cyan and its interacting residues in orange. VHH is represented in purple and its interacting residues in yellow. The interacting surface is represented in purple. (B and C) Identification of the residues on the structure MTase and VHH respectively. (D) Sequences of the heterodimers topped with their secondary structures, and underlined with solvent accessibility, the hydrophobic character of the sequences and contact information. In the structure file, MTase is assigned chainID A and VHH chainID B. The contact is reported with the chainID character in red if the distance is less than 3.2 Å and in black if it is in the range 3.2-4.0 Å. Disulfide bridge is marked by green linked S'.

Figure S4 - Further structural comparison with several other viral MTase domains

(A) Superimposition of *Sudan ebolavirus* (SUDV) methyltransferase (MTase, blue) with vesicular stomatitis virus (VSV, PDB: 5A22, top of the panel, purple), parainfluenza 5 virus (PIV5, PDB: 6V85, middle of the panel, yellow) and Zika virus (ZIKV, PDB: 5M5B, bottom of the panel, green) MTase domains. Central β-sheet of all structures are represented in dark blue, dark purple, dark yellow and

dark green for SUDV, VSV, PIV5 and ZIKV MTase domains, respectively. Compared to other mononegaviruses, a supplementary N-terminal α -helix in SUDV MTase is also found in ZIKV MTase structure (red). **(B)** Left column represents structures of SUDV MTase+CTD (modelled using Phyre 2.0), human metapneumovirus (hMPV) MTase+CTD (PDB: 4UCI), dengue 2 virus (DENV2) non-structural protein 5 (NS5, PDB: 5ZQK) and human 2'O MTase (CMTr1, PDB: 4N49), respectively from top to bottom. MTase domains are depicted in orange, mononegavirus L C-terminal domains (CTD) in yellow, and other domains in white. Right column represents cartography of positive patches on these same structures using PatchFinderPlus predictor (1). Positive patches are depicted in dark blue.

Table S1 - Structure/function comparison of SUDV MTase domain with other mononegavirus MTases

This table summarizes phenotypes of single mutation analysis in both *Sudan ebolavirus* (SUDV) and human metapneumovirus (hMPV) methyltransferase domains (MTase, PDB: 6YU8 and 4UCI, respectively) and homologous loop sequences of the different functional pockets including those of vesicular stomatitis virus (VSV, PDB: 5A22). As previously described for hMPV (2), loops were classified according to the functional pocket they belong to: the S-adenosylmethionine (SAM)-binding pocket (^{SAMP}), the RNA-binding pocket (^{SUBP}) and the nucleoside-binding pocket (^{NSP}). For each virus, location, sequence and mutants of each loop (if available) were reported. Mutations are highlighted in blue in the loop sequence and phenotypes are indicated as percentages of cap-N7, cap-2'O and internal 2'O activities according to wild-type proteins, respectively. The color code reflect the impact of the mutation (from dark green (>150% of WT activity) to red (0-25% of WT activity)).

Supplementary References

1. Shazman,S., Celniker,G., Haber,O., Glaser,F. and Mandel-Gutfreund,Y. (2007) Patch Finder Plus (PFplus): A web server for extracting and displaying positive electrostatic patches on protein surfaces. *Nucleic Acids Res.*, **35**, W526–W530.
2. Paesen,G.C., Collet,A., Sallamand,C., Debart,F., Vasseur,J.-J., Canard,B., Decroly,E. and Grimes,J.M. (2015) X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat. Commun.*, **6**, 8749.

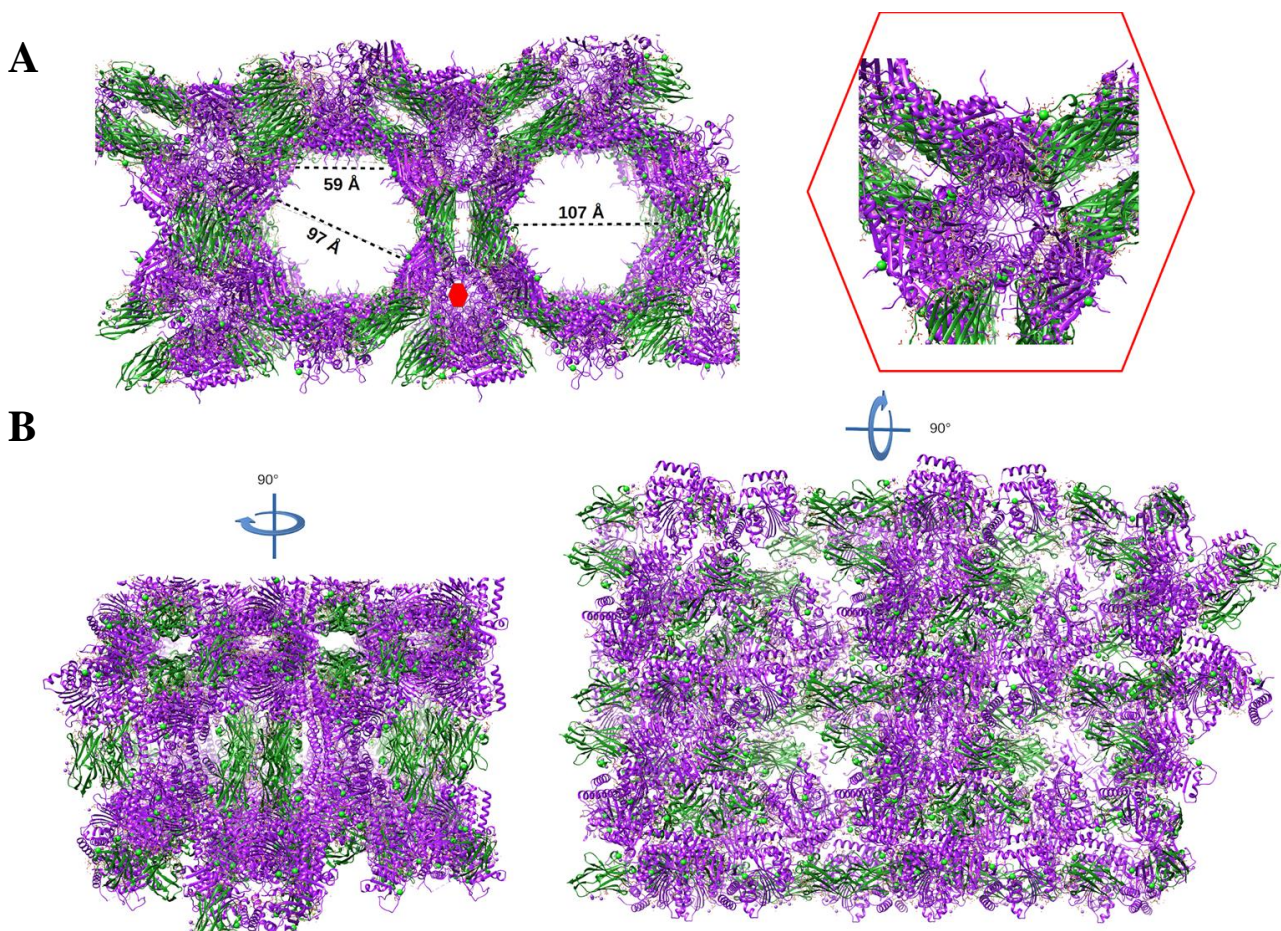


Figure S1 - Analysis of the crystal packing

Ribbon representation of the heterodimer MTase-VHH colored in purple and in green, respectively. (A) Orientation along the P6 axis marked by a red hexagone. Right close caption presents the hexameric shape solvent channel of large dimension with a side ~ 60 Å. (B) Left and right panels show 90° rotation of panel (A) following different axes, both axes presenting a tight packing.

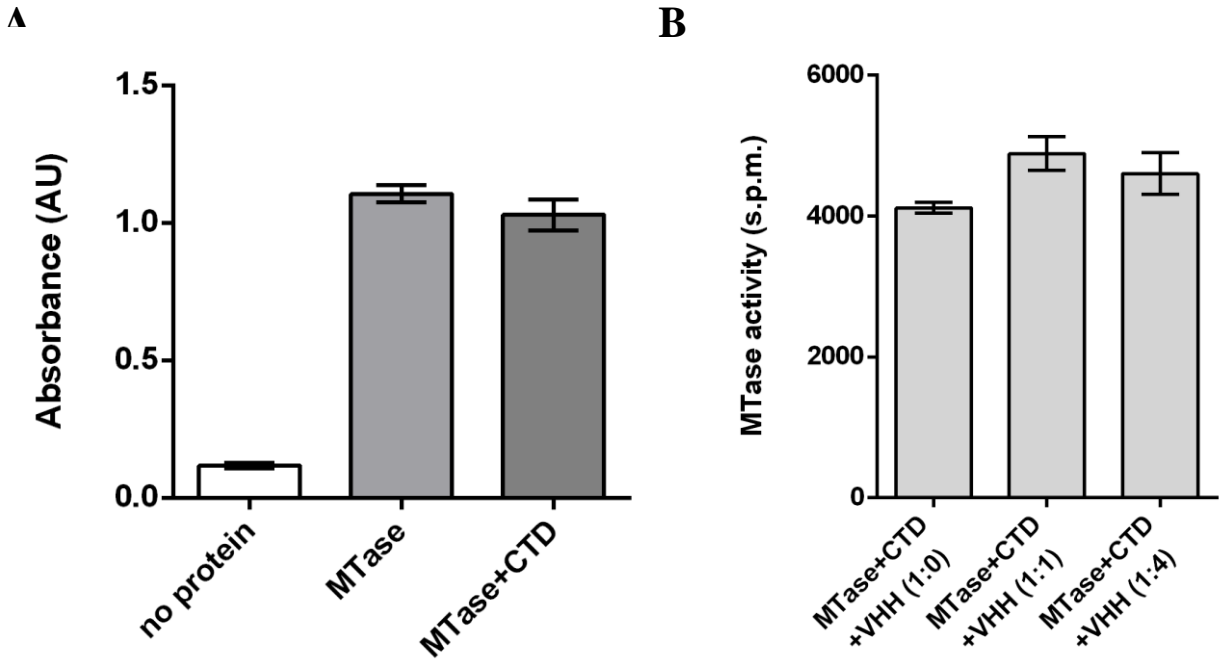


Figure S2 - Characterization of nanobody interaction with SUDV MTase+CTD domain

(A) Evaluation of *Sudan ebolavirus* (SUDV) methyltransferase alone (MTase) or along with the L C-terminal domain (MTase+CTD) with the nanobody by enzyme-linked immunosorbent assay (ELISA). Data represent mean \pm standard deviation (n=5). (B) MTase activity of SUDV MTase+CTD incubated with GpppG-SUDV₁₂, a radiolabeled methyl donor (³H-SAM) and different concentration ratios of nanobody (VHH). Data represent mean \pm standard deviation (n=3).

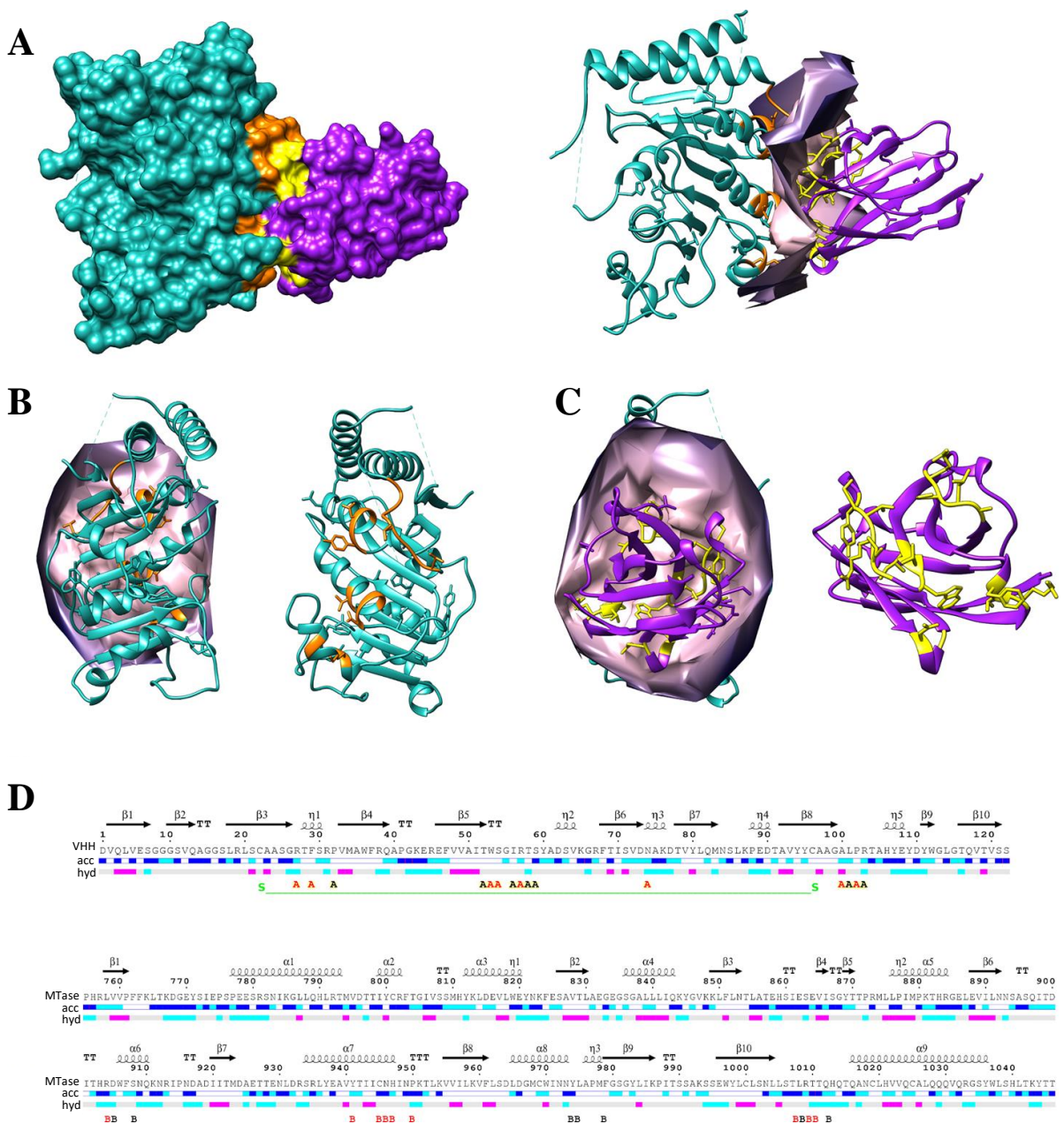


Figure S3 - Surface interaction between SUDV MTase domain and the nanobody (VHH)

(A) Surface (left panel) and ribbon (right panel) representations of the *Sudan ebolavirus* (SUDV) methyltransferase (MTase) domain in complex with a nanobody (VHH). MTase is depicted in cyan and its interacting residues in orange. VHH is represented in purple and its interacting residues in yellow. The interacting surface is represented in purple. (B and C) Identification of the residues on the structure MTase and VHH respectively. (D) Sequences of the heterodimers topped with their secondary structures, and underlined with solvent accessibility, the hydrophobic character of the sequences and contact information. In the structure file, MTase is assigned chainID A and VHH chainID B. The contact is reported with the chainID character in red if the distance is less than 3.2 Å and in black if it is in the range 3.2-4.0 Å. Disulfide bridge is marked by green linked S'.

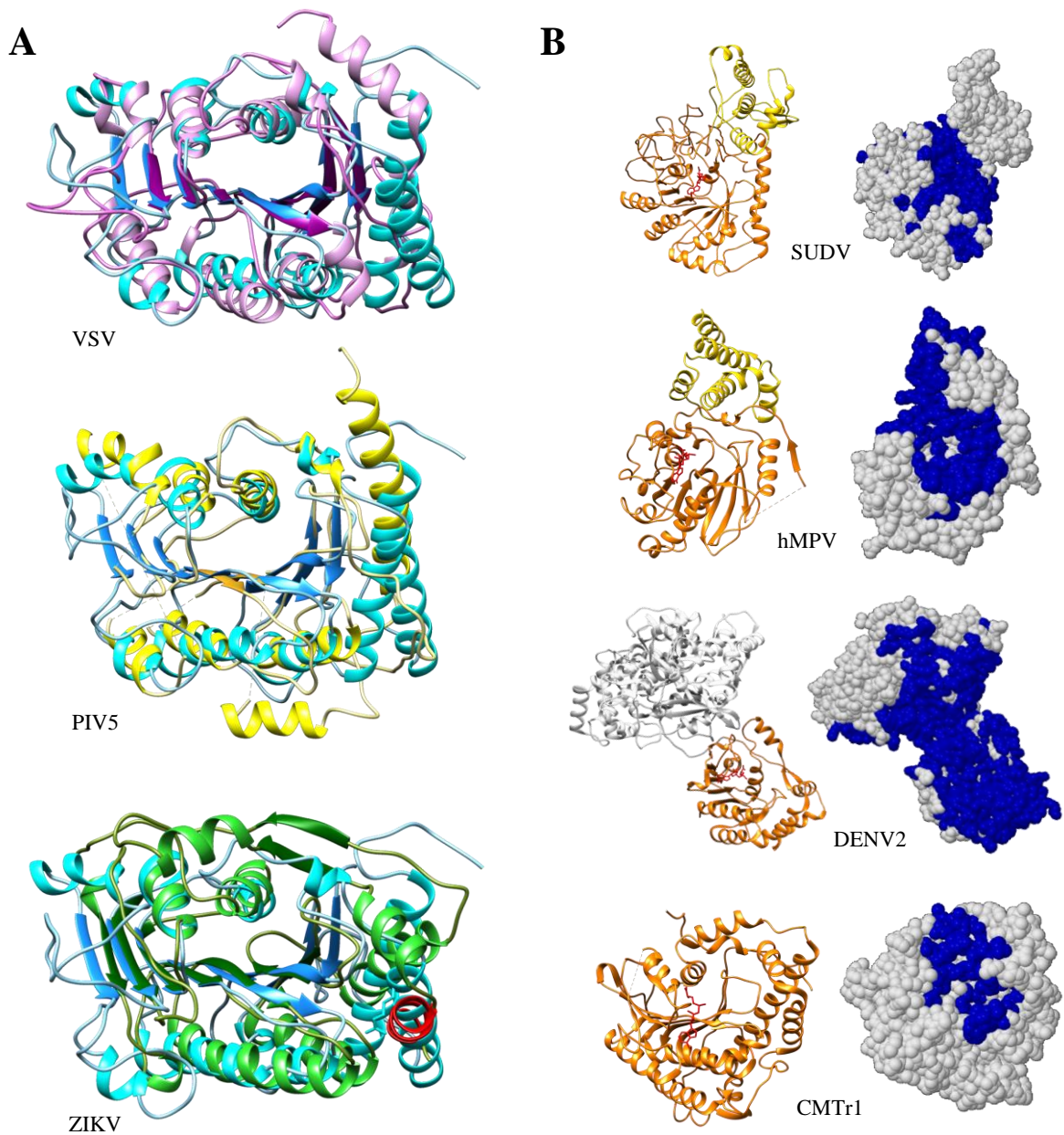


Figure S4 - Further structural comparison with several other viral MTase domains

(A) Superimposition of *Sudan ebolavirus* (SUDV) methyltransferase (MTase, blue) with vesicular stomatitis virus (VSV, PDB: 5A22, top of the panel, purple), parainfluenza 5 virus (PIV5, PDB: 6V85, middle of the panel, yellow) and Zika virus (ZIKV, PDB: 5M5B, bottom of the panel, green) MTase domains. Central β -sheet of all structures are represented in dark blue, dark purple, dark yellow and dark green for SUDV, VSV, PIV5 and ZIKV MTase domains, respectively. Compared to other mononegaviruses, a supplementary N-terminal α -helix in SUDV MTase is also found in ZIKV MTase structure (red). (B) Left column represents structures of SUDV MTase+CTD (modelled using Phyre 2.0), human metapneumovirus (hMPV) MTase+CTD (PDB: 4UCI), dengue 2 virus (DENV2) non structural protein 5 (NS5, PDB: 5ZQK) and human 2' O MTase (CMTr1, PDB: 4N49), respectively from top to bottom. MTase domains are depicted in orange, mononegavirus L C-terminal domains (CTD) in yellow, and other domains in white. Right column represents cartography of positive patches on these same structures using PatchFinderPlus predictor (1). Positive patches are depicted in dark blue.

Table S1 - Structure/function comparison of SUDV MTase domain with other mononegavirus MTases

		SUDV				hMPV				VSV	
		location	sequence	mutant	phenotype	location	sequence	mutant	phenotype	location	sequence
SAMP	loop 1	1831-1835	A-E-G- E -G	E1834A	124% / 198% / 241%	1696-1700	G- E -G-A-G	E1697C	32% / 15%	1670-1674	G-D-G-S-G
				G1835S	17% / 9% / 3%						
	loop 2	1854-1875	T- L -A-T-E-H-S-I-E-S-E-V-I-S-G-Y-T-T-P-R-M-R	T1854A	6% / 1% / 0%	1718-1729	R-S-L-K-D-D-L-D-H- H -Y-P	H1727A	130% / 105%	1693-1710	S-L-L-E-L-S-G-S-V-M-R-G-A-S-P-E-P-P
				L1855A	11% / 8% / 24%						
SUBP	loop 1	1803-1811	F-T-G- I - V - S - S -M-H	I1806A	34% / 8% / 7%	1660-1669	V-N- R -Y-N- F - V - F - S - S	R1662E	29% / 18%	1640-1648	R-L-G-Q-L-P-T-G-A
				V1807A	13% / 33% / 25%			R1662Q	40% / 15%		
				S1808A	78% / 94% / 41%			R1662A	35% / 20%		
				S1809A	4% / 3% / 1%			F1665A	105% / 140%		
				S1990A	24% / 23% / 19%			F1667A	92% / 92%		
	loop 2	1987-1995	P-I-T- S - S -A- K -S-S	S1991A	99% / 45% / 44%	1840-1847	Q-G- S - K -L-S-G	S1668A	27% / 48%	1825-1832	E-F-S-S-S-Q-T-S
				K1993A	79% / 8% / 2%			S1669A	10% / 2%		
				S1842A	70% / 25%			K1843A	70% / 25%		
	NSP	loop1	D-A- E - T -E-N-L-D-R	E1926A	17% / 33% / 37%	1779-1784	D-A- E -F-K-D	D1779A	10% / 0%	1762-1767	D-M-E-V-R-D
				T1927A	51% / 37% / 16%			D1779Q	28% / 12%		
					E1781A			75% / 77%			
					E1781Q			132% / 73%			

This table summarizes phenotypes of single mutation analysis in both *Sudan ebolavirus* (SUDV) and human metapneumovirus (hMPV) methyltransferase domains (MTase, PDB: 6YU8 and 4UCI, respectively) and homologous loop sequences of the different functional pockets including those of vesicular stomatitis virus (VSV, PDB: 5A22). As previously described for hMPV (2), loops were classified according to the functional pocket they belong to: the S-adenosylmethionine (SAM)-binding pocket (^{SAMP}), the RNA-binding pocket (^{SUBP}) and the nucleoside-binding pocket (^{NSP}). For each virus, location, sequence and mutants of each loop (if available) were reported. Mutations are highlighted in blue in the loop sequence and phenotypes are indicated as percentages of cap-N7, cap-2'O and internal 2'O activities according to wild-type proteins, respectively. The color code reflect the impact of the mutation (from dark green (>150% of WT activity) to red (0-25% of WT activity)).

3. Les MTases comme cibles antivirales

Parallèlement au travail réalisé pour la caractérisation des domaines C-terminaux de la protéine L des *Mononegavirales*, j'ai également été impliquée dans divers projets de criblage d'inhibiteurs de l'activité MTase dont un exemple est présenté ci-dessous.

Article 4: Design, Synthesis and Discovery of N,N'-Carbazoyl-aryl-urea Inhibitors of Zika NS5 Methyltransferase and Virus Replication.

Auteurs: Sharon Spizzichino, Giulio Mattedi, Kate Lauder, **Coralie Valle**, Wahiba Aouadi, Bruno Canard, Etienne Decroly, Suzanne J. F. Kaptein, Johan Neyts, Carl Graham, Zakary Sule, David J. Barlow, Romano Silvestri, and Daniele Castagnolo

Référence : ChemMedChem. 2020 Feb 17;15(4):385-390. doi: 10.1002/cmdc.201900533.

Les virus ont développé des mécanismes de subversion des défenses de l'hôte comme l'ajout d'une structure coiffe en 5' de leurs transcrits. De plus, de nombreux virus se répliquant dans le cytoplasme ont évolué et codent pour des MTases leur permettant de limiter leur détection par le système immunitaire. Comme présenté en introduction, certains senseurs de l'immunité innée sont sensibles aux 2'O méthylations. En effet, les récepteurs de type RLRs (MDA5 et RIG-I) ainsi que les effecteurs cellulaires, tels que IFIT, produits en réponse à l'activation de la voie de l'IFN sont capables de discriminer les ARNs coiffés 2'O méthylés, des ARNs coiffés ne présentant pas cette modification (*voir chapitre 5*). Le rôle de cette structure coiffe dans la stabilité des ARNm ainsi que dans l'initiation de la traduction est bien décrit. Cibler les MTases virales responsables de l'ajout des groupements méthyl en position 2'O du premier nucléotide et en position N7 de la structure coiffe permettrait d'une part, de limiter la réplication virale et d'autre part pourrait booster la reconnaissance des ARNs exogènes par les senseurs de l'immunité innée conduisant à la clairance virale. Des études se sont intéressées au développement d'inhibiteurs ciblant les MTases virales (299,359,360), notamment chez des flavivirus tels que le virus de la Dengue (DENV) et le virus Zika (ZIKV), dont les structures cristallographiques ont été résolues (361,362).

Dans cette étude, un criblage virtuel de chimiothèque sur la base de la structure MTase de ZIKV (NS5) a été mené dans le but d'identifier des molécules capables d'accommoder le site de fixation du SAM. Dans ce projet, mon travail a consisté à tester le potentiel inhibiteur des composés sélectionnés sur l'activité MTase de ZIKV et DENV. Ce travail illustre l'importance des données structurales dans le développement de stratégies antivirales et le criblage de molécules à potentiel antiviral, pouvant être par la suite validé *in vitro* par des essais enzymatiques optimisés. La résolution de la structure du domaine MTase de SUDV

ouvre la voie vers des campagnes de criblage virtuel permettant le développement d'inhibiteurs spécifiques des activités MTases des filovirus. De plus, la caractérisation de l'activité enzymatique du domaine MTase-CTD de SUDV permettrait de caractériser *in vitro* le potentiel inhibiteur des molécules sélectionnées lors des études *in silico*.

Design, Synthesis and Discovery of *N,N'*-Carbazoyl-aryl-urea Inhibitors of Zika NS5 Methyltransferase and Virus Replication

Sharon Spizzichino,^[a, b] Giulio Mattedi,^[a] Kate Lauder,^[a] Coralie Valle,^[c] Wahiba Aouadi,^[c] Bruno Canard,^[c] Etienne Decroly,^[c] Suzanne J. F. Kaptein,^[d] Johan Neyts,^[d] Carl Graham,^[a] Zakary Sule,^[a] David J. Barlow,^[a] Romano Silvestri,^[b] and Daniele Castagnolo^{*[a]}

The recent outbreaks of Zika virus (ZIKV) infection worldwide make the discovery of novel antivirals against flaviviruses a research priority. This work describes the identification of novel inhibitors of ZIKV through a structure-based virtual screening approach using the ZIKV NS5-MTase. A novel series of molecules with a carbazoyl-aryl-urea structure has been discovered and a library of analogues has been synthesized. The new compounds inhibit ZIKV MTase with IC₅₀ between 23–48 μM. In addition, carbazoyl-aryl-ureas also proved to inhibit ZIKV replication activity at micromolar concentration.

Zika virus (ZIKV)^[1,2] is a single positive-strand RNA mosquito-borne pathogen belonging to the family of *Flaviviridae*, genus *Flavivirus* and it causes mild-severe diseases in humans and animals. ZIKV is closely related to Dengue virus (DENV), and just like DENV, is mainly transmitted to humans by bites of an

infected *Aedes* species mosquitoes (*Ae. aegypti* and *Ae. albopictus*).^[3] However, ZIKV can also be transmitted between humans through contact with blood or other body fluids, sexual transmission^[2,4] and by mother-to-fetus transmission during the pregnancy.^[6] While global epidemics of DENV have spread over the past few decades causing more than 20,000 deaths per year, ZIKV infections have emerged as a major public health concern only in the last few years. Before 2007, only sporadic human disease cases were reported in Africa and Asia. However, following the recent outbreaks in the Americas in 2015, ZIKV was declared by the World Health Organization (WHO) as a Public Health Emergency of International Concern (PHEIC) and the first case of sexually-transmitted ZIKV was reported in USA in 2008. While ZIKV infections generally cause a mild fever, headache, malaise, skin rashes and joint pain, they have been associated in several cases with severe neurological and fetal complications leading to microcephaly in new-borns^[4] and neurological diseases, such as the Guillain-Barré syndrome (GBS), in adults.^[5] The recent *Flavivirus* outbreaks, as well as the increasing number of cases of ZIKV infections worldwide, have raised the attention of pharmaceutical industries and healthcare providers toward the identification and development of efficient treatments against these diseases. Although a vaccine against DENV has recently been commercialized (DENVAXIA[®]),^[6] there is as yet no vaccine against ZIKV available. In addition, there are no drugs available to treat or prevent ZIKV infections, especially in the event of an outbreak. A limited number of early-phase discovery studies have identified few inhibitors of DENV and ZIKV replication,^[7–15] however there are no clinically approved drugs yet available to target flaviviruses directly nor any that may serve as vaccine adjuvants. The development of new antivirals thus represents a research priority.

Within this context, viral proteins represent appealing targets for the development of novel antiviral therapies. Examples of current flavivirus inhibitors, such as 1–5 (Figure 1), have been designed to target the DENV viral proteins NS3 (protease domain,^[16] helicase domain,^[17] and full-length NS3), NS5 (N-terminal Methyltransferase domain (MTase)^[18–19] and C-terminal RNA-dependent RNA polymerase (RdRp).^[20] The NS3 and NS5 proteins of DENV and ZIKV show a high degree of homology, and their crystal structures have recently been determined.^[21–23] While DENV NS5-polymerase has been investigated by a number of research groups as a potential target for

[a] S. Spizzichino, G. Mattedi, K. Lauder, C. Graham, Z. Sule, Dr. D. J. Barlow, Dr. D. Castagnolo

School of Cancer and Pharmaceutical Sciences
King's College London
London, SE1 9NH (UK)

E-mail: daniele.castagnolo@kcl.ac.uk

[b] S. Spizzichino, Prof. R. Silvestri

Department of Drug Chemistry and Technologies
Sapienza University of Rome

Laboratory Affiliated to Istituto Pasteur Italia – Fondazione Cenci Bolognetti
Piazzale Aldo Moro 5
00185, Roma (Italy)

[c] C. Valle, Dr. W. Aouadi, Dr. B. Canard, Dr. E. Decroly

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

[d] Dr. S. J. F. Kaptein, Prof. J. Neyts

Department of Microbiology, Immunology and Transplantation
Rega Institute for Medical Research
Laboratory of Virology and Chemotherapy
KU Leuven

Minderbroedersstraat 10
3000 Leuven (Belgium)

Supporting information for this article is available on the WWW under <https://doi.org/10.1002/cmdc.201900533>

© 2019 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Examples of DENV and ZIKV inhibitors

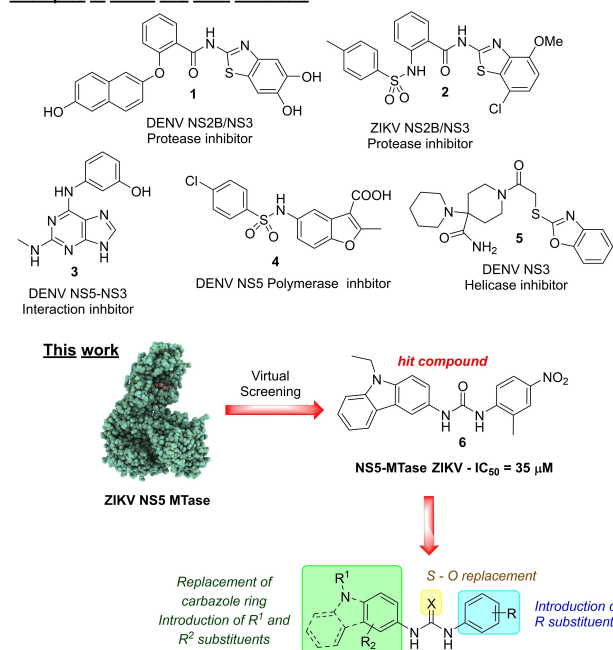


Figure 1. Examples of small molecules active against DENV and/or ZIKV, and overview of this work.

development of new antivirals, there has been very little work carried out on DENV and ZIKV NS5-MTase.^[23–24,18] The NS5-MTase is responsible for maturation of the viral RNA cap and catalyzes the methylation of the N7 position of a guanine and the 2'-OH of the first ribonucleotide to yield¹⁷MeGpppA₂OMe-RNA.^[25] These methylations of the RNA cap structure play a key role during virus replication and are critical to virus survival in infected animals.^[26] Indeed, biochemical studies and reverse genetic analysis have shown that N7-MTase activity is essential for mRNA translation into viral protein, and so viruses devoid of N7-MTase show a strongly reduced replication phenotype.^[18] By contrast, 2'-O-MTase defective viruses can replicate moderately well in infected cells, but are highly attenuated in mice or rhesus monkey and induce a strong antiviral response.^[27] Thus, the inhibition of both N7- and 2'-O-MTase activities should restrain viral replication, making the NS5-MTase a promising target for the development of new anti-ZIKV, and potentially anti-DENV, antivirals.^[24]

Herein, a structure-based virtual screening on a set of chemical libraries using the ZIKV NS5-MTase was performed, with the aim to identify structurally novel flaviviruses inhibitors.

The structure of the ZIKV NS5 MTase was built by homology modelling using SWISS-MODEL^[28–29] and the compounds in the NCI Diversity Set V database docked in the S-adenosyl methionine (SAM) binding site in order to identify molecules that might disrupt the activity of the enzyme. A set of 40 best ranking molecules were initially identified, among which the urea **6** emerged as the best candidate due to its favorable interactions with the SAM binding site. The urea **6** was thus evaluated in enzymatic inhibition assays against ZIKV NS5-MTase, showing an IC₅₀ of 35 μM (Table 1). A similar IC₅₀ of

Table 1. IC₅₀ values of compounds **6**, **9**, **11** and **21** against host and flavivirus MTases.

Compd	hRNMT IC ₅₀ [μM]	ZIKV NS5-MTase	DENV NS5-MTase
6	13.9 ± 0.7	35 ± 5	38 ± 6.7
9a	≈ 122	46 ± 1.1	≈ 133
9b	≈ 253	≈ 144	≈ 267
9c	27.4 ± 1.9	NA ^[b]	≈ 62
9d	5.3 ± 0.4	≈ 397	60 ± 32
9f	6.8 ± 0.5	NA ^[b]	NA ^[b]
11b	≈ 112	70 ± 1.3	≈ 236
21b	nd ^[a]	114 ± 1.3	NA ^[b]
21c	nd ^[a]	23 ± 1.2	NA ^[b]
21e	nd ^[a]	48 ± 1.3	NA ^[b]
21f	nd ^[a]	26 ± 1.2	NA ^[b]
Sinefungin	nd ^[a]	1.18 ± 0.05	0.63 ± 0.04

[a] Activity not determined. [b] No activity observed above 50 μM.

38 μM was observed when **6** was assayed against DENV NS5-MTase. Due to the novelty of the structure when compared to known flavivirus inhibitors, the urea **6** was thus selected as a hit compound for further studies. Figure 2b shows the most favorable binding pose of the hit compound **6** in the ZIKV NS5-MTase binding pocket. The urea spacer connects the carbazole with a 2-methyl-4-nitro-phenyl group, which occupies the region that interacts with the methionine backbone of SAM (Figure 2a). The nitro group of **6** interacts with the side chain of S62 in a similar way to the SAM carboxyl group, possibly stabilizing the interaction of the compound with the protein. The high sequence and structural identity between ZIKV and DENV NS5-MTases led to the identical binding of compound **6** in both proteins (Figure 2c).

Next, a library of analogues of compound **6** was designed and synthesized to explore the chemical space around the urea scaffold (Scheme 1). The carbazole derivatives **9a–f** were first synthesized through reaction of carbazole **7** with the appropriate isocyanate **8** in toluene at 60 °C with the aim to investigate the effect of different substituents (electron-donating and electron-withdrawing) on the phenyl ring bound to the urea moiety (Scheme 1a). Similarly, thioureas **11a–b** were synthesized from **10a–b** to investigate the role of the isosteric sulfur in place of the urea oxygen atom and to evaluate the

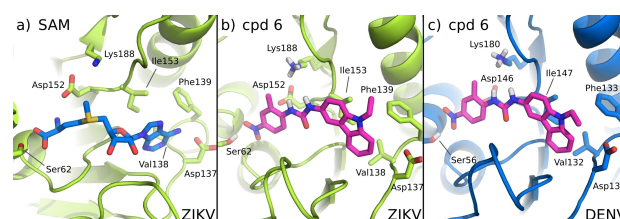
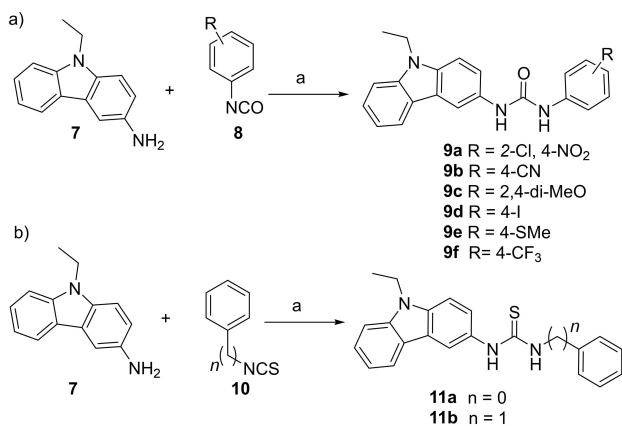


Figure 2. a) SAM bound ZIKV NS5 MTase in PDB 5M5B.^[23] b) Docking of compound **6** in ZIKV NS5 MTase. The carbazole pocket is accommodated in the hydrophobic environment that binds the adenine portion of SAM. The *p*-nitro group interacts with Ser62, mimicking the carboxy moiety of SAM. The carbamide spacer allows the two domains to be spaced appropriately for favorably interacting with the key residues of the binding pocket. c) Docking of compound **6** in DENV NS5-MTase (PDB 5E9Q).^[24]



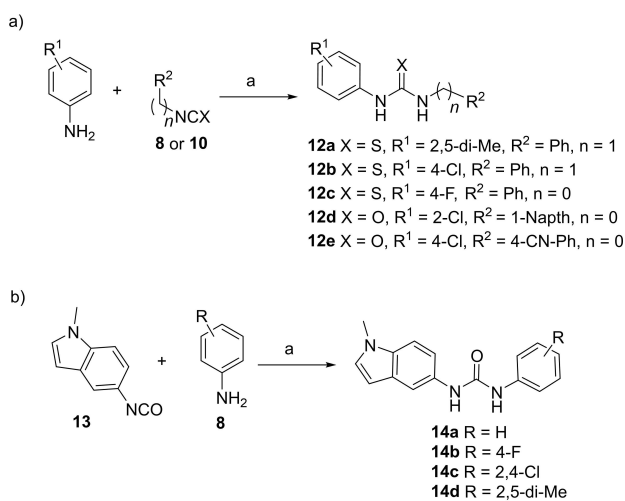
Reagents and conditions: (a) Toluene, 60 °C, 24h

Scheme 1. Synthesis of urea derivatives 9a–f and 11a,b.

influence of an extra methylene group in 11b on MTase inhibition (Scheme 1b).

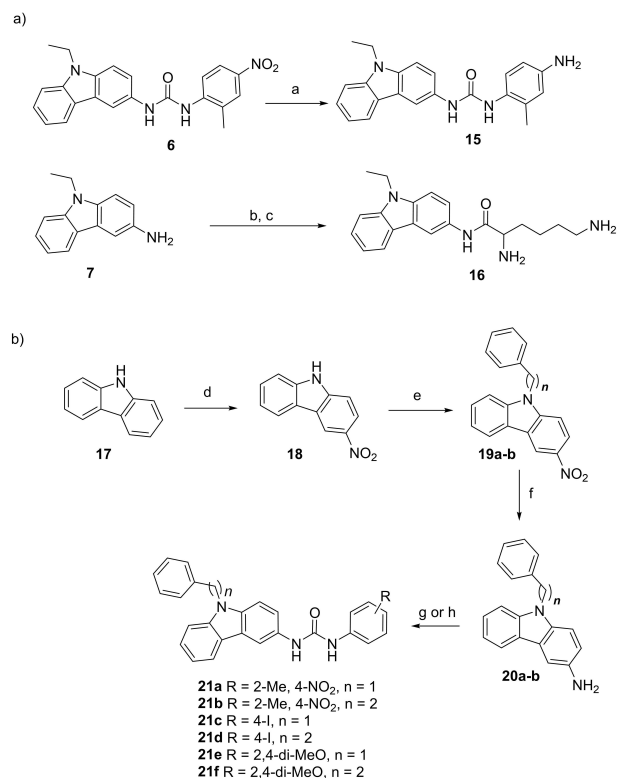
The role of the carbazole ring on the antiviral activity was also explored through its replacement with various phenyl moieties (derivatives 12a–e) and indole ring (derivatives 14a–d). Ureas 12a–e and 14a–d were obtained from appropriate isocyanates 8 and 10 in toluene according to Scheme 2.

In order to evaluate the role of the nitro group on antiviral activity, the amino derivatives 15 and 16 were also synthesized; compound 6 was treated with Fe⁽⁰⁾ in conc. HCl leading to 15 in high yields, while carbazole 7 was reacted under microwave irradiation with Cbz,Boc-lysine to give, after deprotection in HCl/AcOEt, the pure 16 (Scheme 3a). Finally, we also tried to modify the *N*-ethyl substituent of 6 in order to increase the selectivity of our inhibitors targeting the MTase SAM binding site. For this purpose, we compared the structure of the human



Reagents and conditions: (a) Toluene, 60 °C, 24h

Scheme 2. Synthesis of urea derivatives 12a–e and 14a–d.



Reagents and conditions: (a) Fe, HCl conc., MeOH, DCM, rt; (b) *N*-Cbz-*N*-Boc-Lysine, DIPEA, HOBT, EDC, DMF, MW, 60 °C, 10 min; (c) HCl/AcOEt, 48 h; (d) HNO₃ 60%, H₂O, 80 °C, 3h; (e) Benzylchloride or 2-phenylethylbromide, NaH, DMF, 90 °C, 2h; (f) SnCl₂, HCl conc., AcOH, 100 °C, 12h; (g) 2-Me,4-NO₂-isocyanate (prepared *in situ* from 2-Me,4-NO₂-aniline, triphosgene, Et₃N, 0 °C, THF, 4h), toluene, 60 °C, 12h; (h) isocyanate, toluene, 60 °C, 12h

Scheme 3. Synthesis of urea derivatives 15, 16 and 21a–f.

N7 MTase (hRNMT) involved in the capping of cellular mRNA, with that of both viral MTases. The X-ray structure revealed structural similarities between the viral and the human MTases (NS5 and hRNMT). However, both ZIKV and DENV NS5-MTase possess a hydrophobic pocket located near the exocyclic amide of SAM,^[18] which is absent in the hRNMT. In the ZIKV MTase, the region is defined by the amino acids F139, I153, G164, E155, R166, and V170, compared to residues F133, I147, G148, E149, R160, and V164 in the DENV MTase. We anticipated that the absence of this hydrophobic pocket in hRNMT could help to increase compound selectivity for the viral MTases and, hence, limit the interference of inhibitors with host MTases. In the docking snapshot, the ethyl group of compounds 6 is located in proximity of this pocket, clearly showing that its replacement with a benzyl or a phenylethyl moiety could lead to derivatives able to selectively recognize and inhibit only the viral enzymes over the host enzymes (Figure 3).

Thus, the synthesis of derivatives 21a–f bearing a phenyl-alkyl substituent on the carbazole nitrogen was planned. Carbazole 17 was treated with HNO₃ and converted into the nitro derivative 18, which was in turn alkylated on the aromatic nitrogen with appropriate benzyl or alkyl halide in the presence of NaH to give compounds 19a–b. Reduction of the nitro group of 19 with SnCl₂ led to amino derivatives 20a–b. The latter

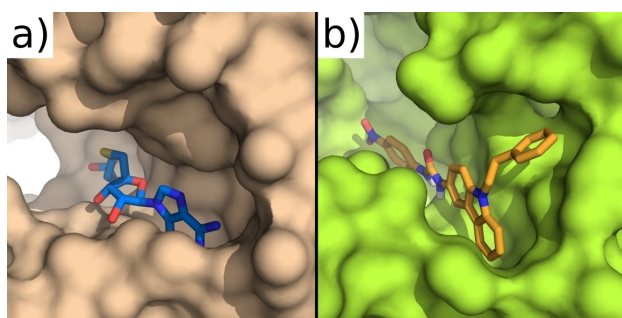


Figure 3. a) SAH bound to hRNTM (PDB 5E8J).^[30] b) Docking of compound **21b** in ZIKV NS5 MTase. The ethyl spacer on the carbazole nitrogen allows the phenyl ring to interact effectively with the hydrophobic pocket near the SAM binding site. The absence of the region in hRNTM is a potential determinant of target selectivity.

were then treated with 4-*l*-phenyl- or 2,4-(MeO)₂-phenyl-isocyanate in toluene at 60 °C affording ureas **21c–f** in high yields. A different strategy was adopted to synthesize the derivatives **21a–b**, due to the commercial lack of the appropriate isocyanate reagent. The 2-methyl-4-nitro-aniline was treated with triphosgene at 0 °C in the presence of Et₃N leading to the formation of the corresponding isocyanate, which was reacted *in situ* with carbazole **7** to give the desired ureas **21a–b** in good yields (Scheme 3).

The library of compounds **9**, **11**, **12**, **14**, **15**, **16** and **21** was then assessed on different purified recombinant MTases. All compounds were initially screened at 50 μM against ZIKV NS5-MTase and hRNTM as well as against DENV NS5-MTase due to its similarity with ZIKV protein. The SAM mimetic sinefungin was included as control.

The MTases were incubated with radiolabeled [³H]-SAM together with the GpppAC₄ RNA substrate and the different compounds at 50 μM. The reaction was stopped after 30 mins at 30 °C. The sample products were filtered on a DEAE membrane to remove non-incorporated [³H]-SAM, and radioactivity transferred on the RNA substrate was counted. Compounds **6**, **9a**, **9d**, **21b**, **21c**, **21e** and **21f** showed inhibition of ZIKV NS5-MTase higher than ~30%. The compound **21e** and **21f** showed the more potent reduction of ZIKV NS5-MTase activity of ~70%.

Conversely the compounds **15** and **16** and those of series **11** and **14** barely inhibited the ZIKV MTase. Interestingly compound **6** and those of the series **9** showed a similar inhibition profile on DENV NS5-MTase, but the compounds of series **21** barely inhibit the DENV MTase suggesting some specificity of this family. A dose-response assay was then performed for the most promising compounds and the IC₅₀ values deduced from titration curves after curve fitting are shown in Table 1. The results indicate that the replacement of the methyl group of **6** with a chlorine atom in **9a** did not affect the inhibitory activity of the compound against ZIKV MTase, whilst it proved to be detrimental for inhibiting the DENV MTase. Similarly, the replacement of the electron withdrawing nitro group of **6** with other substituents (i.e. the electron donating iodine or methoxy in **9c–d**, or electron withdrawing

–CN or –CF₃ in **9b** and **9f**) negatively affected the compound's inhibitory activity mainly against the ZIKV MTase. The different activity of compounds **6** and **9** can be explained by taking in consideration the crucial H-bond interaction between nitro groups of **6** with the hydroxyl group of S62, which is not possible for derivatives **9b–f**. The low activity of compounds **11b** can be attributed to the lack of appropriate substituents on the phenyl ring able to interact with S62. On the other hand, the carbazole ring proved to be crucial for the antiviral activity since its replacement with other rings in **12** and **14** led to inactive compounds. Derivatives **21c**, **21e** and **21f** showed good activity against ZIKV MTase at concentrations similar to **6** (IC₅₀ = 23–48 μM). It is likely that the *N*-benzyl or *N*-phenylethyl substituents in compounds **21** are helping these compounds to bind to the viral proteins even if no substantial improvement from **6** was detected. A dose-response curve for hit compound **6** is reported in Figure 4.

Finally, the ability of all the derivatives synthesized in this work to inhibit ZIKV replication was assessed. Results of the most active compounds, as well as their effect on cell viability, are reported in Table 2. Compound **6** and **9a–d**, which showed MTase inhibitory activity, were devoid of any antiviral activity against ZIKV.

Possible reasons for the lack of antiviral activity may be the inability of the molecules to enter the cell, or their conversion into inactive metabolites in infected cells. However, since compound **6** proved to be cytotoxic (CC₅₀ > 3.99 μM), it is likely

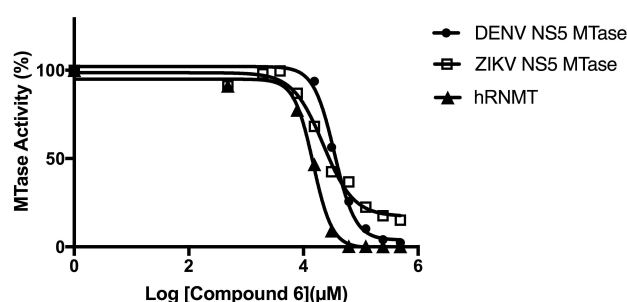


Figure 4. Dose-response curve for hit compound **6**.

Table 2. EC ₅₀ and CC ₅₀ values against ZIKV of the most active compounds of the library.		
Compd	ZIKV EC ₅₀ [μM]	CC ₅₀ [μM]
6	> 3.99	3.99
9a	> 50	30.7
9b	> 10	> 10
9c	> 10	> 10
9d	> 50	> 50
11b	4.78	19.57
15	1.67	> 12.5
21b	12.5	20
21d	25	100
21f	> 20	20
Compd	DENV EC ₅₀ [μM]	CC ₅₀ [μM]
15	9.75	20

that this compound and analogues thereof can enter the cells. Moreover, the inhibitory activity of **6** and **9a–d** against the host MTase (Table 1) hints toward an antimetabolic effect. Another possible explanation is that compounds **6** and **9a–d** mainly inhibit the 2'-O MTase activity without affecting the N7 MTase activity. This could explain the very poor inhibition of ZIKV replication in the cell-based assay, as the 2'-O methylation of the cap is not essential for viral replication in most cell lines. Derivative **11b**, which showed some MTase inhibitory activity, also exhibited antiviral activity against ZIKV in the cell-based assay ($EC_{50} = 4.78 \mu\text{M}$). Interestingly, derivative **15**,^[31] which did not show any activity against ZIKV NS5-MTase, was able to inhibit ZIKV replication with an EC_{50} of $1.67 \mu\text{M}$. It is plausible that **15** inhibits ZIKV replication via a different mode of action than via NS5-MTase inhibition, or that this compound is partially metabolized in the treated cells.

Interestingly, compound **15** also showed some activity against DENV ($EC_{50} = 9.75 \mu\text{M}$), even if the antiviral activity is clearly linked to an adverse toxic effect on the host cell ($CC_{50} = 20 \mu\text{M}$). Finally, compound **21b** which showed no enhanced activity against ZIKV MTases, proved to inhibit ZIKV replication at good concentration ($EC_{50} = 12.5 \mu\text{M}$). However, compound **21b** showed toxicity in cellular assay as well as **21f** which was found active on ZIKV MTase instead. This could suggest that these compounds might target another MTase (i.e. cellular MTase) involved in virus replication. Interestingly, compound **21d**, which showed poor inhibition of ZIKV MTase, inhibits ZIKV with good $EC_{50} = 25 \mu\text{M}$ and no toxicity ($CC_{50} = 100 \mu\text{M}$).

In summary, a structure-based virtual screening protocol on a set of chemical libraries using the ZIKV NS5-MTase was performed leading to the identification of a novel class of carbazoyl-urea inhibitors. Compounds **6** and **21c**, **21e**, **21f** showed inhibitory activity against ZIKV MTase as well as the related DENV MTase. Conversely, the urea derivative **15** showed an antiviral effect against ZIKV with EC_{50} of $1.67 \mu\text{M}$, despite its poor inhibitory activity toward the NS5-MTase *in vitro*. Studies to identify novel MTase inhibitors, to optimize the structure-activity relationships of carbazoyl-urea analogues of **6** and to fully unravel their mode of action are currently in progress in our laboratories.

Experimental Section

Experimental details are reported in the Supporting Information.

Acknowledgements

This work was supported by the European program H2020 under the ZIKAlliance project (grant agreement 734548), the EVAg Research Infrastructure (grant agreement 653316) and by the French research agency ANR (VMTaseln, grant ANR-ST14-ASTR-0026). DC gratefully acknowledges EPSRC (Global Challenges Competition King's College London) and Royal Society (RG160870) for research funding and financial support. We thank Caroline Collard, Charlotte Vanderheydt and Elke Maas for their assistance

in the acquisition of the antiviral data. C.V. get a fellowship from the National Research Agency ANR under the program RAB-Cap and W.A. from the infectiopoie Sud.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: Zika · flavivirus · methyltransferase · antiviral agents · urea

- [1] D. Baud, D. J. Gubler, B. Schaub, M. C. Lanteri, D. Musso, *Lancet* **2017**, *390*, 2099–109.
- [2] L. R. Petersen, D. J. Jamieson, A. M. Powers, M. A. Honein, *N. Engl. J. Med.* **2016**, *374*, 1552–1563.
- [3] European Centre for Disease Prevention and Control, Zika Virus disease. <https://ecdc.europa.eu/en/zika-virus-disease>
- [4] A. Hajra, D. Bandyopadhyay, L. R. Bhadra, S. Ball, S. K. Hajra, *Am. J. Reprod. Immunol. (1980-1984)* **2017**, *77*, e12607.
- [5] C. A. Fontes, A. A. Dos Santos, E. Marchiori, *Neuroradiology* **2016**, *58*, 837–838.
- [6] European Medicines Agency, Dengvaxia. <https://www.ema.europa.eu/en/medicines/human/EPAR/dengvaxia>.
- [7] Y. S. Tian, Y. Zhou, T. Takagi, M. Kameoka, N. Kawashita, *Chem. Pharm. Bull.* **2018**, *66*, 191–206.
- [8] V. Boldescu, M. A. M. Behnam, N. Vasilakis, C. D. Klein, *Nat. Rev. Drug Discovery* **2017**, *16*, 565–586.
- [9] L. Botta, M. Rivara, V. Zuliani, M. Radi, *Front. Biosci.* **2018**, *23*, 997–1019.
- [10] M. A. Behnam, C. Nitsche, V. Boldescu, C. D. Klein, *J. Med. Chem.* **2016**, *59*, 5622–5649.
- [11] M. Xu, E. M. Lee, Z. Wen, Y. Cheng, W. K. Huang, X. Qian, J. Tcw, J. Kouznetsova, S. C. Ogden, C. Hammack, F. Jacob, H. N. Nguyen, M. Itkin, C. Hanna, P. Shinn, C. Allen, S. G. Michael, A. Simeonov, W. Huang, K. M. Christian, A. Goate, K. J. Brennan, R. Huang, M. Xia, G. L. Ming, W. Zheng, H. Song, H. Tang, *Nat. Med.* **2016**, *22*, 1101–1107.
- [12] H. Lee, J. Ren, S. Nocadello, A. J. Rice, I. Ojeda, S. Light, G. Minasov, J. Vargas, D. Nagarathnam, W. F. Anderson, M. E. Johnson, *Antiviral Res.* **2017**, *139*, 49–58.
- [13] E. D. Micewicz, R. Khachatourian, S. W. French, P. Ruchala, *Bioorg. Med. Chem. Lett.* **2018**, *28*, 452–458.
- [14] S. J. F. Kaptein, P. Vincetti, E. Crespan, J. I. A. Rivera, G. Costantino, G. Maga, J. Neyts, M. Radi, *ChemMedChem* **2018**, *13*, 1371–1376.
- [15] P. Vincetti, S. J. F. Kaptein, G. Costantino, J. Neyts, M. Radi, *ACS Med. Chem. Lett.* **2019**, *10*, 558–563.
- [16] H. Wu, S. Bock, M. Snitko, T. Berger, T. Weidner, S. Holloway, M. Kanitz, W. E. Diederich, H. Steuber, C. Walter, D. Hofmann, B. Weißbrich, R. Spannaus, E. G. Acosta, R. Bartenschlager, B. Engels, T. Schirmeister, J. Bodem, *Antimicrob. Agents Chemother.* **2015**, *59*, 1100–1109.
- [17] C. M. Byrd, D. W. Grosenbach, A. Berhanu, D. Dai, K. F. Jones, K. B. Cardwell, C. Schneider, G. Yang, S. Tyavanagimatt, C. Harver, K. A. Wineinger, J. Page, E. Stavale, M. A. Stone, K. P. Fuller, C. Lovejoy, J. M. Leeds, D. E. Hruby, R. Jordan, *Antimicrob. Agents Chemother.* **2013**, *57*, 1902–1912.
- [18] S. P. Lim, L. S. Sonntag, C. Noble, S. H. Nilar, R. H. Ng, G. Zou, P. Monaghan, K. Y. Chung, H. Dong, B. Liu, C. Bodenreider, G. Lee, M. Ding, W. L. Chan, G. Wang, Y. L. Jian, A. T. Chao, J. Lescar, Z. Yin, T. R. Vedananda, T. H. Keller, P. Y. Shi, *J. Biol. Chem.* **2011**, *286*, 6233–6240.
- [19] J. Hernandez, L. Hoffer, B. Coutard, G. Querat, P. Roche, X. Morelli, E. Decroly, K. Barral, *Eur. J. Med. Chem.* **2019**, *161*, 323–333.
- [20] P. Vincetti, F. Caporuscio, S. Kaptein, A. Gioiello, V. Mancino, Y. Suzuki, N. Yamamoto, E. Crespan, A. Lossani, G. Maga, G. Rastelli, D. Castagnolo, J. Neyts, P. Leyssen, G. Costantino, M. Radi, *J. Med. Chem.* **2015**, *58*, 4964–4975.
- [21] B. Zhao, G. Yi, F. Du, Y. C. Chuang, R. C. Vaughan, B. Sankaran, C. C. Kao, P. Li, *Nat. Commun.* **2017**, *8*, 14762.
- [22] Z. Zhang, Y. Li, Y. R. Loh, W. W. Phoo, A. W. Hung, C. Kang, D. Luo, *Science* **2016**, *354*, 1597–1600.

- [23] B. Coutard, K. Barral, J. Lichière, B. Selisko, B. Martin, W. Aouadi, M. O. Lombardia, F. Debart, J. J. Vasseur, J. C. Guillemot, B. Canard, E. Decroly, *J. Virol.* **2017**, *91*, e02202-e02216.
- [24] F. Benmansour, I. Trist, B. Coutard, E. Decroly, G. Querat, A. Brancale, K. Barral, *Eur. J. Med. Chem.* **2017**, *125*, 865–880.
- [25] H. Dong, D. C. Chang, X. Xie, Y. X. Toh, K. Y. Chung, G. Zou, J. Lescar, S. P. Lim, P. Y. Shi, *Virology* **2010**, *405*, 568–578.
- [26] E. Decroly, B. Canard, *Curr. Opin. Virol.* **2017**, *24*, 87–96.
- [27] R. Züst, H. Dong, X. F. Li, D. C. Chang, B. Zhang, T. Balakrishnan, Y. X. Toh, T. Jiang, S. H. Li, Y. Q. Deng, B. R. Ellis, E. M. Ellis, M. Poidinger, F. Zolezzi, C. F. Qin, P. Y. Shi, K. Fink, *PLoS Pathog.* **2013**, *9*, e1003521.
- [28] When the work was carried out, the 3D structure of ZIKV NS5-MTase was not elucidated yet. Thus, a homology model based on the DENV NS5 MTase was built. The homology model was then compared to the 3D structure of ZIKV NS5MTase once it was released, showing full similarity.
- [29] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gummienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, *Nucleic Acids Res.* **2018**, *46*, W296–W303.
- [30] D. Varshney, A. P. Petit, J. A. Bueren-Calabuig, C. Jansen, D. A. Fletcher, M. Peggie, S. Weidlich, P. Scullion, A. V. Pislakov, V. H. Cowling, *Nucleic Acids Res.* **2016**, *44*, 10423–10436.
- [31] Compound **15** did not show any effect on DENV MTase and poorly inhibited the ZIKV MTase (25% of inhibition). For this reason, the IC₅₀ of **15** against ZIKV MTase was not determined.

Manuscript received: September 17, 2019
 Revised manuscript received: November 5, 2019
 Accepted manuscript online: December 5, 2019
 Version of record online: January 7, 2020

CONCLUSION GÉNÉRALE ET PERSPECTIVES

Ce projet de thèse a eu pour objectif de caractériser le mécanisme de mise en place de la structure coiffe des ARNs chez des virus de l'ordre des *Mononégavirales*. Le génome des mononégavirus code pour une protéine multifonctionnelle, la protéine L responsable de la synthèse de l'ARN viral et de l'ajout de la structure coiffe en 5' des ARNm viraux. L'extrémité C-terminale présente le domaine MTase suivi du domaine CTD dont la fonction est encore inconnue. Au cours de ma thèse, je me suis focalisée sur la caractérisation des domaines C-terminaux de la protéine L du virus de la rage et du virus Ebola. Dans un premier temps, nous avons tenté de produire et purifier les domaines CD, MTase et CTD de la protéine L du virus de la rage. Au moment où ce projet a débuté, aucune données structurales ni fonctionnelles n'étaient encore connues pour ce virus, ouvrant la voie vers une étude structure/fonction complète. En dépit des efforts et des nombreuses stratégies engagées, nous n'avons pas pu obtenir de protéines du virus de la Rage solubles fonctionnellement actives nous permettant de caractériser les domaines C-terminaux de la protéine L du virus de la rage. Le mécanisme de synthèse de la structure coiffe chez les filovirus est encore peu documenté, nous avons redirigé les objectifs de cette thèse dans ce modèle.

Une étude préliminaire menée au sein du laboratoire a permis de caractériser, pour la première fois, un domaine actif catalysant une activité MTase chez un virus de la famille des filovirus. Le domaine MTase-CTD du virus SUDV a été exprimé en système bactérien et les conditions de purification étaient connues. Les analyses biochimiques avaient par ailleurs permis de mettre en évidence que le domaine MTase de SUDV présentait une activité originale dans la mesure où, en plus des activités MTases canoniques sur la coiffe, il catalyse la méthylation de résidus internes à la séquence d'ARN *in vitro*. Bien que d'autres virus codent pour des activités MTase pouvant méthyler des positions internes des ARN (299,300,363), l'efficacité de ces méthylations sont très faibles en comparaison avec celle de SUDV. Cette observation questionne d'une part les mécanismes moléculaires impliqués dans ce processus de méthylation original chez SUDV et d'autre part le rôle fonctionnel de ces méthylations internes des ARN. L'une des hypothèses quant au rôle de ces modifications épitranscriptomiques est qu'elles limiteraient la détection par les senseurs de l'immunité innée tels que MDA5. En effet, une récente étude menée sur le virus HIV a permis de montrer que le HIV recrute la MTase cellulaire FTSJ3 afin de méthyler son ARN et échapper à la détection par MDA5 (300). Toutefois l'efficacité de méthylation des ARNs du virus Ebola et le rôle de ces modifications restent à démontrer dans le contexte d'infection viral.

Fort de cette étude, nous nous sommes intéressés à caractériser le rôle du domaine CTD dans l'activité MTase de SUDV. Le domaine MTase-CTD du virus SUDV, délété du domaine CTD, a été cloné et exprimé en système bactérien. En comparant l'activité MTase du domaine MTase et MTase+CTD, nous avons mis en évidence le rôle crucial du domaine CTD dans le recrutement de l'ARN et dans l'activité MTase de SUDV. L'étude bio-informatique des domaines CTD des *Mononegavirales* a montré que ces domaines, bien que peu conservé en séquence et en longueur, présentaient un enrichissement en acides aminés basiques. L'alignement des domaines CTD des filovirus a permis d'identifier des résidus basiques et aromatiques conservés qui pourraient être impliqués dans le recrutement du substrat dans la poche catalytique à l'interface entre le domaine MTase et le domaine CTD. En conclusion, cette étude a démontré que le domaine CTD contribue au recrutement des ARN substrat et régule ainsi des différentes activités de la MTases. Bien que le domaine CTD soit l'un des domaines les plus variables en séquence et en longueur au sein des virus de l'ordre des *Mononegavirales*, l'analyse bio-informatique a révélé quelques caractéristiques communes. En effet, une étude précédente des domaines C-terminaux de la protéine L de hMPV et VSV avait déjà révélé un sillon à l'intervalle entre le domaine MTase et le domaine CTD pouvant accommoder le substrat. La fonction du domaine CTD dans l'activité MTase, à savoir sa participation au recrutement de l'ARN pourrait être une caractéristique commune au sein des *Mononegavirales*, rendant ce domaine une cible attractive pour le développement de molécules à potentiel antiviral pan-mononégavirus.

La caractérisation biochimique du domaine MTase-CTD de SUDV a également révélé un profil d'activité MTase original, jamais encore observé chez des virus de l'ordre des *Mononegavirales*. Afin de mieux comprendre les bases moléculaires et d'établir les relations structure/fonction des MTases des *Mononegavirales*, une campagne de cristallogénèse a été menée. La structure du domaine MTase de SUDV en complexe avec un VHH a été résolue à 1,8 Å. La comparaison de ce domaine avec les domaines MTase de hMPV et VSV, dont les structures sont résolues (266,364), nous a permis d'engager une caractérisation structurale et fonctionnelle du domaine MTase de SUDV. En plus des caractéristiques communes observées chez d'autres MTases virales, certaines originalités de la structure SUDV MTase ont été observées et pourraient être impliquées dans la régulation des activités MTases. En effet, nous avons mis en évidence l'existence d'une hélice α supplémentaire en N-terminal du domaine MTase de SUDV, non retrouvée chez d'autres *Mononegavirales* tels que hMPV, VSV et du virus parainfluenza 5 (PIV5). De façon intéressante, cette hélice α supplémentaire est retrouvée

chez les flavivirus (ZIKV et DENV) et la 2'O MTase cellulaire CMTr1. La caractérisation biochimique des domaines MTases de ZIKV et DENV a par ailleurs également montré que ces virus, tout comme SUDV présentés, en plus de l'activité MTase canonique, une activité MTase non conventionnelle et méthylent en 2'O des adénosines internes à la séquence d'ARN. Ces observations laissent supposer que cette hélice α additionnelle pourrait participer au positionnement de l'ARN dans le site catalytique de sorte à assurer une activité de méthylation des résidus en interne de la séquence nucléotidique. L'étude de mutagenèse dirigée supporte cette hypothèse, dans la mesure où la mutation d'un résidu présent sur cette hélice permet de découpler l'activité coiffe dépendante de l'activité coiffe indépendante (interne). Bien qu'une étude récente ait mis en évidence le rôle de ces modifications épitranscriptomiques dans les subversions des défenses de l'hôte, notamment en limitant la détection des ARNs viraux par MDA5 (300), le rôle de ces modifications épitranscriptomiques dans le cycle viral reste peu documenté dans un contexte d'infection à virus Ebola.

La superposition des domaines MTase de SUDV et de hMPV nous a par ailleurs permis de positionner les molécules de GTP, de SAM et d'ATP dans la structure MTase de SUDV. Nous avons ainsi pu identifier une poche accommodant le substrat (^{SUB}P) et une poche accommodant le donneur de méthyle (^{SAM}P) conservées structurellement et fonctionnellement. De plus, hMPV présente une cavité hydrophobe pouvant accommoder l'ATP ou de SAM (^{NS}P), remplacée par une boucle flexible fixée dans une position dite « fermée » chez SUDV. De plus, l'étude comparative des structures nous a permis d'identifier des boucles flexibles au sein des différentes poches ^{SAMP} et ^{SUB}P pouvant participer à la régulation de l'activité MTase. Par une étude de mutagenèse dirigée, nous avons mis en évidence des résidus impliqués dans la régulation des différentes activités MTase de SUDV. L'ensemble des résultats nous a ainsi permis d'établir un modèle illustrant la régulation des différentes activités MTase de SUDV liée au positionnement de l'ARN dans le site catalytique. En effet nous avons mis en évidence l'existence de deux régions potentiellement impliquées dans le positionnement de l'ARNs pour les différentes activités MTase de la protéine L du virus Soudan. La première région, proches du site actif, comprend les résidus de la ^{SAMP}, la ^{NSP} et la ^{SUB}P via une de ses boucles flexibles, alors que la seconde région inclut l'hélice additionnelle observée chez SUDV et la seconde boucle flexible participant à la ^{SUB}P. Dans ce scénario, la méthylation de la coiffe en position N7 se produirait lorsque l'ARN est recruté par la première région, tandis que les méthylations 2'O de la coiffe et internes seraient catalysées

lorsque l'ARN est positionné dans la seconde région, projetant le 2'OH du ribose dans le site catalytique. Cette compréhension du mode de fixation alternatif du substrat dans la poche catalytique ouvre la voie vers le développement de molécules ciblant spécifiquement les différentes activités MTase.

Une question sous-jacente importante qui découle de la relation structurale/fonction établie pour l'activité MTase est : comment cette activité est interconnectée avec les autres activités catalytiques de la protéine L ? La résolution récente de la structure de la protéine L de PIV5 a permis d'établir un modèle basé sur positionnement du domaine MTase+CTD autour de l'extrémité N-terminale de la protéine L favorisant les évènements de transcription ou de réplication (271) (**Fig. 36**).

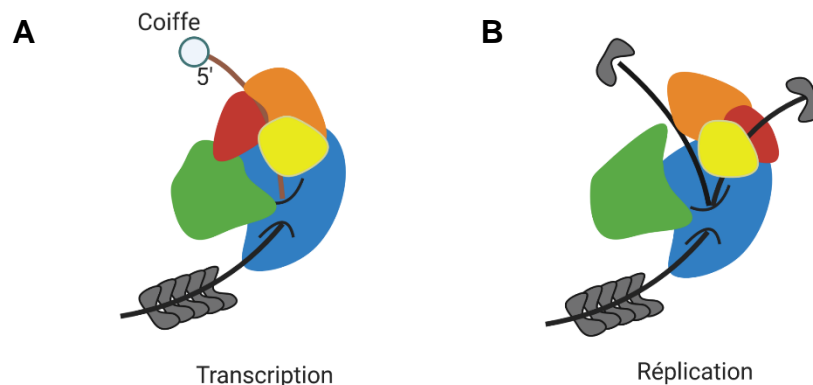


Figure 36 : Modèle des remaniements structuraux de la protéine L

Représentation schématique de la protéine L des *Mononegavirales* avec en bleu le domaine RdRp, en vert le domaine PRNTase, en jaune le domaine CD, en orange le domaine MTase et enfin en rouge le domaine CTD. L'ARN génomique ou antigénomique en noir encapsidé par la nucléoprotéine (en gris). L'ARNm néosynthétisé (en marron) porte une structure coiffe (bleu) à son extrémité 5'. **(A)** Au cours de la transcription le domaine MTase+CTD avoisine le domaine PRNTase. L'ARNm néosynthétisé transite à travers le sillon formé par les domaines MTase et CTD permettant sa méthylation. **(B)** Dans la conformation associée à la réplication, le domaine MTase+CTD est éloigné du domaine PRNTase et l'ARN ne transite pas par le domaine par le domaine MTase. Cette conformation suggère une absence de transfert de méthyl sur les ARNs génomiques et antigénomique.

Inspirée de *Abdella et al. 2020*

Dans le modèle proposé par *Abdella et al.*, la transcription est associée à un positionnement du domaine MTase (en orange) et CTD (en rouge) proche du domaine PRNTase (en vert). Dans cette conformation, l'ARNm néosynthétisé transite dans le sillon formé par les domaines MTase et CTD, compatible à l'activité MTase. Par opposition, au cours de la réplication le domaine MTase+CTD est éloigné du domaine PRNTase et l'ARN ne transite pas par le domaine MTase. Ici, ces observations suggèrent qu'au cours de la transcription, le

positionnement des domaines C-terminaux de la protéine L permet la méthylation efficace de l'ARNm alors que la réplication ne le permet pas et conduit à l'encapsidation de l'ARN génomique et antigénomique, supportant l'hypothèse d'une absence de méthylation de ces ARNs.

Ce modèle met en évidence l'importance des remaniements structuraux de la protéine L et la flexibilité des domaines C-terminaux dans le switch entre réplication et transcription, et suggère que le niveau de méthylation des ARN viraux en mode réplication ou transcription pourrait être très différent. Néanmoins les mécanismes permettant de favoriser une conformation au dépend de l'autre restent encore mal connus. L'implication de partenaires viraux pourrait participer à la régulation de ces événements.

La caractérisation biochimique de l'activité MTase d'un filovirus, à révéler une activité de méthylation interne chez les filovirus. Cependant, confirmer l'existence de ces méthylations sur des ARNs viraux issues de cellules infectées reste un enjeu important. Des essais Nm-seq, basés sur des ARNs issus de cellules infectées par le virus Ebola, permettraient d'une part de confirmer les méthylations des ARNs viraux et d'autre part identifier quels ARNs sont ciblés (les messagers ou génomiques) (365). De plus, l'identification de résidus au sein des domaines MTase et CTD de SUDV permettant de découpler les différentes activités MTase et notamment les activités MTase associées à la coiffe et l'activité MTase interne, permettrait d'évaluer l'effet de ces mutations *in vivo* sur des cellules par le virus Ebola. Cette étude est actuellement en cours en collaboration avec le laboratoire P4 de Lyon. Il serait également intéressant d'étudier l'impact de ces mutations sur la détection du virus par les senseurs de l'immunité innée afin de confirmer le rôle de ces modifications épitranscriptomiques dans les subversions des défenses de l'hôte dans un contexte d'infection à virus Ebola.

Bien que le rôle des méthylations portées par la coiffe soit bien documenté, l'implication des méthylations internes dans le cycle répliatif des virus reste à démontrer et plusieurs hypothèses sont envisageables (**Fig. 37**).

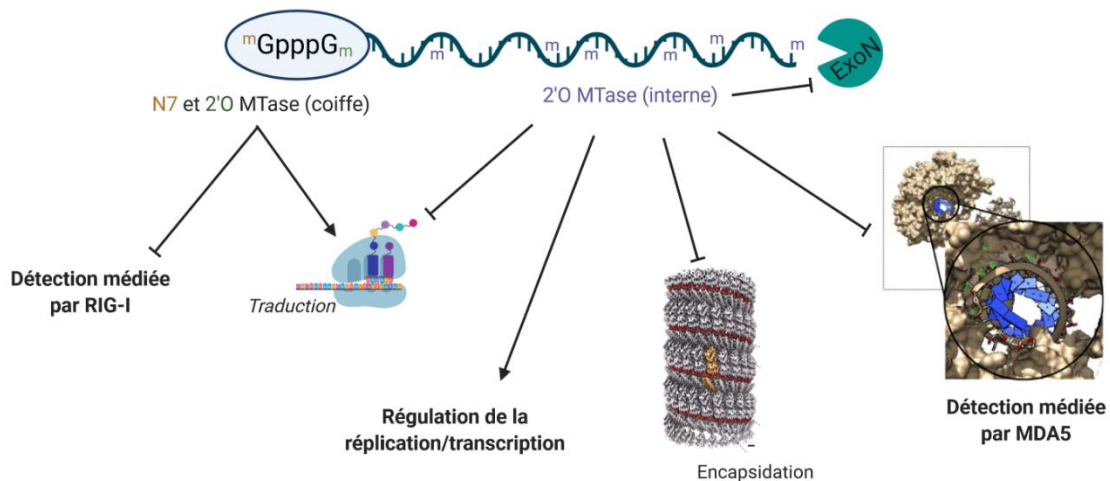


Figure 37 : Rôles potentiels des méthylations de l'ARN chez les filovirus

La structure coiffe (GpppG) en 5' de l'ARN présente une méthylation en N7 (en marron) permettant le recrutement de facteur d'initiation de la traduction et conduisant à une synthèse efficace des protéines virales. La 2'O méthylation du ribose du premier nucléotide (en vert) limite la détection de l'ARN par le senseur de l'immunité innée (RIG-I). Les 2'O méthylations internes (en violet) peuvent avoir plusieurs effets, en régulant la réplication et transcription, en limitant le phénomène d'encapsidation, en inhibant la détection par le senseur de l'immunité innée MDA5 ou encore en limitant la dégradation par des 3'-5' exonucléases.

Comme précédemment évoqué, les méthylations internes pourraient participer à l'échappement de la réponse immunitaire médiée par MDA5 (300), ainsi une inhibition de l'activité MTase interne permettrait de booster la reconnaissance par MDA5 et entraîner l'activation de la réponse immunitaire innée. Ces modifications épitranscriptomiques peuvent également avoir un impact sur la régulation cycle de réplication/transcription. Il a été montré chez les filovirus que ces méthylations diminuent l'efficacité de la traduction des ARNm viraux et atténuent la réplication de l'ARN (298) et il semble également que la présence de méthyle sur les ARN soit un obstacle à la réplication du génome de *Mononegavirales* (366). D'autre part, les 2'O méthylations internes pourraient interférer avec des effecteurs activés des suites de la production d'IFN. Des études préliminaires conduites au laboratoire (non présentées), mettent en évidence le rôle de telles modifications dans la limitation de la dégradation des ARNs par des exonucléases (ExoN) induites par l'interféron. Ces observations suggèrent qu'un excès de méthylation permettrait de limiter la dégradation des ARNs mais ne conduirait pas à une traduction efficace de ces derniers altérant ainsi la multiplication virale. Enfin, ces méthylations internes pourraient jouer un rôle dans la régulation de l'encapsidation. Comme présenté en introduction, l'ARN est intimement lié à la NP pour former la nucléocapside. Il est donc envisageable que les 2'O méthylations portées par les ARNs des filovirus altèrent l'interaction ARN/NP. Cette hypothèse laisse penser que

ces méthylations pourraient permettre à la NP de discriminer les ARNm qui sont non encapsidés, des ARNs génomiques qui sont associés à la NP au cours de leur synthèse. De plus, les remaniements structuraux de la protéine L, la flexibilité des domaines C-terminaux autour des domaines N-terminaux et le modèle qui en découle (**Fig. 36**), confortent cette hypothèse, qui reste cependant à démontrer.

En conclusion, de toutes évidences ces méthylations internes peuvent avoir des effets pro- et antiviraux. Les processus de méthylations sont donc certainement soumis à des contrôles stricts dans les cellules infectées, non encore identifiés à ce jour. Il serait donc nécessaire d'assurer la spécificité de certaines catégories d'ARN (ARNm versus ARN génomique), de contrôler où ces ARNs sont méthyles (coiffe, interne, poly-A) et quand ils sont méthylés. La balance entre un excès et une absence de méthylation semble être la clef de voûte dans la propagation de l'infection à virus Ebola.

RÉFÉRENCES BIBLIOGRAPHIQUES

1. LWOFF A. Lysogeny. *Bacteriol Rev* [Internet]. 1953 Dec [cited 2019 Nov 25];17(4):269–337. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/13105613>
2. Lwoff A. The Concept of Virus. *Microbiology* [Internet]. 1957 Oct 1 [cited 2019 Nov 25];17(2):239–53. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/13481308>
3. Scola BL, Audic S, Robert C, Jungang L, de Lamballerie X, Drancourt M, et al. A Giant Virus in Amoebae. *Science* (80-) [Internet]. 2003 Mar 28 [cited 2019 Nov 25];299(5615):2033–2033. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12663918>
4. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, et al. The 1.2-megabase genome sequence of Mimivirus. *Science* [Internet]. 2004 Nov 19 [cited 2019 Nov 25];306(5700):1344–50. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15486256>
5. Raoult D, Forterre P. Redefining viruses: lessons from Mimivirus. *Nat Rev Microbiol* [Internet]. 2008 Apr 3 [cited 2019 Nov 27];6(4):315–9. Available from: <http://www.nature.com/articles/nrmicro1858>
6. Lwoff A, Horne R, Tournier P. A System of Viruses. *Cold Spring Harb Symp Quant Biol* [Internet]. 1962 Jan 1 [cited 2019 Nov 25];27(0):51–5. Available from: <http://symposium.cshlp.org/cgi/doi/10.1101/SQB.1962.027.001.008>
7. Baltimore D. Evolution of RNA viruses. *Ann N Y Acad Sci* [Internet]. 1980 Nov [cited 2020 Apr 10];354(1 Genetic Varia):492–7. Available from: <http://doi.wiley.com/10.1111/j.1749-6632.1980.tb27988.x>
8. Baltimore D. Expression of animal virus genomes. *Bacteriol Rev* [Internet]. 1971 Sep [cited 2020 Apr 10];35(3):235–41. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/4329869>
9. The order Mononegavirales. *Arch Virol* [Internet]. 1991 [cited 2020 Mar 25];117(1–2):137–40. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/2006902>
10. Afonso CL, Amarasinghe GK, Bányai K, Bào Y, Basler CF, Bavari S, et al. Taxonomy of the order Mononegavirales: update 2016. *Arch Virol* [Internet]. 2016 Aug [cited 2018 Jul 18];161(8):2351–60. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27216929>

11. Amarasinghe GK, Ayllón MA, Bào Y, Basler CF, Bavari S, Blasdel KR, et al. Taxonomy of the order Mononegavirales: update 2019. *Arch Virol* [Internet]. 2019 Jul 14 [cited 2020 Mar 25];164(7):1967–80. Available from: <http://link.springer.com/10.1007/s00705-019-04247-4>
12. Kuhn JH, Adachi T, Adhikari NKJ, Arribas JR, Bah IE, Bausch DG, et al. New filovirus disease classification and nomenclature. *Nat Rev Microbiol* [Internet]. 2019 May 29 [cited 2020 Apr 10];17(5):261–3. Available from: <http://www.nature.com/articles/s41579-019-0187-4>
13. Kuhn JH, Amarasinghe GK, Basler CF, Bavari S, Bukreyev A, Chandran K, et al. ICTV Virus Taxonomy Profile: Filoviridae. *J Gen Virol* [Internet]. 2019 Jun 1 [cited 2020 Apr 10];100(6):911–2. Available from: <https://www.microbiologyresearch.org/content/journal/jgv/10.1099/jgv.0.001252>
14. Vet VG-AM, 1879 undefined. Études sur la rage.
15. The History of Rabies. 2007 Jan 1;1–22.
16. PASTEUR AND HYDROPHOBIA. *Science* (80-) [Internet]. 1885 Nov 27 [cited 2019 Dec 2];ns-6(147):467–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17830016>
17. Pasteur L. Méthode pour prévenir la rage après morsure. 1885;
18. Hampson K, Coudeville L, Lembo T, Sambo M, Kieffer A, Attlan M, et al. Estimating the Global Burden of Endemic Canine Rabies on behalf of the Global Alliance for Rabies Control Partners for Rabies Prevention. [cited 2017 Sep 1]; Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4400070/pdf/pntd.0003709.pdf>
19. Meslin F-X, Briggs DJ. Eliminating canine rabies, the principal source of human infection: what will it take? *Antiviral Res* [Internet]. 2013 May [cited 2020 Mar 25];98(2):291–6. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S016635421300065X>
20. BAER GM, ABELSETH MK, DEBBIE JG. ORAL VACCINATION OF FOXES AGAINST RABIES. *Am J Epidemiol* [Internet]. 1971 Jun [cited 2020 Apr 10];93(6):487–90. Available from: <https://academic.oup.com/aje/article/147138/ORAL>
21. Freuling CM, Hampson K, Selhorst T, Schröder R, Meslin FX, Mettenleiter TC, et al.

- The elimination of fox rabies from Europe: determinants of success and lessons for the future. *Philos Trans R Soc Lond B Biol Sci* [Internet]. 2013 Aug 5 [cited 2020 Apr 10];368(1623):20120142. Available from: <https://royalsocietypublishing.org/doi/10.1098/rstb.2012.0142>
22. Stahl J-P, Gautret P, Ribadeau-Dumas F, Strady C, Le Moal G, Souala F, et al. Update on human rabies in a dog- and fox-rabies-free country. *Médecine Mal Infect* [Internet]. 2014 Jul 1 [cited 2020 Apr 10];44(7):292–301. Available from: <https://www.sciencedirect.com/science/article/pii/S0399077X14001498?via%3Dihub#sec0020>
 23. Rupprecht CE, Barrett J, Briggs D, Cliquet F, Fooks AR, Lumlerdacha B, et al. Can rabies be eradicated? *Dev Biol (Basel)* [Internet]. 2008 [cited 2020 Apr 10];131:95–121. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18634470>
 24. Arambulo P. International programs and veterinary public health in the Americas-- success, challenges, and possibilities. *Prev Vet Med* [Internet]. 2008 Sep 15 [cited 2020 Apr 10];86(3–4):208–15. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0167587708000536>
 25. Ribadeau-Dumas F, Cliquet F, Gautret P, Robardet E, Le Pen C, Bourhy H. Travel-Associated Rabies in Pets and Residual Rabies Risk, Western Europe. *Emerg Infect Dis* [Internet]. 2016 Jul [cited 2020 Apr 10];22(7):1268–71. Available from: http://wwwnc.cdc.gov/eid/article/22/7/15-1733_article.htm
 26. Abela-Ridder B, Knopf L, Martin S, Taylor L, Torres G, De Balogh K. 2016: the beginning of the end of rabies? *Lancet Glob Heal* [Internet]. 2016 Nov [cited 2020 Apr 10];4(11):e780–1. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2214109X16302455>
 27. Ugolini G. Rabies virus as a transneuronal tracer of neuronal connections. *Adv Virus Res* [Internet]. 2011 [cited 2020 Apr 10];79:165–202. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B978012387040700010X>
 28. Davis BM, Rall GF, Schnell MJ. Everything You Always Wanted to Know About Rabies Virus (But Were Afraid to Ask). *Annu Rev Virol* [Internet]. 2015 Nov 9 [cited 2020 Apr 10];2(1):451–71. Available from: <http://www.annualreviews.org/doi/10.1146/annurev-virology-100114-055157>

29. Hemachudha T, Phuapradit P. Rabies. *Curr Opin Neurol* [Internet]. 1997 Jun [cited 2020 Apr 10];10(3):260–7. Available from: <http://journals.lww.com/00019052-199706000-00016>
30. Hemachudha T. Human rabies: clinical aspects, pathogenesis, and potential therapy. *Curr Top Microbiol Immunol* [Internet]. 1994 [cited 2020 Apr 10];187:121–43. Available from: http://link.springer.com/10.1007/978-3-642-78490-3_7
31. Hemachudha T, Tirawatnpong S, Phanthumchinda K. Seizures as the initial manifestation of paralytic rabies. *J Neurol Neurosurg Psychiatry* [Internet]. 1989 Jun 1 [cited 2020 Apr 10];52(6):808–10. Available from: <http://jnnp.bmj.com/cgi/doi/10.1136/jnnp.52.6.808>
32. Hemachudha T, Laothamatas J, Rupprecht CE. Human rabies: a disease of complex neuropathogenetic mechanisms and diagnostic challenges. *Lancet Neurol* [Internet]. 2002 Jun [cited 2020 Apr 10];1(2):101–9. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1474442202000418>
33. Guo Y, Duan M, Wang X, Gao J, Guan Z, Zhang M. Early events in rabies virus infection—Attachment, entry, and intracellular trafficking. *Virus Res* [Internet]. 2019 Apr [cited 2020 Apr 22];263:217–25. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0168170218305380>
34. Gaudin Y, Ruigrok RWH, Knossow M, Flamand A. Low-pH Conformational Changes of Rabies Virus Glycoprotein and Their Role in Membrane Fusion Downloaded from [Internet]. Vol. 67, *JOURNAL OF VIROLOGY*. 1993 [cited 2020 Mar 10]. Available from: <http://jvi.asm.org/>
35. Mire C, White J, pathogens MW-Pl, 2010 undefined. A spatio-temporal analysis of matrix protein and nucleocapsid trafficking during vesicular stomatitis virus uncoating. *ncbi.nlm.nih.gov* [Internet]. [cited 2020 Mar 10]; Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2904772/>
36. Functional characterization of Negri bodies in rabies... - Google Scholar [Internet]. [cited 2020 Mar 10]. Available from: [https://scholar.google.com/scholar?q=Functional characterization of Negri bodies in rabies virus-infected cells: evidence that NBs are sites of viral transcription and replication](https://scholar.google.com/scholar?q=Functional+characterization+of+Negri+bodies+in+rabies+virus-infected+cells:evidence+that+NBs+are+sites+of+viral+transcription+and+replication)
37. Gaudin Y. Folding of Rabies Virus Glycoprotein: Epitope Acquisition and Interaction

- with Endoplasmic Reticulum Chaperones [Internet]. Vol. 71, JOURNAL OF VIROLOGY. 1997 [cited 2020 Mar 11]. Available from: <http://jvi.asm.org/>
38. Albertini A, Ruigrok R, research DB-A in virus, 2011 undefined. Rabies virus transcription and replication. Elsevier [Internet]. [cited 2020 Mar 10]; Available from: <https://www.sciencedirect.com/science/article/pii/B9780123870407000019>
 39. WHO Guide for Rabies Pre and Post Exposure Prophylaxis in Humans [Internet]. 2014 [cited 2020 Mar 25]. Available from: https://www.who.int/rabies/PEP_Prophylaxis_guideline_15_12_2014.pdf
 40. DEAN DJ, BAER GM, THOMPSON WR. Studies on the local treatment of rabies-infected wounds. Bull World Health Organ [Internet]. 1963 [cited 2020 Mar 25];28(4):477–86. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14026136>
 41. Quiambao BP, Dy-Tioco HZ, Dizon RM, Crisostomo ME, Teuwen DE. Rabies post-exposure prophylaxis with purified equine rabies immunoglobulin: one-year follow-up of patients with laboratory-confirmed category III rabies exposure in the Philippines. Vaccine [Internet]. 2009 Nov 27 [cited 2020 Mar 25];27(51):7162–6. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0264410X09013607>
 42. Jackson AC, Warrell MJ, Rupprecht CE, Ertl HCJ, Dietzschold B, O'Reilly M, et al. Management of rabies in humans. Clin Infect Dis [Internet]. 2003 Jan 1 [cited 2020 Mar 25];36(1):60–3. Available from: <https://academic.oup.com/cid/article-lookup/doi/10.1086/344905>
 43. Strady C, Hung Nguyen V, Jaussaud R, Lang J, Lienard M, Strady A. Pre-exposure rabies vaccination: strategies and cost-minimization study. Vaccine [Internet]. 2001 Dec [cited 2020 Mar 25];19(11–12):1416–24. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0264410X00003686>
 44. Martini GA. Marburg Virus Disease. Clinical Syndrome. In: Marburg Virus Disease [Internet]. Berlin, Heidelberg: Springer Berlin Heidelberg; 1971 [cited 2019 Jun 24]. p. 1–9. Available from: http://link.springer.com/10.1007/978-3-662-01593-3_1
 45. Gear JS, Cassel GA, Gear AJ, Trappler B, Clausen L, Meyers AM, et al. Outbreak of Marburg virus disease in Johannesburg. Br Med J [Internet]. 1975 Nov 29 [cited 2019 Sep 16];4(5995):489–93. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/811315>

46. Commission R of an I. Ebola haemorrhagic fever in Zaire, 1976. Bull World Health Organ [Internet]. 1978 [cited 2019 Sep 16];56(2):271–93. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/307456>
47. Breman JG, Heymann DL, Lloyd G, McCormick JB, Miatudila M, Murphy FA, et al. Discovery and Description of Ebola Zaire Virus in 1976 and Relevance to the West African Epidemic During 2013-2016. J Infect Dis [Internet]. 2016 [cited 2019 Sep 16];214(suppl 3):S93–101. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27357339>
48. Emond RT, Evans B, Bowen ET, Lloyd G. A case of Ebola virus infection. Br Med J [Internet]. 1977 Aug 27 [cited 2019 Sep 16];2(6086):541–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/890413>
49. Bowen ET, Lloyd G, Harris WJ, Platt GS, Baskerville A, Vella EE. Viral haemorrhagic fever in southern Sudan and northern Zaire. Preliminary studies on the aetiological agent. Lancet (London, England) [Internet]. 1977 Mar 12 [cited 2019 Sep 16];1(8011):571–3. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/65662>
50. Jahrling PB, Geisbert TW, Dalgard DW, Johnson ED, Ksiazek TG, Hall WC, et al. Preliminary report: isolation of Ebola virus from monkeys imported to USA. Lancet (London, England) [Internet]. 1990 Mar 3 [cited 2019 Sep 17];335(8688):502–5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1968529>
51. Centers for Disease Control (CDC). Ebola virus infection in imported primates--Virginia, 1989. MMWR Morb Mortal Wkly Rep [Internet]. 1989 Dec 8 [cited 2019 Sep 17];38(48):831–2, 837–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/2511410>
52. Miranda MEG, Miranda NLJ. Reston ebolavirus in humans and animals in the Philippines: a review. J Infect Dis [Internet]. 2011 Nov [cited 2019 Sep 17];204 Suppl 3(suppl_3):S757-60. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jir296>
53. Barrette RW, Metwally SA, Rowland JM, Xu L, Zaki SR, Nichol ST, et al. Discovery of Swine as a Host for the Reston ebolavirus. Science (80-) [Internet]. 2009 Jul 10 [cited 2019 Sep 17];325(5937):204–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19590002>

54. Le Guenno B, Formenty P, Formentry P, Wyers M, Gounon P, Walker F, et al. Isolation and partial characterisation of a new strain of Ebola virus. *Lancet* (London, England) [Internet]. 1995 May 20 [cited 2019 Sep 17];345(8960):1271–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7746057>
55. Towner JS, Sealy TK, Khristova ML, Albariño CG, Conlan S, Reeder SA, et al. Newly discovered ebola virus associated with hemorrhagic fever outbreak in Uganda. *PLoS Pathog* [Internet]. 2008 Nov [cited 2019 Sep 17];4(11):e1000212. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19023410>
56. Quetglas J. Estudian la extraña mortandad de miles de murciélagos de cueva (*Miniopterus schreibersii*) / Odd mass mortality in *Miniopterus schreibersii* under study. *Quercus* 203:50 [Internet]. [cited 2019 Sep 17]; Available from: https://www.academia.edu/4794993/Estudian_la_extraña_mortandad_de_miles_de_murciélagos_de_cueva_Miniopterus_schreibersii_Odd_mass_mortality_in_Miniopterus_schreibersii_under_study
57. Negredo A, Palacios G, Vázquez-Morón S, González F, Dopazo H, Molero F, et al. Discovery of an Ebolavirus-Like Filovirus in Europe. Basler CF, editor. *PLoS Pathog* [Internet]. 2011 Oct 20 [cited 2019 Sep 17];7(10):e1002304. Available from: <https://dx.plos.org/10.1371/journal.ppat.1002304>
58. Goldstein T, Anthony SJ, Gbakima A, Bird BH, Bangura J, Tremeau-Bravard A, et al. The discovery of Bombali virus adds further support for bats as hosts of ebolaviruses. *Nat Microbiol* [Internet]. 2018 Oct 27 [cited 2020 Apr 10];3(10):1084–9. Available from: <http://www.nature.com/articles/s41564-018-0227-2>
59. Yang X-L, Tan CW, Anderson DE, Jiang R-D, Li B, Zhang W, et al. Characterization of a filovirus (*Měnglà virus*) from *Rousettus* bats in China. *Nat Microbiol* [Internet]. 2019 Mar 7 [cited 2019 Sep 17];4(3):390–5. Available from: <http://www.nature.com/articles/s41564-018-0328-y>
60. Shi M, Lin X-D, Chen X, Tian J-H, Chen L-J, Li K, et al. The evolutionary history of vertebrate RNA viruses. *Nature* [Internet]. 2018 Apr 4 [cited 2020 Apr 22];556(7700):197–202. Available from: <http://www.nature.com/articles/s41586-018-0012-7>
61. WHO | Ebola outbreak 2014-2016. WHO [Internet]. 2019 [cited 2019 Sep 17];

- Available from: <https://www.who.int/csr/disease/ebola/en/>
62. Marí Saéz A, Weiss S, Nowak K, Lapeyre V, Zimmermann F, Düx A, et al. Investigating the zoonotic origin of the West African Ebola epidemic. *EMBO Mol Med* [Internet]. 2015 Jan 30 [cited 2019 Sep 17];7(1):17–23. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25550396>
 63. Kaner J, Schaack S. Understanding Ebola: the 2014 epidemic. *Global Health* [Internet]. 2016 Dec 13 [cited 2019 Sep 17];12(1):53. Available from: <http://globalizationandhealth.biomedcentral.com/articles/10.1186/s12992-016-0194-4>
 64. WHO | Statement on the 1st meeting of the IHR Emergency Committee on the 2014 Ebola outbreak in West Africa. WHO [Internet]. 2014 [cited 2019 Sep 17]; Available from: <https://www.who.int/mediacentre/news/statements/2014/ebola-20140808/en/>
 65. Inungu J, Iheduru-Anderson K, Odio OJ. Recurrent Ebolavirus disease in the Democratic Republic of Congo: update and challenges. *AIMS public Heal* [Internet]. 2019 [cited 2020 Apr 14];6(4):502–13. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31909070>
 66. Ebola | Ebola situation reports: Democratic Republic of the Congo. WHO [Internet]. 2019 [cited 2019 May 23]; Available from: <https://www.who.int/ebola/situation-reports/drc-2018/en/>
 67. 2018 Eastern Democratic Republic of the Congo Outbreak (ongoing) | Democratic Republic of Congo | Outbreaks | Ebola (Ebola Virus Disease) | CDC [Internet]. [cited 2019 Sep 17]. Available from: <https://www.cdc.gov/vhf/ebola/outbreaks/drc/2018-august.html>
 68. Ebola outbreak in the Democratic Republic of the Congo declared a Public Health Emergency of International Concern [Internet]. [cited 2019 Aug 6]. Available from: <https://www.who.int/news-room/detail/17-07-2019-ebola-outbreak-in-the-democratic-republic-of-the-congo-declared-a-public-health-emergency-of-international-concern>
 69. Lawrence P, Danet N, Reynard O, Volchkova V, Volchkov V. Human transmission of Ebola virus. *Curr Opin Virol* [Internet]. 2017 Feb [cited 2020 May 7];22:51–8. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1879625716301973>
 70. Ksiazek TG, West CP, Rollin PE, Jahrling PB, Peters CJ. ELISA for the Detection of

- Antibodies to Ebola Viruses. *J Infect Dis* [Internet]. 1999 Feb [cited 2019 Oct 1];179(s1):S192–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9988184>
71. Rodriguez LL, De Roo A, Guimard Y, Trappier SG, Sanchez A, Bressler D, et al. Persistence and genetic stability of Ebola virus during the outbreak in Kikwit, Democratic Republic of the Congo, 1995. *J Infect Dis* [Internet]. 1999 Feb [cited 2019 Oct 1];179 Suppl 1(s1):S170-6. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1086/514291>
72. Zaki SR, Shieh W, Greer PW, Goldsmith CS, Ferebee T, Katshitshi J, et al. A Novel Immunohistochemical Assay for the Detection of Ebola Virus in Skin: Implications for Diagnosis, Spread, and Surveillance of Ebola Hemorrhagic Fever. *J Infect Dis* [Internet]. 1999 Feb 1 [cited 2019 Sep 30];179(s1):S36–47. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1086/514319>
73. Dowell SF, Mukunu R, Ksiazek TG, Khan AS, Rollin PE, Peters CJ. Transmission of Ebola hemorrhagic fever: a study of risk factors in family members, Kikwit, Democratic Republic of the Congo, 1995. Commission de Lutte contre les Epidémies à Kikwit. *J Infect Dis* [Internet]. 1999 Feb [cited 2019 Oct 1];179 Suppl 1(s1):S87-91. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1086/514284>
74. Khan AS, Tshioko FK, Heymann DL, Le Guenno B, Nabeth P, Kerstiëns B, et al. The Reemergence of Ebola Hemorrhagic Fever, Democratic Republic of the Congo, 1995. *J Infect Dis* [Internet]. 1999 Feb [cited 2019 Oct 1];179(s1):S76–86. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9988168>
75. Groseth A, Feldmann H, Strong JE. The ecology of Ebola virus. *Trends Microbiol* [Internet]. 2007 Sep [cited 2019 Sep 26];15(9):408–16. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17698361>
76. Morvan JM, Deubel V, Gounon P, Nakouné E, Barrière P, Murri S, et al. Identification of Ebola virus sequences present as RNA or DNA in organs of terrestrial small mammals of the Central African Republic. *Microbes Infect* [Internet]. 1999 Dec [cited 2019 Sep 26];1(14):1193–201. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10580275>
77. Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, Yaba P, et al. Fruit bats as reservoirs of Ebola virus. *Nature* [Internet]. 2005 Dec 30 [cited 2019 Sep

- 26];438(7068):575–6. Available from: <http://www.nature.com/articles/438575a>
78. Towner JS, Amman BR, Sealy TK, Carroll SAR, Comer JA, Kemp A, et al. Isolation of Genetically Diverse Marburg Viruses from Egyptian Fruit Bats. Fouchier RAM, editor. *PLoS Pathog* [Internet]. 2009 Jul 31 [cited 2019 Sep 26];5(7):e1000536. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19649327>
79. Leligdowicz A, Fischer WA, Uyeki TM, Fletcher TE, Adhikari NKJ, Portella G, et al. Ebola virus disease and critical illness. *Crit Care* [Internet]. 2016 Dec 29 [cited 2019 Sep 27];20(1):217. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27468829>
80. Kortepeter MG, Bausch DG, Bray M. Basic Clinical and Laboratory Features of Filoviral Hemorrhagic Fever. *J Infect Dis* [Internet]. 2011 Nov [cited 2019 Oct 1];204(suppl_3):S810–6. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jir299>
81. Feldmann H, Geisbert TW. Ebola haemorrhagic fever. *Lancet* [Internet]. 2011 Mar 5 [cited 2019 Sep 26];377(9768):849–62. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21084112>
82. Keita M, Diallo B, Mesfin S, Marega A, Nebie KY, Magassouba N, et al. Subsequent mortality in survivors of Ebola virus disease in Guinea: a nationwide retrospective cohort study. *Lancet Infect Dis* [Internet]. 2019 Sep 4 [cited 2019 Sep 27]; Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1473309919303135>
83. Baskerville A, Fisher-Hoch SP, Neild GH, Dowsett AB. Ultrastructural pathology of experimental ebola haemorrhagic fever virus infection. *J Pathol* [Internet]. 1985 Nov 1 [cited 2019 Sep 30];147(3):199–209. Available from: <http://doi.wiley.com/10.1002/path.1711470308>
84. Baseler L, Chertow DS, Johnson KM, Feldmann H, Morens DM. The Pathogenesis of Ebola Virus Disease. *Annu Rev Pathol Mech Dis* [Internet]. 2017 Jan 24 [cited 2019 Dec 10];12(1):387–418. Available from: <http://www.annualreviews.org/doi/10.1146/annurev-pathol-052016-100506>
85. Geisbert TW, Hensley LE, Larsen T, Young HA, Reed DS, Geisbert JB, et al. Pathogenesis of Ebola hemorrhagic fever in cynomolgus macaques: evidence that dendritic cells are early and sustained targets of infection. *Am J Pathol* [Internet]. 2003 Dec [cited 2019 Sep 30];163(6):2347–70. Available from:

- <http://www.ncbi.nlm.nih.gov/pubmed/14633608>
86. Jaax NK, Davis KJ, Geisbert TJ, Vogel P, Jaax GP, Topper M, et al. Lethal experimental infection of rhesus monkeys with Ebola-Zaire (Mayinga) virus by the oral and conjunctival route of exposure. *Arch Pathol Lab Med* [Internet]. 1996 Feb [cited 2019 Sep 30];120(2):140–55. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8712894>
 87. Zaki SR, Goldsmith CS. Pathologic Features of Filovirus Infections in Humans. In: *Current topics in microbiology and immunology* [Internet]. 1999 [cited 2019 Sep 30]. p. 97–116. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9893381>
 88. Ryabchikova EI, Kolesnikova LV, Luchko SV. An Analysis of Features of Pathogenesis in Two Animal Models of Ebola Virus Infection. *J Infect Dis* [Internet]. 1999 Feb 1 [cited 2019 Sep 30];179(s1):S199–202. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1086/514293>
 89. Schnittler H, Feldmann H. Marburg and Ebola Hemorrhagic Fevers: Does the Primary Course of Infection Depend on the Accessibility of Organ-Specific Macrophages? *Clin Infect Dis* [Internet]. 1998 Aug [cited 2019 Oct 1];27(2):404–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9709901>
 90. Levi M. Disseminated intravascular coagulation. *Crit Care Med* [Internet]. 2007 Sep [cited 2019 Oct 1];35(9):2191–5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17855836>
 91. Geisbert TW, Young HA, Jahrling PB, Davis KJ, Kagan E, Hensley LE. Mechanisms underlying coagulation abnormalities in ebola hemorrhagic fever: overexpression of tissue factor in primate monocytes/macrophages is a key event. *J Infect Dis* [Internet]. 2003 Dec 1 [cited 2019 Oct 1];188(11):1618–29. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1086/379724>
 92. Fisher-Hoch SP, Platt GS, Neild GH, Southee T, Baskerville A, Raymond RT, et al. Pathophysiology of Shock and Hemorrhage in a Fulminating Viral Infection (Ebola). *J Infect Dis* [Internet]. 1985 Nov 1 [cited 2019 Oct 1];152(5):887–94. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/152.5.887>
 93. Sugita Y, Matsunami H, Kawaoka Y, Noda T, Wolf M. Cryo-EM structure of the Ebola virus nucleoprotein-RNA complex at 3.6 Å resolution. *Nature* [Internet]. 2018

- Nov 17 [cited 2020 Apr 15];563(7729):137–40. Available from: <http://www.nature.com/articles/s41586-018-0630-0>
94. Noda T, Sagara H, Suzuki E, Takada A, Kida H, Kawaoka Y. Ebola Virus VP40 Drives the Formation of Virus-Like Filamentous Particles Along with GP. *J Virol* [Internet]. 2002 May 15 [cited 2020 Apr 15];76(10):4855–65. Available from: <https://jvi.asm.org/content/76/10/4855>
95. Volchkov VE, Feldmann H, Volchkova VA, Klenk H-D. Processing of the Ebola virus glycoprotein by the proprotein convertase furin. *Proc Natl Acad Sci* [Internet]. 1998 May 12 [cited 2020 Apr 15];95(10):5762–7. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.95.10.5762>
96. Aleksandrowicz P, Marzi A, Biedenkopf N, Beimforde N, Becker S, Hoenen T, et al. Ebola virus enters host cells by macropinocytosis and clathrin-mediated endocytosis. *J Infect Dis* [Internet]. 2011 Nov [cited 2020 Apr 15];204 Suppl 3(suppl_3):S957-67. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jir326>
97. Nanbo A, Imai M, Watanabe S, Noda T, Takahashi K, Neumann G, et al. Ebolavirus Is Internalized into Host Cells via Macropinocytosis in a Viral Glycoprotein-Dependent Manner. Rey FA, editor. *PLoS Pathog* [Internet]. 2010 Sep 23 [cited 2020 Apr 15];6(9):e1001121. Available from: <http://dx.plos.org/10.1371/journal.ppat.1001121>
98. Davey RA, Shtanko O, Anantpadma M, Sakurai Y, Chandran K, Maury W. Mechanisms of Filovirus Entry. *Curr Top Microbiol Immunol* [Internet]. 2017 [cited 2020 Apr 15];411:323–52. Available from: http://link.springer.com/10.1007/82_2017_14
99. Chandran K, Sullivan NJ, Felbor U, Whelan SP, Cunningham JM. Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. *Science* [Internet]. 2005 Jun 10 [cited 2020 Apr 15];308(5728):1643–5. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1110656>
100. Mühlberger E, Weik M, Volchkov VE, Klenk HD, Becker S. Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J Virol* [Internet]. 1999 Mar [cited 2020 Apr 15];73(3):2333–42. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9971816>
101. Hoenen T, Shabman RS, Groseth A, Herwig A, Weber M, Schudt G, et al. Inclusion

- bodies are a site of ebolavirus replication. *J Virol* [Internet]. 2012 Nov 1 [cited 2020 Apr 15];86(21):11779–88. Available from: <http://jvi.asm.org/cgi/doi/10.1128/JVI.01525-12>
102. Schudt G, Dolnik O, Kolesnikova L, Biedenkopf N, Herwig A, Becker S. Transport of Ebolavirus Nucleocapsids Is Dependent on Actin Polymerization: Live-Cell Imaging Analysis of Ebolavirus-Infected Cells. *J Infect Dis* [Internet]. 2015 Oct 1 [cited 2020 Apr 15];212 Suppl 2(suppl 2):S160-6. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jiv083>
103. Adu-Gyamfi E, Digman MA, Gratton E, Stahelin R V. Investigation of Ebola VP40 assembly and oligomerization in live cells using number and brightness analysis. *Biophys J* [Internet]. 2012 Jun 6 [cited 2020 Apr 15];102(11):2517–25. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0006349512004675>
104. Hoenen T, Biedenkopf N, Zielecki F, Jung S, Groseth A, Feldmann H, et al. Oligomerization of Ebola Virus VP40 Is Essential for Particle Morphogenesis and Regulation of Viral Transcription. *J Virol* [Internet]. 2010 Jul 15 [cited 2020 Apr 15];84(14):7053–63. Available from: <https://jvi.asm.org/content/84/14/7053>
105. Ro Y-T, Ticer A, Carrion R, Patterson JL. Rapid detection and quantification of Ebola Zaire virus by one-step real-time quantitative reverse transcription-polymerase chain reaction. *Microbiol Immunol* [Internet]. 2017 Apr [cited 2019 Oct 1];61(3–4):130–7. Available from: <http://doi.wiley.com/10.1111/1348-0421.12475>
106. Park S-W, Lee Y-J, Lee W-J, Jee Y, Choi W. One-Step Reverse Transcription-Polymerase Chain Reaction for Ebola and Marburg Viruses. *Osong Public Heal Res Perspect* [Internet]. 2016 Jun [cited 2019 Oct 1];7(3):205–9. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S2210909916300182>
107. Cherpillod P, Schibler M, Vieille G, Cordey S, Mamin A, Vetter P, et al. Ebola virus disease diagnosis by real-time RT-PCR: A comparative study of 11 different procedures. *J Clin Virol* [Internet]. 2016 Apr [cited 2019 Oct 1];77:9–14. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1386653216000408>
108. Rowe AK, Bertolli J, Khan AS, Mukunu R, Muyembe-Tamfum JJ, Bressler D, et al. Clinical, virologic, and immunologic follow-up of convalescent Ebola hemorrhagic fever patients and their household contacts, Kikwit, Democratic Republic of the Congo.

- Commission de Lutte contre les Epidémies à Kikwit. *J Infect Dis* [Internet]. 1999 Feb [cited 2019 Oct 1];179 Suppl 1(s1):S28-35. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1086/514318>
109. Ksiazek TG, Rollin PE, Williams AJ, Bressler DS, Martin ML, Swanepoel R, et al. Clinical virology of Ebola hemorrhagic fever (EHF): virus, virus antigen, and IgG and IgM antibody findings among EHF patients in Kikwit, Democratic Republic of the Congo, 1995. *J Infect Dis* [Internet]. 1999 Feb [cited 2019 Oct 1];179 Suppl 1(s1):S177-87. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1086/514321>
110. Towner JS, Rollin PE, Bausch DG, Sanchez A, Crary SM, Vincent M, et al. Rapid diagnosis of Ebola hemorrhagic fever by reverse transcription-PCR in an outbreak setting and assessment of patient viral load as a predictor of outcome. *J Virol* [Internet]. 2004 Apr [cited 2019 Oct 1];78(8):4330–41. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15047846>
111. Baize S, Leroy EM, Georges-Courbot MC, Capron M, Lansoud-Soukate J, Debré P, et al. Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients. *Nat Med* [Internet]. 1999 Apr [cited 2019 Oct 1];5(4):423–6. Available from: http://www.nature.com/articles/nm0499_423
112. Formenty P, Hatz C, Le Guenno B, Stoll A, Rogenmoser P, Widmer A. Human infection due to Ebola virus, subtype Côte d'Ivoire: clinical and biologic presentation. *J Infect Dis* [Internet]. 1999 Feb [cited 2019 Oct 1];179 Suppl 1(s1):S48-53. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1086/514285>
113. Wauquier N, Becquart P, Gasquet C, Leroy EM. Immunoglobulin G in Ebola Outbreak Survivors, Gabon. *Emerg Infect Dis* [Internet]. 2009 Jul [cited 2019 Oct 1];15(7):1136–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19624943>
114. Gear JS, Cassel GA, Gear AJ, Trappler B, Clausen L, Meyers AM, et al. Outbreak of Marburg virus disease in Johannesburg. *BMJ* [Internet]. 1975 Nov 29 [cited 2019 Oct 1];4(5995):489–93. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/811315>
115. Mulangu S, Dodd LE, Davey RT, Tshiani Mbaya O, Proschan M, Mukadi D, et al. A Randomized, Controlled Trial of Ebola Virus Disease Therapeutics. *N Engl J Med*

- [Internet]. 2019 Nov 27 [cited 2019 Dec 11];NEJMoa1910993. Available from: <http://www.nejm.org/doi/10.1056/NEJMoa1910993>
116. Garbutt M, Liebscher R, Wahl-Jensen V, Jones S, Möller P, Wagner R, et al. Properties of replication-competent vesicular stomatitis virus vectors expressing glycoproteins of filoviruses and arenaviruses. *J Virol* [Internet]. 2004 May 15 [cited 2019 Dec 12];78(10):5458–65. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10482652>
117. Mueller S, Coleman JR, Papamichail D, Ward CB, Nimmual A, Futcher B, et al. Live attenuated influenza virus vaccines by computer-aided rational design. *Nat Biotechnol*. 2010;28(7):723–6.
118. Second Ebola vaccine to complement “ring vaccination” given green light in DRC [Internet]. [cited 2020 Apr 10]. Available from: <https://www.who.int/news-room/detail/23-09-2019-second-ebola-vaccine-to-complement-ring-vaccination-given-green-light-in-drc>
119. Major milestone for WHO-supported Ebola vaccine [Internet]. [cited 2020 Apr 10]. Available from: <https://www.who.int/news-room/detail/18-10-2019-major-milestone-for-who-supported-ebola-vaccine>
120. Ollmann Saphire E. A Vaccine against Ebola Virus. *Cell* [Internet]. 2020 Apr 2 [cited 2020 Apr 10];181(1):6. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867420302737>
121. Wensman JJ, Munir M, Thaduri S, Hörnaeus K, Rizwan M, Blomström A-L, et al. The X proteins of bornaviruses interfere with type I interferon signalling. *J Gen Virol* [Internet]. 2013 Feb 1 [cited 2020 Apr 21];94(Pt 2):263–9. Available from: <https://www.microbiologyresearch.org/content/journal/jgv/10.1099/vir.0.047175-0>
122. Poenisch M, Wille S, Ackermann A, Staeheli P, Schneider U. The X protein of borna disease virus serves essential functions in the viral multiplication cycle. *J Virol* [Internet]. 2007 Jul 1 [cited 2020 Apr 21];81(13):7297–9. Available from: <https://jvi.asm.org/content/81/13/7297>
123. Timmins J, Scianimanico S, Schoehn G, Weissenhorn W. Vesicular release of ebola virus matrix protein VP40. *Virology* [Internet]. 2001 Apr 25 [cited 2020 Apr 21];283(1):1–6. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0042682201908608>

124. Biedenkopf N, Hartlieb B, Hoenen T, Becker S. Phosphorylation of Ebola virus VP30 influences the composition of the viral nucleocapsid complex: impact on viral transcription and replication. *J Biol Chem* [Internet]. 2013 Apr 19 [cited 2020 Apr 21];288(16):11165–74. Available from: <http://www.jbc.org/lookup/doi/10.1074/jbc.M113.461285>
125. Groseth A, Charton JE, Sauerborn M, Feldmann F, Jones SM, Hoenen T, et al. The Ebola virus ribonucleoprotein complex: A novel VP30–L interaction identified. *Virus Res* [Internet]. 2009 Mar [cited 2020 Apr 21];140(1–2):8–14. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0168170208003626>
126. Biedenkopf N, Lier C, Becker S. Dynamic Phosphorylation of VP30 Is Essential for Ebola Virus Life Cycle. Lyles DS, editor. *J Virol* [Internet]. 2016 May 15 [cited 2020 Apr 21];90(10):4914–25. Available from: <http://jvi.asm.org/lookup/doi/10.1128/JVI.03257-15>
127. Feldmann H, Klenk H-D, Sanchez A. Molecular biology and evolution of filoviruses. In 1993 [cited 2020 Apr 21]. p. 81–100. Available from: http://link.springer.com/10.1007/978-3-7091-9300-6_8
128. Bamberg S, Kolesnikova L, Möller P, Klenk H-D, Becker S. VP24 of Marburg virus influences formation of infectious particles. *J Virol* [Internet]. 2005 Nov 1 [cited 2020 Apr 21];79(21):13421–33. Available from: <https://jvi.asm.org/content/79/21/13421>
129. Han Z, Boshra H, Sunyer JO, Zwiers SH, Paragas J, Harty RN. Biochemical and functional characterization of the Ebola virus VP24 protein: implications for a role in virus assembly and budding. *J Virol* [Internet]. 2003 Feb [cited 2020 Apr 21];77(3):1793–800. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12525613>
130. Licata JM, Johnson RF, Han Z, Harty RN. Contribution of ebola virus glycoprotein, nucleoprotein, and VP24 to budding of VP40 virus-like particles. *J Virol* [Internet]. 2004 Jul 15 [cited 2020 Apr 21];78(14):7344–51. Available from: <https://jvi.asm.org/content/78/14/7344>
131. Reid SP, Leung LW, Hartman AL, Martinez O, Shaw ML, Carbonnelle C, et al. Ebola virus VP24 binds karyopherin alpha1 and blocks STAT1 nuclear accumulation. *J Virol* [Internet]. 2006 Jun 1 [cited 2020 Apr 21];80(11):5156–67. Available from: <https://jvi.asm.org/content/80/11/5156>

132. Ebihara H, Takada A, Kobasa D, Jones S, Neumann G, Theriault S, et al. Molecular Determinants of Ebola Virus Virulence in Mice. *PLoS Pathog* [Internet]. 2006 [cited 2020 Apr 21];2(7):e73. Available from: <http://dx.plos.org/10.1371/journal.ppat.0020073>
133. Grosfeld H, Hill MG, Collins PL. RNA replication by respiratory syncytial virus (RSV) is directed by the N, P, and L proteins; transcription also occurs under these conditions but requires RSV superinfection for efficient synthesis of full-length mRNA. *J Virol* [Internet]. 1995 Sep [cited 2020 Apr 21];69(9):5677–86. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7637014>
134. Lo MS, Brazas RM, Holtzman MJ. Respiratory syncytial virus nonstructural proteins NS1 and NS2 mediate inhibition of Stat2 expression and alpha/beta interferon responsiveness. *J Virol* [Internet]. 2005 Jul [cited 2020 Apr 21];79(14):9315–9. Available from: <https://jvi.asm.org/content/79/14/9315>
135. Green TJ, Zhang X, Wertz GW, Luo M. Structure of the vesicular stomatitis virus nucleoprotein-RNA complex. *Science* [Internet]. 2006 Jul 21 [cited 2020 Apr 21];313(5785):357–60. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1126953>
136. Ge P, Tsao J, Schein S, Green TJ, Luo M, Zhou ZH. Cryo-EM Model of the Bullet-Shaped Vesicular Stomatitis Virus. *Science* (80-) [Internet]. 2010 Feb 5 [cited 2020 Apr 21];327(5966):689–93. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1181766>
137. Arnheiter H, Davis NL, Wertz G, Schubert M, Lazzarini RA. Role of the nucleocapsid protein in regulating vesicular stomatitis virus RNA synthesis. *Cell* [Internet]. 1985 May [cited 2020 Apr 21];41(1):259–67. Available from: <https://linkinghub.elsevier.com/retrieve/pii/0092867485900790>
138. Mellon MG, Emerson SU. Rebinding of transcriptase components (L and NS proteins) to the nucleocapsid template of vesicular stomatitis virus. *J Virol* [Internet]. 1978 Sep [cited 2020 Apr 21];27(3):560–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/212581>
139. Emerson SU. Reconstitution studies detect a single polymerase entry site on the vesicular stomatitis virus genome. *Cell* [Internet]. 1982 Dec [cited 2020 Apr

- 21];31(3):635–42. Available from:
<https://linkinghub.elsevier.com/retrieve/pii/0092867482903191>
140. Ogino T, Kobayashi M, Iwama M, Mizumoto K. Sendai virus RNA-dependent RNA polymerase L protein catalyzes cap methylation of virus-specific mRNA. *J Biol Chem* [Internet]. 2005 Feb 11 [cited 2020 Apr 21];280(6):4429–35. Available from: <http://www.jbc.org/lookup/doi/10.1074/jbc.M411167200>
141. Stillman EA, Whitt MA. Transcript initiation and 5'-end modifications are separable events during vesicular stomatitis virus transcription. *J Virol* [Internet]. 1999 Sep [cited 2020 Apr 21];73(9):7199–209. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10438807>
142. Wang JT, McElvain LE, Whelan SPJ. Vesicular stomatitis virus mRNA capping machinery requires specific cis-acting signals in the RNA. *J Virol* [Internet]. 2007 Oct 15 [cited 2020 Apr 22];81(20):11499–506. Available from: <https://jvi.asm.org/content/81/20/11499>
143. Kuo L, Grosfeld H, Cristina J, Hill MG, Collins PL. Effects of mutations in the gene-start and gene-end sequence motifs on transcription of monocistronic and dicistronic minigenomes of respiratory syncytial virus. *J Virol* [Internet]. 1996 [cited 2020 Apr 22];70(10):6892. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC190738/>
144. Fearn R, Collins PL. Model for polymerase access to the overlapped L gene of respiratory syncytial virus. *J Virol* [Internet]. 1999 Jan [cited 2020 Apr 22];73(1):388–97. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9847343>
145. Iverson LE, Rose JK. Localized attenuation and discontinuous synthesis during vesicular stomatitis virus transcription. *Cell* [Internet]. 1981 Feb [cited 2020 Apr 23];23(2):477–84. Available from: <https://linkinghub.elsevier.com/retrieve/pii/0092867481901434>
146. Noton SL, Fearn R. Initiation and regulation of paramyxovirus transcription and replication. *Virology* [Internet]. 2015 May [cited 2018 Apr 11];479–480:545–54. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25683441>
147. Blumberg BM, Giorgi C, Kolakofsky D. N protein of vesicular stomatitis virus selectively encapsidates leader RNA in vitro. *Cell* [Internet]. 1983 Feb [cited 2020 Apr

- 23];32(2):559–67. Available from:
<https://linkinghub.elsevier.com/retrieve/pii/0092867483904750>
148. Blumberg BM, Leppert M, Kolakofsky D. Interaction of VSV leader RNA and nucleocapsid protein may control VSV genome replication. *Cell* [Internet]. 1981 Mar [cited 2020 Apr 23];23(3):837–45. Available from:
<https://linkinghub.elsevier.com/retrieve/pii/0092867481904487>
149. McGivern DR, Collins PL, Fearn R. Identification of internal sequences in the 3' leader region of human respiratory syncytial virus that enhance transcription and confer replication processivity. *J Virol* [Internet]. 2005 Feb 15 [cited 2020 Apr 23];79(4):2449–60. Available from: <https://jvi.asm.org/content/79/4/2449>
150. Modrof J, Möritz C, Kolesnikova L, Konakova T, Hartlieb B, Randolph A, et al. Phosphorylation of Marburg virus VP30 at serines 40 and 42 is critical for its interaction with NP inclusions. *Virology* [Internet]. 2001 Aug 15 [cited 2020 Apr 23];287(1):171–82. Available from:
<https://linkinghub.elsevier.com/retrieve/pii/S004268220191027X>
151. Ilinykh PA, Tigabu B, Ivanov A, Ammosova T, Obukhov Y, Garron T, et al. Role of protein phosphatase 1 in dephosphorylation of Ebola virus VP30 protein and its targeting for the inhibition of viral transcription. *J Biol Chem* [Internet]. 2014 Aug 15 [cited 2020 Apr 23];289(33):22723–38. Available from:
<http://www.jbc.org/lookup/doi/10.1074/jbc.M114.575050>
152. Groseth A, Charton JE, Sauerborn M, Feldmann F, Jones SM, Hoenen T, et al. The Ebola virus ribonucleoprotein complex: a novel VP30-L interaction identified. *Virus Res* [Internet]. 2009 Mar [cited 2020 Apr 23];140(1–2):8–14. Available from:
<https://linkinghub.elsevier.com/retrieve/pii/S0168170208003626>
153. Mondal A, Victor KG, Pudupakam RS, Lyons CE, Wertz GW. Newly identified phosphorylation site in the vesicular stomatitis virus P protein is required for viral RNA synthesis. *J Virol* [Internet]. 2014 Feb 1 [cited 2020 Apr 23];88(3):1461–72. Available from: <http://jvi.asm.org/cgi/doi/10.1128/JVI.02384-13>
154. Sun D, Luthra P, Xu P, Yoon H, He B. Identification of a phosphorylation site within the P protein important for mRNA transcription and growth of parainfluenza virus 5. *J Virol* [Internet]. 2011 Aug 15 [cited 2020 Apr 23];85(16):8376–85. Available from:

- <http://jvi.asm.org/cgi/doi/10.1128/JVI.00618-11>
155. Saikia P, Gopinath M, Shaila MS. Phosphorylation status of the phosphoprotein P of rinderpest virus modulates transcription and replication of the genome. *Arch Virol* [Internet]. 2008 Apr 29 [cited 2020 Apr 23];153(4):615–26. Available from: <http://link.springer.com/10.1007/s00705-008-0034-9>
 156. Finke S, Mueller-Waldeck R, Conzelmann K-K. Rabies virus matrix protein regulates the balance of virus transcription and replication. *J Gen Virol* [Internet]. 2003 Jun 1 [cited 2020 Apr 23];84(6):1613–21. Available from: <https://www.microbiologyresearch.org/content/journal/jgv/10.1099/vir.0.19128-0>
 157. Finke S, Conzelmann K-K. Dissociation of rabies virus matrix protein functions in regulation of viral RNA synthesis and virus assembly. *J Virol* [Internet]. 2003 Nov [cited 2020 Apr 23];77(22):12074–82. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14581544>
 158. Duan Z, Deng S, Ji X, Zhao J, Yuan C, Gao H. Nuclear localization of Newcastle disease virus matrix protein promotes virus replication by affecting viral RNA synthesis and transcription and inhibiting host cell transcription. *Vet Res* [Internet]. 2019 Mar 20 [cited 2020 Apr 23];50(1):22. Available from: <https://veterinaryresearch.biomedcentral.com/articles/10.1186/s13567-019-0640-4>
 159. Ahmed M, McKenzie MO, Puckett S, Hojnacki M, Poliquin L, Lyles DS. Ability of the matrix protein of vesicular stomatitis virus to suppress beta interferon gene expression is genetically correlated with the inhibition of host RNA and protein synthesis. *J Virol* [Internet]. 2003 Apr [cited 2020 Apr 23];77(8):4646–57. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12663771>
 160. Yu X, Shahriari S, Li H-M, Ghildyal R. Measles Virus Matrix Protein Inhibits Host Cell Transcription. Palazzo AF, editor. *PLoS One* [Internet]. 2016 Aug 23 [cited 2020 Apr 23];11(8):e0161360. Available from: <http://dx.plos.org/10.1371/journal.pone.0161360>
 161. Rajani KR, Pettit Kneller EL, McKenzie MO, Horita DA, Chou JW, Lyles DS. Complexes of vesicular stomatitis virus matrix protein with host Rae1 and Nup98 involved in inhibition of host transcription. Whelan SPJ, editor. *PLoS Pathog* [Internet]. 2012 Sep 27 [cited 2020 Apr 23];8(9):e1002929. Available from:

- <http://dx.plos.org/10.1371/journal.ppat.1002929>
162. Ahmed M, Lyles DS. Effect of vesicular stomatitis virus matrix protein on transcription directed by host RNA polymerases I, II, and III. *J Virol* [Internet]. 1998 Oct [cited 2020 Apr 23];72(10):8413–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9733895>
 163. Curran J, Marq JB, Kolakofsky D. An N-terminal domain of the Sendai paramyxovirus P protein acts as a chaperone for the NP protein during the nascent chain assembly step of genome replication. *J Virol* [Internet]. 1995 Feb [cited 2020 Apr 25];69(2):849–55. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7815552>
 164. Moyer SA, Smallwood-Kentro S, Haddad A, Prevec L. Assembly and transcription of synthetic vesicular stomatitis virus nucleocapsids. *J Virol* [Internet]. 1991 May [cited 2020 Apr 25];65(5):2170–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1850004>
 165. Whelan SPJ, Barr JN, Wertz GW. Transcription and replication of nonsegmented negative-strand RNA viruses. *Curr Top Microbiol Immunol* [Internet]. 2004 [cited 2019 May 23];283:61–119. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15298168>
 166. Saijo M, Niikura M, Morikawa S, Kurane I. Immunofluorescence method for detection of Ebola virus immunoglobulin g, using HeLa cells which express recombinant nucleoprotein. *J Clin Microbiol* [Internet]. 2001 Feb 1 [cited 2020 Apr 25];39(2):776–8. Available from: <http://jcm.asm.org/cgi/doi/10.1128/JCM.39.2.776-778.2001>
 167. Schoehn G, Iseni F, Mavrakis M, Blondel D, Ruigrok RW. Structure of recombinant rabies virus nucleoprotein-RNA complex and identification of the phosphoprotein binding site. *J Virol* [Internet]. 2001 Jan 1 [cited 2020 Apr 25];75(1):490–8. Available from: <https://jvi.asm.org/content/75/1/490>
 168. Ribeiro E de A, Leyrat C, Gérard FCA, Albertini AA V, Falk C, Ruigrok RWH, et al. Binding of rabies virus polymerase cofactor to recombinant circular nucleoprotein-RNA complexes. *J Mol Biol* [Internet]. 2009 Dec 4 [cited 2020 Apr 25];394(3):558–75. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0022283609011693>
 169. Iseni F, Ruigrok RW, Blondel D, Barge A, Baudin F. Characterization of rabies virus nucleocapsids and recombinant nucleocapsid-like structures. *J Gen Virol* [Internet].

- 1998 Dec 1 [cited 2020 Apr 25];79(12):2909–19. Available from: <https://www.microbiologyresearch.org/content/journal/jgv/10.1099/0022-1317-79-12-2909>
170. Mavrakis M, Iseni F, Mazza C, Schoehn G, Ebel C, Gentzel M, et al. Isolation and characterisation of the rabies virus N degrees-P complex produced in insect cells. *Virology* [Internet]. 2003 Jan 20 [cited 2020 Apr 25];305(2):406–14. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12573586>
171. Mavrakis M, Méhouas S, Réal E, Iseni F, Blondel D, Tordo N, et al. Rabies virus chaperone: identification of the phosphoprotein peptide that keeps nucleoprotein soluble and free from non-specific RNA. *Virology* [Internet]. 2006 Jun 5 [cited 2020 Apr 25];349(2):422–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16494915>
172. Liu P, Yang J, Wu X, Fu ZF. Interactions amongst rabies virus nucleoprotein, phosphoprotein and genomic RNA in virus-infected and transfected cells. *J Gen Virol* [Internet]. 2004 Dec [cited 2020 Apr 25];85(Pt 12):3725–34. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15557246>
173. Yang J, Hooper DC, Wunner WH, Koprowski H, Dietzschold B, Fu ZF. The specificity of rabies virus RNA encapsidation by nucleoprotein. *Virology* [Internet]. 1998 Mar 1 [cited 2020 Apr 25];242(1):107–17. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9501055>
174. Kawai A, Toriumi H, Tochikura TS, Takahashi T, Honda Y, Morimoto K. Nucleocapsid formation and/or subsequent conformational change of rabies virus nucleoprotein (N) is a prerequisite step for acquiring the phosphatase-sensitive epitope of monoclonal antibody 5-2-26. *Virology* [Internet]. 1999 Oct 25 [cited 2020 Apr 25];263(2):395–407. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10544112>
175. Wu X, Gong X, Foley HD, Schnell MJ, Fu ZF. Both viral transcription and replication are reduced when the rabies virus nucleoprotein is not phosphorylated. *J Virol* [Internet]. 2002 May [cited 2020 Apr 25];76(9):4153–61. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11932380>
176. Toriumi H, Kawai A. Association of rabies virus nominal phosphoprotein (P) with viral nucleocapsid (NC) is enhanced by phosphorylation of the viral nucleoprotein (N).

- Microbiol Immunol [Internet]. 2004 [cited 2020 Apr 25];48(5):399–409. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15215627>
177. Yang J, Koprowski H, Dietzschold B, Fu ZF. Phosphorylation of rabies virus nucleoprotein regulates viral RNA transcription and replication by modulating leader RNA encapsidation. *J Virol* [Internet]. 1999 Feb [cited 2020 Apr 25];73(2):1661–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9882376>
178. Iseni F, Barge A, Baudin F, Blondel D, Ruigrok RW. Characterization of rabies virus nucleocapsids and recombinant nucleocapsid-like structures. *J Gen Virol* [Internet]. 1998 Dec 1 [cited 2020 Apr 25];79 (Pt 12)(12):2909–19. Available from: <https://www.microbiologyresearch.org/content/journal/jgv/10.1099/0022-1317-79-12-2909>
179. Albertini AA V, Wernimont AK, Muziol T, Ravelli RBG, Clapier CR, Schoehn G, et al. Crystal structure of the rabies virus nucleoprotein-RNA complex. *Science* [Internet]. 2006 Jul 21 [cited 2020 Apr 25];313(5785):360–3. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1125280>
180. Le Blanc I, Luyet P-P, Pons V, Ferguson C, Emans N, Petiot A, et al. Endosome-to-cytosol transport of viral nucleocapsids. *Nat Cell Biol* [Internet]. 2005 Jul 12 [cited 2020 Apr 25];7(7):653–64. Available from: <http://www.nature.com/articles/ncb1269>
181. Akira S, Uematsu S, Takeuchi O. Pathogen Recognition and Innate Immunity. *Cell* [Internet]. 2006 Feb 24 [cited 2018 Aug 20];124(4):783–801. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16497588>
182. Espert L, Rey C, Gonzalez L, Degols G, Chelbi-Alix MK, Mechti N, et al. The exonuclease ISG20 is directly induced by synthetic dsRNA via NF-kappaB and IRF1 activation. *Oncogene* [Internet]. 2004 Jun 3 [cited 2020 Apr 25];23(26):4636–40. Available from: <http://www.nature.com/articles/1207586>
183. Albertini AA V, Clapier CR, Wernimont AK, Schoehn G, Weissenhorn W, Ruigrok RWH. Isolation and crystallization of a unique size category of recombinant Rabies virus Nucleoprotein-RNA rings. *J Struct Biol* [Internet]. 2007 Apr [cited 2020 Apr 25];158(1):129–33. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1047847706003200>
184. Alayyoubi M, Leser GP, Kors CA, Lamb RA. Structure of the paramyxovirus

- parainfluenza virus 5 nucleoprotein-RNA complex. *Proc Natl Acad Sci U S A* [Internet]. 2015 Apr 7 [cited 2020 Apr 25];112(14):E1792-9. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1503941112>
185. Jamin M, Yabukarski F. Nonsegmented Negative-Sense RNA Viruses-Structural Data Bring New Insights Into Nucleocapsid Assembly. *Adv Virus Res* [Internet]. 2017 [cited 2020 Apr 25];97:143–85. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0065352716300495>
186. Alvarez Paggi D, Esperante SA, Salgueiro M, Camporeale G, de Oliveira GAP, Prat Gay G. A conformational switch balances viral RNA accessibility and protection in a nucleocapsid ring model. *Arch Biochem Biophys* [Internet]. 2019 Aug [cited 2020 Apr 25];671:77–86. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0003986119302899>
187. Realegeno S, Niezgoda M, Yager PA, Kumar A, Hoque L, Orciari L, et al. An ELISA-based method for detection of rabies virus nucleoprotein-specific antibodies in human antemortem samples. Rupperecht CE, editor. *PLoS One* [Internet]. 2018 Nov 7 [cited 2020 Apr 25];13(11):e0207009. Available from: <http://dx.plos.org/10.1371/journal.pone.0207009>
188. Goto H, Minamoto N, Ito H, Ito N, Sugiyama M, Kinjo T, et al. Mapping of epitopes and structural analysis of antigenic sites in the nucleoprotein of rabies virus. *J Gen Virol* [Internet]. 2000 Jan 1 [cited 2020 Apr 25];81(Pt 1):119–27. Available from: <https://www.microbiologyresearch.org/content/journal/jgv/10.1099/0022-1317-81-1-119>
189. Rahmahani J, Suwarno S, Yuniarti WM, Rantam FA. Antigenic site of nucleoprotein gene from Indonesian rabies virus isolates. *Vet World* [Internet]. 2019 May 31 [cited 2020 Apr 25];12(5):724–8. Available from: <http://www.veterinaryworld.org/Vol.12/May-2019/16.html>
190. Mühlberger E, Lötfering B, Klenk HD, Becker S. Three of the four nucleocapsid proteins of Marburg virus, NP, VP35, and L, are sufficient to mediate replication and transcription of Marburg virus-specific monocistronic minigenomes. *J Virol* [Internet]. 1998 Nov [cited 2020 Apr 25];72(11):8756–64. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9765419>

191. Kolesnikova L, Mühlberger E, Ryabchikova E, Becker S. Ultrastructural Organization of Recombinant Marburg Virus Nucleoprotein: Comparison with Marburg Virus Inclusions. *J Virol* [Internet]. 2000 Apr 15 [cited 2020 Apr 25];74(8):3899–904. Available from: <https://jvi.asm.org/content/74/8/3899>
192. Noda T, Hagiwara K, Sagara H, Kawaoka Y. Characterization of the Ebola virus nucleoprotein-RNA complex. *J Gen Virol* [Internet]. 2010 Jun 1 [cited 2020 Apr 25];91(Pt 6):1478–83. Available from: <https://www.microbiologyresearch.org/content/journal/jgv/10.1099/vir.0.019794-0>
193. Watanabe S, Noda T, Kawaoka Y. Functional mapping of the nucleoprotein of Ebola virus. *J Virol* [Internet]. 2006 Apr 15 [cited 2020 Apr 25];80(8):3743–51. Available from: <https://jvi.asm.org/content/80/8/3743>
194. Sanchez A, Kiley MP, Holloway BP, McCormick JB, Auperin DD. The nucleoprotein gene of ebola virus: Cloning, sequencing, and in vitro expression. *Virology* [Internet]. 1989 May [cited 2020 Apr 25];170(1):81–91. Available from: <https://linkinghub.elsevier.com/retrieve/pii/0042682289903541>
195. Dong S, Yang P, Li G, Liu B, Wang W, Liu X, et al. Insight into the Ebola virus nucleocapsid assembly mechanism: crystal structure of Ebola virus nucleoprotein core domain at 1.8 Å resolution. *Protein Cell* [Internet]. 2015 May 25 [cited 2020 Apr 25];6(5):351–62. Available from: <http://link.springer.com/10.1007/s13238-015-0163-3>
196. Dziubańska PJ, Derewenda U, Ellena JF, Engel DA, Derewenda ZS. The structure of the C-terminal domain of the Zaire ebolavirus nucleoprotein. *Acta Crystallogr D Biol Crystallogr* [Internet]. 2014 Sep 1 [cited 2020 Apr 25];70(Pt 9):2420–9. Available from: <http://scripts.iucr.org/cgi-bin/paper?S1399004714014710>
197. Sanchez A, Kiley MP, Klenk HD, Feldmann H. Sequence analysis of the Marburg virus nucleoprotein gene: comparison to Ebola virus and other non-segmented negative-strand RNA viruses. *J Gen Virol* [Internet]. 1992 Feb 1 [cited 2020 Apr 25];73 (Pt 2)(2):347–57. Available from: <https://www.microbiologyresearch.org/content/journal/jgv/10.1099/0022-1317-73-2-347>
198. Leung DW, Borek D, Luthra P, Binning JM, Anantpadma M, Liu G, et al. An Intrinsically Disordered Peptide from Ebola Virus VP35 Controls Viral RNA Synthesis

- by Modulating Nucleoprotein-RNA Interactions. *Cell Rep* [Internet]. 2015 Apr [cited 2020 Apr 25];11(3):376–89. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2211124715003034>
199. DiCarlo A, Möller P, Lander A, Kolesnikova L, Becker S. Nucleocapsid formation and RNA synthesis of Marburg virus is dependent on two coiled coil motifs in the nucleoprotein. *Virology* [Internet]. 2007 Oct 24 [cited 2020 Apr 25];4(1):105. Available from: <http://virologyj.biomedcentral.com/articles/10.1186/1743-422X-4-105>
200. Leung DW, Borek D, Luthra P, Binning JM, Anantpadma M, Liu G, et al. An Intrinsically Disordered Peptide from Ebola Virus VP35 Controls Viral RNA Synthesis by Modulating Nucleoprotein-RNA Interactions. *Cell Rep* [Internet]. 2015 Apr 21 [cited 2020 Apr 25];11(3):376–89. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2211124715003034>
201. Beniac DR, Melito PL, Devarenes SL, Hiebert SL, Rabb MJ, Lamboo LL, et al. The organisation of Ebola virus reveals a capacity for extensive, modular polyploidy. Park M-S, editor. *PLoS One* [Internet]. 2012 Jan 11 [cited 2020 Apr 25];7(1):e29608. Available from: <http://dx.plos.org/10.1371/journal.pone.0029608>
202. Sanchez A, Kiley MP, Holloway BP, McCormick JB, Auperin DD. The nucleoprotein gene of Ebola virus: cloning, sequencing, and in vitro expression. *Virology* [Internet]. 1989 May [cited 2020 Apr 25];170(1):81–91. Available from: <https://linkinghub.elsevier.com/retrieve/pii/0042682289903541>
203. Noda T, Watanabe S, Sagara H, Kawaoka Y. Mapping of the VP40-binding regions of the nucleoprotein of Ebola virus. *J Virol* [Internet]. 2007 Apr 1 [cited 2020 Apr 25];81(7):3554–62. Available from: <https://jvi.asm.org/content/81/7/3554>
204. Baker LE, Ellena JF, Handing KB, Derewenda U, Utepbergenov D, Engel DA, et al. Molecular architecture of the nucleoprotein C-terminal domain from the Ebola and Marburg viruses. *Acta Crystallogr Sect D Struct Biol* [Internet]. 2016 Jan 1 [cited 2020 Apr 25];72(1):49–58. Available from: <http://scripts.iucr.org/cgi-bin/paper?S2059798315021439>
205. Sherwood LJ, Hayhurst A. Ebolavirus nucleoprotein C-termini potently attract single domain antibodies enabling monoclonal affinity reagent sandwich assay (MARSA) formulation. Ikegami T, editor. *PLoS One* [Internet]. 2013 Apr 5 [cited 2020 Apr

- 25];8(4):e61232. Available from: <http://dx.plos.org/10.1371/journal.pone.0061232>
206. Baker LE, Ellena JF, Handing KB, Derewenda U, Utepbergenov D, Engel DA, et al. Molecular architecture of the nucleoprotein C-terminal domain from the Ebola and Marburg viruses. *Acta Crystallogr Sect D, Struct Biol* [Internet]. 2016 Jan 1 [cited 2020 Apr 25];72(Pt 1):49–58. Available from: <http://scripts.iucr.org/cgi-bin/paper?S2059798315021439>
207. Changula K, Yoshida R, Noyori O, Marzi A, Miyamoto H, Ishijima M, et al. Mapping of conserved and species-specific antibody epitopes on the Ebola virus nucleoprotein. *Virus Res* [Internet]. 2013 Sep [cited 2020 Apr 25];176(1–2):83–90. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S016817021300169X>
208. Niikura M, Ikegami T, Saijo M, Kurata T, Kurane I, Morikawa S. Analysis of linear B-cell epitopes of the nucleoprotein of ebola virus that distinguish ebola virus subtypes. *Clin Diagn Lab Immunol* [Internet]. 2003 Jan [cited 2020 Apr 25];10(1):83–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12522044>
209. Huang Y, Xu L, Sun Y, Nabel GJ. The Assembly of Ebola Virus Nucleocapsid Requires Virion-Associated Proteins 35 and 24 and Posttranslational Modification of Nucleoprotein. *Mol Cell* [Internet]. 2002 Aug [cited 2020 Apr 25];10(2):307–16. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276502005889>
210. Becker S, Mühlberger E. Co- and Posttranslational Modifications and Functions of Marburg Virus Proteins. In 1999 [cited 2020 Apr 25]. p. 23–34. Available from: http://link.springer.com/10.1007/978-3-642-59949-1_2
211. Lötfering B, Mühlberger E, Tamura T, Klenk HD, Becker S. The nucleoprotein of Marburg virus is target for multiple cellular kinases. *Virology* [Internet]. 1999 Mar 1 [cited 2020 Apr 25];255(1):50–62. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0042682298995770>
212. DiCarlo A, Biedenkopf N, Hartlieb B, Klussmeier A, Becker S. Phosphorylation of Marburg virus NP region II modulates viral RNA synthesis. *J Infect Dis* [Internet]. 2011 Nov [cited 2020 Apr 25];204 Suppl 3(suppl_3):S927-33. Available from: <https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jir319>
213. Rahmeh AA, Morin B, Schenk AD, Liang B, Heinrich BS, Brusica V, et al. Critical phosphoprotein elements that regulate polymerase architecture and function in

- vesicular stomatitis virus. *Proc Natl Acad Sci* [Internet]. 2012 Sep 4 [cited 2020 Apr 26];109(36):14628–33. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.1209147109>
214. Brzózka K, Finke S, Conzelmann K-K. Identification of the Rabies Virus Alpha/Beta Interferon Antagonist: Phosphoprotein P Interferes with Phosphorylation of Interferon Regulatory Factor 3. *J Virol* [Internet]. 2005 Jun 15 [cited 2020 Apr 26];79(12):7673–81. Available from: <https://jvi.asm.org/content/79/12/7673>
215. Rieder M, Brzozka K, Pfaller CK, Cox JH, Stitz L, Conzelmann K-K. Genetic Dissection of Interferon-Antagonistic Functions of Rabies Virus Phosphoprotein: Inhibition of Interferon Regulatory Factor 3 Activation Is Important for Pathogenicity. *J Virol* [Internet]. 2011 Jan 15 [cited 2020 Apr 26];85(2):842–52. Available from: <http://jvi.asm.org/cgi/doi/10.1128/JVI.01427-10>
216. Brzózka K, Finke S, Conzelmann K-K. Inhibition of Interferon Signaling by Rabies Virus Phosphoprotein P: Activation-Dependent Binding of STAT1 and STAT2. *J Virol* [Internet]. 2006 Mar 15 [cited 2020 Apr 26];80(6):2675–83. Available from: <https://jvi.asm.org/content/80/6/2675>
217. Vidy A, Chelbi-Alix M, Blondel D. Rabies Virus P Protein Interacts with STAT1 and Inhibits Interferon Signal Transduction Pathways. *J Virol* [Internet]. 2005 Nov 15 [cited 2020 Apr 26];79(22):14411–20. Available from: <https://jvi.asm.org/content/79/22/14411>
218. Vidy A, El Bougrini J, Chelbi-Alix MK, Blondel D. The Nucleocytoplasmic Rabies Virus P Protein Counteracts Interferon Signaling by Inhibiting both Nuclear Accumulation and DNA Binding of STAT1. *J Virol* [Internet]. 2007 Apr 15 [cited 2020 Apr 26];81(8):4255–63. Available from: <https://jvi.asm.org/content/81/8/4255>
219. Ding H, Green TJ, Lu S, Luo M. Crystal Structure of the Oligomerization Domain of the Phosphoprotein of Vesicular Stomatitis Virus. *J Virol* [Internet]. 2006 Mar 15 [cited 2020 Apr 26];80(6):2808–14. Available from: <https://jvi.asm.org/content/80/6/2808>
220. Gerard FCA, Ribeiro E de A, Albertini AA V, Gutsche I, Zaccai G, Ruigrok RWH, et al. Unphosphorylated rhabdoviridae phosphoproteins form elongated dimers in solution. *Biochemistry* [Internet]. 2007 Sep 11 [cited 2020 Apr 26];46(36):10328–38.

- Available from: <https://pubs.acs.org/doi/10.1021/bi7007799>
221. Horwitz JA, Jenni S, Harrison SC, Whelan SPJ. Structure of a rabies virus polymerase complex from electron cryo-microscopy. *Proc Natl Acad Sci* [Internet]. 2020 Jan 28 [cited 2020 Apr 21];117(4):2099–107. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1918809117>
222. Jacob Y, Real E, Tordo N. Functional Interaction Map of Lyssavirus Phosphoprotein: Identification of the Minimal Transcription Domains. *J Virol* [Internet]. 2001 Oct 15 [cited 2020 Apr 26];75(20):9613–22. Available from: <https://jvi.asm.org/content/75/20/9613>
223. Gigant B, Iseni F, Gaudin Y, Knossow M, Blondel D. Neither phosphorylation nor the amino-terminal part of rabies virus phosphoprotein is required for its oligomerization. *J Gen Virol* [Internet]. 2000 Jul 1 [cited 2020 Apr 26];81(7):1757–61. Available from: <https://www.microbiologyresearch.org/content/journal/jgv/10.1099/0022-1317-81-7-1757>
224. Ivanov I, Crépin T, Jamin M, Ruigrok RWH. Structure of the Dimerization Domain of the Rabies Virus Phosphoprotein. *J Virol* [Internet]. 2010 Apr 1 [cited 2020 Apr 26];84(7):3707–10. Available from: <https://jvi.asm.org/content/84/7/3707>
225. Fu ZF, Zheng Y, Wunner WH, Koprowski H, Dietzschold B. Both the N- and the C-terminal domains of the nominal phosphoprotein of rabies virus are involved in binding to the nucleoprotein. *Virology* [Internet]. 1994 May 1 [cited 2020 Apr 26];200(2):590–7. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0042682284712220>
226. Chenik M, Chebli K, Gaudin Y, Blondel D. In Vivo Interaction of Rabies Virus Phosphoprotein (P) and Nucleoprotein (N): Existence of Two N-binding Sites on P Protein. *J Gen Virol* [Internet]. 1994 Nov 1 [cited 2020 Apr 26];75(11):2889–96. Available from: <https://www.microbiologyresearch.org/content/journal/jgv/10.1099/0022-1317-75-11-2889>
227. Gerard FCA, Ribeiro E de A, Leyrat C, Ivanov I, Blondel D, Longhi S, et al. Modular Organization of Rabies Virus Phosphoprotein. *J Mol Biol* [Internet]. 2009 May [cited 2020 Apr 26];388(5):978–96. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0022283609003684>

228. Emerson SU, Wagner RR. L Protein Requirement for In Vitro RNA Synthesis by Vesicular Stomatitis Virus. *J Virol* [Internet]. 1973 Dec 1 [cited 2020 Apr 26];12(6):1325–35. Available from: <https://jvi.asm.org/content/12/6/1325.long>
229. Morin B, Rahmeh AA, Whelan SP. Mechanism of RNA synthesis initiation by the vesicular stomatitis virus polymerase. *EMBO J* [Internet]. 2012 Mar 7 [cited 2020 Apr 26];31(5):1320–9. Available from: <http://emboj.embopress.org/cgi/doi/10.1038/emboj.2011.483>
230. Rahmeh AA, Morin B, Schenk AD, Liang B, Heinrich BS, Brusica V, et al. Critical phosphoprotein elements that regulate polymerase architecture and function in vesicular stomatitis virus. *Proc Natl Acad Sci U S A* [Internet]. 2012 Sep 4 [cited 2020 Apr 26];109(36):14628–33. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.1209147109>
231. Morin B, Liang B, Gardner E, Ross RA, Whelan SPJ. An In Vitro RNA Synthesis Assay for Rabies Virus Defines Ribonucleoprotein Interactions Critical for Polymerase Activity. *J Virol*. 2017 Jan;91(1).
232. Chenik M, Chebli K, Blondel D. Translation initiation at alternate in-frame AUG codons in the rabies virus phosphoprotein mRNA is mediated by a ribosomal leaky scanning mechanism. *J Virol* [Internet]. 1995 Feb 1 [cited 2020 Apr 26];69(2):707–12. Available from: <https://jvi.asm.org/content/69/2/707.long>
233. Mapping the interacting domains between the rabies virus polymerase and phosphoprotein. - PubMed - NCBI [Internet]. [cited 2020 Apr 26]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/?term=9499045>
234. Marschalek A, Drechsel L, Conzelmann K-K. The importance of being short: the role of rabies virus phosphoprotein isoforms assessed by differential IRES translation initiation. *Eur J Cell Biol* [Internet]. 2012 Jan [cited 2020 Apr 26];91(1):17–23. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0171933511000197>
235. Yamaoka S, Ito N, Ohka S, Kaneda S, Nakamura H, Agari T, et al. Involvement of the Rabies Virus Phosphoprotein Gene in Neuroinvasiveness. *J Virol* [Internet]. 2013 Nov 15 [cited 2020 Apr 26];87(22):12327–38. Available from: <http://jvi.asm.org/cgi/doi/10.1128/JVI.02132-13>
236. Okada K, Ito N, Yamaoka S, Masatani T, Ebihara H, Goto H, et al. Roles of the Rabies

- Virus Phosphoprotein Isoforms in Pathogenesis. Lyles DS, editor. *J Virol* [Internet]. 2016 Sep 15 [cited 2020 Apr 26];90(18):8226–37. Available from: <http://jvi.asm.org/lookup/doi/10.1128/JVI.00809-16>
237. Gupta AK, Blondel D, Choudhary S, Banerjee AK. The Phosphoprotein of Rabies Virus Is Phosphorylated by a Unique Cellular Protein Kinase and Specific Isomers of Protein Kinase C. *J Virol* [Internet]. 2000 Jan 1 [cited 2020 Apr 26];74(1):91–8. Available from: <https://jvi.asm.org/content/74/1/91>
238. Chen J-L, Das T, Banerjee AK. Phosphorylated States of Vesicular Stomatitis Virus P Protein in Vitro and in Vivo. *Virology* [Internet]. 1997 Feb [cited 2020 Apr 26];228(2):200–12. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0042682296984019>
239. Constitutive phosphorylation of the vesicular stomatitis virus P protein modulates polymerase complex formation but is not essential for transcript... - PubMed - NCBI [Internet]. [cited 2020 Apr 26]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/?term=8676480>
240. Chattopadhyay D, Banerjee AK. Phosphorylation within a specific domain of the phosphoprotein of vesicular stomatitis virus regulates transcription in vitro. *Cell* [Internet]. 1987 May [cited 2020 Apr 26];49(3):407–14. Available from: <https://linkinghub.elsevier.com/retrieve/pii/0092867487902935>
241. Hartman AL, Towner JS, Nichol ST. A C-terminal basic amino acid motif of Zaire ebolavirus VP35 is essential for type I interferon antagonism and displays high identity with the RNA-binding domain of another interferon antagonist, the NS1 protein of influenza A virus. *Virology* [Internet]. 2004 Oct [cited 2020 Apr 26];328(2):177–84. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0042682204004659>
242. Leung DW, Prins KC, Basler CF, Amarasinghe GK. *Ebolavirus* VP35 is a multifunctional virulence factor. *Virulence* [Internet]. 2010 Nov 27 [cited 2020 Apr 26];1(6):526–31. Available from: <http://www.tandfonline.com/doi/abs/10.4161/viru.1.6.12984>
243. Schumann M, Gantke T, Mühlberger E. Ebola Virus VP35 Antagonizes PKR Activity through Its C-Terminal Interferon Inhibitory Domain. *J Virol* [Internet]. 2009 Sep 1 [cited 2020 Apr 26];83(17):8993–7. Available from:

- <https://jvi.asm.org/content/83/17/8993>
244. Trunschke M, Conrad D, Enterlein S, Olejnik J, Brauburger K, Mühlberger E. The L–VP35 and L–L interaction domains reside in the amino terminus of the Ebola virus L protein and are potential targets for antivirals. *Virology* [Internet]. 2013 Jul [cited 2020 Apr 26];441(2):135–45. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0042682213001530>
 245. Möller P, Pariente N, Klenk H-D, Becker S. Homo-Oligomerization of Marburgvirus VP35 Is Essential for Its Function in Replication and Transcription. *J Virol* [Internet]. 2005 Dec 15 [cited 2020 Apr 26];79(23):14876–86. Available from: <https://jvi.asm.org/content/79/23/14876>
 246. Reid SP, Cárdenas WB, Basler CF. Homo-oligomerization facilitates the interferon-antagonist activity of the ebolavirus VP35 protein. *Virology* [Internet]. 2005 Oct [cited 2020 Apr 26];341(2):179–89. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S004268220500382X>
 247. Basler CF, Wang X, Muhlberger E, Volchkov V, Paragas J, Klenk H-D, et al. The Ebola virus VP35 protein functions as a type I IFN antagonist. *Proc Natl Acad Sci* [Internet]. 2000 Oct 24 [cited 2020 Apr 26];97(22):12289–94. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.220398297>
 248. Basler CF, Mikulasova A, Martinez-Sobrido L, Paragas J, Mühlberger E, Bray M, et al. The Ebola virus VP35 protein inhibits activation of interferon regulatory factor 3. *J Virol* [Internet]. 2003 Jul [cited 2020 Apr 26];77(14):7945–56. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12829834>
 249. Bujnicki JM, Rychlewski L. In silico identification, structure prediction and phylogenetic analysis of the 2'-O-ribose (cap 1) methyltransferase domain in the large structural protein of ssRNA negative-strand viruses. *Protein Eng.* 2002 Feb;15(2):101–8.
 250. Ferron F, Longhi S, Henrissat B, Canard B. Viral RNA-polymerases -- a predicted 2'-O-ribose methyltransferase domain shared by all Mononegavirales. *Trends Biochem Sci.* 2002 May;27(5):222–4.
 251. Poch O, Blumberg BM, Bougueleret L, Tordo N. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical

- assignment of functional domains. *J Gen Virol*. 1990 May;71 (Pt 5):1153–62.
252. Svenda M, Berg M, Moreno-López J, Linné T. Analysis of the large (L) protein gene of the porcine rubulavirus LPMV: identification of possible functional domains. *Virus Res* [Internet]. 1997 Apr [cited 2020 Apr 26];48(1):57–70. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0168170296014268>
253. Rahmeh AA, Schenk AD, Danek EI, Kranzusch PJ, Liang B, Walz T, et al. Molecular architecture of the vesicular stomatitis virus RNA polymerase. *Proc Natl Acad Sci U S A* [Internet]. 2010 Nov 16 [cited 2019 May 23];107(46):20075–80. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21041632>
254. Liang B, Li Z, Jenni S, Rahmeh AA, Morin BM, Grant T, et al. Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy. *Cell* [Internet]. 2015 Jul 16 [cited 2019 May 23];162(2):314–27. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867415007023>
255. Transcriptional activity and mutational analysis of recombinant vesicular stomatitis virus RNA polymerase. - PubMed - NCBI [Internet]. [cited 2020 Apr 26]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/?term=8382299>
256. cis-Acting signals involved in termination of vesicular stomatitis virus mRNA synthesis include the conserved AUAC and the U7 signal for polyadenyl... - PubMed - NCBI [Internet]. [cited 2020 Apr 26]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/?term=9343230>
257. Ogino T, Yadav SP, Banerjee AK. Histidine-mediated RNA transfer to GDP for unique mRNA capping by vesicular stomatitis virus RNA polymerase. *Proc Natl Acad Sci* [Internet]. 2010 Feb 23 [cited 2020 Apr 26];107(8):3463–8. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.0913083107>
258. Neubauer J, Ogino M, Green TJ, Ogino T. Signature motifs of GDP polyribonucleotidyltransferase, a non-segmented negative strand RNA viral mRNA capping enzyme, domain in the L protein are required for covalent enzyme–pRNA intermediate formation. *Nucleic Acids Res* [Internet]. 2016 Jan 8 [cited 2020 Apr 26];44(1):330–41. Available from: <https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkv1286>
259. Li J, Rahmeh A, Morelli M, Whelan SPJ. A Conserved Motif in Region V of the Large

- Polymerase Proteins of Nonsegmented Negative-Sense RNA Viruses That Is Essential for mRNA Capping. *J Virol* [Internet]. 2008 Jan [cited 2020 Apr 20];82(2):775. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18003731>
260. Ogino T. Capping of vesicular stomatitis virus pre-mRNA is required for accurate selection of transcription stop–start sites and virus propagation. *Nucleic Acids Res* [Internet]. 2014 Oct 29 [cited 2020 Apr 26];42(19):12112–25. Available from: <http://academic.oup.com/nar/article/42/19/12112/2903128/Capping-of-vesicular-stomatitis-virus-premRNA-is>
261. Ogino M, Gupta N, Green TJ, Ogino T. A dual-functional priming-capping loop of rhabdoviral RNA polymerases directs terminal *de novo* initiation and capping intermediate formation. *Nucleic Acids Res* [Internet]. 2019 Jan 10 [cited 2020 Apr 26];47(1):299–309. Available from: <https://academic.oup.com/nar/article/47/1/299/5160996>
262. Horwitz JA, Jenni S, Harrison SC, Whelan SPJ. Structure of a rabies virus polymerase complex from electron cryo-microscopy. *Proc Natl Acad Sci U S A* [Internet]. 2020 Jan 28 [cited 2020 Apr 21];117(4):2099–107. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1918809117>
263. Jenni S, Bloyet L-M, Diaz-Avalos R, Liang B, Whelan SPJ, Grigorieff N, et al. Structure of the Vesicular Stomatitis Virus L Protein in Complex with Its Phosphoprotein Cofactor. *Cell Rep* [Internet]. 2020 Jan 7 [cited 2020 Apr 21];30(1):53-60.e5. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2211124719316742>
264. Gilman MSA, Liu C, Fung A, Behera I, Jordan P, Rigaux P, et al. Structure of the Respiratory Syncytial Virus Polymerase Complex. *Cell* [Internet]. 2019 Sep 30 [cited 2019 Sep 10]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31495574>
265. Pan J, Qian X, Lattmann S, El Sahili A, Yeo TH, Jia H, et al. Structure of the human metapneumovirus polymerase phosphoprotein complex. *Nature*. 2020 Jan 9;577(7789):275–9.
266. Paesen GC, Collet A, Sallamand C, Debart F, Vasseur J-J, Canard B, et al. X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat Commun* [Internet]. 2015 Nov 9 [cited 2018 Jul 18];6:8749. Available from:

- <http://www.ncbi.nlm.nih.gov/pubmed/26549102>
267. Paesen GC, Collet A, Sallamand C, Debart F, Vasseur J-J, Canard B, et al. X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat Commun* [Internet]. 2015 Dec 9 [cited 2018 Oct 17];6(1):8749. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26549102>
268. Rahmeh AA, Li J, Kranzusch PJ, Whelan SPJ. Ribose 2'-O Methylation of the Vesicular Stomatitis Virus mRNA Cap Precedes and Facilitates Subsequent Guanine-N-7 Methylation by the Large Polymerase Protein. *J Virol* [Internet]. 2009 Nov 1 [cited 2018 Jul 18];83(21):11043–50. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19710136>
269. Li J, Wang JT, Whelan SPJ. A unique strategy for mRNA cap methylation used by vesicular stomatitis virus. *Proc Natl Acad Sci* [Internet]. 2006 May 30 [cited 2020 Apr 26];103(22):8493–8. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.0509821103>
270. Martin B, Coutard B, Guez T, Paesen GC, Canard B, Debart F, et al. The methyltransferase domain of the Sudan ebolavirus L protein specifically targets internal adenosines of RNA substrates, in addition to the cap structure. *Nucleic Acids Res* [Internet]. 2018 Sep 6 [cited 2019 May 23];46(15):7902–12. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30192980>
271. Abdella R, Aggarwal M, Okura T, Lamb RA, He Y. Structure of a paramyxovirus polymerase complex reveals a unique methyltransferase-CTD conformation. *Proc Natl Acad Sci* [Internet]. 2020 Mar 3 [cited 2020 Mar 23];117(9):4931–41. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1919837117>
272. CHANDRIKA R, HORIKAMI SM, SMALLWOOD S, MOYER SA. Mutations in Conserved Domain I of the Sendai Virus L Polymerase Protein Uncouple Transcription and Replication. *Virology* [Internet]. 1995 Nov [cited 2020 Apr 27];213(2):352–63. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0042682285700086>
273. Qiu S, Ogino M, Luo M, Ogino T, Green TJ. Structure and Function of the N-Terminal Domain of the Vesicular Stomatitis Virus RNA Polymerase. Lyles DS, editor. *J Virol* [Internet]. 2016 Jan 15 [cited 2020 Apr 27];90(2):715–24. Available from: <http://jvi.asm.org/lookup/doi/10.1128/JVI.02317-15>

274. Cevik B, Holmes DE, Vrotsos E, Feller JA, Smallwood S, Moyer SA. The phosphoprotein (P) and L binding sites reside in the N-terminus of the L subunit of the measles virus RNA polymerase. *Virology* [Internet]. 2004 Oct [cited 2020 Apr 27];327(2):297–306. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0042682204004581>
275. Nakagawa K, Kobayashi Y, Ito N, Suzuki Y, Okada K, Makino M, et al. Molecular Function Analysis of Rabies Virus RNA Polymerase L Protein by Using an L Gene-Deficient Virus. García-Sastre A, editor. *J Virol* [Internet]. 2017 Oct 15 [cited 2020 Apr 26];91(20). Available from: <https://jvi.asm.org/lookup/doi/10.1128/JVI.00826-17>
276. Perry KL, Watkins KP, Agabian N. Trypanosome mRNAs have unusual ‘cap 4’ structures acquired by addition of a spliced leader. *Proc Natl Acad Sci U S A* [Internet]. 1987 Dec 1 [cited 2020 Apr 16];84(23):8190–4. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.84.23.8190>
277. Shatkin AJ. Capping of eucaryotic mRNAs. *Cell* [Internet]. 1976 Dec [cited 2019 May 23];9(4 PT 2):645–53. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1017010>
278. Myette JR, Niles EG. Characterization of the vaccinia virus RNA 5'-triphosphatase and nucleotide triphosphate phosphohydrolase activities. Demonstrate that both activities are carried out at the same active site. *J Biol Chem* [Internet]. 1996 May 17 [cited 2020 Apr 16];271(20):11945–52. Available from: <http://www.jbc.org/lookup/doi/10.1074/jbc.271.20.11945>
279. Cong P, Shuman S. Mutational analysis of mRNA capping enzyme identifies amino acids involved in GTP binding, enzyme-guanylate formation, and GMP transfer to RNA. *Mol Cell Biol* [Internet]. 1995 Nov [cited 2020 Apr 16];15(11):6222–31. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7565775>
280. Werner M, Purta E, Kaminska KH, Cymerman IA, Campbell DA, Mitra B, et al. 2'-O-ribose methylation of cap2 in human: Function and evolution in a horizontally mobile family. *NAR*. 2011;39(11):4756–68.
281. Smietanski M, Werner M, Purta E, Kaminska KH, Stepinski J, Darzynkiewicz E, et al. Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nat Commun*. 2014;5:3004.
282. Decroly E, Ferron F, Lescar J, Canard B. Conventional and unconventional

- mechanisms for capping viral mRNA. *Nat Rev Microbiol* [Internet]. 2011 Dec 5 [cited 2019 May 23];10(1):51–65. Available from: <http://www.nature.com/articles/nrmicro2675>
283. Goodfellow I, Chaudhry Y, Gioldasi I, Gerondopoulos A, Natoni A, Labrie L, et al. Calicivirus translation initiation requires an interaction between VPg and eIF 4 E. *EMBO Rep* [Internet]. 2005 Oct [cited 2020 Apr 16];6(10):968–72. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1038/sj.embor.7400510>
284. Cougot N, van Dijk E, Babajko S, Séraphin B. “Cap-tabolism”. *Trends Biochem Sci* [Internet]. 2004 Aug [cited 2020 Apr 16];29(8):436–44. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0968000404001549>
285. 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* [Internet]. 1995 Jan 17 [cited 2020 Apr 16];92(2):507–11. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.92.2.507>
286. Vasiljeva L, Merits A, Auvinen P, Kääriäinen L. Identification of a Novel Function of the *Alphavirus* Capping Apparatus. *J Biol Chem* [Internet]. 2000 Jun 9 [cited 2020 Apr 16];275(23):17281–7. Available from: <http://www.jbc.org/lookup/doi/10.1074/jbc.M910340199>
287. 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* [Internet]. 2014 Feb 14 [cited 2020 Apr 16];343(6172):783–7. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1248465>
288. Ogino T, Banerjee AK. Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Mol Cell*. 2007 Jan;25(1):85–97.
289. Rahmeh AA, Li J, Kranzusch PJ, Whelan SPJ. Ribose 2'-O methylation of the vesicular stomatitis virus mRNA cap precedes and facilitates subsequent guanine-N-7 methylation by the large polymerase protein. *J Virol* [Internet]. 2009 Nov 1 [cited 2019 May 23];83(21):11043–50. Available from: <http://jvi.asm.org/cgi/doi/10.1128/JVI.01426-09>
290. Fischer PM. Cap in hand: Targeting eIF4E. *Cell Cycle* [Internet]. 2009 Aug 15 [cited

- 2020 Apr 16];8(16):2535–41. Available from: <http://www.tandfonline.com/doi/abs/10.4161/cc.8.16.9301>
291. Borden KLB. The eukaryotic translation initiation factor eIF4E wears a ‘cap’ for many occasions. *Transl (Austin, Tex)* [Internet]. 2016 Jul 2 [cited 2020 Apr 16];4(2):e1220899. Available from: <https://www.tandfonline.com/doi/full/10.1080/21690731.2016.1220899>
292. Beelman CA, Parker R. Degradation of mRNA in eukaryotes. *Cell* [Internet]. 1995 Apr 21 [cited 2020 Apr 16];81(2):179–83. Available from: <https://linkinghub.elsevier.com/retrieve/pii/0092867495903267>
293. Liu H, Kiledjian M. Decapping the message: a beginning or an end. *Biochem Soc Trans* [Internet]. 2006 Feb 1 [cited 2020 Apr 16];34(1):35. Available from: <http://www.biochemsoctrans.org/bst/034/bst0340035.htm>
294. Devarkar SC, Wang C, Miller MT, Ramanathan A, Jiang F, Khan AG, et al. Structural basis for m7G recognition and 2'-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proc Natl Acad Sci U S A* [Internet]. 2016 Jan 19 [cited 2019 Jun 19];113(3):596–601. Available from: <http://www.pnas.org/lookup/doi/10.1073/pnas.1515152113>
295. Züst R, Cervantes-Barragan L, Habjan M, Maier R, Neuman BW, Ziebuhr J, et al. Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat Immunol* [Internet]. 2011 Feb 9 [cited 2019 Jun 19];12(2):137–43. Available from: <http://www.nature.com/articles/ni.1979>
296. Kumar P, Sweeney TR, Skabkin MA, Skabkina O V., Hellen CUT, Pestova T V. 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* [Internet]. 2014 Mar 1 [cited 2018 Jun 29];42(5):3228–45. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24371270>
297. Daffis S, Szretter KJ, Schriewer J, Li J, Youn S, Errett J, et al. 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature*. 2010 Nov;468(7322):452–6.
298. Dong H, Chang DC, Hua MHC, Lim SP, Chionh YH, Hia F, et al. 2'-O methylation of

- internal adenosine by flavivirus NS5 methyltransferase. *PLoS Pathog* [Internet]. 2012 [cited 2019 Jun 19];8(4):e1002642. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22496660>
299. Coutard B, Barral K, Lichièrè J, Selisko B, Martin B, Aouadi W, et al. Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. *J Virol* [Internet]. 2017 Mar 1 [cited 2018 Jul 19];91(5):e02202-16. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28031359>
300. Ringeard M, Marchand V, Decroly E, Motorin Y, Bennasser Y. FTSJ3 is an RNA 2'-O-methyltransferase recruited by HIV to avoid innate immune sensing. *Nature* [Internet]. 2019 Jan 9 [cited 2019 May 23];565(7740):500–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/30626973>
301. Drake JW, Holland JJ. Mutation rates among RNA viruses. *Proc Natl Acad Sci U S A* [Internet]. 1999 Nov 23 [cited 2020 Apr 15];96(24):13910–3. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.96.24.13910>
302. Sanjuán R, Nebot MR, Chirico N, Mansky LM, Belshaw R. Viral mutation rates. *J Virol* [Internet]. 2010 Oct 1 [cited 2016 Apr 19];84(19):9733–48. Available from: <http://jvi.asm.org/content/84/19/9733.full>
303. Janeway CA, Medzhitov R. INNATE IMMUNE RECOGNITION. *Annu Rev Immunol* [Internet]. 2002 Apr 28 [cited 2018 Aug 20];20(1):197–216. Available from: <http://www.annualreviews.org/doi/10.1146/annurev.immunol.20.083001.084359>
304. Goubau D, Deddouche S, Reis e Sousa C. Cytosolic sensing of viruses. *Immunity* [Internet]. 2013 May 23 [cited 2020 Apr 15];38(5):855–69. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1074761313002045>
305. Yoneyama M, Onomoto K, Jogi M, Akaboshi T, Fujita T. Viral RNA detection by RIG-I-like receptors. *Curr Opin Immunol* [Internet]. 2015 Feb [cited 2020 Apr 15];32:48–53. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0952791514001605>
306. Yoneyama M, Fujita T. Recognition of viral nucleic acids in innate immunity. *Rev Med Virol* [Internet]. 2010 Jan [cited 2020 Apr 15];20(1):4–22. Available from: <http://doi.wiley.com/10.1002/rmv.633>

307. Takeuchi O, Akira S. Pattern recognition receptors and inflammation. *Cell* [Internet]. 2010 Mar 19 [cited 2020 Apr 15];140(6):805–20. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867410000231>
308. Kato H, Sato S, Yoneyama M, Yamamoto M, Uematsu S, Matsui K, et al. Cell type-specific involvement of RIG-I in antiviral response. *Immunity* [Internet]. 2005 Jul [cited 2020 Apr 21];23(1):19–28. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1074761305001421>
309. Luo D, Kohlway A, Pyle AM. Duplex RNA activated ATPases (DRAs): platforms for RNA sensing, signaling and processing. *RNA Biol* [Internet]. 2013 Jan 28 [cited 2020 Apr 16];10(1):111–20. Available from: <http://www.tandfonline.com/doi/abs/10.4161/rna.22706>
310. Hodel AE, Gershon PD, Quioco FA. Structural basis for sequence-nonspecific recognition of 5'-capped mRNA by a cap-modifying enzyme. *Mol Cell*. 1998 Feb;1(3):443–7.
311. Yoneyama M, Kikuchi M, Natsukawa T, Shinobu N, Imaizumi T, Miyagishi M, et al. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* [Internet]. 2004 Jul 20 [cited 2020 Apr 19];5(7):730–7. Available from: <http://www.nature.com/articles/ni1087>
312. Egloff M-P, Decroly E, Malet H, Selisko B, Benarroch D, Ferron F, et al. Structural and Functional Analysis of Methylation and 5'-RNA Sequence Requirements of Short Capped RNAs by the Methyltransferase Domain of Dengue Virus NS5. *J Mol Biol* [Internet]. 2007 Sep 21 [cited 2019 Jun 18];372(3):723–36. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17686489>
313. Rehwinkel J, Tan CP, Goubau D, Schulz O, Pichlmair A, Bier K, et al. RIG-I detects viral genomic RNA during negative-strand RNA virus infection. *Cell* [Internet]. 2010 Feb 5 [cited 2020 Apr 19];140(3):397–408. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867410000218>
314. Pichlmair A, Schulz O, Tan CP, Näslund TI, Liljeström P, Weber F, et al. RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science* [Internet]. 2006 Nov 10 [cited 2020 Apr 19];314(5801):997–1001. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1132998>

315. Hornung V, Ellegast J, Kim S, Brzozka K, Jung A, Kato H, et al. 5'-Triphosphate RNA Is the Ligand for RIG-I. *Science* (80-) [Internet]. 2006 Nov 10 [cited 2020 Apr 19];314(5801):994–7. Available from: <https://www.sciencemag.org/lookup/doi/10.1126/science.1132505>
316. Yoneyama M, Kikuchi M, Natsukawa T, Shinobu N, Imaizumi T, Miyagishi M, et al. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* [Internet]. 2004 Jul 20 [cited 2020 Apr 19];5(7):730–7. Available from: <http://www.nature.com/articles/ni1087>
317. Shatkin AJ. Capping of eucaryotic mRNAs. *Cell* [Internet]. 1976 Dec [cited 2020 Apr 16];9(4 PT 2):645–53. Available from: <https://linkinghub.elsevier.com/retrieve/pii/0092867476901288>
318. Pichlmair A, Schulz O, Tan C-P, Rehwinkel J, Kato H, Takeuchi O, et al. Activation of MDA5 Requires Higher-Order RNA Structures Generated during Virus Infection. *J Virol* [Internet]. 2009 Oct 15 [cited 2020 Apr 26];83(20):10761–9. Available from: <https://jvi.asm.org/content/83/20/10761>
319. Kato H, Takeuchi O, Mikamo-Satoh E, Hirai R, Kawai T, Matsushita K, et al. Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. *J Exp Med* [Internet]. 2008 Jul 7 [cited 2020 Apr 26];205(7):1601–10. Available from: <https://rupress.org/jem/article/205/7/1601/46724/Lengthdependent-recognition-of-doublestranded>
320. Li X, Lu C, Stewart M, Xu H, Strong RK, Igumenova T, et al. Structural basis of double-stranded RNA recognition by the RIG-I like receptor MDA5. *Arch Biochem Biophys* [Internet]. 2009 Aug [cited 2020 Apr 26];488(1):23–33. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0003986109001854>
321. Berke IC, Modis Y. MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding double-stranded RNA. *EMBO J* [Internet]. 2012 Apr 4 [cited 2020 Apr 26];31(7):1714–26. Available from: <http://emboj.emboPress.org/cgi/doi/10.1038/emboj.2012.19>
322. Childs KS, Andrejeva J, Randall RE, Goodbourn S. Mechanism of mda-5 Inhibition by paramyxovirus V proteins. *J Virol* [Internet]. 2009 Feb 1 [cited 2020 Apr

- 26];83(3):1465–73. Available from: <https://jvi.asm.org/content/83/3/1465>
323. Berke IC, Yu X, Modis Y, Egelman EH. MDA5 assembles into a polar helical filament on dsRNA. *Proc Natl Acad Sci [Internet]*. 2012 Nov 6 [cited 2020 Apr 26];109(45):18437–41. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.1212186109>
324. Hou F, Sun L, Zheng H, Skaug B, Jiang Q-X, Chen ZJ. MAVS Forms Functional Prion-like Aggregates to Activate and Propagate Antiviral Innate Immune Response. *Cell [Internet]*. 2011 Aug [cited 2020 Apr 26];146(3):448–61. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867411007161>
325. Rodriguez KR, Bruns AM, Horvath CM. MDA5 and LGP2: Accomplices and Antagonists of Antiviral Signal Transduction. *J Virol [Internet]*. 2014 Aug 1 [cited 2020 Apr 26];88(15):8194–200. Available from: <http://jvi.asm.org/cgi/doi/10.1128/JVI.00640-14>
326. Zhu Z, Zhang X, Wang G, Zheng H. The Laboratory of Genetics and Physiology 2: Emerging Insights into the Controversial Functions of This RIG-I-Like Receptor. *Biomed Res Int [Internet]*. 2014 [cited 2020 Apr 26];2014:1–7. Available from: <http://www.hindawi.com/journals/bmri/2014/960190/>
327. Childs KS, Randall RE, Goodbourn S. LGP2 Plays a Critical Role in Sensitizing mda-5 to Activation by Double-Stranded RNA. Mossman KL, editor. *PLoS One [Internet]*. 2013 May 9 [cited 2020 Apr 26];8(5):e64202. Available from: <https://dx.plos.org/10.1371/journal.pone.0064202>
328. Bruns AM, Pollpeter D, Hadizadeh N, Myong S, Marko JF, Horvath CM. ATP Hydrolysis Enhances RNA Recognition and Antiviral Signal Transduction by the Innate Immune Sensor, Laboratory of Genetics and Physiology 2 (LGP2). *J Biol Chem [Internet]*. 2013 Jan 11 [cited 2020 Apr 26];288(2):938–46. Available from: <http://www.jbc.org/lookup/doi/10.1074/jbc.M112.424416>
329. Liniger M, Summerfield A, Zimmer G, McCullough KC, Ruggli N. Chicken Cells Sense Influenza A Virus Infection through MDA5 and CARDIF Signaling Involving LGP2. *J Virol [Internet]*. 2012 Jan 15 [cited 2020 Apr 26];86(2):705–17. Available from: <http://jvi.asm.org/cgi/doi/10.1128/JVI.00742-11>
330. Uchikawa E, Lethier M, Malet H, Brunel J, Gerlier D, Cusack S. Structural Analysis of

- dsRNA Binding to Anti-viral Pattern Recognition Receptors LGP2 and MDA5. *Mol Cell* [Internet]. 2016 May [cited 2020 Apr 26];62(4):586–602. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1097276516300971>
331. Ahmad S, Hur S. Helicases in Antiviral Immunity: Dual Properties as Sensors and Effectors. *Trends Biochem Sci* [Internet]. 2015 Oct [cited 2020 Apr 26];40(10):576–85. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0968000415001413>
332. Dixit E, Boulant S, Zhang Y, Lee ASY, Odendall C, Shum B, et al. Peroxisomes Are Signaling Platforms for Antiviral Innate Immunity. *Cell* [Internet]. 2010 May [cited 2020 Apr 26];141(4):668–81. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867410004356>
333. Saito T, Hirai R, Loo Y-M, Owen D, Johnson CL, Sinha SC, et al. Regulation of innate antiviral defenses through a shared repressor domain in RIG-I and LGP2. *Proc Natl Acad Sci* [Internet]. 2007 Jan 9 [cited 2020 Apr 26];104(2):582–7. Available from: <http://www.pnas.org/cgi/doi/10.1073/pnas.0606699104>
334. Kawai T, Takahashi K, Sato S, Coban C, Kumar H, Kato H, et al. IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nat Immunol* [Internet]. 2005 Oct 28 [cited 2020 Apr 26];6(10):981–8. Available from: <http://www.nature.com/articles/ni1243>
335. Belgnaoui SM, Paz S, Hiscott J. Orchestrating the interferon antiviral response through the mitochondrial antiviral signaling (MAVS) adapter. *Curr Opin Immunol* [Internet]. 2011 Oct [cited 2020 Apr 26];23(5):564–72. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0952791511001038>
336. West AP, Shadel GS, Ghosh S. Mitochondria in innate immune responses. *Nat Rev Immunol* [Internet]. 2011 Jun 20 [cited 2020 Apr 26];11(6):389–402. Available from: <http://www.nature.com/articles/nri2975>
337. Saha SK, Pietras EM, He JQ, Kang JR, Liu S-Y, Oganessian G, et al. Regulation of antiviral responses by a direct and specific interaction between TRAF3 and Cardif. *EMBO J* [Internet]. 2006 Jul 26 [cited 2020 Apr 26];25(14):3257–63. Available from: <http://emboj.embopress.org/cgi/doi/10.1038/sj.emboj.7601220>
338. Seth RB, Sun L, Ea C-K, Chen ZJ. Identification and Characterization of MAVS, a Mitochondrial Antiviral Signaling Protein that Activates NF- κ B and IRF3. *Cell*

- [Internet]. 2005 Sep [cited 2020 Apr 26];122(5):669–82. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867405008160>
339. Kumar H, Kawai T, Akira S. Pathogen Recognition by the Innate Immune System. *Int Rev Immunol* [Internet]. 2011 Jan 14 [cited 2020 Apr 26];30(1):16–34. Available from: <http://www.tandfonline.com/doi/full/10.3109/08830185.2010.529976>
340. Melchjorsen J. Learning from the messengers: innate sensing of viruses and cytokine regulation of immunity - clues for treatments and vaccines. *Viruses* [Internet]. 2013 Jan 31 [cited 2020 Apr 22];5(2):470–527. Available from: <http://www.mdpi.com/1999-4915/5/2/470>
341. Gambin A, Charzyńska A, Ellert-Miklaszewska A, Rybiński M. Computational models of the JAK1/2-STAT1 signaling. *JAK-STAT* [Internet]. 2013 Jul 15 [cited 2020 Apr 22];2(3):e24672. Available from: <http://www.tandfonline.com/doi/abs/10.4161/jkst.24672>
342. Nguyen LH, Espert L, Mehti N, Wilson DM. The Human Interferon- and Estrogen-Regulated *ISG20/HEM45* Gene Product Degrades Single-Stranded RNA and DNA in Vitro †. *Biochemistry* [Internet]. 2001 Jun [cited 2020 Apr 22];40(24):7174–9. Available from: <https://pubs.acs.org/doi/10.1021/bi010141t>
343. Zheng Z, Wang L, Pan J. Interferon-stimulated gene 20-kDa protein (ISG20) in infection and disease: Review and outlook. *Intractable Rare Dis Res* [Internet]. 2017 [cited 2020 Apr 26];6(1):35–40. Available from: https://www.jstage.jst.go.jp/article/irdr/6/1/6_2017.01004/_article
344. Pichlmair A, Lassnig C, Eberle C-A, Górna MW, Baumann CL, Burkard TR, et al. IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. *Nat Immunol* [Internet]. 2011 Jul 5 [cited 2020 Apr 26];12(7):624–30. Available from: <http://www.nature.com/articles/ni.2048>
345. Schuberth-Wagner C, Ludwig J, Bruder AK, Herzner A-M, Zillinger T, Goldeck M, et al. A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2'O-Methylated Self RNA. *Immunity* [Internet]. 2015 Jul 21 [cited 2019 May 23];43(1):41–51. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1074761315002599>
346. Menachery VD, Yount BL, Josset L, Gralinski LE, Scobey T, Agnihothram S, et al.

- Attenuation and restoration of severe acute respiratory syndrome coronavirus mutant lacking 2'-o-methyltransferase activity. *J Virol* [Internet]. 2014 Apr [cited 2020 Apr 16];88(8):4251–64. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24478444>
347. Decroly E, Canard B. Biochemical principles and inhibitors to interfere with viral capping pathways. *Curr Opin Virol* [Internet]. 2017 Jun [cited 2020 Apr 25];24:87–96. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1879625717300196>
348. Ferron F, Decroly E, Selisko B, Canard B. The viral RNA capping machinery as a target for antiviral drugs. *Antiviral Res* [Internet]. 2012 Oct [cited 2020 Apr 25];96(1):21–31. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22841701>
349. Sievers F, Higgins DG. Clustal Omega. *Curr Protoc Bioinforma* [Internet]. 2014 Dec 12 [cited 2020 May 2];48(1):3.13.1-16. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25501942>
350. Gouy M, Guindon S, Gascuel O. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Mol Biol Evol* [Internet]. 2010 Feb 1 [cited 2020 May 2];27(2):221–4. Available from: <https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msp259>
351. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* [Internet]. 2014 Jul 1 [cited 2020 May 2];42(W1):W320–4. Available from: <http://academic.oup.com/nar/article/42/W1/W320/2435247/Deciphering-key-features-in-protein-structures>
352. Lieutaud P, Canard B, Longhi S. MeDor: a metaserver for predicting protein disorder. *BMC Genomics* [Internet]. 2008 [cited 2020 May 2];9(Suppl 2):S25. Available from: <http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-9-S2-S25>
353. Esposito D, Chatterjee DK. Enhancement of soluble protein expression through the use of fusion tags. *Curr Opin Biotechnol* [Internet]. 2006 Aug [cited 2020 May 2];17(4):353–8. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0958166906000875>
354. Turchetto J, Sequeira AF, Ramond L, Peysson F, Brás JLA, Saez NJ, et al. High-throughput expression of animal venom toxins in *Escherichia coli* to generate a large library of oxidized disulphide-reticulated peptides for drug discovery. *Microb Cell Fact*

- [Internet]. 2017 Dec 17 [cited 2020 May 2];16(1):6. Available from: <http://microbialcellfactories.biomedcentral.com/articles/10.1186/s12934-016-0617-1>
355. Chhetri G, Kalita P, Tripathi T. An efficient protocol to enhance recombinant protein expression using ethanol in *Escherichia coli*. *MethodsX* [Internet]. 2015 [cited 2020 May 2];2:385–91. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S2215016115000485>
356. Nishihara K, Kanemori M, Kitagawa M, Yanagi H, Yura T. Chaperone Coexpression Plasmids: Differential and Synergistic Roles of DnaK-DnaJ-GrpE and GroEL-GroES in Assisting Folding of an Allergen of Japanese Cedar Pollen, Cryj2, in *Escherichia coli*. *Appl Environ Microbiol* [Internet]. 1998 [cited 2020 May 2];64(5):1694. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC106217/>
357. Ciccarone VC, Polayes DA, Luckow VA. Generation of Recombinant Baculovirus DNA in *E. coli* Using a Baculovirus Shuttle Vector. In: *Molecular Diagnosis of Infectious Diseases* [Internet]. New Jersey: Humana Press; [cited 2020 May 2]. p. 213–36. Available from: <http://link.springer.com/10.1385/0-89603-485-2:213>
358. Reece-Hoyes JS, Walhout AJM. Gateway Recombinational Cloning. *Cold Spring Harb Protoc* [Internet]. 2018 Jan [cited 2020 May 2];2018(1):pdb.top094912. Available from: <http://www.cshprotocols.org/lookup/doi/10.1101/pdb.top094912>
359. Lim SP, Sonntag LS, Noble C, Nilar SH, Ng RH, Zou G, et al. Small Molecule Inhibitors That Selectively Block Dengue Virus Methyltransferase. *J Biol Chem* [Internet]. 2011 Feb 25 [cited 2020 May 2];286(8):6233–40. Available from: <http://www.jbc.org/lookup/doi/10.1074/jbc.M110.179184>
360. Benmansour F, Trist I, Coutard B, Decroly E, Querat G, Brancale A, et al. Discovery of novel dengue virus NS5 methyltransferase non-nucleoside inhibitors by fragment-based drug design. *Eur J Med Chem* [Internet]. 2017 Jan [cited 2020 May 2];125:865–80. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0223523416308698>
361. Coloma J, Jain R, Rajashankar KR, García-Sastre A, Aggarwal AK. Structures of NS5 Methyltransferase from Zika Virus. *Cell Rep* [Internet]. 2016 Sep 20 [cited 2020 May 2];16(12):3097–102. Available from: [https://www.cell.com/cell-reports/fulltext/S2211-1247\(16\)31200-1?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS2](https://www.cell.com/cell-reports/fulltext/S2211-1247(16)31200-1?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS2)

- 211124716312001%3Fshowall%3Dtrue
362. Egloff M-P, Benarroch D, Selisko B, Romette J-L, Canard B. An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. *EMBO J* [Internet]. 2002 Jun 3 [cited 2019 May 23];21(11):2757–68. Available from: <http://emboj.embopress.org/cgi/doi/10.1093/emboj/21.11.2757>
363. Dong H, Chang DC, Hua MHC, Lim SP, Chionh YH, Hia F, et al. 2'-O methylation of internal adenosine by flavivirus NS5 methyltransferase. Kuhn RJ, editor. *PLoS Pathog* [Internet]. 2012 Apr 5 [cited 2019 Jun 19];8(4):e1002642. Available from: <https://dx.plos.org/10.1371/journal.ppat.1002642>
364. Liang B, Li Z, Jenni S, Rahmeh AA, Morin BM, Grant T, et al. Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy. *Cell* [Internet]. 2015 Jul [cited 2019 May 23];162(2):314–27. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0092867415007023>
365. Dai Q, Moshitch-Moshkovitz S, Han D, Kol N, Amariglio N, Rechavi G, et al. Nm-seq maps 2'-O-methylation sites in human mRNA with base precision. *Nat Methods* [Internet]. 2017 Jul 15 [cited 2020 May 9];14(7):695–8. Available from: <http://www.nature.com/articles/nmeth.4294>
366. Morin B, Whelan SPJ. Sensitivity of the polymerase of vesicular stomatitis virus to 2' substitutions in the template and nucleotide triphosphate during initiation and elongation. *J Biol Chem* [Internet]. 2014 Apr 4 [cited 2019 Jun 19];289(14):9961–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24526687>

ANNEXES

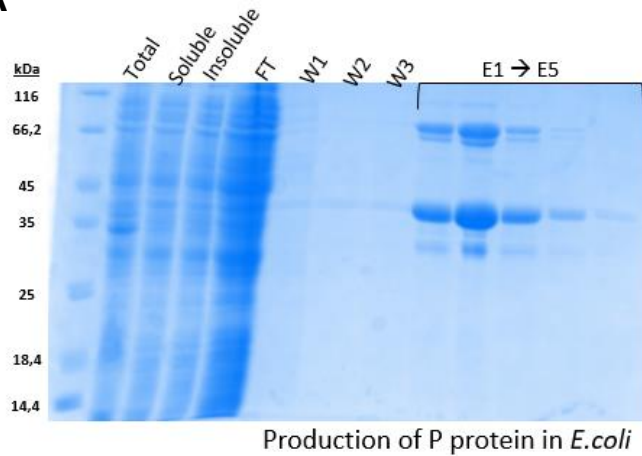
VSV	1380	1390	1400	1410	1420	1430	1440	1450	1460	1470	1480																																																																																												
RABV	LSGRE	SRHIGS	AGILYS	ILVA	HD	SGYN	DS	IF	FPVN	ISK	VP	RD	YREGL	AR	GV	LV	GS	IC	FL	RM	NI	NP	LE	IS	CV	YI	LLR	L	DN	HP	SL	YI	M	IR	EP	SL																																																																			
VSV	LAPAE	S	YOV	GR	CI	E	LY	GD	LY	GD	L	AYR	K	S	HA	E	D	S	L	F	FL	S	D	GR	I	R	GR	GE	GL	LD	GL	MR	AS	CC	OV	I	R	K	S	L	A	H	L	R	S	L	A	H	L	R	P	AN	AV	Y	GL	L	D	L	R	S	V	P	P	F	L	S	L	R	S	G	P																																
VSV	1490	1500	1510	1520	1530	1540	1550	1560	1570	1580	1590																																																																																												
RABV	RGE	IF	TP	OK	IP	AK	YPT	IM	KE	GR	NS	IL	Q	H	VR	Y	R	E	V	T	A	S	P	E	N	D	IL	W	I	F	S	D	F	R	S	K	M	T	Y	L	L	I	Y	O	S	H	L	L	O	R	V	E	R	N	L	S	K	S	M	R	T	L	R	O	M	S	S	L	M	R	O	V	L	G	H	C	H	E																									
VSV	RDE	LE	T	F	H	K	I	P	T	S	N	R	D	M	G	V	I	R	N	F	K	Y	O	C	R	L	I	E	K	G	R	S	H	Y	S	O	L	W	F	S	D	V	L	S	I	D	F	I	G	F	S	I	T	I	L	L	O	L	L	Y	K	F	F	L	S	G	K	D	K	N	L	R	E	L	A	N	L	S	L	L	S	L	S	S	G	E	G																
VSV	1590	1600	1610	1620	1630	1640	1650	1660	1670	1680	1690																																																																																												
RABV	DT	LE	S	D	D	I	O	R	L	K	D	S	L	R	R	T	R	W	D	C	E	V	R	H	A	R	T	M	T	G	G	S	E	N	K	V	S	R	K	A	G	C	S	E	W	V	C	S	A	Q	O	V	A	S	T	S	A	N	A	P	V	S	E	L	D	I	R	A	L	S	K	K	E	O	N	P	L	I	S	G	R	V	V	Q	W	A	T	G	A	R	K	L	K	P									
VSV	E	...	D	I	H	V	K	F	T	K	D	L	...	C	P	E	E	I	R	H	A	C	K	F	G	I	A	K	D	N	K	N	K	D	M	S	Y	P	P	R	G	R	E	S	R	G	T	I	T	I	P	V	Y	T	I	P	Y	...	K	M	L	E	M	P	R	I	O	N	P	L	I	S	G	H	R	L	G	O	L	P	T	G	A	H	K	I	R																
VSV	1700	1710	1720	1730	1740	1750	1760	1770	1780	1790	1800																																																																																												
RABV	IL	D	L	N	V	P	S	L	C	L	V	V	G	D	S	G	G	G	G	G	G	G	I	S	R	A	V	L	N	M	F	P	D	A	K	L	V	F	N	S	L	L	E	V	N	D	L	M	A	S	G	T	H	P	I	P	P	S	A	L	M	S	G	G	D	I	I	S	R	V	V	D	E	S	I	M	E	R	P	S	D	L	R	N	L	T	T	W	R	Y	F	O	S	V	K	Q	V	N	N	S	Y	D	L
VSV	I	L	H	G	M	G	I	H	R	D	F	L	S	C	D	G	S	G	G	M	T	A	L	L	R	E	N	V	H	R	G	I	F	N	S	L	L	E	L	S	G	S	V	M	R	G	A	S	E	P	P	S	A	E	T	I	G	G	K	.S	R	C	V	N	G	E	I	C	W	E	Y	P	S	D	L	C	D	P	R	T	W	D	Y	F	R	L	K	A	G	L	G	L	O	I									
VSV	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900																																																																																												
RABV	IC	DA	E	V	D	I	A	S	I	N	H	I	L	L	M	S	D	F	A	L	S	H	.D	G	P	L	V	F	K	T	Y	G	T	M	L	V	N	P	E	Y	R	A	I	Q	H	L	S	R	A	F	S	V	T	G	F	I	T	O	M	T	S	S	F	S	E	F	L	R	F	S	K	R	G	F	R	D	A	E	Y	L	T	S	T	L	R	E	M	S	I	V	L	F	N										
VSV	V	M	D	E	V	E	D	S	S	T	S	L	K	E	T	N	V	R	N	Y	V	H	R	L	D	E	O	G	L	Y	K	T	Y	G	T	V	I	C	E	S	E	K	N	A	T	L	G	P	M	E	K	T	V	D	L	V	O	T	E	F	S	S	O	T	S	.P	Y	M	.V	C	K	L	I	D	E	P	N	D	M	S	I	N	E	S	K	N	L	I	A														
VSV	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010																																																																																												
RABV	CS	SP	K	S	E	M	O	R	A	R	S	L	N	Y	Q	D	L	V	R	Q	E	P	E	E	I	I	S	N	P	N	E	M	I	I	T	L	D	S	D	V	E	S	F	L	V	H	K	M	V	D	D	L	E	Q	R	G	T	L	S	K	N	S	I	I	A	I	M	I	V	F	S	N	R	V	F	N	V	S	K	P	L	T	D	P	L	F	P	P	S	D	P	K	I	L	R	H	F	N					
VSV	F	O	S	S	E	E	F	A	R	A	K	K	V	S	T	F	L	T	G	P	S	O	F	I	P	D	E	V	N	I	E	T	M	L	O	I	F	G	V	P	T	G	V	S	H	A	A	L	K	S	S	R	P	A	D	L	L	I	S	L	F	Y	M	A	I	S	Y	.Y	N	I	N	H	I	R	V	G	P	I	P	N	P	S	D	G	I	L	Q	N	V	G													
VSV	2010	2020	2030	2040	2050	2060	2070	2080	2090	2100	2110																																																																																												
RABV	CC	S	T	M	Y	.L	S	M	A	L	G	V	P	S	F	A	R	L	H	D	L	Y	N	R	P	I	T	Y	F	K	Q	V	I	R	C	N	I	L	S	M	S	D	S	T	S	V	F	K	R	V	A	C	N	S	L	S	S	H	W	I	R	L	I	K	I	V	K	T	T	R	F	I	G	S	I	E	D	L	P	G	E	V	A	R	H	L	Q	G	Y	N	R	W											
VSV	A	I	T	G	I	F	W	L	S	L	M	E	K	D	I	P	L	Y	Q	C	L	A	V	I	Q	S	F	P	R	E	A	V	S	R	G	G	K	Q	K	W	S	I	R	G	D	.G	L	P	K	O	T	R	I	S	D	L	A	P	I	G	N	W	I	R	S	L	E	L	V	R	N	Q	V	R	N	P	F	N	E	L	F	N	Q	L	C	R	T	V	D	N	H	L	K	W									
VSV	2100	2110	2120																																																																																																				
RABV	T	L																																																																																												
VSV	N	L	R	R	T	G	M	I	E	N	I	S	K	E	D	S	L	M	L	K	S	D	L	H	E	N	S	W	R	D																																																																									

Annexe 1 : Alignement séquences des domaines C-terminaux de VSV et RABV

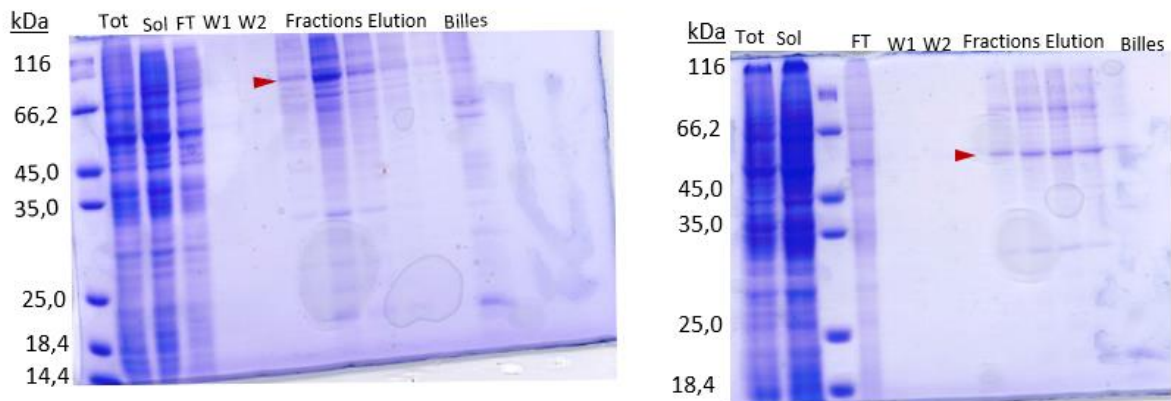
L'extrémité C-terminale de la protéine L de VSV et RABV est constituée du domaine CD (en jaune), du domaine MTase (en orange) et du domaine CTD (en rouge). Le motif conservé GxGxG de fixation du SAM est indiqué en jaune et les résidus de la tétrade catalytique KDKE sont surlignés en violet dans l'alignement. Les structures secondaires au-dessus de l'alignement correspondent aux structures secondaires de VSV dont la structure est résolue.

ANNEXE 2

A



B



Annexe 2 : Analyse de l'expression de protéines recombinantes

(A) Gel SDS-PAGE (12%) coloré au bleu de Coomassie de la purification par chromatographie d'affinité (Strep-tag/Strep-tactin chromatography) de la protéine P du virus de la rage. La protéine P est retrouvée sous forme de dimère en solution. (B) Gels SDS-PAGE (10%) colorés au bleu de Coomassie de la purification du domaine CD-MTase-CTD (à gauche) et MTase-CTD (à droite) de RABV exprimé en cellules d'insectes.

ANNEXE 3

Fw_1378	GAAAACCTGTA CTCCAGGGTAGCCGTCATATTGGTAGCGCACAGGG
Fw_1581	GAAAACCTGTA CTCCAGGGTCATGGTGAAGATACCCTGGAATCAGATGATGATA
Fw_1629	GAAAACCTGTA CTCCAGGGTCGTA AAGCAGGTTGTAGCGAATGGGTTTGC
Fw_1639	GAAAACCTGTA CTCCAGGGTAGCGCACAGCAGGTTGCAGTTAGCAC
Fw_1929	GAAAACCTGTA CTCCAGGGTTCCAATCCGTATAACGAAATGATCATCACCCCTGA
Rv_1579	GGGACCAC TTTGTACAAGAAAGCTGGGTCTTATTAACCTAACACCTGACGCATCAGGCTGC
Rv_1628	GGGACCAC TTTGTACAAGAAAGCTGGGTCTTATTAGCTCACTTTTTTGTTCGGGCTATAACCACCG
Rv_1915	GGGACCAC TTTGTACAAGAAAGCTGGGTCTTATTAATTCAGGCTACGTGCACGCTGCATTTTCG
Rv_1921	GGGACCAC TTTGTACAAGAAAGCTGGGTCTTATTAACGAACCAGATCCTGATAATTCAGGCTACGTG
Rv_2126	GGGACCAC TTTGTACAAGAAAGCTGGGTCTTATTAGCTATAATCCAGCAGACTGCTACGGCTAC
Fw-attB1TEV	GGGACAAG TTTGTACAAAAAAGCAGGCTTAGAAAACCTGTA CTCCAGGGT

Tableau des amorces utilisées pour les PCR d'assemblage

ANNEXE 4

AUTRES PROJETS



The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade

B. Coutard^a, C. Valle^b, X. de Lamballerie^a, B. Canard^b, N.G. Seidah^c, E. Decroly^{b,*}

^a Unité des Virus Émergents (UVE: Aix-Marseille Univ – IRD 190 – Inserm 1207 – IHU Méditerranée Infection), Marseille, France

^b Aix Marseille Université, CNRS, AFMB UMR 7257, Marseille, France

^c Laboratory of Biochemical Neuroendocrinology, Montreal Clinical Research Institute (IRCM, Affiliated to the University of Montreal), 110 Pine Ave West, Montreal, QC, H2W1R7, Canada

ARTICLE INFO

Keywords:

2019-nCoV
SARS-CoV
Spike protein
Maturation protease
Furin
Antivirals

ABSTRACT

In 2019, a new coronavirus (2019-nCoV) infecting Humans has emerged in Wuhan, China. Its genome has been sequenced and the genomic information promptly released. Despite a high similarity with the genome sequence of SARS-CoV and SARS-like CoVs, we identified a peculiar furin-like cleavage site in the Spike protein of the 2019-nCoV, lacking in the other SARS-like CoVs. In this article, we discuss the possible functional consequences of this cleavage site in the viral cycle, pathogenicity and its potential implication in the development of antivirals.

Human coronaviruses (CoV) are enveloped positive-stranded RNA viruses belonging to the order *Nidovirales*, and are mostly responsible for upper respiratory and digestive tract infections. Among them SARS-CoV and MERS-CoV that spread in 2002 and 2013 respectively, have been associated with severe human illnesses, such as severe pneumonia and bronchiolitis, and even meningitis in more vulnerable populations (de Wit et al., 2016). In December 2019, a new CoV (2019-nCoV) has been detected in the city of Wuhan, and this emerging viral infection was associated with severe human respiratory disease with a ~2–3% fatality rate (Li et al., 2020). The virus that was presumed to have initially been transmitted from an animal reservoir to humans possibly via an amplifying host. However human-to-human transmission has been reported, leading to a sustained epidemic spread with > 31,000 confirmed human infections, including > 640 deaths, reported by the WHO in early February 2020. The estimated effective reproductive number (R) value of ~2.90 (95%: 2.32–3.63) at the beginning of the outbreak raises the possibility of a pandemic (Zhao et al., 2020). This prompted WHO to declare it as a Public Health Emergency of International Concern. This is especially relevant because so far there are no specific antiviral treatments available or vaccine. Based on its genome sequence, 2019-nCoV belongs to lineage b of Betacoronavirus (Fig. 1A), which also includes the SARS-CoV and bat CoV ZXC21, the latter and CoV ZC45 being the closest to 2019-nCoV. 2019-nCoV shares ~76% amino acid sequence identity in the Spike (S)-protein sequence with SARS-CoV and 80% with CoV ZXC21 (Chan et al., 2020). In this article,

we focus on a specific furin-like protease recognition pattern present in the vicinity of one of the maturation sites of the S protein (Fig. 1B) that may have significant functional implications for virus entry.

The proprotein convertases (PCs; genes PCSKs) constitute a family of nine serine secretory proteases that regulate various biological processes in both healthy and disease states (Seidah and Prat, 2012). By proteolysis, PCs are responsible for the activation of a wide variety of precursor proteins, such as growth factors, hormones, receptors and adhesion molecules, as well as cell surface glycoproteins of infectious viruses (Seidah and Chretien, 1999) (Table 1). Seven PCs cleave precursor proteins at specific single or paired basic amino acids (aa) within the motif (R/K)-(2X)n-(R/K)↓, where n = 0, 1, 2, or 3 spacer aa (Seidah and Chretien, 1999). Because of their role in the processing of many critical cell surface proteins PCs, especially furin, have been implicated in viral infections. They have the potential to cleave specifically viral envelope glycoproteins, thereby enhancing viral fusion with host cell membranes (Izaguirre, 2019; Moulard and Decroly, 2000). In the case of human-infecting coronaviruses such as HCoV-OC43 (Le Coupanec et al., 2015), MERS-CoV (Millet and Whittaker, 2014), and HKU1 (Chan et al., 2008) the spike protein has been demonstrated to be cleaved at an S1/S2 cleavage site (Fig. 2) generating the S1 and S2 subunits. The above three viruses display the canonical (R/K)-(2X)n-(R/K)↓ motif (Table 1). Additionally, it has been demonstrated that variation around the viral envelope glycoprotein cleavage site plays a role in cellular tropism and pathogenesis. For instance, the pathogenesis of some CoV

* Corresponding author.

E-mail address: etienne.decroly@afmb.univ-mrs.fr (E. Decroly).

<https://doi.org/10.1016/j.antiviral.2020.104742>

Received 3 February 2020; Received in revised form 7 February 2020; Accepted 8 February 2020

Available online 10 February 2020

0166-3542/ © 2020 Elsevier B.V. All rights reserved.

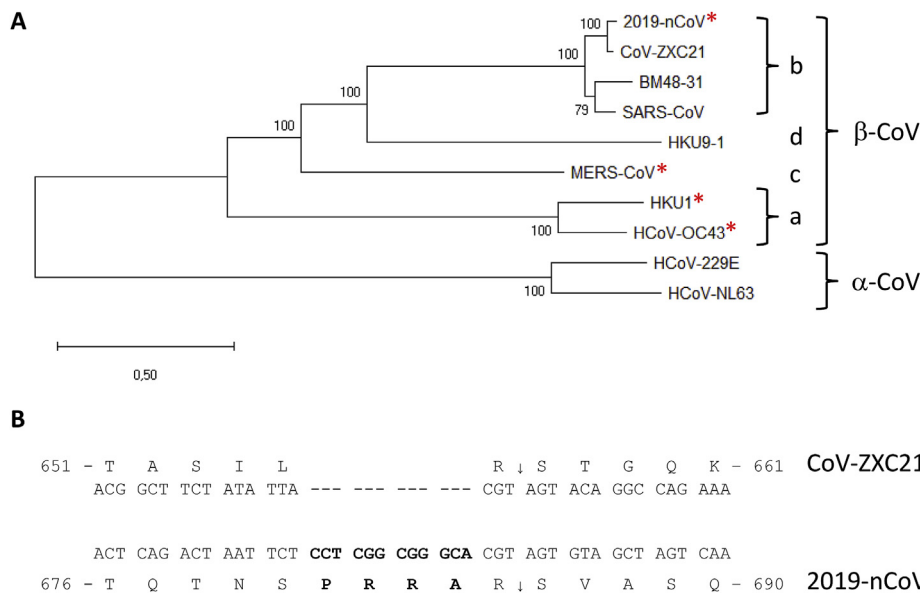


Fig. 1. Characterization of an nCoV-peculiar sequence at the S1/S2 cleavage site in the S-protein sequence, compared SARS-like CoV. (A) Phylogenetic tree of selected coronaviruses from genera alphacoronavirus (α -CoV) and betacoronavirus (β -CoV), lineages a, b, c and d: 2019-nCoV (NC_045512.2), CoV-ZXC21 (MG772934), SARS-CoV (NC_004718.3), SARS-like BM4821 (MG772934), HCoV-OC43 (AY391777), HKU9-1 (EF065513), HCoV-NL63 (KF530114.1), HCoV229E (KF514433.1), MERS-CoV (NC019843.3), HKU1 (NC_006577.2). The phylogenetic tree was obtained on the Orf1ab amino acid sequence using the Maximum Likelihood method by Mega X software. Red asterisks indicate the presence of a canonical furin-like cleavage motif at site 1; (B) Alignment of the coding and amino acid sequences of the S-protein from CoV-ZXC21 and 2019-nCoV at the S1/S2 site. The 2019-nCoV-specific sequence is in bold. The sequence of CoV-ZXC21 S-protein at this position is representative of the sequence of the other betacoronaviruses belonging to lineage b, except the one of 2019-nCoV. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 1

Comparative sequences of envelope protein cleavage site(s) in coronaviruses (above) and in other RNA viruses (below). Empty boxes: no consensus motif detected..

Coronavirus	S1/S2, site 1	S1/S2, site 2	S2'
2019-nCoV	SPRRAR↓SVAS	IAY↓TMS	SKPSKR↓SF
CoV-ZX21	TASILR↓STGQ	IAY↓TMS	SKPSKR↓SF
Bat-AC45	TASILR↓STGQ	IAY↓TMS	SKPSKR↓SF
SARS-CoV	TVSLLR↓STGQ	IAY↓TMS	LKPTKR↓SF
BM48-31	SSTLVR↓SGGH	LAY↓TMS	LKPTKR↓SF
HKU9-1	ADSLPR↓LQLV	VNY↓DPL	GATTYR↓SA
MERS-CoV	TPRSCR↓SVPG		GSR SAR↓SA
HKU1	SRKRKR↓SISA		CGSSSR↓SF
HCoV-OC43	KNRRSR↓GAITT		SKASSR↓SA
HCoV-229E	IAVQPR↓NVSVD		SRVAGR↓SA
HCoV-NL63	IPVRPR↓NSSDN		SRIAGR↓SA

Virus	Envelope Protein	Cleavage site
HIV	Gp160	VQREKR↓AV
Influenza Virus H5	HA	RKRKR↓GL
Avian H5N1 A/HK/98	HA	REKRKR↓GL
Avian H5N1 TKY/ENG	HA	NTPQKR↓GL
Human CMV	gB	HKRTR↓ST
Human RSV	F protein	KRKRKR↓FL
Yellow Fever Virus	PrM	SRRSR↓AI
Zika Virus	PrM	ARRSR↓AV
Ebola virus	GP	GRTRR↓EA
2019-nCoV (S1/S2) site	Spike Protein	TNSPRRAR↓SV
2019-nCoV (S2') site	Spike Protein	SKPSKR↓SF

has been previously related to the presence of a furin-like cleavage site in the S-protein sequence. For example, the insertion of a similar cleavage site in the infectious bronchitis virus (IBV) S-protein results in higher pathogenicity, pronounced neural symptoms and neurotropism in infected chickens (Cheng et al., 2019).

Similarly, in the case of influenza virus, low-pathogenicity forms of influenza virus contain a single basic residue at the cleavage site, which is cleaved by trypsin-like proteases and the tissue distribution of the activating protease(s) typically restricts infections to the respiratory and/or intestinal organs (Sun et al., 2010). Conversely, the highly pathogenic forms of influenza have a furin-like cleavage site cleaved by different cellular proteases, including furin, which are expressed in a wide variety of cell types allowing a widening of the cell tropism of the virus (Kido et al., 2012). Furthermore the insertion of a multibasic motif

RERRRKR↓GL at the H5N1 hemagglutinin HA cleavage site was likely associated with the hyper-virulence of the virus during the Hong Kong 1997 outbreak (Claas et al., 1998). This motif exhibits the critical Arg at P1 and basic residues at P2 and P4, as well as P6 and P8 and an aliphatic Leu at P2' positions (Table 1) (Schechter and Berger nomenclature (Schechter and Berger, 1968)), typical of a furin-like cleavage specificity (Braun and Sauter, 2019; Izaguirre, 2019; Seidah and Prat, 2012).

The coronavirus S-protein is the structural protein responsible for the crown-like shape of the CoV viral particles, from which the original name "coronavirus" was coined. The ~1200 aa long S-protein belongs to class-I viral fusion proteins and contributes to the cell receptor binding, tissue tropism and pathogenesis (Lu et al., 2015; Millet and Whittaker, 2014). It contains several conserved domains and motifs

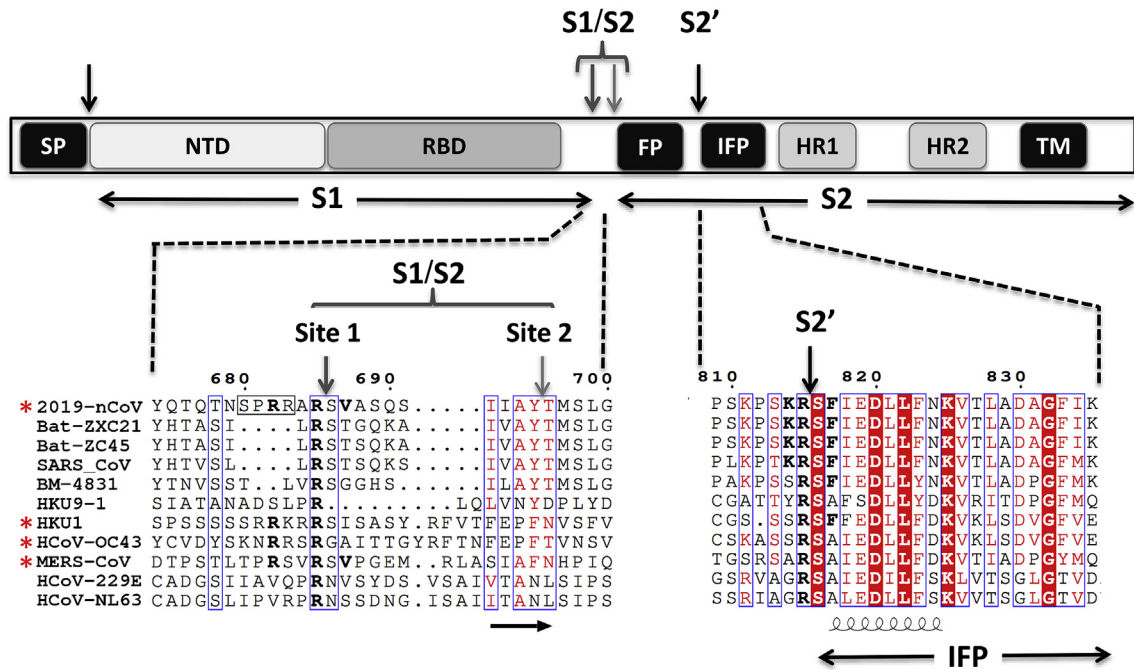


Fig. 2. Schematic representation of the human 2019-nCoV S-protein with a focus on the putative maturation sites. The domains were previously characterized in SARS-CoV and MERS-CoV: Signal peptide (SP), N-terminal domain (NTD), receptor-binding domain (RBD), fusion peptide (FP), internal fusion peptide (IFP), heptad repeat 1/2 (HR1/2), and the transmembrane domain (TM). The SP, S1/S2 and S2' cleavage sites are indicated by arrows. The sequence of different CoV S1/S2 and S2' cleavage sites were aligned using Multalin webserver (<http://multalin.toulouse.inra.fr/multalin/>) with manual adjustments and the figure prepared using ESPript 3 (<http://esprict.ibcp.fr/ESPript/ESPript/>) presenting the secondary structure of SARS-CoV S-protein at the bottom of the alignment (PDB 5X58) (Yuan et al., 2017). Insertion of furin like cleavage site is surrounded by a black frame. Red asterisks indicate the presence of a canonical furin-like cleavage motif at the S1/S2 site. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(Fig. 2). The trimetric S-protein is processed at the S1/S2 cleavage site by host cell proteases, during infection. Following cleavage, also known as priming, the protein is divided into an N-terminal S1-ectodomain that recognises a cognate cell surface receptor and a C-terminal S2-membrane-anchored protein involved in viral entry. The SARS-CoV S1-protein contains a conserved Receptor Binding Domain (RBD), which recognises the angiotensin-converting enzyme 2 (ACE2) (Li et al., 2003). The SARS-CoV binds to both bat and human cells, and the virus can infect both organisms (Ge et al., 2013; Kuhn et al., 2004). The RBD surface of S1/ACE2 implicates 14 aa in the S1 of SARS-CoV (Li et al., 2005). Among them, 8 residues are strictly conserved in 2019-nCoV, supporting the hypothesis that ACE2 is also the receptor of the newly emerged nCoV (Wan et al., 2020). The S2-protein contains the fusion peptide (FP), a second proteolytic site (S2'), followed by an internal fusion peptide (IFP) and two heptad-repeat domains preceding the transmembrane domain (TM) (Fig. 2). Notably, the IFPs of the 2019-nCoV and SARS-CoV are identical, displaying characteristics of viral fusion peptides (Fig. 2). While the molecular mechanism involved in cell entry is not yet fully understood, it is likely that both FP and IFP participate in the viral entry process (Lu et al., 2015) and thus the S-protein must likely be cleaved at both S1/S2 and S2' cleavage sites for virus entry. The furin-like S2' cleavage site at KR↓SF with P1 and P2 basic residues and a P2' hydrophobic Phe (Seidah and Prat, 2012), downstream of the IFP is identical between the 2019-nCoV and SARS-CoV (Fig. 2). In the MERS-CoV and HCoV-OC43 the S1/S2 site is replaced by RXXR↓SA, with P1 and P4 basic residues, and an Ala (not aliphatic) at P2', suggesting a somewhat less favourable cleavage by furin. However, in the other less pathogenic circulating human CoV, the S2' cleavage site only exhibits a monobasic R↓S sequence (Fig. 2) with no basic residues at either P2 and/or P4 needed to allow furin cleavage, suggesting a less efficient cleavage or higher restriction at the entry step depending on the cognate proteases expressed by target cells. Even though processing at S2' in 2019-nCoV is expected to be a key event for

the final activation of the S-protein, the protease(s) involved in this process have not yet been conclusively identified. Based on the 2019-nCoV S2' sequence and the above arguments, we propose that one or more furin-like enzymes would cleave the S2' site at KR↓SF. In contrast to the S2', the first cleavage between the RBD and the FP (S1/S2 cleavage site, Fig. 2) has been extensively studied for many CoVs (Lu et al., 2015). Interestingly the S1/S2 processing site exhibits different motifs among coronaviruses (Fig. 2, site 1 & site 2), with many of them displaying cleavage after a basic residue. It is thus likely that the priming process is ensured by different host cell proteases depending on the sequence of the S1/S2 cleavage site. Accordingly the MERS-CoV S-protein, which contains a RSVR↓SV motif is cleaved during virus egress, probably by furin (Mille and Whittaker, 2014). Conversely the S-protein of SARS-CoV remains largely uncleaved after biosynthesis, possibly due to the lack of a favourable furin-like cleavage site (SLLR-ST). In this case, it was reported that following receptor binding the S-protein is cleaved at a conserved sequence AYT↓M (located 10 aa downstream of SLLR-ST) by target cells' proteases such as elastase, cathepsin L or TMPRSS2 (Bosch et al., 2008; Matsuyama et al., 2010, 2005; Millet and Whittaker, 2015). As the priming event is essential for virus entry, the efficacy and extent of this activation step by the proteases of the target cells should regulate cellular tropism and viral pathogenesis. In the case of the 2019-nCoV S-protein, the conserved site 2 sequence AYT↓M may still be cleaved, possibly after the preferred furin-cleavage at the site 1 (Fig. 2).

Since furin is highly expressed in lungs, an enveloped virus that infects the respiratory tract may successfully exploit this convertase to activate its surface glycoprotein (Bassi et al., 2017; Mbikay et al., 1997). Before the emergence of the 2019-nCoV, this important feature was not observed in the lineage b of betacoronaviruses. However, it is shared by other CoV (HCoV-OC43, MERS-CoV, MHV-A59) harbouring furin-like cleavage sites in their S-protein (Fig. 2; Table 1), which were shown to be processed by furin experimentally (Le Coupanec et al.,

2015; Mille and Whittaker, 2014). Strikingly, the 2019-nCoV S-protein sequence contains 12 additional nucleotides upstream of the single Arg↓ cleavage site 1 (Figs. 1B and 2) leading to a predictively solvent-exposed PRRAR↓SV sequence, which corresponds to a canonical furin-like cleavage site (Braun and Sauter, 2019; Izaguirre, 2019; Seidah and Prat, 2012). This furin-like cleavage site, is supposed to be cleaved during virus egress (Mille and Whittaker, 2014) for S-protein “priming” and may provide a gain-of-function to the 2019-nCoV for efficient spreading in the human population compared to other lineage b betacoronaviruses. This possibly illustrates a convergent evolution pathway between unrelated CoVs. Interestingly, if this site is not processed, the S-protein is expected to be cleaved at site 2 during virus endocytosis, as observed for the SARS-CoV.

Obviously much more work is needed to demonstrate experimentally our assertion, but the inhibition of such processing enzyme(s) may represent a potential antiviral strategy. Indeed, it was recently shown that in an effort to limit viral infections, host cells that are infected by a number of viruses provoke an interferon response to inhibit the enzymatic activity of furin-like enzymes. It was also demonstrated that HIV infection induces the expression of either the protease activated receptor 1 (PAR1) (Kim et al., 2015) or guanylate binding proteins 2 and 5 (GBP2,5) (Braun and Sauter, 2019) that restrict the trafficking of furin to the *trans* Golgi network (PAR1) or to early Golgi compartments (GBP2,5) where the proprotein convertase remains inactive. Altogether, these observations suggest that inhibitors of furin-like enzymes may contribute to inhibiting virus propagation.

A variety of approaches have been proposed to inhibit furin activity to limit tumour growth, viral and bacterial infection. Thus, a variant of the naturally occurring serine protease inhibitor α -1 antitrypsin harbouring a consensus furin cleavage, called α -1 antitrypsin Portland (α 1-PDX), inhibits furin and prevents the processing of HIV-1 Env (Anderson et al., 1993). The addition of a chloromethylketone (CMK) moiety to the C-terminus of a polybasic cleavage motif and a decanoyl group at the N-terminus to favour cell penetration (dec-RVKR-cmk) irreversibly blocked the enzymatic activity of furin, PC7, PC5, PACE4 and PC7 (Decroly et al., 1996; Garten et al., 1994). Finally, the elucidation of the crystal structure of furin resulted in the design of a 2,5-dideoxystreptamine-derived inhibitor, where two molecules of the inhibitor form a complex with furin (Dahms et al., 2017). As furin-like enzymes are involved in a multitude of cellular processes, one important issue would be to avoid systemic inhibition that may result in some toxicity. Accordingly, it is likely that such small molecule inhibitors, or other more potent orally active ones, possibly delivered by inhalation and exhibiting a slow dissociation rate from furin to allow for sustained inhibition, deserve to be rapidly tested to assess their antiviral effect against 2019-nCoV.

Acknowledgments

This work was supported by a CIHR Foundation grant # 148363 (NGS), a Canada Research Chairs in Precursor Proteolysis (NGS; # 950-231335), and by the European Virus Archive Global (BCo; EVA GLOBAL) funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 871029.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.antiviral.2020.104742>.

References

Anderson, E.D., Thomas, L., Hayflick, J.S., Thomas, G., 1993. Inhibition of HIV-1 gp160-dependent membrane fusion by a furin-directed α 1-antitrypsin variant. *J. Biol. Chem.* 268, 24887–24891.

Bassi, D.E., Zhang, J., Renner, C., Klein-Szanto, A.J., 2017. Targeting proprotein

convertases in furin-rich lung cancer cells results in decreased *in vitro* and *in vivo* growth. *Mol. Carcinog.* 56, 1182–1188. <https://doi.org/10.1002/mc.22550>.

Bosch, B.J., Bartelink, W., Rottier, P.J.M., 2008. Cathepsin L functionally cleaves the severe acute respiratory syndrome coronavirus class I fusion protein upstream of rather than adjacent to the fusion peptide. *J. Virol.* 82, 8887–8890. <https://doi.org/10.1128/jvi.00415-08>.

Braun, E., Sauter, D., 2019. Furin-mediated protein processing in infectious diseases and cancer. *Clin. Transl. Immunol.* 8, e1073. <https://doi.org/10.1002/cti2.1073>.

Chan, C.M., Woo, P.C., Lau, S.K., Tse, H., Chen, H.L., Li, F., Zheng, B.J., Chen, L., Huang, J.D., Yuen, K.Y., 2008. Spike protein, S, of human coronavirus HKU1: role in viral life cycle and application in antibody detection. *Exp. Biol. Med.* 233, 1527–1536. <https://doi.org/10.3181/0806-RM-197>.

Chan, J.F., Kok, K.H., Zhu, Z., Chu, H., To, K.K., Yuan, S., Yuen, K.Y., 2020. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg. Microb. Infect.* 9, 221–236. <https://doi.org/10.1080/22221751.2020.1719902>.

Cheng, J., Zhao, Y., Xu, G., Zhang, K., Jia, W., Sun, Y., Zhao, J., Xue, J., Hu, Y., Zhang, G., 2019. The S2 subunit of QX-type infectious bronchitis coronavirus spike protein is an essential determinant of neurotropism. *Viruses* 11. <https://doi.org/10.3390/v11100972>.

Claas, E.C., Osterhaus, A.D., Van Beek, R., De Jong, J.C., Rimmelzwaan, G.F., Senne, D.A., Krauss, S., Shortridge, K.F., Webster, R.G., 1998. Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. *Lancet* 351, 472–477. [https://doi.org/10.1016/S0140-6736\(97\)11212-0](https://doi.org/10.1016/S0140-6736(97)11212-0).

Dahms, S.O., Jiao, G.-S., Than, M.E., 2017. Structural studies revealed active site distortions of human furin by a small molecule inhibitor. *ACS Chem. Biol.* 12, 2474. <https://doi.org/10.1021/acscchembio.7b00633>.

de Wit, E., van Doremalen, N., Falzarano, D., Munster, V.J., 2016. SARS and MERS: recent insights into emerging coronaviruses. *Nat. Publ. Gr.* <https://doi.org/10.1038/nrmicro.2016.81>.

Decroly, E., Wouters, S., Di Bello, C., Lazure, C., Ruyschaert, J.-M., Seidah, N.G., 1996. Identification of the Paired Basic Convertases Implicated in HIV gp160 Processing Based on *In Vitro* Assays and Expression in CD4⁺ Cell Lines. *J. Biol. Chem.* 271, 30442–30450. <https://doi.org/10.1074/jbc.271.48.30442>.

Garten, W., Hallenberger, S., Ortmann, D., Schäfer, W., Vey, M., Anglikler, H., Shaw, E., Klenk, H.D., 1994. Processing of viral glycoproteins by the subtilisin-like endoprotease furin and its inhibition by specific peptidylchloroalkylketones. *Biochimie* 76, 217–225. [https://doi.org/10.1016/0300-9084\(94\)90149-x](https://doi.org/10.1016/0300-9084(94)90149-x).

Ge, X.-Y., Li, J.-L., Yang, X.-L., Chmura, A.A., Zhu, G., Epstein, J.H., Mazet, J.K., Hu, B., Zhang, W., Peng, C., Zhang, Y.-J., Luo, C.-M., Tan, B., Wang, N., Zhu, Y., Cramer, G., Zhang, S.-Y., Wang, L.-F., Daszak, P., Shi, Z.-L., 2013. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 503, 535–538. <https://doi.org/10.1038/nature12711>.

Izaguirre, G., 2019. The proteolytic regulation of virus cell entry by furin and other proprotein convertases. *Viruses* 11. <https://doi.org/10.3390/v11090837>.

Kido, H., Okumura, Y., Takahashi, E., Pan, H.Y., Wang, S., Yao, D., Yao, M., Chida, J., Yano, M., 2012. Role of host cellular proteases in the pathogenesis of influenza and influenza-induced multiple organ failure. *Biochim. Biophys. Acta Protein Proteomics.* <https://doi.org/10.1016/j.bbapap.2011.07.001>.

Kim, W., Zekas, E., Lodge, R., Susan-Resiga, D., Marcinkiewicz, E., Essalmani, R., Mihara, K., Ramachandran, R., Asachop, E., Gelman, B., Cohen, É.A., Power, C., Hollenberg, M.D., Seidah, N.G., 2015. Neuroinflammation-Induced interactions between protease-activated receptor 1 and proprotein convertases in HIV-associated neurocognitive disorder. *Mol. Cell Biol.* 35, 3684–3700. <https://doi.org/10.1128/mcb.00764-15>.

Kuhn, J.H., Li, W., Choe, H., Farzan, M., 2004. Angiotensin-converting enzyme 2: a functional receptor for SARS coronavirus. *Cell. Mol. Life Sci.* 61, 2738–2743. <https://doi.org/10.1007/s00018-004-4242-5>.

Le Coupanec, A., Desforges, M., Meessen-Pinard, M., Dubé, M., Day, R., Seidah, N.G., Talbot, P.J., 2015. Cleavage of a neuroinvasive human respiratory virus spike glycoprotein by proprotein convertases modulates neurovirulence and virus spread within the central nervous system. *PLoS Pathog.* 11. <https://doi.org/10.1371/journal.ppat.1005261>.

Li, F., Li, W., Farzan, M., Harrison, S.C., 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. *Science* 309, 1864–1868. <https://doi.org/10.1126/science.1116480>.

Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K.S.M., Lau, E.H.Y., Wong, J.Y., Xing, X., Xiang, N., Wu, Y., Li, C., Chen, Q., Li, D., Liu, T., Zhao, J., Li, M., Tu, W., Chen, C., Jin, L., Yang, R., Wang, Q., Zhou, S., Wang, R., Liu, H., Luo, Y., Liu, Y., Shao, G., Li, H., Tao, Z., Yang, Y., Deng, Z., Liu, B., Ma, Z., Zhang, Y., Shi, G., Lam, T.T.Y., Wu, J.T.K., Gao, G.F., Cowling, B.J., Yang, B., Leung, G.M., Feng, Z., 2020. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* <https://doi.org/10.1056/NEJMoa2001316>.

Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greeneugh, T.C., Choe, H., Farzan, M., 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 426, 450–454. <https://doi.org/10.1038/nature02145>.

Lu, G., Wang, Q., Gao, G.F., 2015. Bat-to-human: spike features determining “host jump” of coronaviruses SARS-CoV, MERS-CoV, and beyond. *Trends Microbiol.* <https://doi.org/10.1016/j.tim.2015.06.003>.

Matsuyama, S., Nagata, N., Shirato, K., Kawase, M., Takeda, M., Taguchi, F., 2010. Efficient activation of the severe acute respiratory syndrome coronavirus spike protein by the transmembrane protease TMPRSS2. *J. Virol.* 84, 12658–12664. <https://doi.org/10.1128/JVI.01542-10>.

Matsuyama, S., Ujike, M., Morikawa, S., Tashiro, M., Taguchi, F., 2005. Protease-

- mediated enhancement of severe acute respiratory syndrome coronavirus infection. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12543–12547. <https://doi.org/10.1073/pnas.0503203102>.
- Mbikay, M., Sirois, F., Yao, J., Seidah, N.G., Chrétien, M., 1997. Comparative analysis of expression of the proprotein convertases furin, PACE4, PC1 and PC2 in human lung tumours. *Br. J. Canc.* 75, 1509–1514. <https://doi.org/10.1038/bjc.1997.258>.
- Mille, J.K., Whittaker, G.R., 2014. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proc. Natl. Acad. Sci. U.S.A.* 111, 15214–15219. <https://doi.org/10.1073/pnas.1407087111>.
- Millet, J.K., Whittaker, G.R., 2015. Host cell proteases: critical determinants of coronavirus tropism and pathogenesis. *Virus Res.* 202, 120–134. <https://doi.org/10.1016/j.virusres.2014.11.021>.
- Millet, J.K., Whittaker, G.R., 2014. Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proc. Natl. Acad. Sci. U.S.A.* 111, 15214–15219. <https://doi.org/10.1073/pnas.1407087111>.
- Mouillard, M., Decroly, E., 2000. Maturation of HIV envelope glycoprotein precursors by cellular endoproteases. *Biochim. Biophys. Acta Rev. Biomembr.* [https://doi.org/10.1016/S0304-4157\(00\)00014-9](https://doi.org/10.1016/S0304-4157(00)00014-9).
- Schechter, I., Berger, A., 1968. On the active site of proteases. 3. Mapping the active site of papain; specific peptide inhibitors of papain. *Biochem. Biophys. Res. Commun.* 32, 898–902. [https://doi.org/10.1016/0006-291x\(68\)90326-4](https://doi.org/10.1016/0006-291x(68)90326-4).
- Seidah, N.G., Chretien, M., 1999. Proprotein and prohormone convertases: a family of subtilases generating diverse bioactive polypeptides. *Brain Res.* 848, 45–62. [https://doi.org/10.1016/S0006-8993\(99\)01909-5](https://doi.org/10.1016/S0006-8993(99)01909-5).
- Seidah, N.G., Prat, A., 2012. The biology and therapeutic targeting of the proprotein convertases. *Nat. Rev. Drug Discov.* <https://doi.org/10.1038/nrd3699>.
- Sun, X., Tse, L.V., Ferguson, A.D., Whittaker, G.R., 2010. Modifications to the hemagglutinin cleavage site control the virulence of a neurotropic H1N1 influenza virus. *J. Virol.* 84, 8683–8690. <https://doi.org/10.1128/JVI.00797-10>.
- Wan, Y., Shang, J., Graham, R., Baric, R.S., Li, F., 2020. Receptor recognition by novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS. *J. Virol.* <https://doi.org/10.1128/JVI.00127-20>.
- Yuan, Y., Cao, D., Zhang, Y., Ma, J., Qi, J., Wang, Q., Lu, G., Wu, Y., Yan, J., Shi, Y., Zhang, X., Gao, G.F., 2017. Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat. Commun.* <https://doi.org/10.1038/ncomms15092>.
- Zhao, S., Ran, J., Musa, S.S., Yang, G., Wang, W., Lou, Y., Gao, D., Yang, L., He, D., Wang, M.H., 2020. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: a data-driven analysis in the early phase of the outbreak. *Int J Infect Dis* 30053–30059. <https://doi.org/10.1016/j.ijid.2020.01.050>.

29 **HIGHLIGHTS**

30 - The Covid-19 is a major threat to public health in absence of therapeutics and vaccines.

31 - The development of antivirals is urgently needed to treat Covid-19 patients and limit virus
32 dissemination.

33 - Several studies identified compounds limiting coronavirus replication *in vitro* and the mode of
34 action of some promising antiviral is established.

35 - Repurposing therapeutic strategies against SARS-CoV-2 is in progress and we need to clarify
36 the mode of action of the used compounds.

37

38

39

40

41

42

43 1. Introduction to SARS-Coronavirus 2

44 Over the last 20 years, three coronaviruses (CoVs) causing severe pulmonary infections in
45 humans, have crossed the species barrier (Ksiazek et al., 2003; Zaki et al., 2012). The last of
46 these CoVs, named SARS-CoV-2, emerged in December 2019 in the Hubei province (China)
47 and rapidly spread around the world becoming a major public health threat (Wu et al., 2020). In
48 approximately 20% of patients, the disease progresses to severe pneumonia, respiratory and
49 multi-visceral failure often leading to death of patients with comorbidity factors (Zhou et al.,
50 2020). This illness evolution is associated with a deregulated immune response, including
51 exacerbated production of pro-inflammatory cytokines (Huang et al., 2020). The development
52 of antiviral therapies is thus urgently needed especially given the current absence of a vaccine.

53 Coronaviruses belong to the *Nidovirales* order, a group of enveloped viruses with genomic RNA
54 of positive polarity. Their 27-34 kb genome codes for 16 non-structural proteins (nsps) and 4
55 structural proteins (spike protein (S), envelope protein (E), membrane protein (M) and
56 nucleoprotein (N)) (de Wit et al., 2016). The virus life cycle, described in **Figure 1**, begins by the
57 attachment of the viral particle to the ACE2 cell surface receptor mediated by the S protein (Li,
58 2015; Wan et al., 2020). Virus entry is next achieved by endocytosis and/or direct fusion of the
59 cell and viral membranes. The S protein is synthesized as an inactive precursor that is inserted
60 in the viral membrane and requires two subsequent cleavages by cellular proteases to become
61 functionally active. Different cellular protease such as furin-like enzymes or the Transmembrane
62 Protease Serine 2 (TMPRSS2) were proposed to cleave the S protein into two subunits, S1 and
63 S2 in a process called "priming" (Coutard et al., 2020; Matsuyama et al., 2010). Recent study
64 has reported that this cleavage involving furin in lung cells is important to virus entry (Hoffmann
65 Markus et al, 2020). In addition, the virus entry requires a second proteolytic cleavage occurring
66 at a S2' site localized immediately upstream to the fusion peptide (Belouzard et al., 2009; Millet
67 and Whittaker, 2015), which seem to involve at least TMPRSS2 (Bestle et al., 2020). The viral
68 genome is then released into the cytoplasm of infected cells, allowing translation of the viral
69 mRNA into two polyproteins, pp1a and pp1ab, controlled by a -1 ribosomal frameshift. These
70 polyprotein precursors are subsequently cleaved by two viral proteases, chymotrypsin-like 3
71 (3CLpro) and papain-like (PLpro), to generate 16 non-structural proteins (nsp1 to 16). Many of
72 these proteins participate in the formation of the replication and transcription complex (RTC)
73 orchestrating genome replication, mRNA synthesis and capping. At the final stage of viral
74 infection, the N protein assembles into the neo-synthesized viral genome to form the

75 nucleocapsid, which associates with the viral structural proteins to generate new virions that are
76 released by exocytosis (de Wit et al., 2016).

77 Thus far, the ongoing SARS-CoV-2 pandemic has already caused at least 3.2 million confirmed
78 cases and more than 228.000 deaths. In response to the global health emergency, public health
79 measures to control the spread of the virus were established and efforts to identify molecules
80 with antiviral potential and research for vaccine development have been intensified. The antiviral
81 strategies explored "*hic et nunc*" are principally based on the use of existing drugs, initially
82 developed to treat other diseases. However, even more than 800 different clinical trials are
83 currently in progress, they include only a limited set of drugs, tested alone or in combination,
84 (Thorlund et al., 2020). Although drugs used in these therapeutic assays have shown some
85 inhibitory effect in CoV infected cells, the molecular bases of their antiviral activity are often
86 poorly understood, except for nucleoside analogs (e.g. Favipiravir for Influenza virus) and fusion
87 inhibitors (e.g. arbidol). In addition, important concerns have arisen pending the results of
88 clinical studies, such as the risk of inappropriate self-treatment, the potential toxicity or adverse
89 effect of compounds and the risk of depletion of pharmacy stocks making them unavailable for
90 treatment of other pathologies. Conversely, the emergence of SARS-CoV and Middle East
91 respiratory syndrome coronavirus (MERS-CoV) has prompted an increasing number of
92 publications describing new CoV inhibitors and characterizing in depth their mode of action
93 (MoA). In this review, we summarize the promising candidate molecules which could be
94 repurposed, as well as anti-CoV specific drugs derived from long standing research on CoVs,
95 for which we provide additional information pertaining to their potential MoA based on our
96 knowledge of the SARS-COV-2 viral life-cycle.

97 Immunomodulation strategies, vaccination and supporting treatment strategies are not
98 described in detail in this review, although these are also key issues in the combat against
99 Covid-19.

100

101 **2. Classical antiviral approaches**

102 In absence of specific therapeutics, supportive treatments such as oxygen therapy or
103 mechanical ventilation remain the only option for managing severe disease symptoms ("Clinical
104 management of severe acute respiratory infection when COVID-19 is suspected," n.d.). In order
105 to fight against the early virus propagation and limit the illness severity, one important strategy
106 should be to develop therapeutics based on existing antiviral compounds. Following the rapid
107 emergence and spread of SARS-CoV and MERS-CoV in 2003 and 2012 respectively, important

108 efforts were made in order to identify drugs targeting specific steps of the coronavirus life cycle
109 (**Figure 1**) including i) cellular receptor Angiotensin Converting Enzymes 2 (ACE2), ii) spike
110 glycoprotein processing enzymes, iii) viral entry, iv) polyprotein processing, and v) the
111 replication/transcription complex. The latest SARS-CoV-2 pandemic has prompted rapid
112 screening of these compounds in infected cells, with several now pushed to small animal
113 models and/or clinical trials.

114 2.1. Antibodies neutralization

115 An usefull therapeutic strategy consists of interfering with the first step of the viral cycle by
116 targeting the interaction of the viral S protein with the cell surface receptors. It was recently
117 demonstrated that sera of patients contain neutralizing antibodies (Abs) that can limit SARS-
118 CoV-2 viral infection *in vitro* (Zhou and Zhao, 2020). In support of this, treatment of patients
119 showing respiratory distress with convalescent plasma has shown some efficacy, suggesting
120 that the passive immunity might help recovery from SARS-CoV-2 infection (Cheng et al., 2005;
121 Mair-Jenkins et al., 2015; Yeh et al., 2005). Clinical trials using plasmaphereses are thus
122 underway to evaluate the efficacy of this strategy in the treatment of COVID-19 (L. Chen et al.,
123 2020; Shen et al., 2020). The development of monoclonal Abs (MAbs) directed against the viral
124 S protein is also a rational option for therapy (Berry, 2005). This approach was previously used
125 in the latest Ebola outbreaks with the development of ZMapp (Hoenen et al., 2019). Neutralizing
126 MABs targeting the SARS- and MERS-CoV S protein have been evidenced and their antiviral
127 effect was experimentally demonstrated in cells and animals infected with SARS-CoV or MERS-
128 CoV (Greenough et al., 2005; He et al., 2005; Hofmann et al., 2004; Sui et al., 2004; Wang et
129 al., 2018). Thus, the development of MAbs targeting the S protein of MERS-coV-2 is also
130 currently in progress. Alternatively, the development of MAbs targeting ACE2 is also an
131 interesting option to reduce viral entry (Walls et al., 2020; Zhu et al., 2007), but may be more
132 risky due to the potential for interference of the physiological function of the ACE2 receptor,
133 which is involved in angiotensin maturation. Another possibility is to use human recombinant
134 soluble ACE2 in order to block SARS-CoV-2 infection (Monteil et al., 2020). While this strategy
135 seems to limit the SARS-CoV-2 infection in cell culture, the stability of the soluble receptor in the
136 serum of infected patients and potential adverse effects in patients remain to be determined.
137 Interestingly, MAbs targeting other cellular receptors such as CD147 have also been shown to
138 block viral infection in cell cultures, with a direct interaction between CD147 and the S protein
139 demonstrated using co-immunoprecipitation (K. Wang et al., 2020). While these first results
140 have yet to be confirmed, this observation would suggest that the SARS-CoV-2 could use an

141 alternative entry receptor in certain cell types. A clinical trial using an anti-CD147 antibody
142 (Meplazumab) is currently ongoing in China (Bian et al., 2020), however the antiviral effect of
143 this therapeutic is uncertain as the virus is well known to bind with high affinity to the ACE2
144 receptor (Walls et al., 2020).

145 2.2. Antiviral molecules

146 Although SARS-CoV-2 has spread very recently, several antivirals previously tested against
147 other pathogenic CoVs have already tested in infected patients (**Table 1**). The first interesting
148 possibility is to interfere with virus entry by using molecules that block virus-cell fusion. This
149 includes Umifenovir, also known as Arbidol, an antiviral initially developed for the treatment of
150 influenza infection (Kadam and Wilson, 2017). Interestingly, Arbidol has also been shown to
151 have broad spectrum antiviral activity by inhibiting the viral entry other viruses (Vigant et al.,
152 2015) and stimulating the immune response (Boriskin et al., 2008). Arbidol has recently been
153 shown to potently inhibit the SARS-CoV-2 infection *in vitro* with an IC₅₀ of 10 μ M (Touret et al.,
154 2020). Although not yet approved by the FDA, this broad-spectrum antiviral molecule is used in
155 China and Russia for influenza treatment and it was thus quickly included in several clinical
156 trials alone or in combination with other therapeutics. The first comparative study performed on
157 a small set of patients (33) showed a marginally faster decrease in viral load in the cohort
158 treated with Arbidol in combination with protease inhibitors lopinavir/ritonavir than in the patients
159 treated with lopinavir/ritonavir alone (Deng et al., 2020).

160 After virus entry, the polyproteins pp1a and pp1ab are processed into 16 nsps by the viral
161 3CLpro and PLpro. These cleavage events play a pivotal role in the viral life cycle as the
162 processing generates, among others, the RdRp required for viral RNA replication and
163 transcription. The high conservation of the cleavage sites and the protease structures makes
164 these proteases an optimal antiviral target. The X-ray structure of the SARS-CoV-2 3CLpro was
165 recently reported alone and in presence of an α -ketoamide inhibitor (L. Zhang et al., 2020).
166 Based on these structures, the α -ketoamide inhibitor was optimized and showed antiviral activity
167 on SARS-CoV-2 below 10 μ M in Calu3 lung cells. Additionally, protease inhibitors used in HIV
168 therapy such as lopinavir/ritonavir (inhibitor+boster, combination known as kaletra), nelfinavir,
169 and dolutegravir were also demonstrated to limit virus propagation in infected cells (Bhatnagar
170 et al., 2020; Riva et al., 2020; Touret et al., 2020) and several clinical trials are already in
171 progress (Chu et al., 2004; de Wilde et al., 2014; Yamamoto et al., 2004). However, the
172 molecular basis of their inhibition mechanism is still elusive given the structural differences
173 between the HIV protease, which belongs to the aspartyl protease family, and the two CoV

174 proteases. The first clinical trials of lopinavir combined with the ritonavir booster have
175 unfortunately not demonstrated any significant effect on SARS-CoV-2 infection (Cao et al.,
176 2020).

177 Historically, the most attractive class of antiviral compounds are those that block virus
178 replication by inhibiting RdRp activity. Nucleotide or nucleoside analogs are one of the most
179 promising groups of RdRp inhibitors. The antiviral effect of these analogs can be attributed to
180 three non-mutually exclusive mechanisms. Firstly, incorporation of nucleotide analogs (NAs)
181 into viral RNA by the error-prone polymerase can induce early chain termination in either an
182 obligate (immediate) or non-obligate fashion, resulting in incomplete, non-infectious viral RNA.
183 The second mechanism, named error catastrophe, is associated with ready insertion and
184 extension of NAs throughout the the viral RNA, subsequently resulting in errors during the RNA
185 synthesis (Crotty et al., 2000; Pruijssers and Denison, 2019). Furthermore, several NAs have
186 been shown to deplete cytoplasmic levels of their equivalent native nucleotide, causing
187 nucleotide pool imbalances which further impacts RdRp fidelity (Feld and Hoofnagle, 2005).
188 While nucleoside analogs are successfully used in the treatment of other viral diseases, the
189 situation is complicated in CoVs due to the presence of a viral exonuclease (nsp14 ExoN)
190 carrying proof-reading activity that can lower the antiviral effect of NAs (Ferron et al., 2018;
191 Smith et al., 2013). *In vitro* studies have shown, for example, that SARS-CoV polymerase can
192 incorporate ribavirin triphosphate during replication, but this purine NA is detected by nsp14
193 ExoN and eliminated by the repair mechanism (Ferron et al., 2018). This observation can be
194 associated to the poor antiviral effect of ribavirin alone observed for SARS- and MERS-CoV
195 infections (Falzarano et al., 2013; Ströher et al., 2004). However, molecules such as remdesivir,
196 initially developed against Ebola virus (Warren et al., 2016), appear to have an inhibitory effect
197 against SARS-CoV and MERS-CoV *in vitro* at the sub-micro- to micromolar range, depending
198 on the cell type used for infection (Agostini et al., 2018; Sheahan et al., 2017, 2020a). The
199 antiviral activity of remdesivir was also evaluated in a primate model infected by SARS-CoV,
200 where it was demonstrated that the antiviral activity is dependant on early treatment following
201 infection (de Wit et al., 2020). Remdesivir was shown to be incorporated by SARS-CoV-2
202 RdRp complex in the nascent RNA (Gordon et al., 2020a). Several clinical trials, such as the
203 European DISCOVERY clinical program (“Launch of a European clinical trial against COVID-19
204 | Newsroom | Inserm,” n.d.), are planned to evaluate the safety and efficacy of remdesivir as a
205 therapeutics for hospitalized patients. First results of these ongoing trials show that
206 compassionate use of remdesivir leads to clinical improvement in 68% of patients in a small

207 cohort (53 patients) (Grein et al., 2020). Notably, some adverse events were reported including
208 abnormal liver function, diarrhea, rashes, renal impairment, and hypotension.

209 The cytidine analog β -d-N4-Hydroxycytidine (NHC) has also been identified as a potential anti-
210 CoV therapeutic, with inhibition of MERS-CoV, SARS-CoV and SARS-CoV-2 replication *in vitro*
211 in the micromolar range (0.09 to 0.3 μ M), with no apparent sensitivity for the viral proofreading
212 activity of nsp14 (Sheahan et al., 2020b). The decreased viral replication was associated with
213 an increased mutation frequency, supporting a mechanism of lethal mutagenesis in CoV.
214 Additionally, the orally bioavailable β -D-N4-hydroxycytidine-5'-isopropyl ester was shown to
215 improve pulmonary function and reduce virus titer and body weight loss in mice infected with
216 SARS- and MERS-CoV. These promising results suggest that NHC molecules should be
217 considered for the treatment of SARS-CoV-2 (Agostini et al., 2019; Sheahan et al., 2020a), and
218 need to enter in clinical trials. A third potential compound is the guanine analog Favipiravir (T-
219 705). While initially developed against influenza virus (Furuta et al., 2002), T-705 has been
220 shown to have broad-spectrum activity, with antiviral effects against flaviviruses (Kim et al.,
221 2018; Zmurko et al., 2016), alphaviruses (Delang et al., 2014), noroviruses (Rocha-Pereira et
222 al., 2012) and Ebola virus (Guedj et al., 2018; Oestereich et al., 2014; Smither et al., 2014).
223 Favipiravir was approved in Japan in 2014 to treat influenza infection and is currently being
224 studied for SARS-CoV-2 treatment in several clinical trials (Cai et al., 2020; C. Chen et al.,
225 2020). Preliminary results suggest a shorter viral clearance time for the Favipiravir arm versus
226 the control arm (Cai et al., 2020).

227 A last nucleoside analog potentially active against SARS-CoV-2 is sofosbuvir, which is a broad-
228 acting antiviral approved for the management of hepatitis C virus infection. In a recent study,
229 sofosbuvir triphosphate has been demonstrated to be incorporated during RNA elongation by
230 the recombinant SARS-CoV-2 RdRp (Gordon et al., 2020a). Further work is needed to confirm
231 the antiviral effect in infected cells and in animal models before possibly moving to clinical trial.

232 **3. Repositioning strategies SARS-CoV-2 treatment**

233 The fastest option for the treatment of COVID-19 is the identification of approved drugs
234 originally developed for other diseases that show inhibitory activity against SARS-CoV-2 in
235 infected cells. The main advantages of this strategy are that drugs are already available on the
236 market, and their safety and toxicity profiles are already documented. However, the MoA of
237 such drugs against the virus can be speculative, and the efficacy and clinical doses required for
238 CoV treatment in humans is unknown given the lack of large scale clinical studies. Moreover,
239 one must pay attention to the risk of over-dosage or unexpected increase of viral infection or

240 symptoms. Several teams have already setup platforms in order to screen FDA-approved
241 libraries on SARS-CoV-2 infected cells (Jeon et al., 2020; Riva et al., 2020; Touret et al., 2020).
242 Even though chemical libraries and experimental conditions used in these studies are quite
243 different, a comparative analysis of these results reveals that some compounds, such as
244 chloroquine derivatives, have been identified under these different experimental conditions.
245 Additionally, several groups have also reported antiviral activity of FDA-approved compounds,
246 with some of these already used in clinical trials (Thorlund et al., 2020). In this review, we
247 cannot discuss all drugs identified by this strategy. We have thus selected the key compounds
248 that are already in clinical trials and the most interesting classes of compounds based on their
249 potential MoA on CoV life cycle inferred from the known activity of the compounds.

250 *3.1. Immunomodulators*

251 COVID-19 is associated with an inefficient antiviral response in some patients and thus the use
252 of immunomodulators has been proposed to restore the homeostasis of the immune system.
253 Concerning the immunomodulation strategy, the use of interferon (IFN) during early stages of
254 infection could be used to boost the innate immune response and promote viral clearance. IFN
255 treatments are now tested in several clinical trials, but the benefit of such therapy is not yet
256 known. However, since infection may also be accompanied by a runaway immune response
257 leading to a massive production of pro-inflammatory cytokines, IFN treatments could lead to
258 unexpected effects (Cinatl et al., 2003; Hensley et al., 2004; Huang et al., 2020; PARAGAS et
259 al., 2005; Totura et al., 2015). Indeed, IFN has been recently shown to trigger overexpression
260 of ACE2, even in cell lines with a low basal expression level, leading to a larger dissemination
261 of SARS-CoV-2 (Ziegler et al., 2020). Conversely, in the second phase of the disease, the use
262 of molecules which limit the effect of the cytokine storm could be advantageous. In this
263 hypothesis, the use of MABs such as mepolizumab directed against interleukin (IL)-5,
264 tocilizumab directed against the IL-6 receptor (C. Zhang et al., 2020) or anakinra directed
265 against IL-1 (Conti et al., 2020a), could prove effective in controlling the immune response and
266 be of therapeutic interest in managing the production of pro-inflammatory cytokines in patients
267 with severe infection (Conti et al., 2020b).

268 *3.2. Repositioning drugs interplaying with the viral life cycle*

269 Several therapeutic molecules (**table 2**), selected through antiviral screening performed in
270 infected cells, target the renin/angiotensin pathway (Kuster et al., 2020; Meng et al., 2020;
271 Touret et al., 2020). The MoA of these class of inhibitors is speculative, but it is noteworthy that

272 ACE2 belongs to this pathway. Further work is needed to unravel the molecular basis of the
273 virus inhibition mediated by these compounds before clinical trials.

274 Once the virus binds to the target cell, the S protein is cleaved by cellular proteases into S1 and
275 S2 subunits (priming process) and upstream its fusion peptide. The serine protease TMPRSS2,
276 which is highly expressed in lung cells, has been reported to cleave a variety of CoV spike
277 proteins (Belouzard et al., 2009; Millet and Whittaker, 2015), and seems to play a key role for
278 virus entry (Bestle et al., 2020). Camostat mesylate, a TMPRSS2 inhibitor approved for clinical
279 use, was shown to abrogate the S protein processing *in vitro* (Hoffmann et al., 2020a). As a
280 consequence, cell entry of vesicular stomatitis virus (VSV) pseudotypes bearing SARS-CoV-2 S
281 protein was inhibited in different cell types overexpressing TMPRSS2 by Camostat mesylate
282 and its derivative Nafamostat mesylate in the nM to μ M range . (Hoffmann et al., 2020b, 2020c).
283 This observation suggests that inhibition of S protein priming blocks the ACE2-mediated entry.
284 However, the inhibition observed with Camostat mesylate and Nafamostat mesylate in VeroE6
285 cells infected by SARS-CoV-2 is lower and probably inconsistent with the possibility to use
286 Camostat mesylate as treatment. The picture is further complicated as other cellular proteases
287 have also been suggested to participate in the spike protein processing. Indeed, cathepsin L
288 (Hoffmann et al., 2020a) and furin (Coutard et al., 2020) have been proposed to be involved in
289 the spike protein activation depending on the infected cell type. Thus, we can expect that even if
290 such inhibitors of these different enzymes can limit virus propagation in certain cell types, their
291 efficacy in the treatment of infection in humans could be limited. The best option is perhaps to
292 combine furin and TMPRSS2 inhibitors in order to limit the virus propagation (Bestle et al.,
293 2020).

294 Interestingly, several compounds identified in different screening of infected cells are well
295 known to interplay with the cellular endosome acidification (Kuster et al., 2020; Meng et al.,
296 2020; Touret et al., 2020). We could speculate that the increase of the endosomal pH will limit
297 the viral/endosomal membrane fusion, which is necessary for the release of the viral genome
298 into the cytoplasm of the host cell. Two recent studies have shown that proton pump inhibitors
299 such as omeprazole and vonoprazan reduces the infection of cells by SARS-CoV-2 (Riva et al.,
300 2020; Touret et al., 2020). In addition, chloroquine (CQ) and hydroxychloroquine (HCQ), initially
301 used in the treatment of malaria with known safety efficacy and long-term potential, also limit
302 acidification of endosomes, Golgi vesicles and lysosomes (Savarino et al., 2003). These
303 molecules have shown significant *in vitro* effects on the replication of SARS-CoV and SARS-
304 CoV-2 viruses (Keyaerts et al., 2004; Pantoliano et al., 2001; Vincent et al., 2005; M. Wang et
305 al., 2020; Yao et al., 2020). Additionally, it should be noted that chloroquine derivatives also

306 have immunosuppressive effects and inhibit the activation of innate immune receptors such as
307 TLRs (Schrezenmeier and Dörner, 2020) and were associated with some adverse cardiac
308 effects. Numerous clinical studies are currently underway (Gao et al., 2020; Gautret et al.,
309 2020), and initial results from these studies are controversial (“Statement on IJAA paper |
310 International Society of Antimicrobial Chemotherapy,” n.d.). Randomized, double-blind studies
311 involving a larger number of patients are expected to resolve this issue. Recent studies have
312 also demonstrated the therapeutic effect of the combined use of HCQ with azithromycin
313 (Gautret et al., 2020). Azithromycin is an antibiotic also described to have antiviral activity on
314 respiratory viral infections (Gielen et al., 2010). Interestingly, the MoA of azithromycin and HCQ
315 may be related to the modulation of the endosomal and trans-Golgi network pH (Deretic and
316 Timmins, 2020). In addition, azithromycin plays a role in regulating the production of interleukins
317 (IL), which can help control the immune response and prevent progression towards severe
318 symptoms (Aghai et al., 2007; Bouwman et al., 2004; Taylor et al., 2015). More recently, it has
319 also been proposed that azithromycin may inhibit viral invasion by interfering with CD147-
320 mediated recognition mechanism (Ulrich and Pillat, 2020).

321 More recently, ivermectin was reported to inhibit the SARS-CoV-2 with a ~5,000-fold reduction
322 in viral RNA after 48 h (Caly et al., 2020). Ivermectin is a FDA-approved anti-parasitic molecule,
323 initially identified as an inhibitor of HIV protein nuclear import by interacting with IMP α / β 1
324 (Wagstaff et al., 2012). Its broad-spectrum antiviral activity on other viruses was documented *in*
325 *vitro* (Lundberg et al., 2013; Tay et al., 2013). While ivermectin is known to perturb the nuclear
326 import, mechanism mediated by IMP α / β 1, the role of this machinery during the CoV life cycle is
327 not yet elucidated. However, an observational study reported interesting survival benefits during
328 hospitalization (Patel, 2020) but no data are available regarding viral load, highlighting the
329 necessity validate the observation with clinical trials.

330 Finally, oseltamivir, a neuraminidase inhibitor that prevents influenza viral particle release is
331 also under investigation (Gubareva et al., 2000). Much like ritonavir/lopinavir, there is no
332 molecular basis supporting such trials since, to the best of our knowledge, CoVs (unlike
333 Influenza viruses) do not rely on neuraminidases during their life cycle. Should antiviral effect be
334 observed, off-target mechanisms would still need to be elucidated.

335 **4. Conclusion**

336 The health emergency caused by the recent SARS-CoV-2 pandemic highlights the need to
337 identify effective treatments to combat this new emerging virus. Due to the urgent requirement
338 for therapeutics, initial clinical trials have mainly been conducted with repurposed drugs which

339 have shown promising anti-CoV effects in cell culture. Although more than 800 clinical trials are
340 already underway, the actual number of compounds included in these studies is still limited to
341 around twenty, used alone or in combination. In addition, the understanding of the anti-CoV
342 MoA for most of these therapeutic molecules is currently limited. The first results of these
343 clinical studies reveal some inconsistencies and contradictory results and suggest that the size
344 of cohorts and quality of the control arm is a key issue to obtain more conclusive studies. In
345 addition, we noticed that discrepancies observed in the first studies might also result from
346 important differences in clinical inclusion criteria including the possibility of early treatment that
347 may be essential for efficacy of CoV therapy. Thus, the detection of the early steps of the viral
348 infection in patients is an important issue in order to early administer treatments, leading to a
349 better efficacy of antiviral compounds and limiting viral transmission. Importantly, efforts should
350 also be made in order to push new compounds with documented MoA against SARS-CoV-2 into
351 clinical trials as there is no definitive proof of the existence of efficient compounds to date. The
352 success of CoV antiviral therapies will probably result on multi-therapies based on the use of
353 antiviral compounds targeting different steps of the virus life cycle. Moreover, a multidisciplinary
354 approach combining artificial intelligence, docking of compounds and the set-up of robust *in*
355 *vitro* and *in vivo* assays would speed up the development of antiviral compounds. The
356 implementation of large retrospective studies on hospitalized patients and provided treatments
357 are also needed to evaluate benefits and risks of treatments with robust statistical tools to
358 establish the most efficient and appropriate therapeutic response at each stage of the Covid-19
359 illness.

360

361 **Figure 1: SARS-CoV-2 life cycle and potential mechanism of action of currently evaluated**
362 **therapeutics**

363 SARS-CoV-2 life cycle and the presumed mechanism of action (MoA) of major SARS-CoV-2
364 replication inhibitors. Therapeutic strategies against SARS-CoV-2 are set-up in order to find
365 compounds targeting different steps of the viral life cycle. Molecules currently in clinical trials
366 against SARS-CoV-2 are indicated by a grey star. Therapeutic strategies against SARS-CoV-2
367 based on antiviral compounds are indicated in red and approved drugs used for other diseases
368 or selected by virtual screening are indicated in yellow. The black question marks indicate
369 unknown or elusive MoA in a context of CoV infection. *Corresponding to antibody (Ab)
370 strategies including monoclonal Abs and convalescence patient plasma.

371 *Biorender website was used to generate this figure.*

372

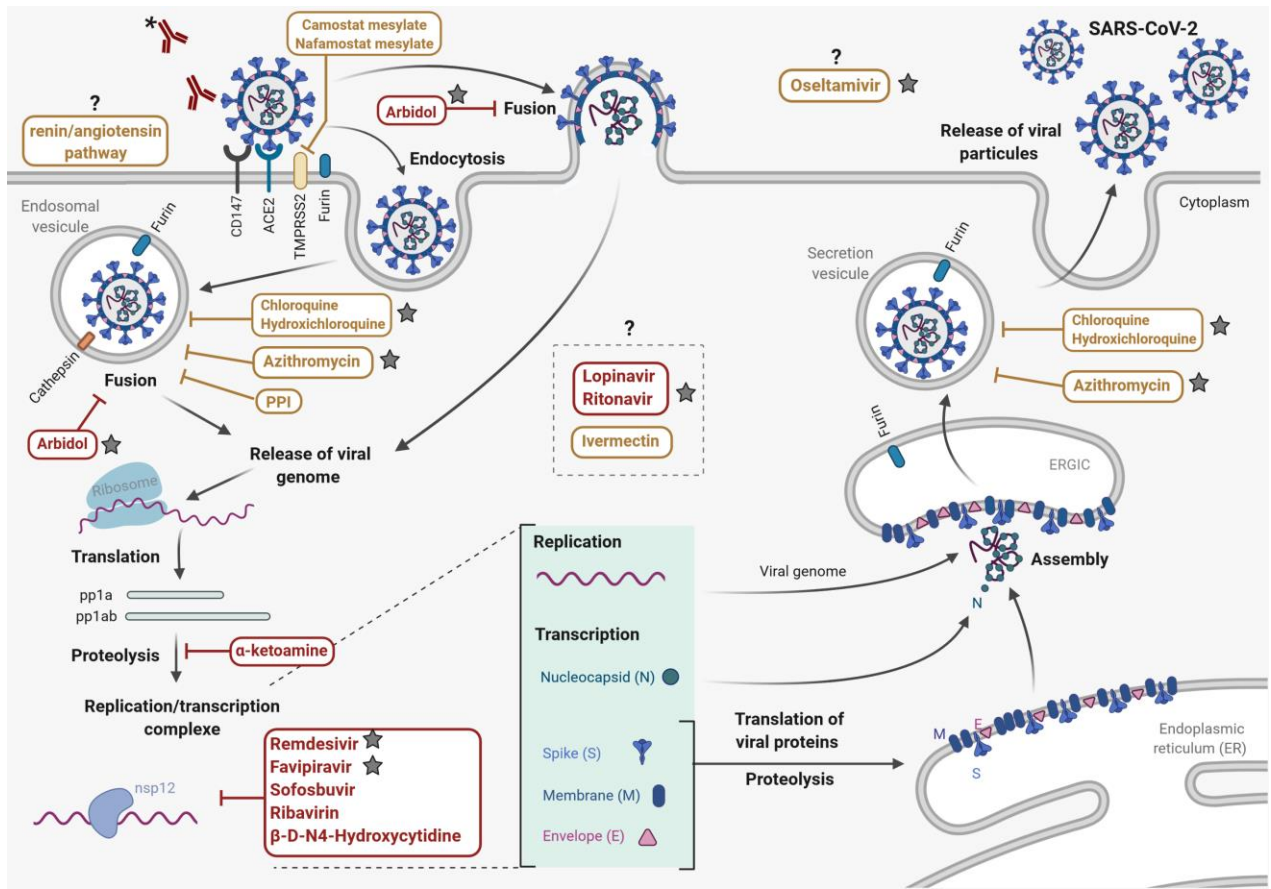
373

374 **Table 1: Main Antiviral molecules currently being tested against SARS-CoV-2 infection**

Compounds	Target/MoA	Tests <i>in vitro</i> (CoV)	Clinical trials SARS-CoV-2
Arbidol hydrochloride	Endocytosis/Inhibition membrane fusion	(Boriskin et al., 2008; Khamitov et al., n.d.; Touret et al., 2020)	(C. Chen et al., 2020; Deng et al., 2020)
Ribavirin	RNA polymerase/Inhibition	(Barnard et al., 2006; Falzarano et al., 2013)	No data
Remdesivir	RNA polymerase/Inhibition	(Agostini et al., 2018; M. Wang et al., 2020; Warren et al., 2016) (de Wit et al., 2020)	("Launch of a European clinical trial against COVID-19 Newsroom Inserm," n.d.) (Grein et al., 2020)
Favipiravir	RNA polymerase/Inhibition	No data	(Cai et al., 2020; C. Chen et al., 2020)
β -d-N4-Hydroxycytidine	RNA polymerase/Inhibition	(Agostini et al., 2019; Sheahan et al., 2020b)	No data
Sofosbuvir	RNA polymerase/Inhibition	(Gordon et al., 2020b)	No data
α -ketoamide inhibitor	Proteases/Inhibition	(L. Zhang et al., 2020)	No data
Ritonavir	Proteases/Inhibition	(Chu et al., 2004)	((Bhatnagar et al., 2020; Cao et al., 2020; Deng et al., 2020; "Launch of a European clinical trial against COVID-19 Newsroom Inserm," n.d.; Lim et al., 2020)
Lopinavir	Proteases/Inhibition	((Chu et al., 2004; de Wilde et al., 2014)	((Bhatnagar et al., 2020; Cao et al., 2020; Deng et al., 2020; "Launch of a European clinical trial against COVID-19 Newsroom Inserm," n.d.; Lim et al., 2020)
Nelfinavir	Proteases/Inhibition	((Yamamoto et al., 2004)	No data
Dolutegravir	Proteases/Inhibition	((Touret et al., 2020)	No data

375
376 **Table 2: Main repositioning molecules currently being tested against SARS-CoV-2**
377 **infection.**

Compounds	Target/MoA	Tests <i>in vitro</i> (CoV)	Clinical trials SARS-CoV-2
Antivirals			
Nafamostat mesylate	TMPRSS2/Inhibitor	(Hoffmann et al., 2020b)	No data
Camostat mesylate	TMPRSS2/Inhibitor	(Hoffmann et al., 2020b; Touret et al., 2020)	No data
Chloroquine Hydroxychloroquine	pH increases in endosomal compartment	(Keyaerts et al., 2004; Liu et al., 2020; Touret et al., 2020; Vincent et al., 2005; M. Wang et al., 2020; Yao et al., 2020)	(Gao et al., 2020; Gautret et al., 2020; "Launch of a European clinical trial against COVID-19 Newsroom Inserm," n.d.)
Azithromycin (antibiotic)	pH increases in endosomal compartment/Immunomodulator	(Touret et al., 2020)	(Gautret et al., 2020)
Omeprazole	Proton pump inhibitor (PPI)	(Touret et al., 2020)	No data
Vonoprazan	Proton pump inhibitor (PPI)	(Touret et al., 2020)	No data
Ivermectin	Limits viral infection	(Caly et al., 2020)	No data
Oseltamivir	Neuraminidase inhibitor	No data	No data



379 **REFERENCES**

- 380 Aghai, Z.H., Kode, A., Saslow, J.G., Nakhla, T., Farhath, S., Stahl, G.E., Eydelman, R., Strande,
381 L., Leone, P., Rahman, I., 2007. Azithromycin Suppresses Activation of Nuclear Factor-
382 kappa B and Synthesis of Pro-inflammatory Cytokines in Tracheal Aspirate Cells From
383 Premature Infants. *Pediatr. Res.* 62, 483–488.
384 <https://doi.org/10.1203/PDR.0b013e318142582d>
- 385 Agostini, M.L., Andres, E.L., Sims, A.C., Graham, R.L., Sheahan, T.P., Lu, X., Smith, E.C.,
386 Case, J.B., Feng, J.Y., Jordan, R., Ray, A.S., Cihlar, T., Siegel, D., Mackman, R.L., Clarke,
387 M.O., Baric, R.S., Denison, M.R., 2018. Coronavirus Susceptibility to the Antiviral
388 Remdesivir (GS-5734) Is Mediated by the Viral Polymerase and the Proofreading
389 Exoribonuclease. *MBio* 9. <https://doi.org/10.1128/mBio.00221-18>
- 390 Agostini, M.L., Pruijssers, A.J., Chappell, J.D., Gribble, J., Lu, X., Andres, E.L., Bluemling, G.R.,
391 Lockwood, M.A., Sheahan, T.P., Sims, A.C., Natchus, M.G., Saindane, M., Kolykhalov,
392 A.A., Painter, G.R., Baric, R.S., Denison, M.R., 2019. Small-Molecule Antiviral β -D- N 4 -
393 Hydroxycytidine Inhibits a Proofreading-Intact Coronavirus with a High Genetic Barrier to
394 Resistance . *J. Virol.* 93. <https://doi.org/10.1128/jvi.01348-19>
- 395 Barnard, D.L., Day, C.W., Bailey, K., Heiner, M., Montgomery, R., Lauridsen, L., Winslow, S.,
396 Hoopes, J., Li, J.K.-K., Lee, J., Carson, D.A., Cottam, H.B., Sidwell, R.W., 2006.
397 Enhancement of the infectivity of SARS-CoV in BALB/c mice by IMP dehydrogenase
398 inhibitors, including ribavirin. *Antiviral Res.* 71, 53–63.
399 <https://doi.org/10.1016/j.antiviral.2006.03.001>
- 400 Belouzard, S., Chu, V.C., Whittaker, G.R., 2009. Activation of the SARS coronavirus spike
401 protein via sequential proteolytic cleavage at two distinct sites. *Proc. Natl. Acad. Sci. U. S.*
402 *A.* 106, 5871–6. <https://doi.org/10.1073/pnas.0809524106>
- 403 Berry, J.D., 2005. Rational monoclonal antibody development to emerging pathogens, biothreat
404 agents and agents of foreign animal disease: The antigen scale. *Vet. J.* 170, 193–211.
405 <https://doi.org/10.1016/j.tvjl.2004.04.021>
- 406 Bestle, D., Heindl, M.R., Limburg, H., van, T.V.L., Pilgram, O., Moulton, H., Stein, D.A., Harges,
407 K., Eickmann, M., Dolnik, O., Rohde, C., Becker, S., Klenk, H.-D., Garten, W., Steinmetzer,
408 T., Böttcher-Friebertshäuser, E., 2020. TMPRSS2 and furin are both essential for

409 proteolytic activation and spread of SARS-CoV-2 in human airway epithelial cells and
410 provide promising drug targets. bioRxiv 2020.04.15.042085.
411 <https://doi.org/10.1101/2020.04.15.042085>

412 Bhatnagar, T., Murhekar, M. V, Soneja, M., Gupta, N., Giri, S., Wig, N., Gangakhedkar, R.,
413 2020. Lopinavir/ritonavir combination therapy amongst symptomatic coronavirus disease
414 2019 patients in India: Protocol for restricted public health emergency use. Indian J. Med.
415 Res. 0, 0. https://doi.org/10.4103/ijmr.IJMR_502_20

416 Bian, H., Zheng, Z.-H., Wei, D., Zhang, Zheng, Kang, W.-Z., Hao, C.-Q., Dong, K., Kang, W.,
417 Xia, J.-L., Miao, J.-L., Xie, R.-H., Wang, B., Sun, X.-X., Yang, X.-M., Lin, P., Geng, J.-J.,
418 Wang, K., Cui, H.-Y., Zhang, K., Chen, X.-C., Tang, H., Du, H., Yao, N., Liu, S.-S., Liu, L.-
419 N., Zhang, Zhe, Gao, Z.-W., Nan, G., Wang, Q.-Y., Lian, J.-Q., Chen, Z.-N., Zhu, P., 2020.
420 Meplazumab treats COVID-19 pneumonia: an open-labelled, concurrent controlled add-on
421 clinical trial. medRxiv 2020.03.21.20040691. <https://doi.org/10.1101/2020.03.21.20040691>

422 Boriskin, Y.S., Leneva, I.A., Pécheur, E.-I., Polyak, S.J., 2008. Arbidol: a broad-spectrum
423 antiviral compound that blocks viral fusion. *Curr. Med. Chem.* 15, 997–1005.
424 <https://doi.org/10.2174/092986708784049658>

425 Bouwman, J.J., Visseren, F.L., Bouter, P.K., Diepersloot, R.J., 2004. Azithromycin inhibits
426 interleukin-6 but not fibrinogen production in hepatocytes infected with cytomegalovirus
427 and chlamydia pneumoniae. *J. Lab. Clin. Med.* 144, 18–26.
428 <https://doi.org/10.1016/j.lab.2004.03.012>

429 Cai, Qingxian, Yang, M., Liu, D., Chen, J., Shu, D., Xia, J., Liao, X., Gu, Y., Cai, Qiue, Yang, Y.,
430 Shen, C., Li, X., Peng, L., Huang, D., Zhang, J., Zhang, S., Wang, F., Liu, J., Chen, L.,
431 Chen, S., Wang, Z., Zhang, Z., Cao, R., Zhong, W., Liu, Y., Liu, L., 2020. Experimental
432 Treatment with Favipiravir for COVID-19: An Open-Label Control Study. *Engineering*.
433 <https://doi.org/10.1016/J.ENG.2020.03.007>

434 Caly, L., Druce, J.D., Catton, M.G., Jans, D.A., Wagstaff, K.M., 2020. The FDA-approved Drug
435 Ivermectin inhibits the replication of SARS-CoV-2 in vitro. *Antiviral Res.* 104787.
436 <https://doi.org/10.1016/J.ANTIVIRAL.2020.104787>

437 Cao, B., Wang, Y., Wen, D., Liu, W., Wang, Jingli, Fan, G., Ruan, L., Song, B., Cai, Y., Wei, M.,
438 Li, X., Xia, J., Chen, N., Xiang, J., Yu, T., Bai, T., Xie, X., Zhang, L., Li, C., Yuan, Y., Chen,

439 H., Li, Huadong, Huang, H., Tu, S., Gong, F., Liu, Y., Wei, Y., Dong, C., Zhou, F., Gu, X.,
440 Xu, J., Liu, Z., Zhang, Y., Li, Hui, Shang, L., Wang, K., Li, K., Zhou, X., Dong, X., Qu, Z.,
441 Lu, S., Hu, X., Ruan, S., Luo, S., Wu, J., Peng, L., Cheng, F., Pan, L., Zou, J., Jia, C.,
442 Wang, Juan, Liu, X., Wang, S., Wu, X., Ge, Q., He, J., Zhan, H., Qiu, F., Guo, L., Huang,
443 C., Jaki, T., Hayden, F.G., Horby, P.W., Zhang, D., Wang, C., 2020. A Trial of Lopinavir-
444 Ritonavir in Adults Hospitalized with Severe Covid-19. *N. Engl. J. Med.* NEJMoa2001282.
445 <https://doi.org/10.1056/NEJMoa2001282>

446 Chen, C., Huang, J., Cheng, Z., Wu, J., Chen, S., Zhang, Y., Chen, B., Lu, M., Luo, Y., Zhang,
447 J., Yin, P., Wang, X., 2020. Favipiravir versus Arbidol for COVID-19: A Randomized
448 Clinical Trial. *medRxiv* 2020.03.17.20037432.
449 <https://doi.org/10.1101/2020.03.17.20037432>

450 Chen, L., Xiong, J., Bao, L., Shi, Y., 2020. Convalescent plasma as a potential therapy for
451 COVID-19. *Lancet. Infect. Dis.* 20, 398–400. [https://doi.org/10.1016/S1473-](https://doi.org/10.1016/S1473-3099(20)30141-9)
452 [3099\(20\)30141-9](https://doi.org/10.1016/S1473-3099(20)30141-9)

453 Cheng, Y., Wong, R., Soo, Y.O.Y., Wong, W.S., Lee, C.K., Ng, M.H.L., Chan, P., Wong, K.C.,
454 Leung, C.B., Cheng, G., 2005. Use of convalescent plasma therapy in SARS patients in
455 Hong Kong. *Eur. J. Clin. Microbiol. Infect. Dis.* 24, 44–6. [https://doi.org/10.1007/s10096-](https://doi.org/10.1007/s10096-004-1271-9)
456 [004-1271-9](https://doi.org/10.1007/s10096-004-1271-9)

457 Chu, C.M., Cheng, V.C.C., Hung, I.F.N., Wong, M.M.L., Chan, K.H., Chan, K.S., Kao, R.Y.T.,
458 Poon, L.L.M., Wong, C.L.P., Guan, Y., Peiris, J.S.M., Yuen, K.Y., HKU/UCH SARS Study
459 Group, 2004. Role of lopinavir/ritonavir in the treatment of SARS: initial virological and
460 clinical findings. *Thorax* 59, 252–6. <https://doi.org/10.1136/thorax.2003.012658>

461 Cinatl, J., Morgenstern, B., Bauer, G., Chandra, P., Rabenau, H., Doerr, H.W., 2003. Treatment
462 of SARS with human interferons. *Lancet (London, England)* 362, 293–4.
463 [https://doi.org/10.1016/s0140-6736\(03\)13973-6](https://doi.org/10.1016/s0140-6736(03)13973-6)

464 Clinical management of severe acute respiratory infection when COVID-19 is suspected [WWW
465 Document], n.d. URL [https://www.who.int/publications-detail/clinical-management-of-](https://www.who.int/publications-detail/clinical-management-of-severe-acute-respiratory-infection-when-novel-coronavirus-(ncov)-infection-is-suspected)
466 [severe-acute-respiratory-infection-when-novel-coronavirus-\(ncov\)-infection-is-suspected](https://www.who.int/publications-detail/clinical-management-of-severe-acute-respiratory-infection-when-novel-coronavirus-(ncov)-infection-is-suspected)
467 (accessed 4.4.20).

468 Conti, P., Gallenga, C.E., Tetè, G., Caraffa, A., Ronconi, G., Younes, A., Toniato, E., Ross, R.,

469 Kritas, S.K., 2020a. How to reduce the likelihood of coronavirus-19 (CoV-19 or SARS-CoV-
470 2) infection and lung inflammation mediated by IL-1. *J. Biol. Regul. Homeost. Agents* 34.
471 <https://doi.org/10.23812/Editorial-Conti-2>

472 Conti, P., Ronconi, G., Caraffa, A., Gallenga, C., Ross, R., Frydas, I., Kritas, S., 2020b.
473 Induction of pro-inflammatory cytokines (IL-1 and IL-6) and lung inflammation by
474 Coronavirus-19 (COVI-19 or SARS-CoV-2): anti-inflammatory strategies. *J. Biol. Regul.*
475 *Homeost. Agents* 34. <https://doi.org/10.23812/CONTI-E>

476 Coutard, B., Valle, C., de Lamballerie, X., Canard, B., Seidah, N.G., Decroly, E., 2020. The
477 spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site
478 absent in CoV of the same clade. *Antiviral Res.* 176, 104742.
479 <https://doi.org/10.1016/j.antiviral.2020.104742>

480 Crotty, S., Maag, D., Arnold, J.J., Zhong, W., Lau, J.Y.N., Hong, Z., Andino, R., Cameron, C.E.,
481 2000. The broad-spectrum antiviral ribonucleoside ribavirin is an RNA virus mutagen. *Nat.*
482 *Med.* 6, 1375–1379. <https://doi.org/10.1038/82191>

483 de Wilde, A.H., Jochmans, D., Posthuma, C.C., Zevenhoven-Dobbe, J.C., van Nieuwkoop, S.,
484 Bestebroer, T.M., van den Hoogen, B.G., Neyts, J., Snijder, E.J., 2014. Screening of an
485 FDA-approved compound library identifies four small-molecule inhibitors of Middle East
486 respiratory syndrome coronavirus replication in cell culture. *Antimicrob. Agents Chemother.*
487 58, 4875–84. <https://doi.org/10.1128/AAC.03011-14>

488 de Wit, E., Feldmann, F., Cronin, J., Jordan, R., Okumura, A., Thomas, T., Scott, D., Cihlar, T.,
489 Feldmann, H., 2020. Prophylactic and therapeutic remdesivir (GS-5734) treatment in the
490 rhesus macaque model of MERS-CoV infection. *Proc. Natl. Acad. Sci. U. S. A.* 117, 6771–
491 6776. <https://doi.org/10.1073/pnas.1922083117>

492 de Wit, E., van Doremalen, N., Falzarano, D., Munster, V.J., 2016. SARS and MERS: recent
493 insights into emerging coronaviruses. *Nat. Rev. Microbiol.* 14, 523–34.
494 <https://doi.org/10.1038/nrmicro.2016.81>

495 Delang, L., Segura Guerrero, N., Tas, A., Quérat, G., Pastorino, B., Froeyen, M., Dallmeier, K.,
496 Jochmans, D., Herdewijn, P., Bello, F., Snijder, E.J., de Lamballerie, X., Martina, B., Neyts,
497 J., van Hemert, M.J., Leyssen, P., 2014. Mutations in the chikungunya virus non-structural
498 proteins cause resistance to favipiravir (T-705), a broad-spectrum antiviral. *J. Antimicrob.*

499 Chemother. 69, 2770–2784. <https://doi.org/10.1093/jac/dku209>

500 Deng, L., Li, C., Zeng, Q., Liu, X., Li, X., Zhang, H., Hong, Z., Xia, J., 2020. Arbidol combined
501 with LPV/r versus LPV/r alone against Corona Virus Disease 2019: A retrospective cohort
502 study. *J. Infect.* <https://doi.org/10.1016/j.jinf.2020.03.002>

503 Deretic, V., Timmins, G.S., 2020. Azithromycin and ciprofloxacin have a chloroquine-like effect
504 on respiratory epithelial cells. *bioRxiv* 2020.03.29.008631.
505 <https://doi.org/10.1101/2020.03.29.008631>

506 Falzarano, D., de Wit, E., Martellaro, C., Callison, J., Munster, V.J., Feldmann, H., 2013.
507 Inhibition of novel β coronavirus replication by a combination of interferon- α 2b and
508 ribavirin. *Sci. Rep.* 3, 1686. <https://doi.org/10.1038/srep01686>

509 Feld, J.J., Hoofnagle, J.H., 2005. Mechanism of action of interferon and ribavirin in treatment of
510 hepatitis C. *Nature* 436, 967–972. <https://doi.org/10.1038/nature04082>

511 Ferron, F., Subissi, L., Silveira De Moraes, A.T., Le, N.T.T., Sevajol, M., Gluais, L., Decroly, E.,
512 Vonrhein, C., Bricogne, G., Canard, B., Imbert, I., 2018. Structural and molecular basis of
513 mismatch correction and ribavirin excision from coronavirus RNA. *Proc. Natl. Acad. Sci. U.*
514 *S. A.* 115, E162–E171. <https://doi.org/10.1073/pnas.1718806115>

515 Furuta, Y., Takahashi, K., Fukuda, Y., Kuno, M., Kamiyama, T., Kozaki, K., Nomura, N., Egawa,
516 H., Minami, S., Watanabe, Y., Narita, H., Shiraki, K., 2002. In vitro and in vivo activities of
517 anti-influenza virus compound T-705. *Antimicrob. Agents Chemother.* 46, 977–81.
518 <https://doi.org/10.1128/aac.46.4.977-981.2002>

519 Gao, J., Tian, Z., Yang, X., 2020. Breakthrough: Chloroquine phosphate has shown apparent
520 efficacy in treatment of COVID-19 associated pneumonia in clinical studies. *Biosci. Trends*
521 14, 72–73. <https://doi.org/10.5582/bst.2020.01047>

522 Gautret, P., Lagier, J.-C., Parola, P., Hoang, V.T., Meddeb, L., Mailhe, M., Doudier, B., Courjon,
523 J., Giordanengo, V., Vieira, V.E., Dupont, H.T., Honoré, S., Colson, P., Chabrière, E., La
524 Scola, B., Rolain, J.-M., Brouqui, P., Raoult, D., 2020. Hydroxychloroquine and
525 azithromycin as a treatment of COVID-19: results of an open-label non-randomized clinical
526 trial. *Int. J. Antimicrob. Agents* 105949. <https://doi.org/10.1016/j.ijantimicag.2020.105949>

527 Gielen, V., Johnston, S.L., Edwards, M.R., 2010. Azithromycin induces anti-viral responses in
528 bronchial epithelial cells. *Eur. Respir. J.* 36, 646–654.
529 <https://doi.org/10.1183/09031936.00095809>

530 Gordon, C.J., Tchesnokov, E.P., Woolner, E., Perry, J.K., Feng, J.Y., Porter, D.P., Gotte, M.,
531 2020a. Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase
532 from severe acute respiratory syndrome coronavirus 2 with high potency. *J. Biol. Chem.*
533 *jbc.RA120.013679*. <https://doi.org/10.1074/jbc.RA120.013679>

534 Gordon, C.J., Tchesnokov, E.P., Woolner, E., Perry, J.K., Feng, J.Y., Porter, D.P., Gotte, M.,
535 2020b. Remdesivir is a direct-acting antiviral that inhibits RNA-dependent RNA polymerase
536 from severe acute respiratory syndrome coronavirus 2 with high potency. *J. Biol. Chem.*
537 *jbc.RA120.013679*. <https://doi.org/10.1074/JBC.RA120.013679>

538 Greenough, T.C., Babcock, G.J., Roberts, A., Hernandez, H.J., Thomas, Jr., W.D., Coccia, J.A.,
539 Graziano, R.F., Srinivasan, M., Lowy, I., Finberg, R.W., Subbarao, K., Vogel, L.,
540 Somasundaran, M., Luzuriaga, K., Sullivan, J.L., Ambrosino, D.M., 2005. Development and
541 Characterization of a Severe Acute Respiratory Syndrome–Associated Coronavirus–
542 Neutralizing Human Monoclonal Antibody That Provides Effective Immunoprophylaxis in
543 Mice. *J. Infect. Dis.* 191, 507–514. <https://doi.org/10.1086/427242>

544 Grein, J., Ohmagari, N., Shin, D., Diaz, G., Asperges, E., Castagna, A., Feldt, T., Green, G.,
545 Green, M.L., Lescure, F.-X., Nicastri, E., Oda, R., Yo, K., Quiros-Roldan, E., Studemeister,
546 A., Redinski, J., Ahmed, S., Bernetti, J., Chelliah, D., Chen, D., Chihara, S., Cohen, S.H.,
547 Cunningham, J., D’Arminio Monforte, A., Ismail, S., Kato, H., Lapadula, G., L’Her, E.,
548 Maeno, T., Majumder, S., Massari, M., Mora-Rillo, M., Mutoh, Y., Nguyen, D., Verweij, E.,
549 Zoufaly, A., Osinusi, A.O., DeZure, A., Zhao, Y., Zhong, L., Chokkalingam, A.,
550 Elboudwarej, E., Telep, L., Timbs, L., Henne, I., Sellers, S., Cao, H., Tan, S.K.,
551 Winterbourne, L., Desai, P., Mera, R., Gaggar, A., Myers, R.P., Brainard, D.M., Childs, R.,
552 Flanigan, T., 2020. Compassionate Use of Remdesivir for Patients with Severe Covid-19.
553 *N. Engl. J. Med.* *NEJMoa2007016*. <https://doi.org/10.1056/NEJMoa2007016>

554 Gubareva, L. V, Kaiser, L., Hayden, F.G., 2000. Influenza virus neuraminidase inhibitors. *Lancet*
555 (London, England) 355, 827–35. [https://doi.org/10.1016/S0140-6736\(99\)11433-8](https://doi.org/10.1016/S0140-6736(99)11433-8)

556 Guedj, J., Piorkowski, G., Jacquot, F., Madelain, V., Nguyen, T.H.T., Rodallec, A., Gunther, S.,

557 Carbonnelle, C., Mentré, F., Raoul, H., de Lamballerie, X., 2018. Antiviral efficacy of
558 favipiravir against Ebola virus: A translational study in cynomolgus macaques. *PLOS Med.*
559 15, e1002535. <https://doi.org/10.1371/journal.pmed.1002535>

560 He, Y., Lu, H., Siddiqui, P., Zhou, Y., Jiang, S., 2005. Receptor-Binding Domain of Severe
561 Acute Respiratory Syndrome Coronavirus Spike Protein Contains Multiple Conformation-
562 Dependent Epitopes that Induce Highly Potent Neutralizing Antibodies. *J. Immunol.* 174,
563 4908–4915. <https://doi.org/10.4049/jimmunol.174.8.4908>

564 Hensley, L.E., Fritz, E.A., Jahrling, P.B., Karp, C., Huggins, J.W., Geisbert, T.W., 2004.
565 Interferon- β 1a and SARS Coronavirus Replication. *Emerg. Infect. Dis.* 10, 317–319.
566 <https://doi.org/10.3201/eid1002.030482>

567 Hoenen, T., Groseth, A., Feldmann, H., 2019. Therapeutic strategies to target the Ebola virus
568 life cycle. *Nat. Rev. Microbiol.* 17, 593–606. <https://doi.org/10.1038/s41579-019-0233-2>

569 Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S.,
570 Schiergens, T.S., Herrler, G., Wu, N.-H., Nitsche, A., Müller, M.A., Drosten, C., Pöhlmann,
571 S., 2020a. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a
572 Clinically Proven Protease Inhibitor. *Cell.* <https://doi.org/10.1016/j.cell.2020.02.052>

573 Hoffmann, M., Schroeder, S., Kleine-Weber, H., Müller, M.A., Drosten, C., Pöhlmann, S., 2020b.
574 Nafamostat mesylate blocks activation of SARS-CoV-2: New treatment option The
575 currently unfolding coronavirus pandemic threatens health systems and economies.
576 <https://doi.org/10.1128/AAC.00754-20>

577 Hoffmann, M., Schroeder, S., Kleine-Weber, H., Müller, M.A., Drosten, C., Pöhlmann, S., 2020c.
578 Nafamostat mesylate blocks activation of SARS-CoV-2: New treatment option for COVID-
579 19. *Antimicrob. Agents Chemother.* <https://doi.org/10.1128/AAC.00754-20>

580 Hoffmann Markus et al, 2020. A multibasic cleavage site in the spike protein of SARS-CoV-2 is
581 essential for infection of human lung cells. <https://doi.org/10.1016/j.molcel.2020.04.022>

582 Hofmann, H., Hattermann, K., Marzi, A., Gramberg, T., Geier, M., Krumbiegel, M., Kuate, S.,
583 Überla, K., Niedrig, M., Pöhlmann, S., 2004. S Protein of Severe Acute Respiratory
584 Syndrome-Associated Coronavirus Mediates Entry into Hepatoma Cell Lines and Is
585 Targeted by Neutralizing Antibodies in Infected Patients. *J. Virol.* 78, 6134–6142.

586 <https://doi.org/10.1128/JVI.78.12.6134-6142.2004>

587 Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., Cheng,
588 Z., Yu, T., Xia, J., Wei, Y., Wu, W., Xie, X., Yin, W., Li, H., Liu, M., Xiao, Y., Gao, H., Guo,
589 L., Xie, J., Wang, G., Jiang, R., Gao, Z., Jin, Q., Wang, J., Cao, B., 2020. Clinical features
590 of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* (London,
591 England) 395, 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)

592 Jeon, S., Ko, M., Lee, J., Choi, I., Byun, S.Y., Park, S., Shum, D., Kim, S., 2020. Identification of
593 antiviral drug candidates against SARS-CoV-2 from FDA-approved drugs. *bioRxiv*
594 2020.03.20.999730. <https://doi.org/10.1101/2020.03.20.999730>

595 Kadam, R.U., Wilson, I.A., 2017. Structural basis of influenza virus fusion inhibition by the
596 antiviral drug Arbidol. *Proc. Natl. Acad. Sci.* 114, 206–214.
597 <https://doi.org/10.1073/pnas.1617020114>

598 Keyaerts, E., Vijgen, L., Maes, P., Neyts, J., Van Ranst, M., 2004. In vitro inhibition of severe
599 acute respiratory syndrome coronavirus by chloroquine. *Biochem. Biophys. Res. Commun.*
600 323, 264–8. <https://doi.org/10.1016/j.bbrc.2004.08.085>

601 Khamitov, R.A., Loginova, S.I., Shchukina, V.N., Borisevich, S. V, Maksimov, V.A., Shuster,
602 A.M., n.d. [Antiviral activity of arbidol and its derivatives against the pathogen of severe
603 acute respiratory syndrome in the cell cultures]. *Vopr. Virusol.* 53, 9–13.

604 Kim, J.-A., Seong, R.-K., Kumar, M., Shin, O.S., 2018. Favipiravir and Ribavirin Inhibit
605 Replication of Asian and African Strains of Zika Virus in Different Cell Models. *Viruses* 10,
606 72. <https://doi.org/10.3390/v10020072>

607 Ksiazek, T.G., Erdman, D., Goldsmith, C.S., Zaki, S.R., Peret, T., Emery, S., Tong, S., Urbani,
608 C., Comer, J.A., Lim, W., Rollin, P.E., Dowell, S.F., Ling, A.-E., Humphrey, C.D., Shieh,
609 W.-J., Guarner, J., Paddock, C.D., Rota, P., Fields, B., DeRisi, J., Yang, J.-Y., Cox, N.,
610 Hughes, J.M., LeDuc, J.W., Bellini, W.J., Anderson, L.J., SARS Working Group, 2003. A
611 novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.*
612 348, 1953–66. <https://doi.org/10.1056/NEJMoa030781>

613 Kuster, G.M., Pfister, O., Burkard, T., Zhou, Q., Twerenbold, R., Haaf, P., Widmer, A.F.,
614 Osswald, S., 2020. SARS-CoV2: should inhibitors of the renin–angiotensin system be

615 withdrawn in patients with COVID-19? *Eur. Heart J.*
616 <https://doi.org/10.1093/eurheartj/ehaa235>

617 Launch of a European clinical trial against COVID-19 | Newsroom | Inserm [WWW Document],
618 n.d. URL [https://presse.inserm.fr/en/launch-of-a-european-clinical-trial-against-covid-](https://presse.inserm.fr/en/launch-of-a-european-clinical-trial-against-covid-19/38737/)
619 [19/38737/](https://presse.inserm.fr/en/launch-of-a-european-clinical-trial-against-covid-19/38737/) (accessed 4.6.20).

620 Li, F., 2015. Receptor recognition mechanisms of coronaviruses: a decade of structural studies.
621 *J. Virol.* 89, 1954–64. <https://doi.org/10.1128/JVI.02615-14>

622 Lim, J., Jeon, S., Shin, H.Y., Kim, M.J., Seong, Y.M., Lee, W.J., Choe, K.W., Kang, Y.M., Lee,
623 B., Park, S.J., 2020. Case of the Index Patient Who Caused Tertiary Transmission of
624 COVID-19 Infection in Korea: the Application of Lopinavir/Ritonavir for the Treatment of
625 COVID-19 Infected Pneumonia Monitored by Quantitative RT-PCR. *J. Korean Med. Sci.*
626 35, e79. <https://doi.org/10.3346/jkms.2020.35.e79>

627 Liu, J., Cao, R., Xu, M., Wang, X., Zhang, H., Hu, H., Li, Y., Hu, Z., Zhong, W., Wang, M., 2020.
628 Hydroxychloroquine, a less toxic derivative of chloroquine, is effective in inhibiting SARS-
629 CoV-2 infection in vitro. *Cell Discov.* 6, 16. <https://doi.org/10.1038/s41421-020-0156-0>

630 Lundberg, L., Pinkham, C., Baer, A., Amaya, M., Narayanan, A., Wagstaff, K.M., Jans, D.A.,
631 Kehn-Hall, K., 2013. Nuclear import and export inhibitors alter capsid protein distribution in
632 mammalian cells and reduce Venezuelan Equine Encephalitis Virus replication. *Antiviral*
633 *Res.* 100, 662–672. <https://doi.org/10.1016/J.ANTIVIRAL.2013.10.004>

634 Mair-Jenkins, J., Saavedra-Campos, M., Baillie, J.K., Cleary, P., Khaw, F.-M., Lim, W.S., Makki,
635 S., Rooney, K.D., Nguyen-Van-Tam, J.S., Beck, C.R., Convalescent Plasma Study Group,
636 2015. The effectiveness of convalescent plasma and hyperimmune immunoglobulin for the
637 treatment of severe acute respiratory infections of viral etiology: a systematic review and
638 exploratory meta-analysis. *J. Infect. Dis.* 211, 80–90. <https://doi.org/10.1093/infdis/jiu396>

639 Matsuyama, S., Nagata, N., Shirato, K., Kawase, M., Takeda, M., Taguchi, F., 2010. Efficient
640 activation of the severe acute respiratory syndrome coronavirus spike protein by the
641 transmembrane protease TMPRSS2. *J. Virol.* 84, 12658–64.
642 <https://doi.org/10.1128/JVI.01542-10>

643 Meng, J., Xiao, G., Zhang, J., He, X., Ou, M., Bi, J., Yang, R., Di, W., Wang, Z., Li, Z., Gao, H.,

644 Liu, L., Zhang, G., 2020. Renin-angiotensin system inhibitors improve the clinical outcomes
645 of COVID-19 patients with hypertension. *Emerg. Microbes Infect.* 9, 757–760.
646 <https://doi.org/10.1080/22221751.2020.1746200>

647 Millet, J.K., Whittaker, G.R., 2015. Host cell proteases: Critical determinants of coronavirus
648 tropism and pathogenesis. *Virus Res.* 202, 120–34.
649 <https://doi.org/10.1016/j.virusres.2014.11.021>

650 Monteil, V., Kwon, H., Prado, P., Hagelkrüys, A., Wimmer, R.A., Stahl, M., Leopoldi, A., Garreta,
651 E., Hurtado del Pozo, C., Prosper, F., Romero, J.P., Wirnsberger, G., Zhang, H., Slutsky,
652 A.S., Conder, R., Montserrat, N., Mirazimi, A., Penninger, J.M., 2020. Inhibition of SARS-
653 CoV-2 Infections in Engineered Human Tissues Using Clinical-Grade Soluble Human
654 ACE2. *Cell.* <https://doi.org/10.1016/j.cell.2020.04.004>

655 Oestereich, L., Lüdtke, A., Wurr, S., Rieger, T., Muñoz-Fontela, C., Günther, S., 2014.
656 Successful treatment of advanced Ebola virus infection with T-705 (favipiravir) in a small
657 animal model. *Antiviral Res.* 105, 17–21. <https://doi.org/10.1016/j.antiviral.2014.02.014>

658 Pantoliano, M.W., Petrella, E.C., Kwasnoski, J.D., Lobanov, V.S., Myslik, J., Graf, E., Carver, T.,
659 Asel, E., Springer, B.A., Lane, P., Salemme, F.R., 2001. High-density miniaturized thermal
660 shift assays as a general strategy for drug discovery. *J. Biomol. Screen.* 6, 429–40.
661 <https://doi.org/10.1177/108705710100600609>

662 PARAGAS, J., BLATT, L., HARTMANN, C., HUGGINS, J., ENDY, T., 2005. Interferon alfacon1
663 is an inhibitor of SARS-corona virus in cell-based models. *Antiviral Res.* 66, 99–102.
664 <https://doi.org/10.1016/j.antiviral.2005.01.002>

665 Patel, A., 2020. Usefulness of Ivermectin in COVID-19 Illness.

666 Pruijssers, A.J., Denison, M.R., 2019. Nucleoside analogues for the treatment of coronavirus
667 infections. *Curr. Opin. Virol.* <https://doi.org/10.1016/j.coviro.2019.04.002>

668 Riva, L., Yuan, S., Yin, X., Martin-Sancho, L., Matsunaga, N., Burgstaller-Muehlbacher, S.,
669 Pache, L., Jesus, P.P. De, Hull, M. V., Chang, M., Chan, J.F.-W., Cao, J., Poon, V.K.-M.,
670 Herbert, K., Nguyen, T.-T., Pu, Y., Nguyen, C., Rubanov, A., Martinez-Sobrido, L., Liu, W.-
671 C., Miorin, L., White, K.M., Johnson, J.R., Benner, C., Sun, R., Schultz, P.G., Su, A.,
672 Garcia-Sastre, A., Chatterjee, A.K., Yuen, K.-Y., Chanda, S.K., 2020. A Large-scale Drug

673 Repositioning Survey for SARS-CoV-2 Antivirals. bioRxiv 2020.04.16.044016.
674 <https://doi.org/10.1101/2020.04.16.044016>

675 Rocha-Pereira, J., Jochmans, D., Dallmeier, K., Leyssen, P., Nascimento, M.S.J., Neyts, J.,
676 2012. Favipiravir (T-705) inhibits in vitro norovirus replication. *Biochem. Biophys. Res.*
677 *Commun.* 424, 777–780. <https://doi.org/10.1016/j.bbrc.2012.07.034>

678 Savarino, A., Boelaert, J.R., Cassone, A., Majori, G., Cauda, R., 2003. Effects of chloroquine on
679 viral infections: an old drug against today's diseases. *Lancet Infect. Dis.* 3, 722–727.
680 [https://doi.org/10.1016/S1473-3099\(03\)00806-5](https://doi.org/10.1016/S1473-3099(03)00806-5)

681 Schrezenmeier, E., Dörner, T., 2020. Mechanisms of action of hydroxychloroquine and
682 chloroquine: implications for rheumatology. *Nat. Rev. Rheumatol.* 16, 155–166.
683 <https://doi.org/10.1038/s41584-020-0372-x>

684 Sheahan, T.P., Sims, A.C., Graham, R.L., Menachery, V.D., Gralinski, L.E., Case, J.B., Leist,
685 S.R., Pirc, K., Feng, J.Y., Trantcheva, I., Bannister, R., Park, Y., Babusis, D., Clarke, M.O.,
686 Mackman, R.L., Spahn, J.E., Palmiotti, C.A., Siegel, D., Ray, A.S., Cihlar, T., Jordan, R.,
687 Denison, M.R., Baric, R.S., 2017. Broad-spectrum antiviral GS-5734 inhibits both epidemic
688 and zoonotic coronaviruses. *Sci. Transl. Med.* 9, eaal3653.
689 <https://doi.org/10.1126/scitranslmed.aal3653>

690 Sheahan, T.P., Sims, A.C., Leist, S.R., Schäfer, A., Won, J., Brown, A.J., Montgomery, S.A.,
691 Hogg, A., Babusis, D., Clarke, M.O., Spahn, J.E., Bauer, L., Sellers, S., Porter, D., Feng,
692 J.Y., Cihlar, T., Jordan, R., Denison, M.R., Baric, R.S., 2020a. Comparative therapeutic
693 efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against
694 MERS-CoV. *Nat. Commun.* 11, 222. <https://doi.org/10.1038/s41467-019-13940-6>

695 Sheahan, T.P., Sims, A.C., Zhou, S., Graham, R.L., Pruijssers, A.J., Agostini, M.L., Leist, S.R.,
696 Schäfer, A., Dinnon, K.H., Stevens, L.J., Chappell, J.D., Lu, X., Hughes, T.M., George,
697 A.S., Hill, C.S., Montgomery, S.A., Brown, A.J., Bluemling, G.R., Natchus, M.G., Saindane,
698 M., Kolykhalov, A.A., Painter, G., Harcourt, J., Tamin, A., Thornburg, N.J., Swanstrom, R.,
699 Denison, M.R., Baric, R.S., 2020b. An orally bioavailable broad-spectrum antiviral inhibits
700 SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice.
701 *Sci. Transl. Med.* eabb5883. <https://doi.org/10.1126/scitranslmed.abb5883>

702 Shen, C., Wang, Z., Zhao, F., Yang, Yang, Li, J., Yuan, J., Wang, F., Li, D., Yang, M., Xing, L.,

703 Wei, J., Xiao, H., Yang, Yan, Qu, J., Qing, L., Chen, L., Xu, Z., Peng, L., Li, Y., Zheng, H.,
704 Chen, F., Huang, K., Jiang, Y., Liu, D., Zhang, Z., Liu, Y., Liu, L., 2020. Treatment of 5
705 Critically Ill Patients With COVID-19 With Convalescent Plasma. JAMA.
706 <https://doi.org/10.1001/jama.2020.4783>

707 Smith, E.C., Blanc, H., Surdel, M.C., Vignuzzi, M., Denison, M.R., 2013. Coronaviruses lacking
708 exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading
709 and potential therapeutics. PLoS Pathog. 9, e1003565.
710 <https://doi.org/10.1371/journal.ppat.1003565>

711 Smither, S.J., Eastaugh, L.S., Steward, J.A., Nelson, M., Lenk, R.P., Lever, M.S., 2014. Post-
712 exposure efficacy of Oral T-705 (Favipiravir) against inhalational Ebola virus infection in a
713 mouse model. Antiviral Res. 104, 153–155. <https://doi.org/10.1016/j.antiviral.2014.01.012>

714 Statement on IJAA paper | International Society of Antimicrobial Chemotherapy [WWW
715 Document], n.d. URL <https://www.isac.world/news-and-publications/official-isac-statement>
716 (accessed 4.30.20).

717 Ströher, U., DiCaro, A., Li, Y., Strong, J.E., Aoki, F., Plummer, F., Jones, S.M., Feldmann, H.,
718 2004. Severe Acute Respiratory Syndrome–Related Coronavirus Is Inhibited by Interferon-
719 α . J. Infect. Dis. 189, 1164–1167. <https://doi.org/10.1086/382597>

720 Sui, J., Li, W., Murakami, A., Tamin, A., Matthews, L.J., Wong, S.K., Moore, M.J., Tallarico,
721 A.S.C., Olurinde, M., Choe, H., Anderson, L.J., Bellini, W.J., Farzan, M., Marasco, W.A.,
722 2004. Potent neutralization of severe acute respiratory syndrome (SARS) coronavirus by a
723 human mAb to S1 protein that blocks receptor association. Proc. Natl. Acad. Sci. U. S. A.
724 101, 2536–41. <https://doi.org/10.1073/pnas.0307140101>

725 Tay, M.Y.F., Fraser, J.E., Chan, W.K.K., Moreland, N.J., Rathore, A.P., Wang, C., Vasudevan,
726 S.G., Jans, D.A., 2013. Nuclear localization of dengue virus (DENV) 1–4 non-structural
727 protein 5; protection against all 4 DENV serotypes by the inhibitor Ivermectin. Antiviral Res.
728 99, 301–306. <https://doi.org/10.1016/J.ANTIVIRAL.2013.06.002>

729 Taylor, S.P., Sellers, E., Taylor, B.T., 2015. Azithromycin for the Prevention of COPD
730 Exacerbations: The Good, Bad, and Ugly. Am. J. Med. 128, 1362.e1–6.
731 <https://doi.org/10.1016/j.amjmed.2015.07.032>

732 Thorlund, K., Dron, L., Park, J., Hsu, G., Forrest, J.I., Mills, E.J., 2020. A real-time dashboard of
733 clinical trials for COVID-19. *Lancet Digit. Heal.* 0. [https://doi.org/10.1016/S2589-](https://doi.org/10.1016/S2589-7500(20)30086-8)
734 [7500\(20\)30086-8](https://doi.org/10.1016/S2589-7500(20)30086-8)

735 Totura, A.L., Whitmore, A., Agnihothram, S., Schäfer, A., Katze, M.G., Heise, M.T., Baric, R.S.,
736 2015. Toll-Like Receptor 3 Signaling via TRIF Contributes to a Protective Innate Immune
737 Response to Severe Acute Respiratory Syndrome Coronavirus Infection. *MBio* 6, e00638-
738 15. <https://doi.org/10.1128/mBio.00638-15>

739 Touret, F., Gilles, M., Barral, K., Nougairede, A., Decroly, E., Lamballerie, X. de, Coutard, B.,
740 2020. In vitro screening of a FDA approved chemical library reveals potential inhibitors of
741 SARS-CoV-2 replication. *bioRxiv* 2020.04.03.023846.
742 <https://doi.org/10.1101/2020.04.03.023846>

743 Ulrich, H., Pillat, M.M., 2020. CD147 as a Target for COVID-19 Treatment: Suggested Effects of
744 Azithromycin and Stem Cell Engagement. *Stem Cell Rev. Reports*.
745 <https://doi.org/10.1007/s12015-020-09976-7>

746 Vigant, F., Santos, N.C., Lee, B., 2015. Broad-spectrum antivirals against viral fusion. *Nat. Rev.*
747 *Microbiol.* 13, 426–437. <https://doi.org/10.1038/nrmicro3475>

748 Vincent, M.J., Bergeron, E., Benjannet, S., Erickson, B.R., Rollin, P.E., Ksiazek, T.G., Seidah,
749 N.G., Nichol, S.T., 2005. Chloroquine is a potent inhibitor of SARS coronavirus infection
750 and spread. *Viol. J.* 2, 69. <https://doi.org/10.1186/1743-422X-2-69>

751 Wagstaff, K.M., Sivakumaran, H., Heaton, S.M., Harrich, D., Jans, D.A., 2012. Ivermectin is a
752 specific inhibitor of importin α/β -mediated nuclear import able to inhibit replication of HIV-1
753 and dengue virus. *Biochem. J.* 443, 851–6. <https://doi.org/10.1042/BJ20120150>

754 Walls, A.C., Park, Y.-J., Tortorici, M.A., Wall, A., McGuire, A.T., Velesler, D., 2020. Structure,
755 Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*.
756 <https://doi.org/10.1016/j.cell.2020.02.058>

757 Wan, Y., Shang, J., Graham, R., Baric, R.S., Li, F., 2020. Receptor Recognition by the Novel
758 Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS
759 Coronavirus. *J. Virol.* 94. <https://doi.org/10.1128/JVI.00127-20>

760 Wang, K., Chen, W., Zhou, Y.-S., Lian, J.-Q., Zhang, Z., Du, P., Gong, L., Zhang, Y., Cui, H.-Y.,
761 Geng, J.-J., Wang, B., Sun, X.-X., Wang, C.-F., Yang, X., Lin, P., Deng, Y.-Q., Wei, D.,
762 Yang, X.-M., Zhu, Y.-M., Zhang, K., Zheng, Z.-H., Miao, J.-L., Guo, T., Shi, Y., Zhang, J.,
763 Fu, L., Wang, Q.-Y., Bian, H., Zhu, P., Chen, Z.-N., 2020. SARS-CoV-2 invades host cells
764 via a novel route: CD147-spike protein. *bioRxiv* 2020.03.14.988345.
765 <https://doi.org/10.1101/2020.03.14.988345>

766 Wang, L., Shi, W., Chappell, J.D., Joyce, M.G., Zhang, Y., Kanekiyo, M., Becker, M.M., van
767 Doremalen, N., Fischer, R., Wang, N., Corbett, K.S., Choe, M., Mason, R.D., Van Galen,
768 J.G., Zhou, T., Saunders, K.O., Tatti, K.M., Haynes, L.M., Kwong, P.D., Modjarrad, K.,
769 Kong, W.-P., McLellan, J.S., Denison, M.R., Munster, V.J., Mascola, J.R., Graham, B.S.,
770 2018. Importance of Neutralizing Monoclonal Antibodies Targeting Multiple Antigenic Sites
771 on the Middle East Respiratory Syndrome Coronavirus Spike Glycoprotein To Avoid
772 Neutralization Escape. *J. Virol.* 92. <https://doi.org/10.1128/JVI.02002-17>

773 Wang, M., Cao, R., Zhang, L., Yang, X., Liu, J., Xu, M., Shi, Z., Hu, Z., Zhong, W., Xiao, G.,
774 2020. Remdesivir and chloroquine effectively inhibit the recently emerged novel
775 coronavirus (2019-nCoV) in vitro. *Cell Res.* 30, 269–271. [https://doi.org/10.1038/s41422-](https://doi.org/10.1038/s41422-020-0282-0)
776 [020-0282-0](https://doi.org/10.1038/s41422-020-0282-0)

777 Warren, T.K., Jordan, R., Lo, M.K., Ray, A.S., Mackman, R.L., Soloveva, V., Siegel, D., Perron,
778 M., Bannister, R., Hui, H.C., Larson, N., Strickley, R., Wells, J., Stuthman, K.S., Van
779 Tongeren, S.A., Garza, N.L., Donnelly, G., Shurtleff, A.C., Retterer, C.J., Gharaibeh, D.,
780 Zamani, R., Kenny, T., Eaton, B.P., Grimes, E., Welch, L.S., Gomba, L., Wilhelmsen, C.L.,
781 Nichols, D.K., Nuss, J.E., Nagle, E.R., Kugelman, J.R., Palacios, G., Doerffler, E., Neville,
782 S., Carra, E., Clarke, M.O., Zhang, L., Lew, W., Ross, B., Wang, Q., Chun, K., Wolfe, L.,
783 Babusis, D., Park, Y., Stray, K.M., Trancheva, I., Feng, J.Y., Barauskas, O., Xu, Y., Wong,
784 P., Braun, M.R., Flint, M., McMullan, L.K., Chen, S.-S., Fearn, R., Swaminathan, S.,
785 Mayers, D.L., Spiropoulou, C.F., Lee, W.A., Nichol, S.T., Cihlar, T., Bavari, S., 2016.
786 Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus
787 monkeys. *Nature* 531, 381–5. <https://doi.org/10.1038/nature17180>

788 Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., Hu, Y., Tao, Z.-W., Tian, J.-H.,
789 Pei, Y.-Y., Yuan, M.-L., Zhang, Y.-L., Dai, F.-H., Liu, Y., Wang, Q.-M., Zheng, J.-J., Xu, L.,
790 Holmes, E.C., Zhang, Y.-Z., 2020. A new coronavirus associated with human respiratory

791 disease in China. *Nature* 579, 265–269. <https://doi.org/10.1038/s41586-020-2008-3>

792 Yamamoto, Norio, Yang, R., Yoshinaka, Y., Amari, S., Nakano, T., Cinatl, J., Rabenau, H.,
793 Doerr, H.W., Hunsmann, G., Otaka, A., Tamamura, H., Fujii, N., Yamamoto, Naoki, 2004.
794 HIV protease inhibitor nelfinavir inhibits replication of SARS-associated coronavirus.
795 *Biochem. Biophys. Res. Commun.* 318, 719–25. <https://doi.org/10.1016/j.bbrc.2004.04.083>

796 Yao, X., Ye, F., Zhang, M., Cui, C., Huang, B., Niu, P., Liu, X., Zhao, L., Dong, E., Song, C.,
797 Zhan, S., Lu, R., Li, H., Tan, W., Liu, D., 2020. In Vitro Antiviral Activity and Projection of
798 Optimized Dosing Design of Hydroxychloroquine for the Treatment of Severe Acute
799 Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). *Clin. Infect. Dis.*
800 <https://doi.org/10.1093/cid/ciaa237>

801 Yeh, K.-M., Chiueh, T.-S., Siu, L.K., Lin, J.-C., Chan, P.K.S., Peng, M.-Y., Wan, H.-L., Chen, J.-
802 H., Hu, B.-S., Perng, C.-L., Lu, J.-J., Chang, F.-Y., 2005. Experience of using convalescent
803 plasma for severe acute respiratory syndrome among healthcare workers in a Taiwan
804 hospital. *J. Antimicrob. Chemother.* 56, 919–22. <https://doi.org/10.1093/jac/dki346>

805 Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D.M.E., Fouchier, R.A.M., 2012.
806 Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *N. Engl. J.*
807 *Med.* 367, 1814–1820. <https://doi.org/10.1056/NEJMoa1211721>

808 Zhang, C., Wu, Z., Li, J.-W., Zhao, H., Wang, G.-Q., 2020. The cytokine release syndrome
809 (CRS) of severe COVID-19 and Interleukin-6 receptor (IL-6R) antagonist Tocilizumab may
810 be the key to reduce the mortality. *Int. J. Antimicrob. Agents* 105954.
811 <https://doi.org/10.1016/J.IJANTIMICAG.2020.105954>

812 Zhang, L., Lin, D., Sun, X., Curth, U., Drosten, C., Sauerhering, L., Becker, S., Rox, K.,
813 Hilgenfeld, R., 2020. Crystal structure of SARS-CoV-2 main protease provides a basis for
814 design of improved α -ketoamide inhibitors. *Science* (80-). eabb3405.
815 <https://doi.org/10.1126/science.abb3405>

816 Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., Guan, L.,
817 Wei, Y., Li, H., Wu, X., Xu, J., Tu, S., Zhang, Y., Chen, H., Cao, B., 2020. Clinical course
818 and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a
819 retrospective cohort study. *Lancet* 395, 1054–1062. [https://doi.org/10.1016/S0140-](https://doi.org/10.1016/S0140-6736(20)30566-3)
820 [6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3)

821 Zhou, G., Zhao, Q., 2020. Perspectives on therapeutic neutralizing antibodies against the Novel
822 Coronavirus SARS-CoV-2. *Int. J. Biol. Sci.* 16, 1718–1723.
823 <https://doi.org/10.7150/ijbs.45123>

824 Zhu, Z., Chakraborti, S., He, Y., Roberts, A., Sheahan, T., Xiao, X., Hensley, L.E., Prabakaran,
825 P., Rockx, B., Sidorov, I.A., Corti, D., Vogel, L., Feng, Y., Kim, J.-O., Wang, L.-F., Baric, R.,
826 Lanzavecchia, A., Curtis, K.M., Nabel, G.J., Subbarao, K., Jiang, S., Dimitrov, D.S., 2007.
827 Potent cross-reactive neutralization of SARS coronavirus isolates by human monoclonal
828 antibodies. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12123–8.
829 <https://doi.org/10.1073/pnas.0701000104>

830 Ziegler, C.G.K., Allon, S.J., Nyquist, S.K., Mbano, I.M., Miao, V.N., Tzouanas, C.N., Cao, Y.,
831 Yousif, A.S., Bals, J., Hauser, B.M., Feldman, J., Muus, C., Wadsworth, M.H., Kazer, S.W.,
832 Hughes, T.K., Doran, B., Gatter, G.J., Vukovic, M., Taliaferro, F., Mead, B.E., Guo, Z.,
833 Wang, J.P., Gras, D., Plaisant, M., Ansari, M., Angelidis, I., Adler, H., Sucre, J.M.S., Taylor,
834 C.J., Lin, B., Waghray, A., Mitsialis, V., Dwyer, D.F., Buchheit, K.M., Boyce, J.A., Barrett,
835 N.A., Laidlaw, T.M., Carroll, S.L., Colonna, L., Tkachev, V., Peterson, C.W., Yu, A., Zheng,
836 H.B., Gideon, H.P., Winchell, C.G., Lin, P.L., Bingle, C.D., Snapper, S.B., Kropski, J.A.,
837 Theis, F.J., Schiller, H.B., Zaragosi, L.-E., Barbry, P., Leslie, A., Kiem, H.-P., Flynn, J.L.,
838 Fortune, S.M., Berger, B., Finberg, R.W., Kean, L.S., Garber, M., Schmidt, A.G., Lingwood,
839 D., Shalek, A.K., Ordovas-Montanes, J., 2020. SARS-CoV-2 receptor ACE2 is an
840 interferon-stimulated gene in human airway epithelial cells and is detected in specific cell
841 subsets across tissues. *Cell*. <https://doi.org/10.1016/J.CELL.2020.04.035>

842 Zmurko, J., Marques, R.E., Schols, D., Verbeken, E., Kaptein, S.J.F., Neyts, J., 2016. The Viral
843 Polymerase Inhibitor 7-Deaza-2'-C-Methyladenosine Is a Potent Inhibitor of In Vitro Zika
844 Virus Replication and Delays Disease Progression in a Robust Mouse Infection Model.
845 *PLoS Negl. Trop. Dis.* 10, e0004695. <https://doi.org/10.1371/journal.pntd.0004695>

846

Résumé

L'ordre des *Mononegavirales* regroupe des virus hautement pathogènes pour l'Homme tels que les virus de la rougeole, de la rage ou encore Ebola. Leur génome code pour une protéine L (« Large ») qui joue un rôle déterminant dans les étapes de transcription et réplication virale. La protéine L catalyse également la synthèse de la structure coiffe. Cette structure chimique est constituée d'une guanosine liée au premier nucléotide de l'extrémité 5' de l'ARNm par une liaison 5'-5' triphosphate. L'extrémité C-terminale de la protéine L porte des activités méthyltransférases (MTases) responsables des méthylation en position 2'O du premier nucléotide de l'ARN et en N7 de la structure coiffe. Ces méthylation de la coiffe jouent un rôle clef pour la réplication virale dans la mesure où elles permettent la traduction des ARNm viraux en protéines tout en empêchant leur détection par les senseurs de l'immunité innée tels que RIG-I, MDA5 ou les protéines IFITs. Le domaine MTase est suivi d'un domaine C-terminal (CTD) de fonction inconnue à ce jour. L'objectif de ce projet de recherche a été d'étudier le mécanisme de coiffage chez les *Mononegavirales*, prenant pour modèle le virus de la rage et le virus Ebola. Une étude récente menée au sein du laboratoire a permis de caractériser la fonction MTase de la protéine L du virus *Sudan ebolavirus* (SUDV). En plus de ses activités MTases sur la coiffe (2'O et N7 MTases), le domaine MTase+CTD de SUDV catalyse des modifications épitranscriptomiques sur les ARNs viraux en méthylant en 2'O des adénosines internes aux séquences d'ARN. Bien que des modifications similaires aient été identifiées chez plusieurs virus, le rôle de telles modifications reste peu documenté. Dans ce projet, nous avons analysé le rôle du domaine CTD dans la régulation des activités MTases de SUDV et le recrutement de l'ARN. Il a été montré que le domaine CTD joue un rôle crucial dans le recrutement de l'ARN, et donc indispensable pour les activités MTases, ouvrant la voie vers le développement de nouvelles thérapies antivirales. Par une étude de mutagenèse dirigée nous avons mis en évidence que certains résidus des domaines MTase et CTD modulent spécifiquement les différentes activités MTases de SUDV. Des résidus dans le sillon de liaison à l'ARN à l'interface entre les domaines MTase et CTD et découplant les différentes activités MTases ont été identifiés. De plus, la structure cristallographique du domaine MTase de SUDV en complexe avec un VHH a été résolue à 1.8 Å. Les résidus catalytiques de SUDV MTase sont exposés à la surface de la structure dépourvue site apparent de liaison à l'ARN, confirmant le rôle du CTD dans le recrutement du substrat. L'ensemble de ces résultats contribue donc à une meilleure compréhension de l'activité MTase de la L de SUDV, un acteur central dans la réplication virale et l'échappement à la réponse immunitaire innée de l'hôte, qui pourra servir de cible pour le développement de futures molécules antivirales.

Abstract

Viruses belonging to the order *Mononegavirales* represent a wide variety of pathogens for humans such as measles virus, rabies virus and Ebola virus. Their genome codes for a large (L) protein which plays an important role in viral transcription and replication. The L protein catalyzes the synthesis of cap structure. This chemical structure consists in a guanosine linked to the first nucleotide of the 5' end of the mRNA by a 5'-5' triphosphate bond. The C-terminus of L protein carries methyltransferase activity (MTase) responsible to methylation of 2'O position of N1 ribose and N7 of guanosine. These methylations of cap structure play a key role in viral replication, insofar as they allow viral translation while preventing their detection by host immune sensor such as MDA5, RIG-I and IFIT. The MTase is followed by a C-terminal domain (CTD) of unknown function. The aim of this research project is to study the capping mechanism of *Mononegavirales* using rabies virus and Ebola virus as models. A recent study carried out in the laboratory lead to characterize the MTase function of the L protein of *Sudan ebolavirus* (SUDV). Surprisingly, SUDV MTase, in addition to its conventional MTase activity (2'O and N7 MTase), also catalyzes also catalyzes internal adenosine 2'-O methylation. Although similar modifications have been identified in several viruses, the role of such modifications remains poorly understood. In this project, we addressed the question of the role of the CTD domain in the regulation of SUDV MTase activities and the recruitment of RNA. The CTD domain has been shown to play a crucial role in the recruitment of RNA, and in turn regulates the different MTase activities, paving the way for the development of new antiviral therapies. By mutagenesis studies, we have demonstrated that the mutation of residues on the MTase and CTD domains specifically modulates the different SUDV MTase activities. In addition, we were able to identify residues in the RNA binding groove, at the interface between the MTase and CTD, decoupling the different MTase activities. The crystallographic structure of the MTase domain of SUDV in complex with a nanobody was also resolved at 1.8 Å. The catalytic residues of SUDV MTase are exposed on the surface of the structure lacking an apparent RNA binding site, confirming the role of CTD in the recruitment of the substrate. All of these results contribute to a better understanding of the MTase activity of SUDV L protein, a central actor in viral replication and the escape to the innate immune response of the host, and could serve as a basis for the development of future antiviral molecules.