



afmb
architecture et fonction
des macromolécules biologiques



Aix-Marseille Université

THESE

Afin d'obtenir le grade de
Docteur de l'Université Aix-Marseille

Domaine d'étude: Pathologies humaines

Spécialité : Maladies infectieuses

Ecole Doctorale: Sciences de la Vie et de la Santé

Présentée et défendue publiquement le 10 Novembre 2017 par

Baptiste MARTIN

METHYLTRANSFERASES DES FILOVIRUS ET AUTRES MONONEGAVIRUS: CARACTERISATION, ORIGINALITES ET DRUG DESIGN

Réalisée au sein du laboratoire « Réplicases virales : Structure, Mécanisme et Drug Design »
Laboratoire Architecture et Fonction des Macromolécules Biologique (AFMB)
UMR7257 – CNRS – Aix Marseille Université

Sous la direction du **Docteur Etienne DECROLY**
et la tutelle DGA en la personne de **Madame Emmanuelle GUILLOT COMBE**

JURY

Dr. Branka HORVAT
Dr. Jean François ELEOUET
Dr. Sonia LONGHI
Dr. Yamina BENNASSER
Dr. Hervé BOURHY
Dr. Olivier REYNARD
Mme Emmanuelle GUILLOT-COMBE
Dr. Etienne DECROLY
Dr. Françoise DEBART
Dr. Christophe PEYREFITTE

Rapporteur
Rapporteur
Examineur
Examineur
Examineur
Examineur
Responsable DGA
Directeur de thèse
Invité
Invité

*A Yvette,
A mes grands-parents,
Et à mes parents.*

“Science never solves a problem without creating ten more.”

George Bernard Shaw

“The book of nature is written in the language of mathematics.”

Galileo Galelei

“Success means having the courage, the determination, and the will to become the person you believe you were meant to be.”

George Sheehan

REMERCIEMENTS

Vous me voyez bienheureux de vous savoir vous aventurer dans mes travaux de thèse. Avant de vous exposer ma science, je me dois de remercier toutes les personnes qui m'ont forgé en tant que jeune scientifique mais aussi en tant que personne.

Mes premiers remerciements vont naturellement à mon directeur de thèse, Dr. Etienne DECROLY. Sa confiance, son expérience et son enthousiasme ont été de véritables moteurs quotidiens dans la conduite de ces travaux. Il m'a accompagné dans l'échange, l'écoute et la mansuétude. Il m'a laissé exprimer ma science comme un chercheur à part entière. Je lui serai éternellement reconnaissant de m'avoir apporté autant, tant sur le plan professionnel que personnel.

Je tiens également à remercier chaleureusement Dr. Yves BOURNE, directeur de l'AFMB, et Dr. Bruno CANARD, responsable de l'équipe « Réplicases virales », de m'avoir accueilli ces trois années au sein du laboratoire pour réaliser ces travaux. Leur soutien et leurs encouragements m'ont apportés motivation et confiance.

Je suis également reconnaissant de l'investissement de la DGA dans ce projet ainsi que ma tutrice DGA, Emmanuelle GUILLOT COMBE, et Dr. Christophe PEYREFITTE pour leur suivi et leur intérêt dans mes recherches.

Je remercie évidemment les membres du jury qui me font l'honneur de juger mes travaux de thèse. Merci aux Drs. Branka HORVAT, Sonia LONGHI, Yamina BENNASSER, Hervé BOURHY, Jean François ELEOUET et Olivier REYNARD pour leur lecture attentive et leur expertise.

Ces travaux de thèse n'auraient pu se faire sans nos collaborateurs. Merci à Françoise DEBART (Montpellier, FR) pour la synthèse des ARNs synthétiques et sa bienveillance, à l'entreprise Prestwick Chemical (Strasbourg, FR) et au laboratoire du Pr. Katherine L. SELEY RADTKE (Baltimore, USA) pour les composés chimiques testés, au Pr. Jonathan M. GRIMES et Dr. Guido C. PAESEN du laboratoire Strubi (Oxford, UK) pour nos échanges scientifiques et leur expertise dans la production de protéines virales et enfin au Dr. Aline DESMYTER de la plateforme Nanobodies de l'AFMB pour son aide dans le développement de VHHs.

Merci à tous les « gens du labo » que j'ai croisé et avec qui j'ai pu partager beaucoup. Pour commencer, mes pensées vont à Joëlle et sa famille. Je resterai admiratif de son courage et son optimisme communicatif. Merci à Bruno COUTARD pour son regard critique et son engouement tout au long de ce travail (et sa patience pour la pêche aux micro-cristaux). Merci à Jean Claude GUILLEMOT pour sa bienveillance et ses « high five » et à Cécilia pour son aide sur la partie criblage (et courage, 2017 c'est bientôt fini !). Je remercie très chaleureusement Barbara pour les petits conseils quotidiens et nos échanges culturels (et tes petits Haribo aussi). Merci aussi à Nadia, qui continue de m'impressionner à savoir parler autant de langues, à Ana en mémoire de nos heures passées ensemble dans l'aquarium, à Véronique et nos salutations matinales à 14h, à François pour ses compétences bioinformatiques, à Nicolas pour ses leçons de philosophie, et à Chantal et Laurianne pour rendre le

travail plus facile au labo. Mes pensées vont aussi aux étudiants qui ont « souffert » avec moi : les thésards, Coralie (et sa délicatesse), Wahiba, Elsie, Jennifer, Yoan, Clémence, Veronica, Sergio, Simon, Sabine, Nhung, Bin, Van, et les Masters 2, Camille, Gabriel, Dafni, Awa. Merci également aux personnels des plateaux techniques de l'AFMB et en particulier Gerlind et Silvia pour la cristallographie mais aussi la team pétanque du mercredi soir, Polo, Véronique, Laurie, Julie, et Jérémy. La liste est encore longue alors merci à tous ceux que je n'ai pu citer de ne pas m'en vouloir. Vous avez tous contribué de près ou de loin au bon déroulement de ces travaux.

Et parce qu'il n'y a pas que la thèse mais aussi le monitorat, je voudrais remercier Dr. Pedro COUTINHO pour m'avoir donné l'opportunité d'être moniteur à Polytech Marseille. Merci également aux étudiants de 3A et de 5A avec qui j'ai eu de très bons échanges. Bon courage pour la suite

Je remercie les personnes qui m'ont donné toute cette envie qui m'a mené jusqu'à cette thèse. Merci à mes anciens maîtres de stage, René FERRERA, Claire CROLA DA SILVA et Odile DAMOUR à Lyon, Svetlana GOROKHOVA et Aziz MOQRICH à Marseille et Jim WELLS à Cincinnati (I trully miss stem cells !). Je remercie également Michelle SERGENT et Magalie CLAEYS BRUNO de m'avoir éclairé avec leurs plans d'expérience. J'ai aussi une pensée pour certains anciens enseignants qui ont marqué mon parcours, mon prof de S.V.T. de 1^{ère} Gilles SIRIU (qui porte bien son nom finalement), Pr. NATAF et ses cours d'histologie en médecine, et Mme LONGIN et sa passion pour la microbiologie. J'espère réaliser de belles choses avec le bagage que vous m'avez chacun offert.

Et parce que je reste quelqu'un d'ordinaire avec une vie sociale tout ce qu'il y a de plus normal, j'aimerais remercier quelques amis. Pour commencer, je pense à Axelle et Marion qui ont été mes plus belles rencontres durant cette thèse. Elles ont beaucoup contribué à ce travail, Axelle pour m'avoir tout appris à la paillasse et pour toutes nos discussions enrichissantes, et Marion pour m'avoir transmis les joies de la cristallographie. Elles ont toujours su trouver les mots pour calmer mes angoisses comme pour me faire mourir de rire. J'attends nos prochaines vacances ensemble avec impatience. Un grand merci également à Isabelle qui m'a fait confiance en me proposant de revenir des Etats Unis pour entamer cette aventure. Ca n'a pas toujours été simple mais tu m'as transmis ton savoir et ta passion pour la virologie et l'enseignement avec amitié. Pour cela, je t'en serai toujours reconnaissant. Merci pour toutes ces heures de TP « exceptionnels » passées avec toi. Je pense aussi à mes anciens camarades, surtout à mes deux « alcoytes » Théo et Philippe, mais également Emma (ma binôme mais pas que), Emilie, Soizic, Romain, Marie, Mathilde, Romy, Morgane et Alice. Je remercie Johan de m'avoir fait l'amitié de relire ce manuscrit mais surtout de toujours voir la vie en jaune. Je tiens à remercier aussi les membres de l'association MUST, en particulier son président, Giampiero, pour leur accueil, leur énergie et tout ce que nous avons partagé et ce qui nous reste à partager ensemble. Merci aux masters de l'ASPTT Natation de Marseille – Eric, Karine, Marie, Charles, Felipe, Maxime, Jérôme, Charles Eric, Emanuele, Rabah, Thomas et Nicolas – qui m'ont réconcilié avec le comptage de carreaux (presque) chaque semaine et qui m'ont fait découvrir les joies de la nage en mer.

Pour finir, je tiens à remercier du fond du cœur mes parents qui ont toujours été là pour moi pendant toutes ses années. Une grande partie du mérite leur revient car je ne serai certainement pas allé si loin sans leur soutien et leurs encouragements. Vous êtes ma bonne étoile. Je pense aussi à mes

sœurs, Tiphaine et Raphaëlle, que je ne vois pas si souvent et dont la complicité me manque. Merci également au reste de ma famille. Tous, restez comme vous êtes, ne changez rien

Je réserve mes dernières lignes de remerciement pour Enrico qui m'a supporté pendant plus de deux ans et qui me donne tout et toujours plus tous les jours. Tu m'as chamboulé dès notre première rencontre et j'espère être bousculé encore longtemps. Nous sommes le parfait contre-exemple du « qui se ressemble s'assemble ». « Parce que c'était lui, parce que c'était moi » comme l'a écrit Montaigne. Et c'est pour ça que je t'aime.

Maintenant que j'ai rendu à César ce qui était à César, je te souhaite bon courage, cher lecteur, dans ta lecture et merci à toi de faire vivre mes travaux.

TABLE DES MATIERES

<u>LISTE DES TABLEAUX</u>	1
<u>TABLE DES FIGURES</u>	3
<u>CODES : ADN, ARN ET ACIDES AMINES</u>	7
<u>LISTE DES ABREVIATIONS</u>	9
<u>AVANT-PROPOS</u>	13
<u>CHAPITRE 1</u>	15
Prolégomènes aux fièvres hémorragiques à filovirus	15
1. Historique de l'émergence et réémergences des filovirus	17
1.1. Découverte du premier filovirus	17
1.2. Premières émergences du virus Ebola	17
1.3. Apparition de nouvelles souches	19
1.4. La dernière épidémie à virus Ebola	19
2. Eco-épidémiologie	20
3. Physiopathologie	22
3.1. Manifestations cliniques	22
3.2. Voies d'infection	22
3.3. Cellules et tissus cibles	23
4. Diagnostic	23
5. Introduction aux stratégies thérapeutiques	24
Bibliographie	27
<u>CHAPITRE 2</u>	29
Protéines des filovirus et drug discovery : Analyse structure/fonction de la glycoprotéine et de l'entrée du virus	29
Résumé de la Revue 1	31
Revue 1	33
<u>CHAPITRE 3</u>	35
Protéines des filovirus et drug discovery :	35
Bases structurales et fonctionnelles du cycle réplcatif	35
Résumé de la Revue 2	37
Revue 2	39
<u>CHAPITRE 4</u>	41
Protéines des filovirus et drug discovery :	41
Bases structurales et fonctionnelles de l'assemblage et du bourgeonnement	41
Résumé de la Revue 3	43
Revue 3	45
<u>CHAPITRE 5</u>	47

Immunité innée, filovirus et structure coiffe: bases moléculaires et drug design	47
1. L'immunité innée antivirale	51
1.1. Détection des acides nucléiques viraux	51
1.2. Activation de la synthèse d'interféron de type I	53
1.3. Réponse à l'interféron	54
2. Les mécanismes d'échappement des filovirus à l'immunité innée	56
2.1. VP35 et inhibition de la voie des senseurs cytosoliques	56
2.2. Subversion de la réponse à l'INF	57
2.3. Subversion d'autres mécanismes cellulaires	58
3. Coiffe des ARNs: implication dans l'échappement de l'immunité innée et cible thérapeutique potentielle	59
3.1. La coiffe chez les eucaryotes	59
3.2. Voies de synthèse de la coiffe chez les virus	62
3.3. Coiffe comme marqueur du « soi »	63
3.4. Inhiber la synthèse des coiffes virales, une stratégie pertinente ?	65
Bibliographie	67
CHAPITRE 6	75
La protéine L chez les virus de l'ordre des <i>Mononegavirales</i>	75
1. Généralités sur les mononégavirus	77
2. Etude bioinformatique et structurale de la protéine L	79
3. Fonctionnalités de la polymérase	80
4. Ajout de la coiffe par la PRNTase	81
5. Activités méthyltransférases de la coiffe	83
6. Conjecture des remaniements structuraux	84
Bibliographie	87
CHAPITRE 7	91
Objectifs	91
Bibliographie	95
CHAPITRE 8	97
Etude bioinformatique des domaines MTase et CTD de la protéine L des mononégavirus	97
1. Etude de la conservation des domaines MTase et CTD de la protéine L des mononégavirus	99
1.1. Alignements multiples et définition des domaines MTase et CTD	99
1.2. Modélisation de la conservation des résidus	100
1.3. Modélisation de la conservation de la structure	101
2. Propriétés biochimiques des domaines MTase et CTD de la protéine L des mononégavirus	102
Bibliographie	105
CHAPITRE 9	107
Développement de la production, de la purification et du dosage enzymatique de l'activité MTase du virus Soudan	107
1. Production et purification de protéines recombinantes	109

2. Optimisation des conditions expérimentales	111
2.1. Optimisation du tampon de lyse par plan d'expériences	111
2.2. Mise au point du test d'activité pour le domaine MTase+CTD de SUDV	113
3. Outils biochimiques pour la caractérisation des domaines MTase+CTD des mononégavirus	115
3.1. Polarisation de fluorescence	115
3.2. Test de thermostabilité	115
3.3. Chromatographie sur couche mince	116
4. Cristallographie aux rayons X des protéines	116
4.1. Principe général	116
4.2. Cristallogenèse	117
4.3. Diffraction des rayons X par un cristal.....	118
Bibliographie	119
CHAPITRE 10	121
Caractérisation d'une activité MTase originale portée par le domaine MTase+CTD du virus Soudan	121
1. Rôle du domaine CTD dans l'activité MTase chez SUDV	123
2. Identification d'une activité MTase non conventionnelle	124
2.1. Validation des ARNs synthétiques	124
2.2. Identification d'une activité MTase non conventionnelle	125
2.3. Une nouvelle activité MTase chez les mononégavirus	126
3. Caractérisation des méthylations internes chez SUDV	127
3.1. Activité interne et séquence ARN	127
3.2. Identification d'une activité adénosine-2'O MTase interne	128
3.3. Identification des résidus catalytiques	129
3.4. Méthylations internes et recrutement de l'ARN.....	130
3.5. Activité sur de l'ARN db.....	131
3.6. Activité de méthylation interne chez les MTases d'autres mononégavirus	131
3.7. Perspective d'identification des A-2'O méthylations internes chez SUDV	133
4. Implication des méthylations internes dans la biologie des filovirus	134
4.1. Une nouvelle stratégie d'échappement à l'immunité innée ?.....	134
4.2. Implication dans les mécanismes de régulation de la réplication virale	136
4.3. Conclusion générale sur l'activité MTase interne de SUDV	137
Bibliographie	139
CHAPITRE 11	143
Etude des méthylations de la coiffe par le domaine MTase+CTD du virus Soudan.....	143
1. Méthylations de la coiffe et pH	145
1.1. Profils de méthylation de différents ARNs en fonction du pH	145
1.2. Activités MTases sur la coiffe	147
1.3. Mécanismes réactionnels des N7 et 2'O MTases de la coiffe	148
2. Identification des activités N7 et 2'O MTases	148
2.1. Caractérisation par TLC des produits de méthylation de la coiffe	148
2.2. N7 méthylation	149
2.3. 2'O méthylation.....	150
2.4. Conclusions	150

3. Implication de la coiffe pour la 2'O méthylation	150
4. Structuration des ARNs de SUDV et méthylations	151
4.1. Extrémités 5' des ARNm des mononégavirus.....	151
4.2. N7 méthylation et structure tige/boucle	152
4.3. Présence ou absence de 2'O méthylation de la coiffe ?	153
5. Vers des mécanismes de méthylation plus complexes	153
Bibliographie	155
CHAPITRE 12	157
Vers la caractérisation structurale du domaine MTase du virus Soudan	157
1. Identification d'un domaine stable par protéolyse ménagée et analyse bioinformatique ...	159
1.1. Protéolyse ménagée	159
1.2. Analyse bioinformatique	161
2. Cristallogénèse du domaine MTase de SUDV seul.....	161
3. Développement de VHHs anti-domaine MTase de SUDV	162
4. Cristallogénèse du domaine MTase de SUDV en complexe avec le VHH10.....	164
5. Perspectives	166
Bibliographie	169
CHAPITRE 13	171
A la recherche d'inhibiteurs contre les MTases des mononégavirus	171
1. Validation du système expérimental	173
1.1. Calcul des paramètres d'évaluation du criblage	173
1.2. Evaluation d'inhibiteurs de MTases connus.....	174
2. Criblage d'une petite banque de composés développée par la société Prestwick Chemical	176
2.1. Criblage virtuel (réalisé par Prestwick Chemical).....	176
2.2. Test d'inhibition <i>in vitro</i>	176
2.3. Conclusions	177
3. Participation au développement de nouvelles molécules innovantes, les « fleximers ».....	178
3.1. Présentation des fleximers	178
3.2. Description des composés	178
3.3. Tests <i>in vitro</i> et IC50	179
3.4. Conclusions	180
4. Perspectives	181
Bibliographie	183
CHAPITRE 14	185
Conclusions et perspectives	185
Bibliographie	189
ANNEXES	191
LISTE DES ANNEXES	193
<u>ANNEXE 1</u> – Alignement de séquences multiples pour l'étude de la conservation de la séquence protéique du domaine MTase+CTD de la protéine L chez les mononégavirus	195
<u>ANNEXE 2</u> – Contrôle des protéines produites par spectrométrie de masse	199
<u>ANNEXE 3</u> – Liste des ARNs synthétiques	201
<u>ANNEXE 4</u> – Comparaison des propriétés électrostatiques du site de fixation du SAM chez les MTases humaines et virales.....	203

<u>ANNEXE 5</u> – Article 1.....	205
<u>ANNEXE 6</u> – Article 2.....	207
<u>Résumé</u>	
<u>Abstract</u>	

LISTE DES TABLEAUX

Tableau 1	18
Chronologie des grandes flambées de fièvres hémorragiques à filovirus	
Tableau 2	102
Analyse des propriétés biochimiques des domaines méthyltransférases et C-terminaux des filovirus et autres virus de l'ordre des <i>Mononegavirales</i>	
Tableau 3	111
Domaine et plan d'expérimentation pour l'optimisation par plan d'expériences du tampon d'extraction des protéines recombinantes du virus Soudan	
Tableau 4	166
Criblage des conditions de cristallisation des complexes 6HN/VHH10 et 6HC/VHH10	

TABLE DES FIGURES

Figure 1	20
Clin d’œil sur l’épidémiologie des filovirus	
Figure 2	21
Cycle infectieux et physiopathologie des filovirus	
Figure 3	52
Représentation structurale des « RIG-I-like receptors »	
Figure 4	54
Détection des acides nucléiques exogènes et stimulation de la synthèse d’interférons	
Figure 5	55
Voie de synthèse d’effecteurs en réponse à une stimulation par des interférons	
Figure 6	59
Structure chimique de la coiffe des ARNs	
Figure 7	60
Voie « conventionnelle » de synthèse de la coiffe chez les eucaryotes	
Figure 8	62
Stratégies de traduction des ARNs viraux	
Figure 9	64
Bases moléculaires de la discrimination des ARNs par la 2’O méthylation	
Figure 10	78
Classification phylogénique des virus de l’ordre des <i>Mononegavirales</i>	
Figure 11	79
Structure de la protéine L des mononégavirus	
Figure 12	81
Structure du domaine ARN polymérase ARN-dépendante du VSV	
Figure 13	82
Voie de synthèse de la coiffe chez les mononégavirus	
Figure 14	82
Structure du domaine Cap du VSV	
Figure 15	84
Structure du domaine méthyltransférase de la protéine L de hMPV	
Figure 16	100
Modélisation de la conservation structure/fonction du domaine méthyltransférase des mononégavirus	

Figure 17	101
Conservation structurale des domaines méthyltransférase et C-terminal de la protéine L des mononégavirus	
Figure 18	109
Procédé global de production/purification du domaine méthyltransférase recombinant du virus Soudan	
Figure 19	112
Quantification de la réponse expérimentale pour l'optimisation par plan d'expériences du tampon d'extraction des protéines recombinantes du virus Soudan	
Figure 20	114
Optimisation des conditions expérimentales pour le dosage enzymatique de l'activité méthyltransférase chez le virus Soudan	
Figure 21	115
Techniques de caractérisation biochimiques utilisées pour l'étude du domaine méthyltransférase du virus Soudan	
Figure 22	117
Principe et mise en œuvre du processus de cristallogénèse	
Figure 23	123
Le CTD participe au recrutement de l'ARN pour assurer l'activité MTase chez SUDV.	
Figure 24	125
Le domaine MTase+CTD de SUDV n'a pas une activité MTase conventionnelle.	
Figure 25	126
Le recrutement de l'ARN ne dépend pas du type de méthylation sur la coiffe.	
Figure 26	127
Le domaine MTase+CTD méthyle des ARNs non coiffés.	
Figure 27	128
L'activité MTase interne ne dépend pas de la séquence consensus des filovirus.	
Figure 28	129
Le domaine MTase+CTD est actif sur les résidus A internes.	
Figure 29	130
La tétrade K-D-K-E catalyse l'activité adénosine-2'O MTase interne.	
Figure 30	130
Les méthylations internes n'ont pas d'effet sur le recrutement de l'ARN.	
Figure 31	131
Le domaine MTase+CTD est capable de méthyler l'ARN double brin.	
Figure 32	132
Les domaines MTase+CTD des mononégavirus présentent une diversité fonctionnelle.	

Figure 33	133
Les 2'O méthylations internes peuvent être détectées chez les virus par différentes techniques biochimiques.	
Figure 34	135
Les méthylations internes pourraient être impliquées dans différents mécanismes de régulation chez les filovirus.	
Figure 35	146
Les activités MTases de coiffe de SUDV sont dépendantes du pH.	
Figure 36	149
Les activités MTases de coiffe sont portées sur les positions N7 de la coiffe et 2'O du premier nucléotide de l'extrémité 5' de l'ARN.	
Figure 37	151
La 2'O méthylation ne précède pas l'ajout de la coiffe chez SUDV.	
Figure 38	152
Les extrémités des ARNm de SUDV forment des structures secondaires en tige/boucle.	
Figure 39	160
La protéolyse ménagée a permis d'identifier un domaine stable pour la cristallogénèse.	
Figure 40	162
Le domaine MTase de SUDV est cristallisable.	
Figure 41	163
Un VHH contre le domaine MTase a été identifié puis produit comme outil de cristallisation.	
Figure 42	165
Le complexe MTase/VHH est cristallisable.	
Figure 43	174
Le test d'activité radioactif est adapté au criblage d'inhibiteurs.	
Figure 44	175
La sinéfungine inhibe les activités MTases humaine et virales.	
Figure 45	177
Deux hits ont été identifiés grâce au criblage virtuel de composés sur la MTase de la dengue.	
Figure 46	178
Les « fleximers » sont des analogues nucléosidiques flexibles.	
Figure 47	179
Les « fleximers » sont des composés intéressants pour l'inhibition des MTases virales.	

CODES : ADN, ARN ET ACIDES AMINÉS

Codes des nucléosides ADN et ARN

A	adénosine
T	thymidine
U	uridine
C	cytidine
G	guanosine
X	n'importe quel résidu

Codes des acides aminés

A	Ala	alanine
C	Cys	cystéine
D	Asp	aspartic acid
E	Glu	glutamic acid
F	Phe	phénylalanine
G	Gly	glycine
H	His	histidine
I	Ile	isoleucine
K	Lys	lysine
L	Leu	leucine
M	Met	méthionine
N	Asn	asparagine
P	Pro	proline
Q	Gln	glutamine
R	Arg	arginine
S	Ser	sérine
T	Thr	thréonine
V	Val	valine
W	Trp	tryptophane
Y	Tyr	tyrosine

LISTE DES ABREVIATIONS

ADAR : déaminase spécifique des adénosines d'ARN double brin
ADN : acide désoxyribonucléique
AFMB : Architecture et Fonction des Macromolécules Biologiques
ARN(m) : acide ribonucléique messenger
ATPase : adénosine triphosphatase
BDBV : virus Bundibugyo, *Bundibugyo ebolavirus* CARD : caspase activation and recruitment domain, domaine d'activation et de recrutement de caspase
CD : connector domain, domaine connecteur
CR : conserved domain, domaine conservé
cryo-EM : cryo-electromicroscopy, cryomicroscopie électronique
CTD : C-terminal domain, domaine C-terminal
db : double brin
DEAE : diéthylaminoéthyle
DENV : dengue virus, virus de la dengue
DGA : Délégation Générale de l'Armement
EBOV : virus Zaïre, *Zaire ebolavirus*
eIF : eucaryotic initiation factor, facteur d'initiation eucaryote
ELISA : Enzyme-linked immunosorbent assay
FR : France
GDP : guanosine diphosphate
GMP : guanosine monophosphate
GP : glycoprotéine
sGP : soluble glycoprotein, glycoprotéine soluble
ssGP : small soluble glycoprotein, petite glycoprotéine soluble
GTase : guanylyl transférase
GTP : guanosine triphosphate
GTPase : guanosine triphosphatase
hMPV : human metapneumovirus, métapneumovirus humain, *Human metapneumovirus*
hN7 : N7 méthyltransférase humaine
HPLC : high performance liquid chromatography, chromatographie liquide haute performance
HRP : horseradish peroxydase, peroxydase du raifort
hRSV : human respiratory syncytial virus, virus respiratoire syncytial, *Human orthopneumovirus*
IFIT : interferon induced protein, protéine induite par les interférons
IgG : immunoglobuline G
IgM : immunoglobuline M
INF : interféron
IID : interferon inhibitory domain, domaine d'inhibition des interférons

IMAC : immobilized metal affinity chromatography, chromatographie d'affinité sur métal immobilisé

IPTG : isopropyl- β -D-1-thiogalactopyranoside

IRF : interferon regulatory facteur, facteur de régulation des interférons

ISG : interferon stimulated gene, gène stimulé par les interférons

L : « large » protéine

LB : milieu Lysogeny Broth

LGP2 : laboratory of genetics and physiology 2

LLOV : virus Lloviu, *Lloviu cuevavirus*

MAMP : motifs moléculaires associés aux pathogènes

MARV : virus Marburg, *Marbug marburgvirus*

MDA5 : melanoma differentiation associated gene, gène associé à la différenciation du mélanome 5

MERS-CoV : Middle East respiratory syndrom coronavirus

MeV : measles virus, virus de la rougeole, *Measles morbillivirus*

MTase : méthyltransférase

MuV : mumps virus, virus des oreillons, *Mumps rubulavirus*

N ou NP : nucléoprotéine

NCBI : National Center for Biotechnology Information

OMS : Organisation Mondiale de la Santé

ONG : Organisation Non Gouvernementale

ORF : open-reading frame, phase ouverte de lecture

P : phosphoprotéine

PF : polarisation de fluorescence

PKR : double-stranded RNA-dependent protein kinase, kinase des ARNs double brin

PRR : pattern recognition receptor, récepteur de reconnaissance de motif

PRNTase : polyribonucléotidyltransférase

(RT)q-PCR : (Real Time) quantitative Polymerization Chain Reaction

RABV : rabies virus, virus de la rage, *Rabies lyssavirus*

RD : repressor domain, domaine répresseur

RDC : République Démocratique du Congo

RdRp : RNA-dependent RNA polymerase, ARN polymérase ARN-dépendante

RESTV : virus Reston, *Reston ebolavirus*

RIG-I : retinoic acid inducible gene 1, gène induit par l'acide rétinoïque 1

RLR : RIG-I-like receptor, récepteur de la famille RIG-I

RNP : complexe ribonucléoprotéique

RPV : Rinderpest virus, virus Rinderpest

RT : reverse transcription, transcription inverse ou rétrotranscription

RTPase : RNA triphosphatase, ARN triphosphates

SAH : S-adénosyl homocystéine

SAM : S-adénosyl méthionine

SARS-CoV : severe acute respiratory syndrom coronavirus

sb : simple brin

SDS-PAGE : sodium dodecylsulfate – polyacrilamide gel electrophoresis

SHV : virus de la fièvre Hémorragique des Singes
SUDV : virus Soudan, *Sudan ebolavirus*
TAFV : virus Taï Forest, *Tai Forest ebolavirus*
TLC : thin layer chromatography, chromatographie sur couche mince
TLR : Toll-like receptor, récepteur de la famille Toll
TSA : thermostability assay, test de thermostabilité
UK : United Kingdom, Royaume Uni
USA : United States of America, Etas Unis
UTR : untranslated region, région non traduite
VEEV : venezuelan equine encephalitis virus, virus de l'encéphalite équine vénézuélienne
VHH : variable domain of heavy chain of the heavy chain-only antibody, domaine variable de la chaîne lourde des anticorps à chaîne lourde
VSV : virus de la stomatite vésiculaire, *Indiana vesiculovirus*
VP : viral protein, protéine virale
VV2'O : 2'O méthyltransférase du virus de la vaccine
WNV : West Nile virus, virus du Nil Occidental
YFV : yellow fever virus, virus de la fièvre jaune

AVANT-PROPOS

L'Homme est-il un virus pour l'homme ?

Les dérives philosophiques d'un doctorant en biologie

Avant de commencer ma démonstration scientifique, j'aimerais te faire part, cher lecteur, de l'état d'esprit dans lequel j'ai entamé cette thèse et ma prise de conscience quant à mon rôle dans la vie de demain. Il est vrai que bien souvent la thèse est un exercice convenu mais ce fut pour moi l'objet d'une vraie réflexion pour trouver ma place de biologiste dans le monde actuel. Tu pourras donc lire dans les paragraphes suivants certaines de mes réflexions et questions qui méritent, à mon sens, d'être réfléchies afin de répondre aux enjeux actuels de notre société.

Commençons par un bref état des lieux de la situation actuelle. Aujourd'hui, l'écart entre les riches et les pauvres n'a jamais été si grand, poussant les systèmes sociaux à se retourner contre eux-mêmes en élisant des magnats du business réactionnaires. Tout comme le serpent qui se mord la queue. Des camps de concentrations ont été établis en Tchétchénie sans que les acteurs internationaux des droits de l'Homme ne s'offusquent outre mesure alors que les camps de la mort nazis ont fermé il n'y a même pas cent ans. D'un point de vue plus biologique, le nombre d'espèces menacées a doublé depuis 2000, le réchauffement de la planète a dépassé la barre critique du +1°C en 2016 et certains nient encore complètement l'impact de l'homme sur son environnement. La liste est encore longue et alarmante, presque honteuse. Je dirai que nous avons pour la première fois dans l'humanité « régressé » dans notre course évolutive.

Concernant mon histoire, je suis revenu des Etats-Unis suite à mon stage de fin d'étude d'école d'ingénieur en septembre 2014, des étoiles plein les yeux, et l'envie de profiter de ce libéralisme excessif typiquement américain une fois la thèse en poche. Je bouillonnais de motivation, pensant finalement à profiter, presque avec vénalité, des bénéfices des grandes entreprises. Pourtant, ces entreprises ne reposent que sur un seul pilier: la propriété intellectuelle. C'est elle qui permet aux entreprises de protéger leurs produits et leurs stratégies. Il est de coutume que cette propriété intellectuelle soit également utilisée pour bloquer les concurrents. Seulement, à force de tout « protéger », des stratégies ou nouvelles technologies sont tuées dans l'œuf alors qu'elles pourraient profiter au plus grand nombre (ou au plus petit) et permettre une meilleure ouverture à la concurrence. Afin d'illustrer mon propos, prenons l'exemple de l'agriculture qui répond au besoin élémentaire de l'homme de se nourrir. Les leaders biotechnologiques du domaine, tels que Monsanto, Syngenta ou DuPont Pioneer, ont investi des millions d'euros dans la construction et la maintenance de la banque mondiale de graines à Svalbard (Norvège) alors qu'ils sont régulièrement accusés d'appauvrir la biodiversité mondiale (« écocide »). En effet, ils commercialisent des semences OGM brevetées et organisent à leur profit un monopole sur les semences. Ayant maintenant main mise sur les ressources sauvages et donc sur la biodiversité végétale, quel scénario envisager ? Les opinions ne manquent pas. Comme le dit Kissinger, « si vous contrôlez le pétrole, vous contrôlez le pays, mais si vous contrôlez les semences, vous contrôlez l'alimentation. Et celui qui contrôle l'alimentation tient la population en son pouvoir ». Ici encore, retour au Moyen-Age. Suite à ce constat, de nombreux échanges, d'autres

lectures et surement avec un peu plus de maturité aussi, je me suis arrêté sur ce terme, « propriété intellectuelle ». Peut-on encore accepter de nos jours que le vivant puisse être breveté ? Quelle position le biologiste devrait-il adopter en son âme et conscience pour défendre la vie ? Ma réflexion est allée de-ci de-là... A ce jour, mon constat est que les besoins de l'homme grandissent toujours plus vite (énergie, eau, nourriture, santé, etc). Les ressources ne peuvent plus être partagées dans un système aussi libéral. Chaque personne, physique ou morale, cherche donc à protéger ce qui lui est sien, à se méfier des autres et à dérober ce qu'il jalouse. L'homme se considère comme l'aboutissement actuel de l'évolution et fuit à tout prix la sélection naturelle. Avec sa prétention d'expansion, de domination et de conquête, l'homme s'est placé au-dessus de l'équilibre naturel et a oublié qu'il dépend de cet équilibre. Finalement, l'homme ne serait-il pas un virus pour l'homme ?

Au cours de ce travail de réflexion, j'ai compris tout le sens et la profondeur des paroles de Gargantua qui enseigne à son fils Pantagruel que « science sans conscience n'est que ruine de l'âme » (Rabelais). Sur ce fondement de bioéthique, je me suis donc constitué un petit code de déontologie personnel que je consigne ici afin que reproche me soit fait si je m'en détourne. Je m'engage donc à adopter un comportement responsable envers l'environnement et de renforcer mon lien à la nature. Je m'engage à travailler pour le bien de la communauté dans la transparence et sans intérêt personnel. Je m'engage à cultiver ma curiosité et continuer à me questionner pour alimenter mon esprit critique. Je m'engage à transmettre mes connaissances et ma passion pour la vie et sa biodiversité dans un but de prévention et de sensibilisation. Je m'engage à m'investir dans des projets de réflexions et d'actions pour la planète. Je m'engage à agir

“Be the change you wish to see in the world.”

Gandhi

CHAPITRE 1

Prolégomènes aux fièvres hémorragiques à filovirus

1. Historique de l'émergence et réémergences des filovirus

1.1. Découverte du premier filovirus

En août et septembre 1967, dans les villes de Marburg (Allemagne), Francfort (Allemagne) et Belgrade (ancienne Yougoslavie et actuelle Serbie), une trentaine de techniciens de laboratoire travaillant pour la société pharmaceutique Behring présentèrent des signes cliniques graves tels qu'une fièvre élevée, de violents maux de tête, des douleurs articulaires et des hémorragies. Sept d'entre eux en moururent (**Tableau 1**). La majorité de ces techniciens travaillaient avec des singes « verts » d'Ouganda appartenant à l'espèce *Cercopithecus aethiops*, utilisés afin de cultiver le virus de la poliomyélite et produire le vaccin. Un virus inconnu de forme filamenteuse a été isolé à partir de ces singes et a été baptisé virus Marburg [1]. Les mécanismes de transmission épidémiques ont pu être reconstitués révélant l'exportation de singes malades ou en incubation suite à leur capture dans la région du lac Victoria en Afrique. Le virus Marburg réapparut en 1975 chez un touriste australien au Zimbabwe qui décéda d'une fièvre hémorragique fulgurante. Sa femme et l'infirmière en charge furent également gravement malades mais en réchappèrent. Deux résurgences majeures du virus Marburg ont été ensuite recensées en 1998-2000 en République Démocratique du Congo (RDC) et en 2005 en Angola avec des taux de mortalité de plus de 80% (**Tableau 1**).

1.2. Premières émergences du virus Ebola

En septembre 1976, un enseignant de retour de mission médicale catholique à Yambuku (anciennement nord-est Zaïre et actuelle RDC) présentant un syndrome grippal et hémorragique dut être hospitalisé d'urgence. Une épidémie, ultérieurement attribuée au virus Ebola, s'étendit à la maternité, différents autres services de l'hôpital et une cinquantaine de villages avoisinants faisant 280 morts sur les 318 patients recensés (**Tableau 1**). En réalité, dès juin 1976, des cas de fièvre hémorragique avaient été observés au sein d'une population d'ouvriers d'une usine de coton dans la petite ville agricole de Nzara au Soudan, à 800 km environ de Yambuku. L'épidémie s'étendit à Maridi en août 1976 suite au transfert d'un patient à l'hôpital de cette ville décimant alors le personnel de l'hôpital et quelques villages alentours avec un taux de mortalité de 50% (**Tableau 1**).

Des prélèvements de patients de Yambuku furent envoyés à l'Institut Pasteur (Paris, France) et à l'Institut de Médecine Tropicale Prince Léopold (Anvers, Belgique). A Anvers, un filovirus a pu être identifié par Peter Piot en collaboration avec Pierre Sureau de l'Institut Pasteur. Cependant, le virus de la maladie de Yambuku se révéla rapidement différent du virus Marburg car le sérum de patients infectés par ce dernier ne protégeait pas de l'infection par ce nouveau virus. Cette nouvelle souche virale a alors été appelé virus Ebola (EBOV) du nom d'une rivière au nord de Yambuku [2]. Dans le même temps, le virus Soudan (SUDV) fut également isolé et s'avéra être une souche différente du virus Ebola [3]. Plusieurs foyers épidémiques émergèrent, à nouveau à Nazra dans la fabrique de coton en 1979 sans que la cause de l'épidémie ne soit identifiée. La maladie réapparut de façon majeure en 1995 au sud de l'ancien Zaïre (actuelle RDC) puis en 2000 en Ouganda (**Tableau 1**). Ces deux souches ont été différenciées ultérieurement [4] et constituent aujourd'hui les deux souches les plus virulentes de fièvre hémorragique à virus Ebola.

Tableau 1 – Chronologie des grandes flambées de fièvres hémorragiques à filovirus

Année	Virus	Pays	Cas	Décès	Taux de létalité
2013-2016	EBOV	Libéria, Sierra Léone, Guinée	28652	11325	40%
2014	EBOV	République Démocratique du Congo	66	49	74%
2012	BDBV	République Démocratique du Congo	36	13	36%
2008-2009	EBOV	République Démocratique du Congo	32	15	47%
2007-2008	BDBV	Ouganda	149	37	25%
2007	EBOV	République Démocratique du Congo	264	187	71%
2005	MARV	Angola	374	329	88%
2003	EBOV	République Démocratique du Congo	35	29	83%
2002-2003	EBOV	République Démocratique du Congo	143	128	89%
2001-2002	EBOV	République Démocratique du Congo	57	43	75%
2001-2002	EBOV	Gabon	65	53	82%
2000-2001	SUDV	Ouganda	425	224	53%
1998-2000	MARV	République Démocratique du Congo	154	128	83%
1996-1997	EBOV	Gabon	60	45	74%
1996	EBOV	Gabon	37	21	57%
1995	EBOV	Zaïre	315	250	81%
1994	EBOV	Gabon	52	31	60%
1979	SUDV	Soudan	34	22	65%
1976	SUDV	Soudan	284	151	53%
1976	EBOV	Zaïre	318	280	88%
1967	MARV	Yougoslavie	2	0	0%
1967	MARV	Allemagne	29	7	24%

Les données épidémiologiques des flambées de fièvres à filovirus ont permis d'évaluer nombre de paramètres sur ces virus. Pour le virus Ebola, il a été montré que les différentes souches n'ont notamment pas la même virulence. Les virus Zaïre (EBOV) et Soudan (SUDV) sont les plus mortels tandis que les virus Forêt de Taï (TAFV) et Bundibugyo (BDBV) sont moins fatals. Le virus Marburg (MARV) semble aussi virulent que EBOV et SUDV. Ces virus ont majoritairement touché l'Afrique Centrale mais la dernière épidémie, la plus dévastatrice de toutes, a frappé pour la première fois l'Afrique de l'Ouest soulignant le large territoire endémique de ces virus.

Données issues de l'OMS (<http://www.who.int>) et du CDC (<http://www.cdc.gov>).

1.3. Apparition de nouvelles souches

En 1989, une première résurgence du virus Ebola eut lieu via l'importation de macaques *Cynomolgus* (*Macaca fascicularis*) au centre de primatologie de Reston (Hazelton, USA) en provenance des Philippines. La mortalité élevée au sein du cheptel de macaques dirigea le diagnostic vers le virus de la fièvre hémorragique des singes (SHV), non pathogène pour l'Homme. Or, l'évolution clinique et la poursuite de l'épidémie firent suspecter le rôle d'un autre agent pathogène qui s'avéra être une nouvelle souche de virus Ebola venue d'Asie et non pathogène pour l'Homme et qui a été nommée virus Reston (RESTV) [5,6]. RESTV a été ultérieurement identifié aux Philippines à d'autres occasions [7,8], notamment au sein de cheptels de cochons [7,9], soulevant ainsi d'importantes inquiétudes pour la santé publique, l'agriculture et la sécurité alimentaire aux Philippines et plus globalement en Asie. Deux autres souches du virus Ebola ont été identifiées par la suite (**Tableau 1**) : le virus de la Forêt de Taï (TAFV) en 1994 en Côte d'Ivoire suite à l'infection d'un ethnologue ayant réalisé une nécropsie sur un chimpanzé infecté [10], et le virus Bundibugyo (BDBV) en 2007 en Afrique équatoriale [11]. A ce jour, ces deux dernières souches n'ont émergé que ponctuellement.

En 2002, des colonies de *Miniopterus schreibersii*, une chauve-souris commune dans certaines régions d'Océanie, du sud de l'Australie, du sud-est de l'Asie et du sud de l'Europe [12], ont été décimées dans des grottes en France, en Espagne et au Portugal [13]. Grâce aux nouvelles techniques de séquençage, qui ont également permises de confirmer l'existence des différentes souches d'Ebola, Negredo *et al.* ont mis en évidence un virus distinct des virus Marburg et Ebola. Ce nouveau virus a été baptisé virus Lloviu (LLOV) du nom d'une des grottes des Asturies (Espagne) où ces populations de chauve-souris mortes ont été retrouvées. Jusqu'à présent, seules des infections de chiroptères ont été reportées.

1.4. La dernière épidémie à virus Ebola

Une épidémie à virus Ebola sans précédent a frappé la Guinée à partir de décembre 2013 et s'est étendue au Libéria et à la Sierra Léone jusqu'en mars 2016. Cette épidémie à EBOV a été la plus meurtrière de toutes les épidémies à filovirus recensées. Le bilan de l'Organisation Mondiale de la Santé (OMS) recense au total au moins 28 000 cas pour plus de 11 000 décès (**Tableau 1 et Figure 1**) [14]. En août 2014, l'OMS qualifie l'épidémie d'« urgence de santé publique de portée mondiale ». Toujours selon l'OMS, le patient zéro, c'est à dire le premier individu infecté, serait un enfant décédé en décembre 2013 près de Gueckédou, un des principaux nœuds commerciaux dans le sud-est de la Guinée, au carrefour des frontières libérienne et sierraléonaise. C'est la première fois que la maladie à virus Ebola est diagnostiquée hors d'Afrique centrale et exportée hors du continent africain. Des efforts internationaux de lutte contre l'épidémie ont permis de réduire la transmission et de parvenir, depuis mars 2015, à limiter sa survenance à des résurgences de faible ampleur. Cette épidémie a révélé le besoin de traitements et de vaccins efficaces contre ces virus.

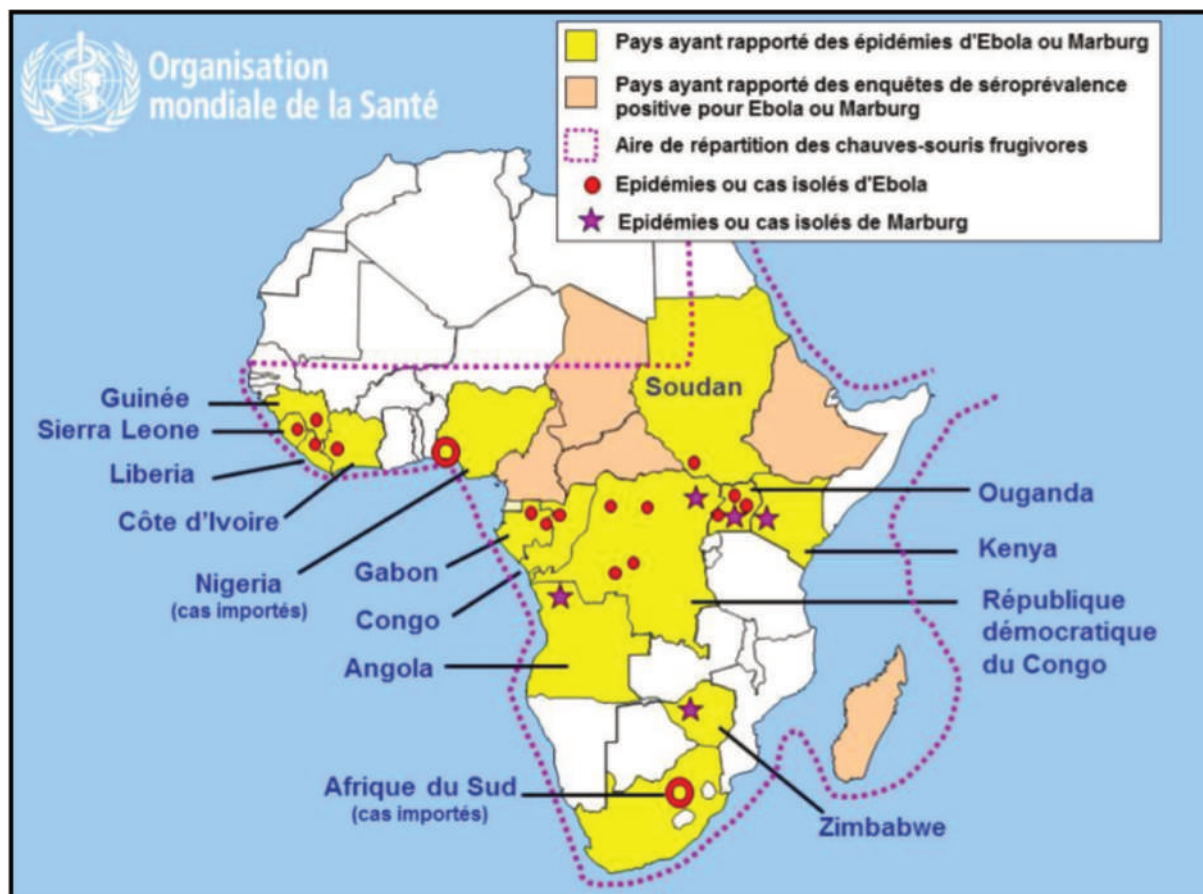


Figure 1 – Clin d'œil sur l'épidémiologie des filovirus

Les épidémies de fièvres hémorragiques à filovirus sont majoritairement concentrées en Afrique Centrale. Les différentes espèces de chauve-souris constituant le réservoir de ces virus peuplent toute l'Afrique subsaharienne. En conséquence, les virus Ebola et Marburg doivent être considérés comme endémiques dans cette région représentant deux tiers du continent du continent africain.

Tiré du rapport de l'OMS WHO/HSE/PED/CED/2014.05

2. Eco-épidémiologie

Les filoviroses, autrement dit les infections à filovirus, sont un groupe de zoonoses capables d'infecter différents hôtes [14-17]. De nombreuses espèces de mammifères (notamment chez les chauves-souris) peuvent être infectés par les filovirus. Toutefois, il semblerait que les oiseaux, reptiles, insectes et autres espèces non mammifères soient résistants à ces virus. Au sein des mammifères, deux modèles d'infection ont été proposés (Figure 2). D'une part, les chauves-souris peuvent être infectées et constituent donc un réservoir infectieux tout au long de leur vie, et ce, sans qu'elles ne présentent de symptômes apparents. Cette infection persistante permettrait la propagation des filovirus sur tout le territoire endémique de ces mammifères volants. D'autre part, l'infection d'autres mammifères cause une fièvre hémorragique aigüe associée à un taux de mortalité élevé. Il semblerait néanmoins qu'une partie des individus, qui réchappent à l'infection, développent une immunité les protégeant d'une seconde infection par la même souche virale, ce qui suggère que les stratégies vaccinales permettront de protéger efficacement les populations touchées.

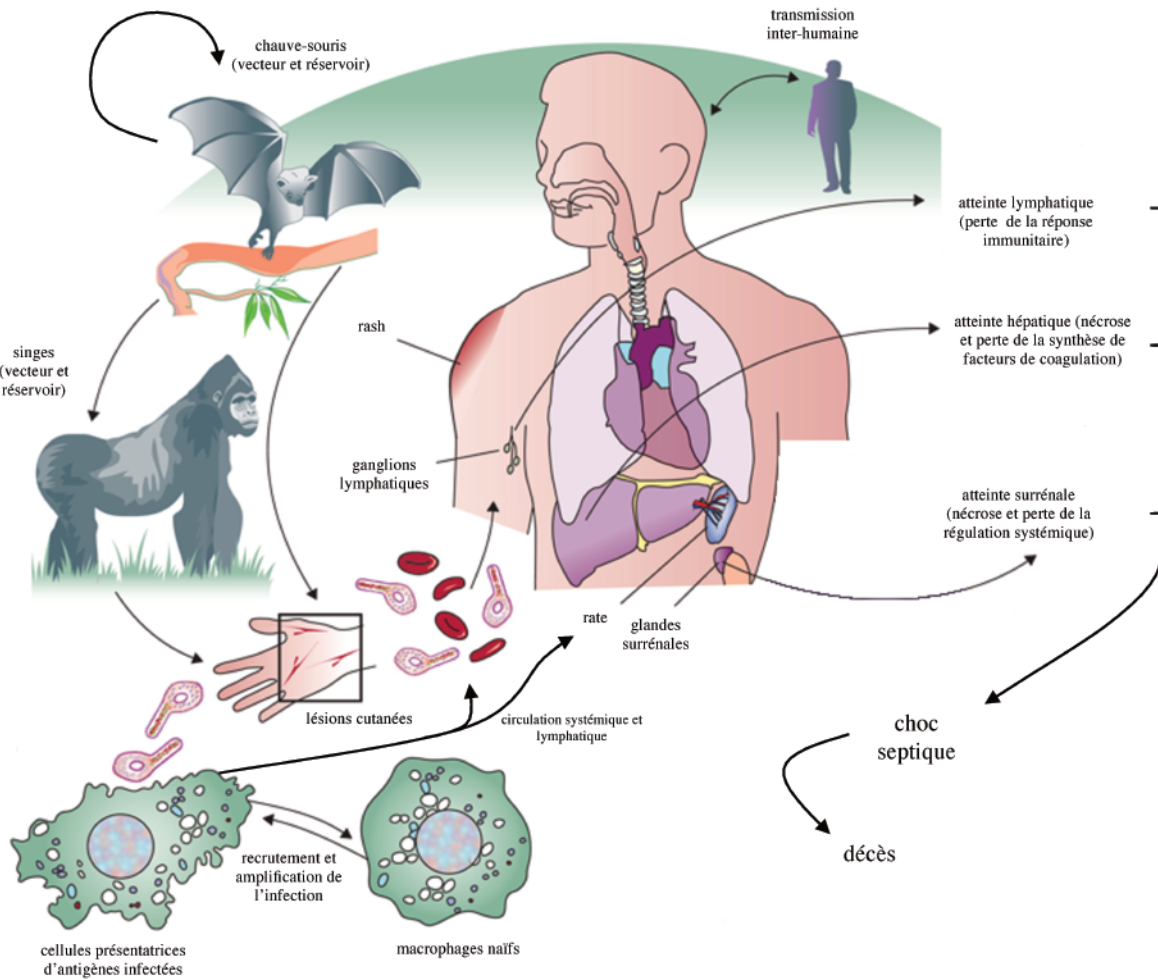


Figure 2 – Cycle infectieux et physiopathologie des filovirus

Le réservoir des filovirus est constitué de différentes espèces de chauve-souris frugivores. Ces chauve-souris peuvent transmettre ces virus aux singes ou, plus rarement, directement à l'Homme. Les singes sont le vecteur principal de transmission à l'Homme via la consommation de leur viande. Le virus entre dans l'organisme par de petites lésions de la peau suite à un contact direct avec des tissus ou liquides infectés. Au site d'infection, les premières cellules infectées sont les cellules présentatrices d'antigènes (monocytes, macrophages, cellules dendritiques). Après infection, ces cellules vont mobiliser des cellules naïves en sécrétant des facteurs pro-inflammatoires et ainsi amplifier et disperser l'infection au niveau systémique. L'infection va alors toucher différents organes dont le foie et les glandes surrénales entraînant leur nécrose et la perte de leurs fonctions, en particulier dans le contrôle de la coagulation et de la régulation vasculaire. L'infection systémique conduit au syndrome hémorragique avec l'apparition de rashs et enfin la mise en place d'un choc septique pouvant conduire à la mort.

Adapté de Feldmann H, et Geisbert TW (2011).

Le principal réservoir de virus Ebola est constitué de trois espèces de chiroptères : *Myonycteris torquata*, une grande chauve-souris frugivore surnommée « renard volant » notamment impliquée dans la dernière épidémie d'Afrique de l'Ouest, et deux autres espèces de chauves-souris, *Hepomops franqueti* et *Hypsignathus monstrosus*, plus rares [18]. Ces espèces vivent dans les forêts tropicales en Afrique ce qui représente un territoire endémique couvrant les deux tiers du continent africain exposant ainsi environ 800 millions de personnes à ce virus [14] (Figure 1). Seul EBOV a été identifié chez ces chiroptères et le réservoir des autres souches reste indéfini. Le virus Marburg a, par contre, été retrouvé chez des chauves-souris différentes, une autre espèce frugivore, *Rousettus aegyptiacus*, et deux espèces insectivores, *Miniopterus inflatus* et *Rhinolophus eloquens* [19,20]. Les contacts entre

ses espèces et l'Homme restent rares, entre autre car ces chauves-souris, nocturnes et silencieuses, sont difficilement localisables donc peu chassées. Plus généralement, les contacts avec d'autres animaux sauvages, notamment les singes, sont plus fréquents puisque les chiroptères font partie de leur régime alimentaire (**Figure 2**). Ces animaux sauvages peuvent être à leur tour chassés et leur viande consommée par l'Homme. Une fois un homme infecté, le virus peut alors facilement être transmis d'un individu à un autre par contact direct déclenchant alors le début d'une potentielle épidémie.

3. Physiopathologie

3.1. Manifestations cliniques

La virulence des filovirus est variable et dépend de la voie d'entrée du virus [2]. Pour les différentes souches du virus Ebola, les souches les plus virulentes sont EBOV (60-90% de mortalité) et SUDV (40-60% de mortalité) alors que BDBV présente un taux de mortalité moins élevé (25%) et TAFV et REST se sont révélés respectivement infectieux mais non létal et non infectieux pour l'Homme (**Tableau 1**). Le virus Marburg est associé à un taux de mortalité semblable à EBOV (70-85%) (**Tableau 1**). Cependant, il est intéressant de relever qu'il n'a présenté que 22% de létalité lors des épidémies européennes. Il semble donc que la prise en charge hospitalière associée à des soins de « soutien » diminueraient le taux de mortalité mais cette hypothèse reste difficile à démontrer à cause des conditions de terrain extrêmement délicates dans les pays où les épidémies à filovirus se propagent.

Les premiers symptômes de la fièvre hémorragique à filovirus apparaissent soudainement entre 2 à 21 jours après l'infection (en moyenne 4 à 10 jours) (résumé dans [21]). Suite à la période d'incubation, la fièvre hémorragique à filovirus est caractérisée par un syndrome grippal sévère avec forte fièvre, frissons, malaise, myalgie et arthralgie. Des signes d'atteintes multiviscérales s'ensuivent au niveau systémique (douleur généralisée et prostration), gastro-intestinal (anorexie, nausée, vomissements, diarrhée), respiratoire (souffle court, toux, rhinorrhée), neurologique (céphalées, confusion, coma) et vasculaire (œdèmes, hypotension posturale, hyperhémie conjonctivale). Les manifestations hémorragiques surviennent dans moins de 50% des cas avec l'apparition dans un premier temps de pétéchies et d'ecchymoses puis d'éruptions érythémateuses 5 à 7 jours après déclaration des symptômes. L'aggravation des signes cliniques entraînent une défaillance multiviscérale traduite par un état de choc, des convulsions et une coagulopathie sévère conduisant à la mort 6 à 16 jours après la survenue des premiers symptômes. Pour les survivants, la convalescence est longue et laisse des séquelles importantes (myélite, hépatite chronique, psychose, uvéite). Le virus persiste même dans certains fluides corporels des convalescents pendant plusieurs mois et particulièrement dans le sperme conduisant à de possibles transmissions sexuelles [22]. Est-ce que le virus reste infectieux durant la convalescence ? Cette question n'a pas encore été clarifiée, mais le corps médical conseille aux convalescents d'être suivis jusqu'à 18 mois après infection.

3.2. Voies d'infection

L'infection par des filovirus survient suite à l'entrée du virus dans le corps via un contact direct entre des liquides ou tissus corporels infectés (sang, sueur, rhinorrhée, sécrétions génitales, sperme, etc.) et les surfaces muqueuses (voies nasale, parentérale, génitales, rectale) ou la peau abîmée (plaies, abrasions) d'un individu sain (**Figure 2**) [2,3, 23-26]. Lors des stades avancés d'infection et même

après la mort, tous les fluides corporels contiennent des concentrations élevées en virus (jusqu'à 10^{10} - 10^{11} particules virales/mL), rendant les patients et les cadavres extrêmement contagieux, entraînant des complications dans la gestion de l'épidémie. En effet, dans les traditions locales, le défunt est lavé et touché par ses proches en guise d'adieu durant le rite mortuaire. Etant très attachés à leurs coutumes et souvent hostiles aux directives des ONGs et autorités sanitaires, ces coutumes/comportements contribuent à la dissémination rapide de l'épidémie. Des cas de contamination en laboratoire ont également été reportés à cause de contacts avec des écouvillons de prélèvement nasal [27,28]. Jusqu'à présent, aucun cas de transmission aérosol n'a été reporté et semble peu probable, permettant un cloisonnement de l'infection par des systèmes de quarantaine et l'identification des « primo-contacts ». Cela est d'autant plus important aujourd'hui puisqu'un vaccin efficace est en fin de développement clinique pour contenir l'infection à EBOV (voir 5.).

3.3. Cellules et tissus cibles

Le tropisme des filovirus est très large car ils infectent aussi bien des cellules épithéliales, que des cellules immunitaires ou encore des cellules muqueuses [25,29-34]. Cependant, des études sur des primates indiquent que l'infection se propage plus efficacement dans les compartiments lymphatique et vasculaire que dans les épithéliums et endothéliums suggérant des cinétiques d'infections différentes [30] (**Figure 2**).

Dans un premier temps, les filovirus infectent principalement les monocytes, macrophages et cellules dendritiques au niveau de la peau et/ou des différentes muqueuses [30]. La création d'une zone inflammatoire et la dérégulation de la sécrétion de médiateurs solubles (interleukines, cytokines, chimiokines) va permettre le recrutement de nouvelles cellules cibles permettant l'amplification de l'infection virale [30,35]. Les cellules infectées vont également migrer à partir du point d'entrée vers le système lymphatique et la rate puis le foie et les autres organes via le système sanguin propageant ainsi l'infection au niveau systémique.

Dans un second temps, les filovirus infectent massivement les hépatocytes et les cellules des glandes surrénales [29,33,36]. Il en résulte des atteintes hépatiques qui entraînent une nécrose du foie accompagnée d'une diminution de la synthèse de certains facteurs de coagulation. Dans le même temps, l'infection des cellules surrénales cause une diminution de la synthèse d'hormones stéroïdes et une diminution du contrôle de la pression sanguine. Ces atteintes multiviscérales induisent une défaillance hépatique sévère et un état hypotendu et hypovolémique qui participent très largement à l'état de choc chez les patients. En outre, les filovirus sont également retrouvés dans différents types de cellules épithéliales, dans les fibroblastes et dans les cellules endothéliales. L'infection de ces cellules endothéliales explique notamment le syndrome hémorragique associé à la coagulopathie sévère due au manque de facteurs de coagulation synthétisés par le foie [37-39]. Cependant, le décès ne survient pas à cause des hémorragies qui restent impressionnantes mais limitées. La cause de la mort des patients résulte principalement de la déferlante de facteurs pro-inflammatoires conduisant au choc septique [38,39].

4. Diagnostic

L'examen clinique est le premier diagnostic réalisé sur les patients suspectés d'être infectés par un filovirus avec la recherche de symptômes (voir 3.1.) associé à la recherche d'un historique de séjour

dans une zone endémique des filovirus. Une des difficultés importantes est que les symptômes associés au début de la maladie peuvent également être signe d'autres maladies endémiques de ces régions (malaria, fièvre typhoïde, fièvre jaune, etc.) rendant l'examen clinique peu informatif sur le type d'infection. Des outils ont donc été développés afin, d'une part, de détecter la particule virale ou un composant du virus pour identifier l'agent pathogène, et d'autre part, de mesurer la réponse immunitaire du patient afin de suivre son état au cours de la maladie.

Trois types de tests permettent de détecter une infection à filovirus sur des prélèvements inactivés [40]. Le statut immunitaire du patient peut être évalué par des tests sérologiques. Les IgM et les IgG sont en effet dosées par des tests ELISA [41,42]. Brièvement, la concentration élevée en IgM étant synonyme d'infection aiguë [42-44] et les IgG apparaissant plus tardivement durant l'infection chez les survivants notamment [45]. Il est donc possible, selon le titre de ces anticorps, de replacer l'infection dans le temps et donner un traitement de soutien adapté. Cependant, les analyses sérologiques ne permettent qu'un diagnostic tardif ce qui explique que les outils de génétique moléculaire soient privilégiés malgré la complexité de l'implémentation de telles techniques dans des régions disposant de systèmes sanitaires rudimentaires. Ces tests reposant sur la technique de RT-qPCR ont été mis en place afin d'identifier la souche virale [46-48]. Bien que cet outil soit tout indiqué pour un diagnostic viral précoce (rapidité, sensibilité et spécificité très élevées pour détecter directement le virus), le coût et les ressources en équipements et la nécessité de formation des personnels restent un obstacle à l'implémentation de ce type de diagnostic dans les zones endémiques. Toutefois, lors de la dernière épidémie à virus Ebola en Afrique de l'Ouest, des kits portables de diagnostic ont été développés avec des réactifs permettant par exemple de réaliser une PCR directement à partir de sérum inactivé de patients avec une incubation à température ambiante et une détection visuelle directement sur le terrain [49]. Enfin, la détection du virus par immunohistochimie a également été utilisée comme outil de validation (diagnostic rétrospectif) au sein des centres de référence afin de confirmer les cas les plus difficiles.

Le diagnostic est une étape clef pour la gestion et l'endiguement de l'épidémie. Malgré le développement d'outils de pointe pour un diagnostic précoce dans les régions reculées d'Afrique, des problèmes logistiques ne sont toujours pas résolus. En effet, l'environnement austère où les ressources manquent rapidement et les différences culturelles rendant la population méfiante voir hostile restent des enjeux majeurs dans ces régions du monde.

5. Introduction aux stratégies thérapeutiques

A ce jour, il n'existe toujours aucun traitement curatif contre les fièvres hémorragiques à filovirus. La mise au point d'un vaccin a donc été un enjeu majeur pour l'avenir. Le vaccin VSV-ZEBOV est en phase finale de développement depuis fin 2015 [50]. Ce vaccin est constitué d'un virus véhicule, cousin d'EBOV, le virus de la stomatite vésiculaire (VSV), pour lequel la glycoprotéine de surface a été remplacée par celle du virus Ebola. Durant les essais de phase 3 du développement clinique, ce vaccin canadien a montré une protection de 100% des patients vaccinés. Il semble que de tels vaccins pourraient aider à juguler les épidémies notamment en utilisant des stratégies de vaccination en ceinture, c'est-à-dire en vaccinant tous les cas contacts de personnes infectées, afin de limiter la propagation des épidémies. La mise au point d'anticorps neutralisants à partir de sérums de patients

convalescents (ZMapp) a également été déterminante mais leur coût et leur disponibilité restent limités [51].

Malgré ces avancées significatives, la vaccination reste un traitement onéreux et sensible pour les pays d'Afrique centrale et de l'Ouest. En effet, les vaccins requièrent une logistique importante notamment liée à leur conditionnement et leur conservation réfrigérée. Il est donc également impératif de développer des inhibiteurs à faible coût de revient valorisant tout le travail de recherche afin de traiter les patients infectés non vaccinés et de simplifier la gestion des crises épidémiques.

Ces dernières années, des découvertes majeures ont fait avancer la recherche dans ce sens. La résolution de structures de la plupart des protéines des filovirus, la caractérisation de nombreuses interactions hôte/pathogène, l'identification d'anticorps neutralisant de patients survivants, le développement d'analogues de nucléotides, l'utilisation d'ARN interférents sont les découvertes (et bien d'autres) qui ont permis de mieux comprendre le fonctionnement du virus au niveau moléculaire. Il a ainsi été possible d'identifier les cibles les plus prometteuses pour le développement de molécules thérapeutiques dont l'enjeu sera à la fois de soigner des patients ayant déclaré la maladie et de prévenir l'infection chez les cas contacts. Dans ce travail, l'ensemble des aspects structure/fonction de chaque protéine virale ainsi que l'avancée des développements thérapeutiques associés ont été résumés dans quatre revues présentées dans les chapitres suivants. Elles te permettront, cher lecteur, de te familiariser avec le virus au niveau moléculaire et d'aborder le contexte de ces travaux de thèse.

Bibliographie

- [1] : Siebert R, Shu HL, Slenczka W, Peters D, Müller G. [On the etiology of an unknown human infection originating from monkeys]. *Dtsch Med Wochenschr.* 1967 Dec 22;92(51):2341-3.
- [2] : World Health Organization. Ebola haemorrhagic fever in Zaire, 1976. *Bull World Health Organ.* 1978;56(2):271-93.
- [3] : World Health Organization. Ebola haemorrhagic fever in Sudan, 1976. Report of a WHO/International Study Team. *Bull World Health Organ.* 1978;56(2):247-70.
- [4] : Cox NJ, McCormick JB, Johnson KM, Kiley MP. Evidence for two subtypes of Ebola virus based on oligonucleotide mapping of RNA. *J Infect Dis.* 1983 Feb;147(2):272-5.
- [5] : Dalgard DW, Hardy RJ, Pearson SL, Pucak GJ, Quander RV, Zack PM, Peters CJ, Jahrling PB. Combined simian hemorrhagic fever and Ebola virus infection in cynomolgus monkeys. *Lab Anim Sci.* 1992 Apr;42(2):152-7.
- [6] : Jahrling PB, Geisbert TW, Dalgard DW, Johnson ED, Ksiazek TG, Hall WC, Peters CJ. Preliminary report: isolation of Ebola virus from monkeys imported to USA. *Lancet.* 1990 Mar 3;335(8688):502-5.
- [7] : Miranda ME, Miranda NL. Reston ebolavirus in humans and animals in the Philippines: a review. *J Infect Dis.* 2011 Nov;204 Suppl 3:S757-60.
- [8] : Miranda ME, Ksiazek TG, Retuya TJ, Khan AS, Sanchez A, Fulhorst CF, Rollin PE, Calaor AB, Manalo DL, Roces MC, Dayrit MM, Peters CJ. Epidemiology of Ebola (subtype Reston) virus in the Philippines, 1996. *J Infect Dis.* 1999 Feb;179 Suppl 1:S115-9.
- [9] : Barrette RW, Metwally SA, Rowland JM, Xu L, Zaki SR, Nichol ST, Rollin PE, Towner JS, Shieh WJ, Batten B, Sealy TK, Carrillo C, Moran KE, Bracht AJ, Mayr GA, Sirios-Cruz M, Catbagan DP, Lautner EA, Ksiazek TG, White WR, McIntosh MT. Discovery of swine as a host for the Reston ebolavirus. *Science.* 2009 Jul 10;325(5937):204-6.
- [10] : Le Guenno B, Formenty P, Wyers M, Gounon P, Walker F, Boesch C. Isolation and partial characterisation of a new strain of Ebola virus. *Lancet.* 1995 May 20;345(8960):1271-4.
- [11] : Towner JS, Sealy TK, Khristova ML, Albariño CG, Conlan S, Reeder SA, Quan PL, Lipkin WI, Downing R, Tappero JW, Okware S, Lutwama J, Bakamutumaho B, Kayiwa J, Comer JA, Rollin PE, Ksiazek TG, Nichol ST. Newly discovered ebola virus associated with hemorrhagic fever outbreak in Uganda. *PLoS Pathog.* 2008 Nov;4(11):e1000212.
- [12] : Appleton BR, McKenzie JA, Christidis L. Molecular systematics and biogeography of the bent-wing bat complex *Miniopterus schreibersii* (Kuhl, 1817) (Chiroptera: Vespertilionidae). *Mol Phylogenet Evol.* 2004 May;31(2):431-9.
- [13] : Quetglas, J., F. González y O. de Paz. 2003. Estudian la extraña mortandad de miles de murciélagos de cueva. *Quercus.* 203:50
- [14] : World Health Organization. Ebola virus disease – Fact sheet n°103; (consulted 2017 May 16). Available from: <http://www.who.int>
- [15] : Centers for Disease Control and Prevention. Ebola Hemorrhagic Fever; (consulted 2017 May 16). Available from: <http://www.cdc.gov>
- [16] : Hartman AI. Ebola and Marburg virus infections. In: Magill AJ, Ryan ET, Hill DR, Solomon T, editors. Hunter's tropical medicine and emerging infectious diseases. London: Saunders Elsevier; 2013. p 330-3
- [17] : MacNeil A, Rollin PE. Ebola and Marburg hemorrhagic fevers: neglected tropical diseases? *PLoS Negl Trop Dis.* 2012 Jun;6(6):e1546.
- [18] : Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, Yaba P, Délicat A, Paweska JT, Gonzalez JP, Swanepoel R. Fruit bats as reservoirs of Ebola virus. *Nature.* 2005 Dec 1;438(7068):575-6.
- [19] : Kuzmin IV, Niezgodna M, Franka R, Agwanda B, Markotter W, Breiman RF, Shieh WJ, Zaki SR, Rupprecht CE. Marburg virus in fruit bat, Kenya. *Emerg Infect Dis.* 2010 Feb;16(2):352-4.
- [20] : Towner JS, Pourrut X, Albariño CG, Nkogue CN, Bird BH, Grard G, Ksiazek TG, Gonzalez JP, Nichol ST, Leroy EM. Marburg virus infection detected in a common African bat. *PLoS One.* 2007 Aug 22;2(8):e764.

- [21] : Goeijenbier M1, van Kampen JJ, Reusken CB, Koopmans MP, van Gorp EC. Ebola virus disease: a review on epidemiology, symptoms, treatment and pathogenesis. *Neth J Med*. 2014 Nov;72(9):442-8.
- [22] : Etard JF, Sow MS, Leroy S, Touré A, Taverne B, Keita AK, Msellati P, Magassouba N, Baize S, Raoul H, Izard S, Kpamou C, March L, Savane I, Barry M, Delaporte E; Postebogui Study Group. Multidisciplinary assessment of post-Ebola sequelae in Guinea (Postebogui): an observational cohort study. *Lancet Infect Dis*. 2017 May;17(5):545-552.
- [23] : Dowell SF, Mukunu R, Ksiazek TG, Khan AS, Rollin PE, Peters CJ. Transmission of Ebola haemorrhagic 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*. 1999; 179 (suppl 1): S87-91.
- [24] : Khan AS, Tshioko FK, Heymann DL et al. The reemergence of Ebola haemorrhagic fever, Democratic Republic of the Congo, 1995. Commission de Lutte contre les Epidémies à Kikwit. *J Infect Dis*. 1999; 179 (suppl 1): S76-86.
- [25] : Rodriguez LL, De Roo A, Guimard Y et al. Persistence and genetic stability of Ebola virus during the outbreak in Kikwit, Democratic Republic of the Congo, 1995. *J Infect Dis*. 1999; 179 (suppl 1): S170-76.
- [26] : Zaki SR, Shieh WJ, Greer PW et al. A novel immunohistochemical assay for the detection of Ebola virus in skin: implications for diagnosis, spread, and surveillance of Ebola haemorrhagic fever. Commission de Lutte contre les Epidémies à Kikwit. *J Infect Dis*. 1999 (suppl 1) : S36-47)
- [27] : Emond RT, Evans B, Bowen ET, Lloyd G. A case of Ebola virus infection. *BMJ*. 1977; 2 : 541-44.
- [28] : International Society for Infectious Diseases. Ebola virus, needlestick injury – Germany (Hamburg) March 17, 2009; Archive number 20090317.1084 <http://www.promedmail.org> (accessed May 2017).
- [29] : Geisbert TW, Hensley LE, Larsen T, Young HA, Reed DS, Geisbert JB, Scott DP, Kagan E, Jahrling PB, Davis KJ. Pathogenesis of Ebola hemorrhagic fever in cynomolgus macaques: evidence that dendritic cells are early and sustained targets of infection. *Am J Pathol*. 2003 Dec;163(6):2347-70.
- [30] : Geisbert TW, Young HA, Jahrling PB, Davis KJ, Larsen T, Kagan E, Hensley LE. Pathogenesis of Ebola hemorrhagic fever in primate models: evidence that hemorrhage is not a direct effect of virus-induced cytolysis of endothelial cells. *Am J Pathol*. 2003 Dec;163(6):2371-82.
- [31] : Ryabchikova EI, Kolesnikova LV, Luchko SV. An analysis of features of pathogenesis in two animal models of Ebola virus infection. *J Infect Dis*. 1999 Feb;179 Suppl 1:S199-202.
- [32] : Zaki SR, Goldsmith CS. Pathologic features of filovirus infections in humans. *Curr Top Microbiol Immunol*. 1999;235:97-116.
- [33] : Jaax NK, Davis KJ, Geisbert TJ, Vogel P, Jaax GP, Topper M, Jahrling PB. Lethal experimental infection of rhesus monkeys with Ebola-Zaire (Mayinga) virus by the oral and conjunctival route of exposure. *Arch Pathol Lab Med*. 1996 Feb;120(2):140-55.
- [34] : Baskerville A, Fisher-Hoch SP, Neild GH, Dowsett AB. Ultrastructural pathology of experimental Ebola haemorrhagic fever virus infection. *J Pathol*. 1985 Nov;147(3):199-209.
- [35] : Schnittler HJ, Feldmann H. Marburg and Ebola hemorrhagic fevers: does the primary course of infection depend on the accessibility of organ-specific macrophages? *Clin Infect Dis*. 1998 Aug;27(2):404-6.
- [36] : Murphy FA. Pathogenesis of Ebola virus infection. In: Pattyn SR, ed. *Ebola Virus Haemorrhagic Fever*. Amsterdam: Elsevier/North-Holland, 1978 : 43-59.
- [37] : Levi M. Disseminated intravascular coagulation. *Crit Care Med*. 2007 Sep;35(9):2191-5.
- [38] : Geisbert TW1, 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*. 2003 Dec 1;188(11):1618-29. Epub 2003 Nov 14.
- [39] : Fisher-Hoch SP, Platt GS, Neild GH, Southee T, Baskerville A, Raymond RT, Lloyd G, Simpson DI. Pathophysiology of shock and hemorrhage in a fulminating viral infection (Ebola). *J Infect Dis*. 1985 Nov;152(5):887-94.
- [40] : Kortepeter MG, Bausch DG, Bray M. Basic clinical and laboratory features of filoviral hemorrhagic fever. *J Infect Dis*. 2011 Nov;204 Suppl 3:S810-6.
- [41] : Rowe AK, Bertolli J, Khan AS, Mukunu R, Muyembe-Tamfum JJ, Bressler D, Williams AJ, Peters CJ, Rodriguez L, Feldmann H, Nichol ST, Rollin PE, Ksiazek TG. 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.* 1999 Feb;179 Suppl 1:S28-35.

[42] : Ksiazek TG, Rollin PE, Williams AJ, Bressler DS, Martin ML, Swanepoel R, Burt FJ, Leman PA, Khan AS, Rowe AK, Mukunu R, Sanchez A, Peters CJ. 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.* 1999 Feb;179 Suppl 1:S177-87.

[43] : Baize S, Leroy EM, Georges-Courbot MC, Capron M, Lansoud-Soukate J, Debré P, Fisher-Hoch SP, McCormick JB, Georges AJ. Defective humoral responses and extensive intravascular apoptosis are associated with fatal outcome in Ebola virus-infected patients. *Nat Med.* 1999 Apr;5(4):423-6.

[44] : 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.* 1999 Feb;179 Suppl 1:S48-53.

[45] : Wauquier N, Becquart P, Gasquet C, Leroy EM. Immunoglobulin G in Ebola outbreak survivors, Gabon. *Emerg Infect Dis.* 2009 Jul;15(7):1136-7.

[46] : Ro YT, Ticer A, Carrion R Jr, Patterson JL. Rapid detection and quantification of Ebola Zaire virus by one-step real-time quantitative reverse transcription-polymerase chain reaction. *Microbiol Immunol.* 2017 Apr;61(3-4):130-137.

[47] : Park SW, Lee YJ, Lee WJ, Jee Y, Choi W. One-Step Reverse Transcription-Polymerase Chain Reaction for Ebola and Marburg Viruses. *Osong Public Health Res Perspect.* 2016 Jun;7(3):205-9.

[48] : Cherpillod P, Schibler M, Vieille G, Cordey S, Mamin A, Vetter P, Kaiser L. Ebola virus disease diagnosis by real-time RT-PCR: A comparative study of 11 different procedures. *J Clin Virol.* 2016 Apr;77:9-14.

[49] : Kurosaki Y, Magassouba N, Oloniniyi OK, Cherif MS, Sakabe S, Takada A, Hirayama K, Yasuda J. Development and Evaluation of Reverse Transcription-Loop-Mediated Isothermal Amplification (RT-LAMP) Assay Coupled with a Portable Device for Rapid Diagnosis of Ebola Virus Disease in Guinea. *PLoS Negl Trop Dis.* 2016 Feb 22;10(2):e0004472.

[50] : Medaglini D, Siegrist CA. Immunomonitoring of human responses to the rVSV-ZEBOV Ebola vaccine. *Curr Opin Virol.* 2017 Apr 28;23:88-94.

[51] : PREVAIL II Writing Group; Multi-National PREVAIL II Study Team, Davey RT Jr, Dodd L, Proschan MA, Neaton J, Neuhaus Nordwall J, Koopmeiners JS, Beigel J, Tierney J, Lane HC, Fauci AS, Massaquoi MBF, Sahr F, Malvy D. A Randomized, Controlled Trial of ZMapp for Ebola Virus Infection. *N Engl J Med.* 2016 Oct 13;375(15):1448-1456.

CHAPITRE 2

**Protéines des filovirus et drug discovery :
Analyse structure/fonction de la
glycoprotéine et de l'entrée du virus**

Résumé de la Revue 1

Titre : Filovirus proteins for antiviral drug discovery: A structure/function analysis of surface glycoproteins and virus entry.

Auteurs : Martin B, Hoenen T, Canard B, Decroly E.

Référence : Antiviral Res. 2016 Nov; 135:1-14. doi: 10.1016/j.antiviral.2016.09.001. Epub 2016 Sep 14.

Revue sélectionnée comme meilleur article du mois de Novembre 2016

Cette première revue présente l'avancée des connaissances sur l'entrée des filovirus dans leurs cellules cibles. Elle se découpe en quatre parties. La première partie résume la récente histoire des filovirus, leur classification dans le règne viral et leur organisation génétique (**Figure 1, 2 et 3**). Ensuite, nous avons détaillé les différents produits du gène codant pour la glycoprotéine et les protéines accessoires sGP, ssGP et peptide Δ en précisant les mécanismes d'édition originaux lors de la transcription chez les différentes souches Ebola et chez LLOV (**Figure 4**). Les caractéristiques structurales et fonctionnelles des glycoprotéines (**Figure 5**) ont ensuite été décrites ainsi que les différents mécanismes d'attachement, de protéolyse endosomale et de fusion qui conduisent à l'entrée du virus et à la libération de sa nucléocapside dans le cytoplasme de la cellule hôte (**Figure 6**). La régulation de l'expression de la GP et son implication dans la modulation de la cytotoxicité et de l'inflammation ont été résumées pour illustrer le rôle de la GP dans la virulence des filovirus. Enfin, la dernière partie récapitule les progrès récents réalisés dans le développement de stratégies antivirales visant l'entrée des filovirus (**Tableau 1**). Différents types d'approches ont été identifiés : des molécules à large spectre visant des interactions ou des mécanismes partagés par de nombreux virus de l'ordre des *Mononegavirales*, des inhibiteurs plus spécifiques des filovirus ciblant directement la GP et ses interactions spécifiques avec d'autres mécanismes cellulaires, ainsi que le repositionnement de médicaments déjà sur le marché ciblant d'autres pathologies.

Revue 1



Review

Filovirus proteins for antiviral drug discovery: A structure/function analysis of surface glycoproteins and virus entry



Baptiste Martin ^a, Thomas Hoenen ^b, Bruno Canard ^a, Etienne Decroly ^{a,*}

^a Laboratoire Architecture et Fonction des Macromolécules Biologiques (AFMB), UMR7257 CNRS, Parc Scientifique de Luminy, Aix-Marseille Université, Marseille, France

^b Friedrich-Loeffler-Institut, Greifswald, Insel Riems, Germany

ARTICLE INFO

Article history:

Received 29 July 2016

Received in revised form

1 September 2016

Accepted 5 September 2016

Available online 14 September 2016

Keywords:

Ebola

Filovirus

Antiviral

Mononegavirales

Entry inhibitors

Glycoprotein

ABSTRACT

This review focuses on the recent progress in our understanding of filovirus protein structure/function and its impact on antiviral research. Here we focus on the surface glycoprotein GP_{1,2} and its different roles in filovirus entry. We first describe the latest advances on the characterization of GP gene-overlapping proteins sGP, ssGP and Δ-peptide. Then, we compare filovirus surface GP_{1,2} proteins in terms of structure, synthesis and function. As they bear potential in drug-design, the discovery of small organic compounds inhibiting filovirus entry is a currently very active field. Although it is at an early stage, the development of antiviral drugs against Ebola and Marburg virus entry might prove essential to reduce outbreak-associated fatality rates through post-exposure treatment of both suspected and confirmed cases.

© 2016 Elsevier B.V. All rights reserved.

Contents

1. Introduction	1
2. Filovirus classification and genomic structure	2
3. The multifaceted aspects of the filovirus GP gene	2
3.1. GP: one gene, many proteins	3
3.2. The viral surface glycoprotein: main structural features	4
4. GP, a fusion protein mediating filovirus cell entry	5
4.1. Filovirus attachment	5
4.2. Uptake and proteolytic processing	6
4.3. Membrane fusion	7
5. Modulation of cytotoxicity and inflammation	8
6. Potential of GP targeting antivirals	8
6.1. Broad-spectrum molecules	8
6.2. Filovirus-specific inhibitors	10
6.3. Repurposing of approved drugs	10
7. Conclusion	10
Acknowledgements	11
References	11

1. Introduction

The filovirus surface glycoprotein GP_{1,2} triggers and drives virus

* Corresponding author.

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

entry into the host cell via a complex mechanism that is, to date, only partially understood. Drawing on the present knowledge, numerous research groups have recently developed antiviral strategies aiming to inhibit viral entry.

This review presents current knowledge on filovirus glycoproteins, and compares their structures and associated functions. After briefly placing filoviruses in viral classification and nomenclature, we describe the soluble GP derivatives sGP, ssGP and Δ -peptide, and review all structural features of the surface GP_{1,2}. We detail step by step the mechanism by which GP_{1,2} triggers virus entry. In this part, the aim is to highlight the relationship between structural features and their role in the various phases of the filovirus entry process, with attention to idiosyncrasies within the family *Filoviridae*. Lastly, we recapitulate existing and ongoing antiviral strategies, in order to connect mechanisms of action to structure/function analysis aiming at potent anti-filovirus therapies.

2. Filovirus classification and genomic structure

The family *Filoviridae* belongs to the negative strand, non-segmented (NNS) RNA viruses of the *Mononegavirales* order. This family groups highly pathogenic viruses such as those found in the *Marburgvirus* and *Ebolavirus* genera (Ascenzi et al., 2008), responsible for severe hemorrhagic fevers, as well as the genus *Cuevavirus* (Negredo et al., 2011), the latter being found so far only in form of RNA sequenced from bats (Fig. 1). The *Marburgvirus* genus is represented by viruses within a single species, *Marburg marburgvirus* (Marburg virus - MARV). It was the first filovirus genus and species discovered in 1967 during related outbreaks in Frankfurt (Germany) and Belgrade (Yugoslavia) upon importation of infected monkeys from Uganda to Marburg (Germany) (Siebert et al., 1967). The *Ebolavirus* genus consists of five virus species. They are known as *Zaire ebolavirus* (Ebola virus - EBOV), which is the first ebolavirus species identified in 1976 in the Democratic Republic of the Congo (formerly northern Zaire) near the Ebola River, *Sudan ebolavirus* (Sudan virus - SUDV), *Tai Forest ebolavirus* (Tai Forest virus - TAFV), *Bundibugyo ebolavirus* (Bundibugyo virus - BDBV) and *Reston ebolavirus* (Reston virus - RESTV) according to the new nomenclature (Kuhn et al., 2010). While RESTV has not been described to cause human disease yet, the other species, including MARV, are highly pathogenic with fatality rates ranging from 25% up to 90% (Feldmann and Geisbert, 2011). The *Cuevavirus* genus was established after the discovery of sequences in 2002 most likely belonging to a new filovirus, *Lloviu cuevavirus* (Lloviu virus - LLOV), presumably infecting bats in Asturias (Spain) (Negredo et al., 2011). Since it is a novel entry in the filovirus phylogeny, only little is known about its biology and putative infectivity in humans.

With their high infectivity and their ability to impair the immune system (Feldmann and Geisbert, 2011; Ramanan et al., 2011), filoviruses trigger an abrupt onset of symptoms including fever,

headache, myalgia and gastrointestinal disorders. Next, hemorrhagic manifestations can arise during the peak of illness. Shock, convulsions, coagulopathy and multi-organ failure appear later and are fatal in many cases (Feldmann and Geisbert, 2011; Nina, 2014). Unfortunately, there are no approved antivirals or vaccines available yet, although significant progress has been made lately in this respect (Mendoza et al., 2016), but supportive treatments such as rehydration and control of fever and pain might help patients to overcome infection. Lately, a lot of efforts have been put together to identify key viral targets in order to inhibit the viral cycle and help to cure the infection (Choi and Croyle, 2013).

Filoviruses share a common genomic organization. Their NNS RNA genome of around 19 kb carries seven main genes leading to the synthesis of the different viral proteins (Figs. 1 and 2) (Ascenzi et al., 2008). All these proteins are essential to establish an infection leading to efficient virus replication (Fig. 3). The sole surface protein GP_{1,2} triggers the first steps of cell infection, which requires attachment to factors present at the surface of target dendritic cells (DCs) and monocytes/macrophages, and on endothelial cells of liver sinusoids and lymph node sinuses. Once attached, the virions are internalized, and endosomal events induce fusion (Feldmann et al., 1999) allowing the release of the viral particle content into the cytoplasm. The nucleocapsid is composed of the genomic RNA in complex with the nucleoprotein NP, the two cofactors VP30 and VP35, and the large protein L, which form a large macromolecular complex protecting the RNA genome and facilitating genome replication/transcription (reviewed by Mühlberger, 2007). The L protein harbors the RNA-dependent RNA polymerase (RdRp) activity, which is essential for both genome replication and transcription. In addition, this protein carries yet uncharacterized enzymatic activities involved in RNA transcriptional modifications such as RNA capping and polyadenylation, protecting viral mRNA from both degradation and detection by the host cell innate immunity guardians (Mühlberger, 2007; Liang et al., 2015). The nucleoprotein NP enwraps and protects the NNS RNA from host nucleases. The VP30 protein acts as a transcription cofactor, while VP35 is the polymerase cofactor (Mühlberger, 2007). After replication of the viral genome and RNA transcription, nascent viral particles are assembled in a process mediated by the matrix protein VP40, and virus budding occurs at the cell surface membrane in a process that involves hijacking the host ESCRT machinery (Hartlieb and Weissenhorn, 2006; Noda et al., 2006).

3. The multifaceted aspects of the filovirus GP gene

Early stages of cell infection have been shown to be mediated by the class I viral glycoprotein GP_{1,2} exposed at the virus membrane surface (reviewed by Lee and Saphire, 2009). This protein is synthesized as a precursor GP₀ after translation of an edited GP open reading frame (Fig. 4), which is cleaved to yield an ectodomain GP₁

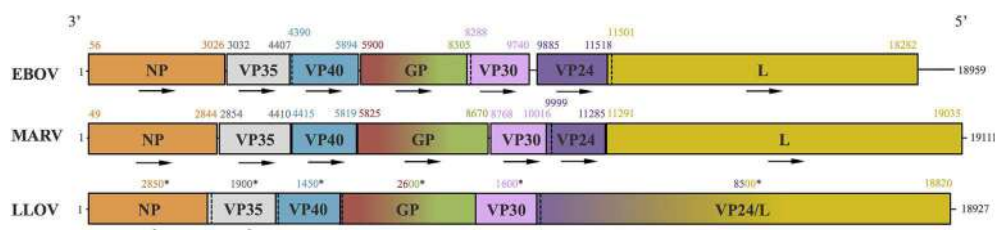


Fig. 1. Filovirus genome organization. Filoviruses are a family of non-segmented negative single stranded RNA viruses, including the genera *Ebolavirus*, *Marburgvirus*, and *Cuevavirus*, with the respective prototype viruses Ebola virus (EBOV), Marburg virus (MARV) and Lloviu virus (LLOV) sharing a common genome organization. Their genome of about 19 kb codes for at least 7 well defined monocistronic mRNAs with the exception of one bicistronic mRNA in the LLOV genome. For EBOV and MARV the first and last nucleotides in the mRNAs are indicated, whereas for LLOV exact mRNA ends are still unclear, but lengths are roughly estimated (*).

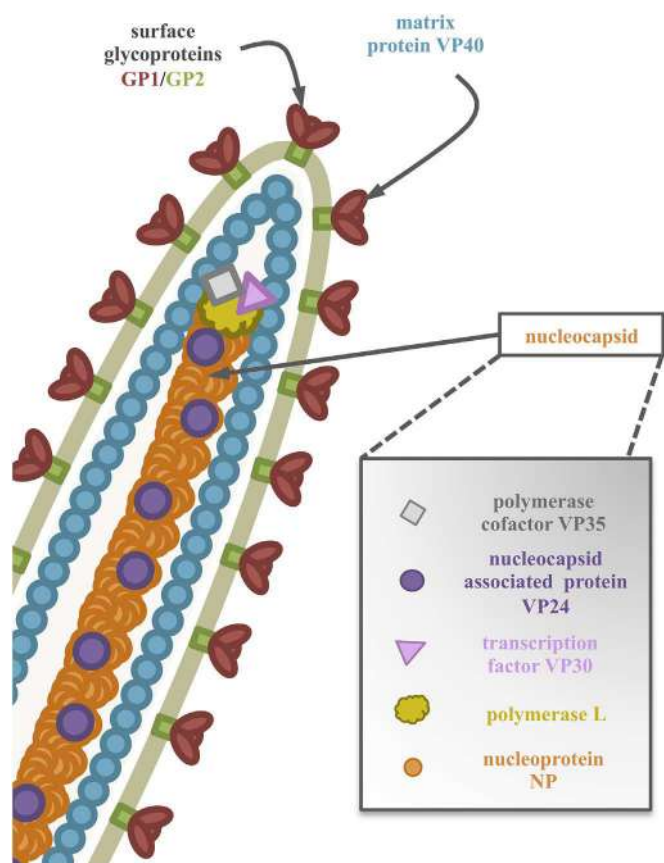


Fig. 2. Schematic ultrastructure of a filoviral particle. The viral protein assembly leads to the formation of filamentous viral particle able to infect host target cell and carrying the required material to complete a viral replication cycle. The surface glycoprotein GP_{1,2} triggers viral attachment and entry. Then, the nucleocapsid components, the RNA-protecting nucleoprotein NP, the viral proteins (VP) 30 and 35 and the “large” (L) polymerase, are released into the cytoplasm for replication and transcription, resulting in synthesis of new viral genomes and proteins. VP24 aids in nucleocapsids assembly, while the matrix protein VP40 orchestrates the formation of new virions.

and a trans-membrane fusion domain GP₂. Viral cell entry being a critical point for infection, this step has been targeted for the design of antiviral molecules. It is also noteworthy that the GP gene codes for additional proteins, whose functions are not completely understood. The mechanism driving the expression of these proteins is described below (Fig. 4).

3.1. GP: one gene, many proteins

The GP gene is the fourth gene along the genome of every filovirus (Sanchez et al., 1993; Negrodo et al., 2011). All filovirus GP genes encode a trans-membrane protein GP_{1,2} localized at the virus surface. In ebolaviruses and presumably in cuevaviruses, unedited and edited transcripts produce several forms of GP, which, together with host furin-dependent proteolysis, lead to the expression of additional proteins: the soluble GP (sGP) described in ebolaviruses and cuevaviruses, the Δ-peptide, and the small soluble GP (ssGP) (Fig. 4). It has been proposed that this edition mechanism limits the surface GP-associated cytotoxicity (see 5.) (Volchkov et al., 2001; Mohan et al., 2015). The RNA editing of ebolaviruses involves a slippage region composed of seven consecutive template uridines where viral polymerase stuttering results in a frameshift in the middle of the GP sequence (Volchkov et al., 1995; Sanchez et al., 1996; Mehedi et al., 2011). Indeed, this editing mechanism has

been recently shown to be regulated by neighboring sequences of the uridine template, probably in synergy with VP30 as a trans-acting factor (Mehedi et al., 2013). Thus, the ebolavirus GP gene is able to generate three different mRNAs coding for protein precursors pre-sGP, pre-GP and ssGP in a ratio of approximately 14:5:1, respectively, albeit the exact ratio is cell type dependent (Fig. 4) (Mehedi et al., 2011).

In ebolaviruses, the most abundant product is the unedited transcript pre-sGP mRNA, which leads to the synthesis of the protein precursor pre-sGP. Remarkably, it is not incorporated into the virion structure *per se*. This precursor of about 60 kDa is cleaved by cellular proteases of the furin family at the C-terminus of the R-X-R↓ conserved motif at position 324 (Volchkova et al., 1999). This cleavage forms the final sGP and the Δ-peptide (Fig. 4). The 40 amino-acid carboxy-terminal Δ-fragment is subsequently highly modified at the post-translational level (O-glycosylation) before being secreted. Although the Δ-peptide function has not completely been understood yet, it has been suggested that this peptide regulates filovirus entry since its expression limits infection on filovirus-permissive cells (Radoshitzky et al., 2011). Moreover, based on *in silico* analysis this peptide was also proposed to act as a virulence factor forming a lytic viroporin, although experimental evidence for such a function is lacking (Gallaher and Garry, 2015). Efforts have been devoted to characterize sGP because it shares its first N-terminal 295 residues with GP_{1,2} and ssGP (Volchkov et al., 1995). After cleavage, sGP monomers bind to each other in a parallel orientation by means of two disulfide bonds involving residues Cys53 and Cys306 (Falzarano et al., 2006). Due to the lack of a transmembrane domain, sGP forms a soluble dimer, mainly N-glycosylated, of 110 kDa. Its roles have been recently investigated *in vitro* and *in vivo*. Indeed, sGP has been suggested as a virulence factor (Volchkova et al., 2015), although currently it is not entirely clear whether or not this is the case, with one study indicating that abolishment of sGP production did not lead to a noticeable attenuation of the virus (Hoenen et al., 2015). However, the fast emergence of revertants both *in vivo* and *in vitro* (in certain cell lines) when sGP production is reduced by introducing an 8A genotype (i.e. 8 uridine residues in the virus genome, leading to production of predominantly 8A mRNAs) in the editing site suggests an important role of sGP in the biology of ebolaviruses (Volchkova et al., 2011; Kugelman et al., 2012; Hoenen et al., 2015; Tsuda et al., 2015). *In vitro*, sGP exerts vascular effects, notably the restoration of the barrier function suggesting an anti-inflammatory role (Wahl-Jensen et al., 2005); however, the relevance of this finding is currently not clear. Also, there is increasing evidence that sGP might reduce viral cytotoxicity by limiting the amount of expressed GP_{1,2} (see 5.) (Iwasa et al., 2011; Mohan et al., 2015). Finally, it has been shown that secreted sGP might also lead to immune subversion, and act as a decoy for antibodies directed against GP_{1,2} (Ito et al., 2001; Mohan et al., 2012). However, for all these roles of sGP (reviewed by de La Vega et al., 2015) further investigations are required to ascertain what relevance they really have for ebolavirus biology, and whether this protein represents a potential antiviral target.

Another protein product of the ebolavirus and cuevavirus GP gene is ssGP, a small protein of 36 kDa that is synthesized from a transcript in which either one adenosine is deleted or two are added during transcriptional editing (Fig. 4). It has been shown that ssGP is secreted as a 100 kDa dimer carrying a disulfide linkage between Cys53 of each monomer, the latter being largely N-glycosylated (Mehedi et al., 2011). Although ssGP shares similar structural properties with sGP (and GP_{1,2}), it does not seem to exert the same anti-inflammatory function on endothelial cells (Mehedi et al., 2011), and its role in viral pathogenicity, as well as its potential as antiviral target, remains unclear.

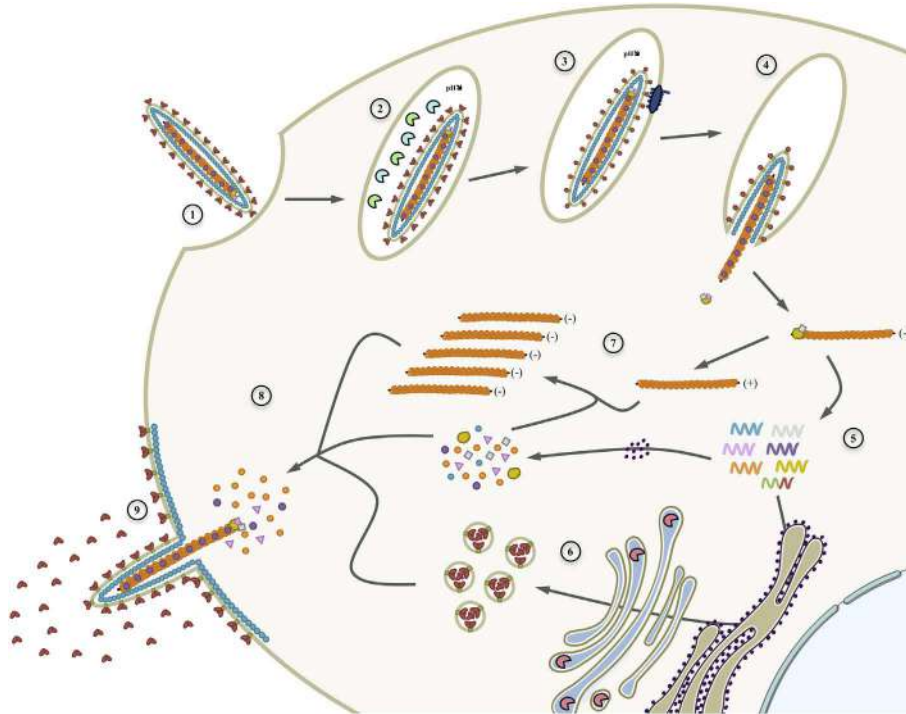


Fig. 3. Viral life cycle of filoviruses. After attachment (1), the viral particle is processed in the endosome by proteases (2) leading to receptor recognition (3) that triggers fusion and release of nucleocapsids into the host cytoplasm (4). Negative strand RNA is transcribed into messenger RNAs (5), allowing translation and protein synthesis to occur, which facilitates further secondary transcription, as well as replication through a complementary positive sense RNA (7). GP_{1,2} transits through the rough endoplasmic reticulum/Golgi apparatus pathway (6). Then, budding occurs by diverting host trafficking machinery (8), leading to the formation of new virions (9).

3.2. The viral surface glycoprotein: main structural features

The GP_{1,2} RNA transcript codes for the GP₀ precursor. This transcript results from a polymerase slippage on its template resulting in an additional adenosine on ebolavirus (and presumably cuevavirus) GP_{1,2} mRNAs, unlike in the case of the unedited marburgvirus mRNAs (Fig. 4) (Volchkov et al., 1995; Sanchez et al., 1996). mRNAs are then translated into the GP₀ precursor, which transits through the endoplasmic reticulum and the Golgi apparatus, where it is cleaved by furin-like protease(s) into two proteins, GP₁ and GP₂. The position of the conserved cleavage site is variable inside the family *Filoviridae*, but gene structure and functional organization are homologous (Volchkov et al., 2000; Manicassamy et al., 2007; Maruyama et al., 2014). Indeed, these two proteins together form a trimeric chalice structure made of three GP₁ and three GP₂ subunits (Figs. 4 and 5) assembled by GP₁/GP₂ and GP₂/GP₂ interactions (Simmons, 2013). The bowl of the chalice is shaped by the GP₁ subunits, while GP₂ organizes and anchors the complex to the membrane. In the trimer, GP_{1,2}s are bound to each other by disulfide bonds - between the Cys53 of GP₁ and the Cys609 of GP₂ as described for ebolaviruses (Lee et al., 2008) - leading to a complex and metastable intricacy.

The ectodomain GP₁ is constituted of a core protein and a mucin-like domain (MLD), which is largely glycosylated (Simmons, 2013). This MLD has not been structurally defined yet, but there are slight differences in terms of sequence, length, and position relative to the cleavage site (Volchkov et al., 1995; Sanchez et al., 1996) between different filovirus genera. Indeed, ebolaviruses hold their MLD as a single unit on GP₁ whereas marburgviruses and cuevaviruses carry it in two blocks, one in GP₁ and the other in GP₂. Despite these differences, it has been proposed that this feature plays a common role both in the attachment of the virus via lectins and in immune escape by hiding potential conserved epitopes

(Simmons, 2013). Actually, known epitopes targeted by neutralizing anti-ebolavirus antibodies such as KZ52 and 16F6 have been shown to lie in an uncovered domain present at the interface between GP₁ and GP₂ (Lee et al., 2008; Dias et al., 2011). Similarly, the antibody MR78, which was identified in a marburgvirus disease survivor, targets key residues for receptor binding, which are masked in case of ebolaviruses by the MLD, but more accessible in case of marburgviruses (Hashiguchi et al., 2015). The fact that such a cross-reactive antibody targeting the receptor-binding site (RBS) was found in a marburgvirus disease survivor, but not in ebolavirus disease survivors, is used as a further argument to strengthen the case for a function of the MLD in immune escape.

The core of GP₁ is subdivided into three domains: the glycan cap, the head, and the base (Lee et al., 2008). The glycan cap is the outer part of GP₁ forming the chalice. The head supposedly helps structuring the metastable pre-fusion conformation. This part is exposed to the host membrane surface carrying the putative RBS. The base subdomain supports the linkage with GP₂ and stabilizes the metastable pre-fusion conformation. Thus, GP₁ has the required structural features to mediate viral attachment to cell receptors (see 4.1).

The trans-membrane GP₂ protein anchors the complex to the viral membrane, but also manages virus entry and fusion (Figs. 4 and 6) (Simmons, 2013). Its structure/function complexity has been well described for ebolaviruses (Lee et al., 2008). Briefly, its structure incorporates a transmembrane domain, a short cytoplasmic tail, an internal fusion loop defined by a disulfide bond between GP₂ Cys511 and Cys556, and two heptad repeat regions (HRR1 and HRR2) surrounding the fusion peptide. This domain constitutes the unstable pre-fusion conformation of GP₂, which rearranges itself at low pH to trigger fusion. To maintain the structure in the pre-fusion state, the GP₁ head packs the GP₂ hydrophobic fusion peptide and stabilizes GP₂. Such features have not

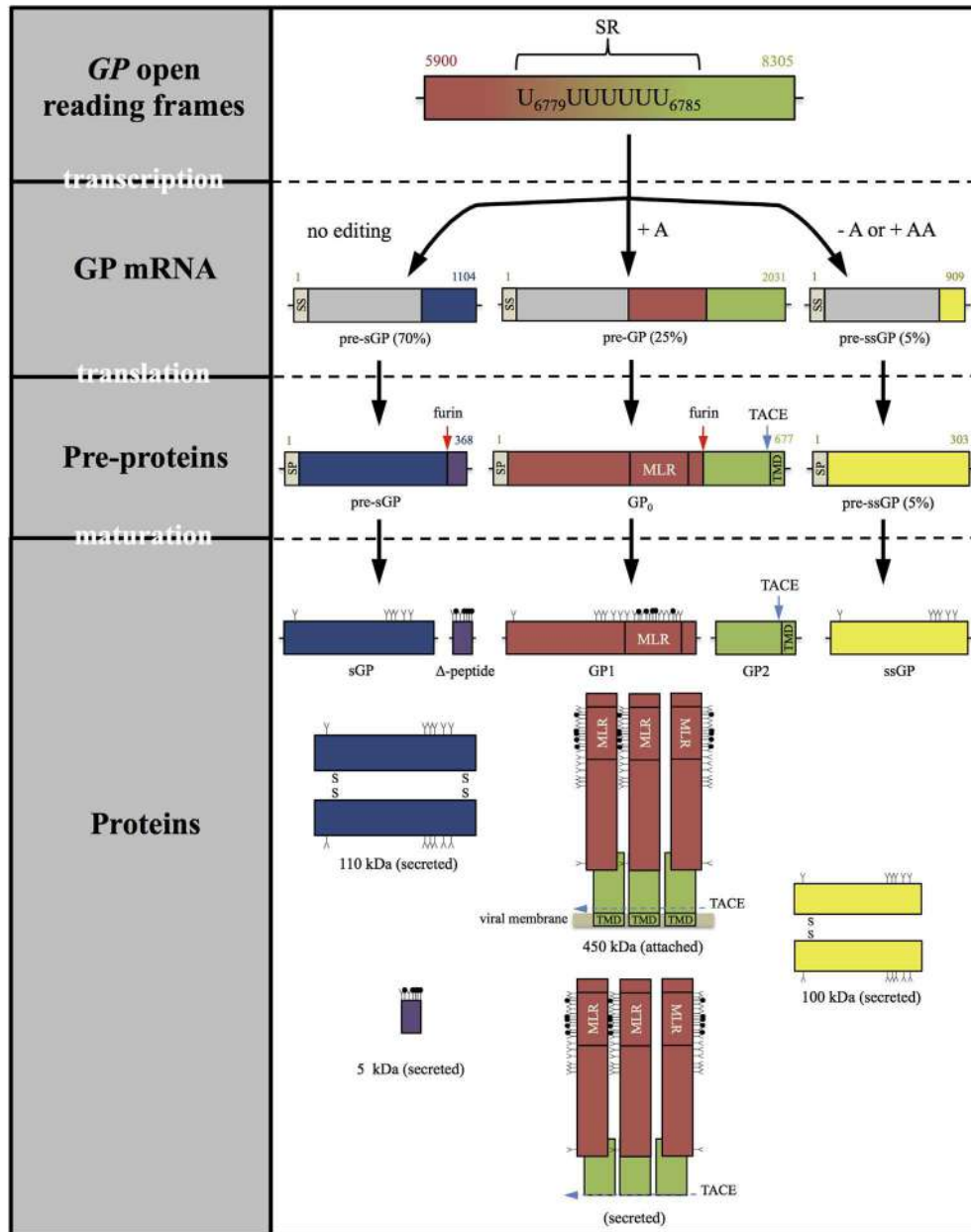


Fig. 4. Ebolavirus GP gene products. GP genes are roughly similar for ebolaviruses, marburgviruses, and cuevoviruses, with the notable exception that the marburgvirus GP gene does not undergo transcriptional editing, but only encodes GP₀. The ebolavirus GP gene, like the cuevovirus GP gene, contains a poly-U repeat (stuttering region, SR), facilitating an editing mechanism that results in the synthesis of three different mRNAs, leading to the synthesis of sGP (shown on the left), GP₀ (shown in the center), and ssGP (shown on the right). All these mRNAs contain a signal sequence (SS) and the coding sequence for the different proteins. They share a common 5' sequence (grey) leading to an identical amino-terminus for all GP-proteins. mRNAs are translated into pre-proteins transiting through the endoplasmic reticulum and the Golgi apparatus. During this intracellular trafficking, the signal peptide (SP) is removed; the protein is glycosylated (N- and O-glycosylations), and GP₀ matures by cleavage by furin-like proteases in GP₁ (red) and GP₂ (green). The mucin-like domain is part of GP₁ (red) for ebolaviruses, whereas for marburgviruses and cuevoviruses it is part of both GP₁ and GP₂. Together, GP₁ and GP₂ form GP_{1,2}, which assembles further into trimers. The surface GP_{1,2} can shed as a soluble trimer upon cleavage by the host TNF α -converting enzyme (TACE). Three other proteins, sGP (blue), the Δ -peptide (purple), and ssGP (yellow) are synthesized by ebolaviruses, and presumably also by cuevoviruses.

been clearly described yet for marburgviruses and cuevoviruses, but sequence alignments and *in vitro* data suggest similar conformations and entry/fusion mechanisms (Manicassamy et al., 2007; Maruyama et al., 2014; Liu et al., 2015).

4. GP, a fusion protein mediating filovirus cell entry

GP_{1,2} orchestrates viral entry, which can be seen as a three step mechanism: attachment, uptake, and fusion (Fig. 6) (Olejnik et al., 2011).

4.1. Filovirus attachment

The attachment between a filovirus and its target cell is mediated by GP₁, but the identity of the main cellular attachment factor remains unclear (reviewed by Takada, 2012), and in fact it appears that there are numerous proteins at the cell surface that can fulfill this function.

First, C-type lectin family members such as hMGL on immature dendritic cells and macrophages (Takada et al., 2004; Matsuno et al., 2010a) or asialo-glycoprotein, DC-SIGN, L-SIGN, L-SECtin

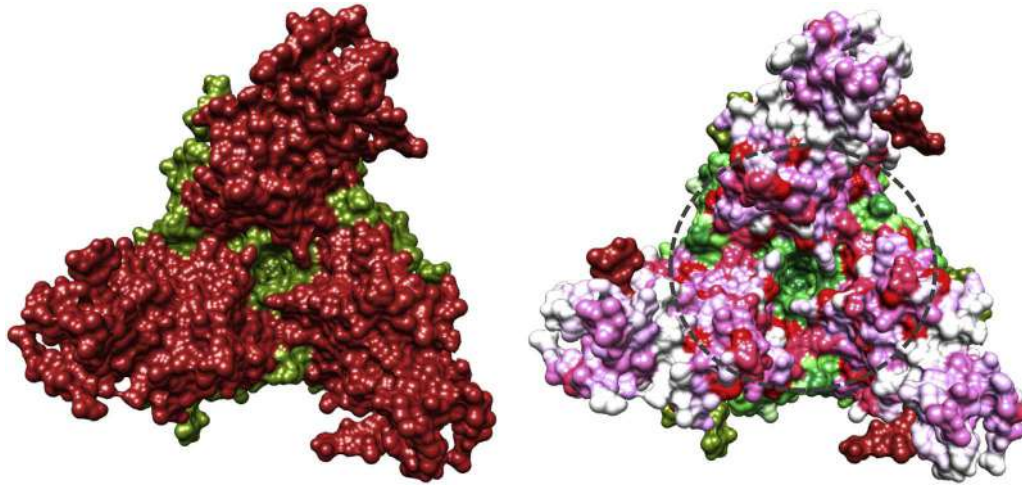


Fig. 5. Ebolavirus surface GP_{1,2} and its conservation among the family *Filoviridae*. Surface GP_{1,2} (PDB: 3CSY) is a trimer composed of three dimers of GP₁ (red) and GP₂ (green), forming a chalice at the viral envelope. The left panel presents a surface representation of the three-dimensional structure of ebolavirus GP_{1,2}, and GP₁ and GP₂ are shown in red and green, respectively. The right panel highlights the conserved residues (from dark red to light red or dark green to light green according to their conservation) derived from a sequence alignment of every filovirus species using hierarchical clustering (MultiAlin server). The figure shows that conserved residues are localized at the center of the trimeric complex (indicated with a grey dotted circle), which contains all features for priming and fusion, as well as in the external domain targeted by the cross-genus neutralizing antibody MR78.

and DC-SIGN(R) on liver endothelial cell and lymphocytes (Becker et al., 1995; Alvarez et al., 2002; Lin et al., 2003; Simmons et al., 2003; Marzi et al., 2004; Gramberg et al., 2005, 2008; Dominguez-Soto et al., 2007; Powlesland et al., 2008; Maruyama et al., 2014) are expected to interact with a set of N- and O-linked glycans on the MLD and the glycan cap of GP₁, since these lectins increase filovirus attachment. However, the binding to such molecules does not seem to be sufficient to trigger virus internalization (Simmons et al., 2003; Marzi et al., 2007; Matsuno et al., 2010b).

Additionally, β_1 -integrins have been proposed to serve as attachment factors for ebolaviruses, since infection decreases in the presence of antibodies targeting these proteins in cell lines and primary cell types (Takada et al., 2000; Simmons et al., 2002). However, no direct interaction between both molecules has been demonstrated yet, and recent studies suggest that $\alpha_5\beta_1$ -integrin is not required for GP-mediated binding of internalization, but rather is a positive regulator of cathepsins, which play an important role in processing GP₁ into its fusion-competent form within the endosomes of infected cells (Schornberg et al., 2009).

Axl, member of the Tyro3/Axl/Mer (TAM) receptor family, has also been proposed as a co-receptor for EBOV attachment but, similarly to β_1 -integrins, it may promote viral entry indirectly (Shimojima et al., 2006; Schornberg et al., 2009; Brindley et al., 2011; Hunt et al., 2011). Other cell-surface molecules, the T-cell immunoglobulin mucin domain-1 and 4 (TIM-1 and TIM-4), have been described to interact with ebolavirus GP_{1,2}, leading to virus internalization (Kondratowicz et al., 2011; Moller-Tank et al., 2013; Yuan et al., 2015; Rhein et al., 2016). Nevertheless, despite this clearly demonstrated role, only epithelial cells and some antigen-presenting cells subsets significantly express TIM-1 and TIM-4 respectively, suggesting that there are other attachment factors involved in other filovirus-susceptible cell types. TAM, TIM-1 and TIM-4 could target phosphatidyl-serine (PtdSer), which is exposed on the outer leaflet of the filovirus membrane, strengthening an interplay promoting efficient attachment (reviewed in Moller-Tank and Maury, 2014).

Finally, a last attachment mechanism has been described, reminiscent to an antibody-dependent enhancement (ADE) process, by which filoviruses divert virus-specific antibodies to get attached to immune system cells through cellular Fc receptors or

via the complement component C1q and its ligands identified in most mammalian cells (Takada and Kawaoka, 2001; Takada et al., 2003, 2007; Nakayama et al., 2010). Interestingly, viral pathogenicity seems to correlate with filovirus ADE (Takada and Kawaoka, 2001; Nakayama et al., 2011).

In all cases, it appears that attachment requires a set of proteins that interact in a complex manner to promote entry.

4.2. Uptake and proteolytic processing

The uptake is a key step in filovirus entry, as it serves to transform the pre-fusion GP_{1,2} conformation into a primed GP_{1,2} that triggers fusion events (Fig. 6).

First, the internalization has been proposed to involve different endocytic pathways. The precise mechanisms were controversial in the past, since clathrin-dependent and caveolin-dependent uptakes have been shown to occur (Bavari et al., 2002; Empig and Goldsmith, 2002; Sanchez, 2007; Bhattacharyya et al., 2010, 2011). However, latest data support that the filovirus uptake mechanism is mainly mediated by macropinocytosis and depends among other factors on the host cell and virus particle size (Nanbo et al., 2010; Saeed et al., 2010; Aleksandrowicz et al., 2011).

After internalization, macropinocytosis vesicles are routed to endosomal vesicles of the host cell, where proteolytic events occur to prime GP_{1,2} (Fig. 6). Once in the endosome, EBOV GP_{1,2} is sequentially processed by the cysteine proteases cathepsin B (catB) and/or cathepsin L (catL) under acidic pH and reducing conditions (Chandran et al., 2005; Schornberg et al., 2006; Brecher et al., 2012). Concisely, it appears that, for EBOV, BDBV, and TAFV, catB removes the major part of GP₁ (glycan cap and MLD). The proteolytic events are slightly different for SUDV, RESTV, MARV and LLOV, as catB has been shown to be dispensable (Chan et al., 2000; Gnirss et al., 2012; Xia et al., 2012; Maruyama et al., 2014; Ng et al., 2014). It is thought that catL partially removes a part of the GP₁ cap and subsequent proteolysis – by catB, catL and/or other proteases – results in a smaller GP₁ form of less than 20 kDa. With latest structural inputs, further information regarding cathepsin cleavage have become available. Indeed, amongst the minor differences between EBOV and MARV GP_{1,2}, a catB cleavage site identified on EBOV has been shown to be disordered, and thus potentially easily accessible,

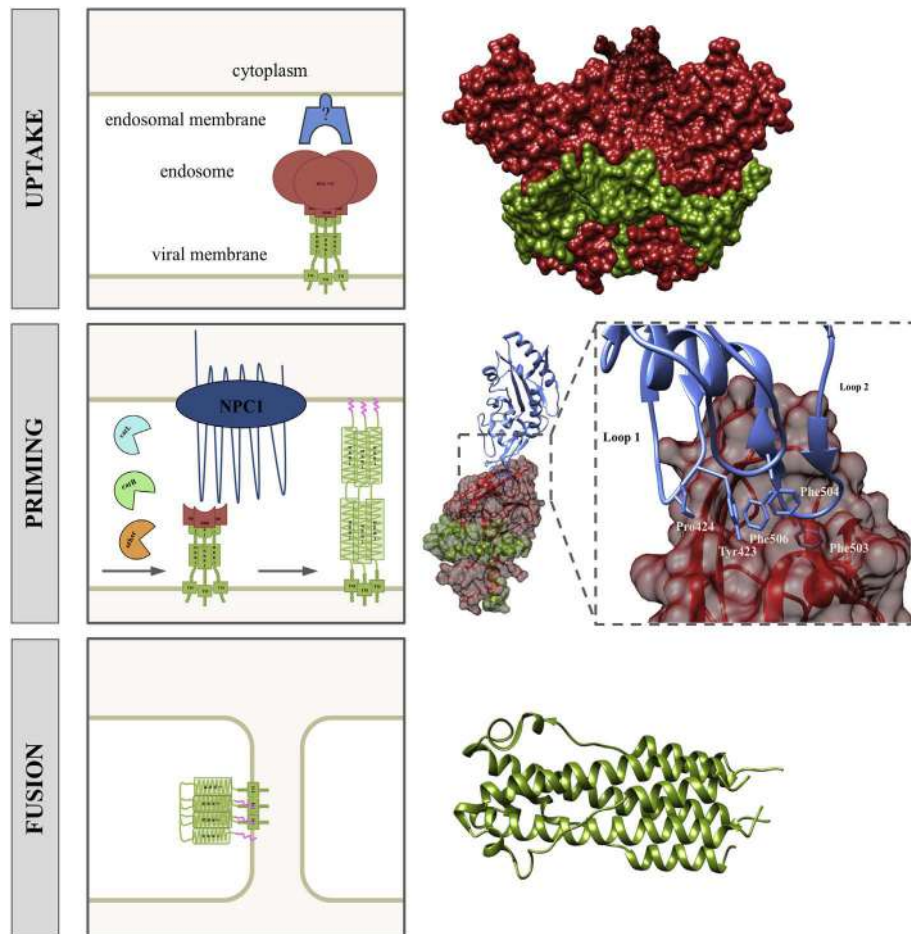


Fig. 6. Surface GP endosomal processing. After attachment mediated by interaction between the filovirus surface protein GP_{1,2} (PDB: 3CSY) and various attachment factors, the complex is internalized and routed to the endosome, where GP_{1,2} is processed to trigger fusion of viral and host membranes. First, in the endosomal low pH environment, cathepsin proteases L&B (catL & catB) and others remove the mucin-like region (MLD) and the glycan cap (GC) domain of GP₁ (red). A receptor binding domain (RBD), also carried by GP₁, is unmasked, leading to the interaction with a mainly hydrophobic pocket in the N-terminal domain (blue) of the endosomal protein Niemann-Pick C1 (NPC1) (PDB: 5F1B). This interaction together with other only partially understood molecular events remove GP₁ constraints on GP₂ (green), forming the primed-GP_{1,2} capable to induce fusion. The GP₂ heptad repeat regions 1 and 2 (HRR1 and HRR2) then rearrange themselves, pushing out the fusion peptide (pink) to anchor it in the host membrane. This intermediate pre-hairpin conformation destabilizes membrane bilayers, and a folding-back into a six-bundle helices conformation (PDB: 2EBO) merges membranes, opening a fusion pore for the release of viral nucleocapsids.

while in MARV the homologous region is an α -helix (Hashiguchi et al., 2015). However, there remain open questions regarding filovirus GP proteolysis processing, since other proteases were also proposed to participate to this process. Further, while much of the cathepsin work was done using GP-pseudotyped VSV or retrovirus particles, work with infectious ebolaviruses challenges the importance of cathepsins for the viral life cycle. Particularly, virus replication *in vitro* as well as *in vivo* was shown to be independent of catB and catL, and cathepsin knockout mice succumbed to ebolavirus challenge, albeit not to VSV infection (Marzi et al., 2012).

In any case, these cleavages trigger an increase in binding and infectivity for Ebola virus by unmasking the potential RBS, which was firstly described as a 6 residue peptide on GP₁ (K114, K115, K140, G143, P146, K147) (Lee et al., 2008). Recently, the Niemann-Pick C1 protein (NPC1), which is an anchored late endosomal/lysosomal protein physiologically implicated in cholesterol absorption and homeostasis, has been shown to interact with this region of GP₁ (Fig. 6) (Carette et al., 2011; Côté et al., 2011) downstream both attachment and uptake events independently from its proper role in cholesterol trafficking (E. H. E.H. Miller et al., 2012; White and Schornberg, 2012). Indeed, the latest structural data define the RBS as a negatively charged crest and a hydrophobic

trough at the apex of GP₁ trimer that interact mostly with the domain C (loops 1 and 2) and partially with the domain NTD of NPC1 (Bornholdt et al., 2016; Gong et al., 2016; Wang et al., 2016; Zhao et al., 2016). This structural characterization of the NPC1/GP₁ interaction provides key information to design promising antiviral strategies. More importantly, further analyses described NPC1 as a crucial receptor that potentially determines the species susceptibility to filoviruses (Ng et al., 2015; Hoffmann et al., 2016; Ndongu et al., 2016), paving the way towards filovirus-specific antiviral molecules (see 6.2.).

4.3. Membrane fusion

Even if deeper investigation is needed into the mechanism of NPC1 during filovirus entry, its interaction with cleaved GP₁ and the endosomal environment triggers a conformational change of GP₂ initiating membrane fusion (Fig. 6) (Lee et al., 2008; Kuroda et al., 2015). Fusion necessitates large conformational changes representing a high-energy barrier. With GP₁ cleavages and a low-pH environment, it is hypothesized that both GP₁ constraint removal from the metastable GP₂ and histidine protonation generate the energy triggering the fusion events (Kampmann et al., 2006; Lee et al.,

2008, 2016; Markosyan et al., 2016). At this point, the disordered HRRs rearrange themselves into α -helices. These helices push through the host membrane, and then anchor the hydrophobic fusion peptide in target membrane. Within the membrane, key residues, amongst them Pro537, trigger bilayer destabilization with an extended pre-hairpin intermediate conformation (Fig. 6). Therefore, elongated GP₂ spans and thus branches both host and viral membranes. The fusion finally occurs by the collapse of this intermediate into a folding-back conformation that distorts both viral and host membranes with their simultaneous rapprochement. GP₂ acts as a clamp to reach its low energy state characterized as a six helices bundle, which is the GP₂ hairpin post-fusion structure (Fig. 6). This novel conformation leads to a merge into a hemi-fusion stalk and then to the opening of a fusion pore allowing the release of the nucleocapsid into the host cytoplasm (Fig. 6).

5. Modulation of cytotoxicity and inflammation

While the principle function of GP_{1,2} is cell infection, several lines of evidence suggest that it might also be involved in pathogenesis. Cellular GP_{1,2}-related cytotoxic events occur in infected cells. This cytotoxicity is reflected in rounding of cells *in vitro*, due to masking of various host cell surface molecules such as β 1-integrins (Takada et al., 2000; Francica et al., 2009). This masking has been shown to be a “glycan umbrella” mediated steric shielding of these adhesion proteins mediated by surface-expressed GP_{1,2}, leading to a loss of accessibility and function of these host surface proteins (Reynard et al., 2009; Francica et al., 2010). The masking of such adhesion molecules leads to cell detachment, and may contribute to the disruption of blood vessel integrity and hemorrhages developed during a filovirus infection. Indeed, in primary endothelial cell cultures overexpression of GP_{1,2} after transduction with an adenovirus vector resulted in a loss of adherence resulting in apoptosis (Ray et al., 2003). Additionally, steric shielding by GP_{1,2} also affects MHC-I surface expression (Francica et al., 2009), a phenomenon also known to occur in infected primary endothelial cells (Harcourt et al., 1999). This might alter immune cells recruitment, and thus may participate in the immune suppression and inflammatory dysfunction linked to a filovirus infection (Harcourt et al., 1999; Reynard et al., 2009).

As mentioned in 3.1., sGP has been shown to regulate GP_{1,2} expression thanks to transcriptional editing in the GP gene (Volchkov et al., 2001; Mohan et al., 2015). Interestingly, when EBOV is serially passaged in certain cell lines like Vero cells, the GP gene gets modified to predominantly generate GP_{1,2} by addition of an 8th U in the editing site (n.b., in contrast to transcriptional editing this occurs on the genome level) (Volchkova et al., 2011). However, when this EBOV is further passaged in animals or in other cell lines like Huh7 cells, the genotype reverts back so that sGP is again produced as the main product of transcription (Volchkova et al., 2011; Hoenen et al., 2015; Tsuda et al., 2015). This observation suggests that either the importance of the function of sGP or the importance of GP gene modulation is dependent on the host cell environment.

Additionally, if GP_{1,2} is over-expressed, the cellular glycosylation machinery is overwhelmed, which competitively inhibits physiological processes (Volchkov et al., 2001; Mohan et al., 2015). Nevertheless, at later stages of viral infection, GP_{1,2} gets highly expressed, which is associated with a massive release of viral particles by infected cells. Thus, cytotoxicity appears even in these situations to be under control by transcriptional editing. However, viral particles produced at this point have been shown to be less pathogenic, suggesting that a quantitative balance in GP_{1,2} expression might also be required to regulate infectivity (Mohan et al., 2015).

Finally, another form of GP_{1,2}, known as shed GP or GP_{1,2} Δ TM, has been shown to be released during ebolavirus infection (Dolnik et al., 2004, 2015). Indeed, GP_{1,2} Δ TM is released from surface-expressed GP_{1,2} by the host TNF α -converting enzyme (TACE), a member of the disintegrin and metalloproteinase (ADAM) family, leaving only the transmembrane and cytoplasmic parts of GP₂ on the cell surface (Fig. 4). This secreted protein may play a role in viral pathogenesis, blocking neutralizing antibodies or stimulating ADE. It might also contribute to an impaired inflammatory response (Escudero-Pérez et al., 2014). A last potential role of shed GP might be the modulation of endothelium homeostasis, thus triggering increase of vascular permeability, and disseminated intravascular coagulation, ultimately causing multi-organ failure and death (Escudero-Pérez et al., 2014).

6. Potential of GP targeting antivirals

Since filovirus GP_{1,2} plays a key role in virus entry, strategies targeting different key points in this process are supposed to block replication at an early stage, thus reducing the chance of virus spread and its potential evolution towards drug resistance. Therefore, numerous methods targeting filovirus entry have been investigated. (i) Immune-based therapies have been rapidly developed (reviewed by Choi and Croyle, 2013) leading to efficient monoclonal antibody cocktails (Olinger et al., 2012; Qiu et al., 2012, 2013a, 2013b) such as the promising antibody-based drug against ebolaviruses ZMapp[®] (Qiu et al., 2014). (ii) Alternatively, peptide-based antiviral molecules like Tat-Ebo and analogs were designed in order to block cell membrane fusion (Miller et al., 2011; Higgins et al., 2013). The inhibition is based on peptides limiting HRR1 and 2 interactions to block GP₂ extension. Such compounds have still to be optimized in terms of treatment window and dosage. (iii) Other strategies based on a broad range of small molecules have been built up to disrupt the entry step. These potential antiviral compounds have been heterogeneously characterized from *in vitro* high throughput screening hits to *in vivo* studies (reviewed by Nyakatura et al., 2015; Rhein and Maury, 2015). Here are reported the latest advances related to small entry inhibitors, which can be sorted as broad-spectrum molecules, filovirus-specific compounds, and FDA-approved therapeutics (summarized in Table 1). With regards to antibody-based therapies the interested reader is referred to (Zeitlin et al., 2016).

6.1. Broad-spectrum molecules

As broad-spectrum molecules we define inhibitors targeting key points of the entry process that are effective against multiple virus multiple RNA viruses. Indeed, filoviruses share entry steps with other viruses using class I envelope glycoproteins such as HIV or influenza.

The first target that has been considered is attachment. Various soluble mannose-specific lectins from plants (e.g., concanavalin A and cyanovirin N) have been proposed because of their potential antiviral effect on HIV-1 (Balzarini et al., 2004). Another example is griffithsin, a lectin of terminal mannose residues of asparagine (N)-linked Man 5-9 GlcNAc₂ structures purified from red-algae, which might also have antiviral potential since these carbohydrate residues are found on HIV-1 and 2, HCV, SARS-CoV and, relevant to our discussion, on ebolaviruses (Barton et al., 2014). Similarly, recombinant human mannose binding lectin has been shown in a mouse model to be protective against ebolaviruses (Michelow et al., 2011); however, the mouse model, while a necessary and important early step in drug evaluation for ebolaviruses, has only limited predictive value regarding the effectivity of a treatment in other, more stringent animal models of EVD, or even human patients.

Table 1
Antiviral compounds acting against filovirus entry.

Category	Target	Drug activity	Molecules	References
Broad spectrum molecules	Attachment	Mannose-specific lectins	griffithsin, concanavalin A, cyanovirin N	(Balzarini et al., 2004; Barton et al., 2014)
	Endosomal processing	Protease inhibitors	Cys&Ser protease inhibitors (leupeptin) Cys protease inhibitors (E-64 and E-64d) catL&catB inhibitors (FY-dmk) catB inhibitors (CA-074, CA-074Me, nafamostat mesilate) catL inhibitors (oxobarzate, ZY(t-Bu)-dmk, triazine derivatives 5705213 and 7402683, and K11777) genistein, tyrphostin AG1478, chloroquine	(Chandran et al., 2005; Schornberg et al., 2006; Barrientos et al., 2007; Shah et al., 2010; Gnirss et al., 2012; Elshabrawy et al., 2014; Nishimura et al., 2015; Zhou et al., 2015; van der Linden et al., 2016)
Filovirus-specific compounds	Fusion	Endosome disruption Intercalating agents Unclear	LJ001, dUY11, arbidol 25HC oxysterols, teicoplanin tridecafullerenes	(Savarino et al., 2001; Keyaerts et al., 2004; Keyaerts et al., 2009; Madrid et al., 2013) (St Vincent et al., 2010; Wolf et al., 2010; Pécheur et al., 2016) (Liu et al., 2013; Schoggins et al., 2013; Wang et al., 2016) (Muñoz et al., 2016)
	Attachment Fusion	Lectin competitors NPC1 inhibitors	U18666A, imipramine, Ro48-8071, compounds 3.47 and 3.0, and MBX2254 and MBX2270	(Cenedella, 2009; Côté et al., 2011; Kolokoltssov et al., 2012; Miller et al., 2012; Lee et al., 2013; Shoemaker et al., 2013; Basu et al., 2015)
Approved drugs against other targets	Attachment	GP2 inhibitors Glycosaminoglycan competitor	compound 7 heparin	(Basu et al., 2011) (Salvador et al., 2013; Cheng et al., 2015; O'Hearn et al., 2015)
	Uptake	Cytoskeleton inhibitors	vinblastine, vincristine, colchicine, nocodazole, cytochalasin B and D, latrunculin A, chondramides	(Kouznetsova et al., 2014; Beck et al., 2016)
	Endosomal processing	G protein-coupled receptor antagonists Endosomal pH increase	benztropine mesylate omeprazol, esomeprazol	(Cheng et al., 2015) (Long et al., 2015)
	Unclear	Estrogen receptor modulators Ion channel inhibitors	clomiphene, toremiphene, raloxifene, taxomifene amiodarone, dronedarone, verapamil, tetrandine, nimodipine, diltiazem, digoxin, rottlerin, noricumazole A)	(Johansen et al., 2013; Shoemaker et al., 2013; Gehring et al., 2014; Kouznetsova et al., 2014; Sakurai et al., 2015; Zhao et al., 2016; Beck et al., 2016)

In addition to serological and peptide-based approaches, numerous studies report potential small anti-filovirus molecules. These compounds can be sorted in three categories (broad-spectrum molecules, filovirus-specific compounds and repurposed FDA-approved therapeutics). In each category, molecules have been arranged according to their targeted entry process and then following their specific activity when described.

The inhibition of catL and catB, involved in EBOV GP_{1,2} endosomal processing, could also have a broad-spectrum antiviral effect. Various more or less specific inhibitors have been described: the unselective Cys and Ser protease inhibitor leupeptin, unselective Cys protease inhibitors E-64 and E-64d and its recent derivatives, the mixed catL and B inhibitor FY-dmk, the specific catB inhibitors CA-074, CA-074Me and nafamostat mesilate (which has a dual action also targeting factor VIIa for an anticoagulation action), and catL-specific inhibitors oxobarzate, ZY(t-Bu)-dmk (also known as CID23631927), triazine derivatives 5705213 and 7402683, and K11777 (Chandran et al., 2005; Schornberg et al., 2006; Barrientos and Rollin, 2007; Shah et al., 2010; Gnirss et al., 2012; Elshabrawy et al., 2014; Nishimura and Yamaya, 2015; Zhou et al., 2015; van der Linden et al., 2016). All these compounds have been characterized in ebolavirus entry modelling systems, such as pseudotypes, and/or studies with infectious virus. They are mechanism-based suicide inhibitors, since they carry an epoxide or diazomethane functional group. They are not only active on filoviruses, but also on SARS-CoV as well as Hendra and Nipah viruses (Elshabrawy et al., 2014). As far as both anti-filovirus attachment and anti-uptake are concerned, these inhibitors need further optimization, especially concerning their toxicity.

Molecules targeting fusion constitute another class of inhibitors. Viruses being unable to repair membrane damage, the use of membrane intercalating agents such as aryl methyldiene rhodamine derivative LJ001 (Wolf et al., 2010), dUY11 (St Vincent et al., 2010), a rigid amphipathic fusion inhibitor (RAFI), the indole based hydrophobic molecule arbidol (Pécheur et al., 2016),

presumably synthetic 25HC oxysterols (Liu et al., 2013; Schoggins and Randall, 2013) and teicoplanin (Y. Wang et al., 2016) may impair virus/host cell membrane fusion. To date, development is at a proof-of-concept stage, but the strategy showed a broad efficiency against ebolaviruses, influenza, HIV, pox-, arena-, bunya-, herpes-, paramyxo- and flaviviruses (Wolf et al., 2010).

Other broad-spectrum compounds can be mentioned such as genistein, a broad tyrosine kinase inhibitor, or tyrphostin AG1478, an epidermal growth factor receptor tyrosine kinase blocker. Both trigger the disruption of endocytosis and endosome formation used by numerous viruses (Chandran et al., 2005). The 9-aminoquinoline, long known as the antimalarial agent chloroquine (Savarino et al., 2003), showed beneficial effects, from endocytosis to exocytosis with increase of endosomal pH, diminution of INF- γ and TNF- α production, on HIV-1, SARS-CoV and ebolaviruses *in vitro* (Savarino et al., 2001; Keyaerts et al., 2004, 2009; Madrid et al., 2013).

Unfortunately, for some of these compounds animal studies have shown that they are not effective against ebolaviruses *in vivo* (Dowall et al., 2015; Falzarano et al., 2015; Akpovwa, 2016), whereas other compounds still need to be tested in such models, before any conclusions regarding their effectiveness can be made. Importantly, different *in vitro* and *in vivo* models have different predictive values regarding the effectiveness of drugs, with *in vitro* experiments providing the weakest evidence (although they remain an essential starting point), followed by the various mouse models, whereas guinea pigs provide a more stringent model, and non-human primates constitute the most stringent model for

ebolavirus and marburgvirus disease. These characteristics have to be kept in mind when evaluating these and other antiviral drugs against filoviruses, or when interpreting study results. Further, all these compounds target a broad spectrum of pathogenic viruses; however, since they often target the cellular machinery, selectivity and toxicity will inevitably remain an important issue.

6.2. Filovirus-specific inhibitors

An alternate strategy would be to select molecules targeting filoviruses specifically. The GP₂/NPC1 interaction that triggers fusion events in filoviruses has been highly considered and recently characterized (Bornholdt et al., 2016; Gong et al., 2016; Wang et al., 2016; Zhao et al., 2016). Such a strategy is challenging, since the interface between NPC1 and GP₂ is large, dynamic and mainly hydrophobic, but also very promising for the development of filovirus specific antiviral molecules.

Development of protein-protein interaction inhibitors towards both GP₂ and NPC1 has been reported. Several compounds have been described targeting NPC1, such as U18666A, imipramine, cationic amphiphiles Ro48-8071 and the terconazole adamantane class of compounds like benzylpiperazine adamantane diamides 3.47 and 3.0, or the lately discovered sulfonamide derivative MBX2254 and the triazole thioether derivative MBX2270 (Rodriguez-Lafrasse et al., 1990; Cenedella, 2009; Côté et al., 2011; Lee et al., 2013; Shoemaker et al., 2013; Basu et al., 2015). Most of these inhibitors helped to identify NPC1 as a critical factor for filovirus entry. Interestingly, the tertiary amine imipramine inhibits both NPC1 and sphingomyelinase, resulting in ebolavirus entry inhibition (M.E. Miller et al., 2012). As NPC1 is involved in other cellular functions, its inhibition disrupts its primary role in cholesterol homeostasis and triggers cholesterol accumulation in endosomes, as is the case in Niemann-Pick disease type C (Lee et al., 2013; Shoemaker et al., 2013). Therefore, further development might be required to uncouple the NPC1 function in cholesterol traffic from its implication in filovirus entry. This unbundling should be tricky, since both cholesterol management and ebolavirus binding are facilitated by the same domains of NPC1. However, as ebolavirus disease constitutes an acute infection, it should be possible to keep the treatment time relatively short, so that this side effect might have limited detrimental effects in practice.

An alternative strategy avoiding unwanted host-directed effects might be to target GP₂, since it is a viral protein with no cellular homolog. Indeed, a benzodiazepine derivative (referred to as compound 7) has been shown - by computer calculation and mutagenesis - to mask the hydrophobic conserved S2 pocket defined at the GP₁/GP₂ interface, inhibiting ebolavirus replication (Basu et al., 2011). These results also show that this S2 pocket does not correspond to the trough of the RBS, suggesting that it might be another binding site or a steric regulation region of the NPC1 binding site. Although the mechanism of inhibition has yet to be clarified, S2 pocket and RBS pharmacophores might be key targets to design filovirus-specific antiviral molecules.

Lately, another filovirus-specific strategy has been described using tridecafullerene derivative compounds (17a and 17c) (Muñoz et al., 2016). These compounds are giant globular multivalent molecules carrying sugar motifs able to compete with lectins. Dependent on the sugar motifs linked on such "superballs", it is possible to target specific lectins. Targeting DC-SIGN for example resulted in inhibition of ebolavirus *in vitro* at subnanomolar concentrations with no cytotoxicity notable, and perfect solubility in water. As a very new technology, these tridecafullerene derivatives require deeper investigations to depict their mechanism of action *in vitro* and their effectiveness *in vivo*, as well as their specificity and their delivery to target organs.

6.3. Repurposing of approved drugs

Given that the development of new therapeutics is a long and expensive process, one approach to rapidly develop drugs against filovirus infection is to reposition compounds that have been already approved by national and international health agencies (Kouznetsova et al., 2014).

Using this strategy, cationic amphiphiles such as the estrogen receptor modulators clomiphene (infertility treatment to induce ovulation), toremiphene (breast cancer chemotherapeutic) raloxifene (osteoporosis prevention) and taxomifene (breast cancer chemotherapeutic), and ion channel inhibitors (cardiac arrhythmias and vascular modulation) such as amiodarone, dronedarone, verapamil, tetrandine, nimodipine, diltiazem, digoxin, and rottlerin have been shown to inhibit filovirus entry (Johansen et al., 2013; Shoemaker et al., 2013; Gehring et al., 2014; Kouznetsova et al., 2014; Sakurai et al., 2015; Nelson et al., 2016). Other channel blocking compounds such as noricumazole A from myxobacteria have been also described as potent inhibitors (Beck et al., 2016). Their inhibition mechanisms are not yet understood, but are distinct from GP₂/NPC1 interaction-based inhibitors.

Another group of therapeutics targeting the cytoskeleton has been described to have anti-EBOV potential (vinblastine, vincristine, colchicine, nocodazole, cytochalasin B and D, latrunculin A and others) (Kouznetsova et al., 2014; Beck et al., 2016). However, these molecules are supposed to be quite toxic, since they are mostly used for anti-cancer chemotherapies.

Additionally, as a large class of approved therapeutics, G protein-coupled receptor (GPCR) antagonists have been screened. Benzotropine mesylate, initially used against Parkinson's disease, exhibits *in vitro* anti-filovirus activity targeting a post-attachment step (Cheng et al., 2015). Such molecules deserve a thorough mechanistic investigation, since there is no other literature mentioning the involvement of GPCR in filovirus entry.

Other drugs such as omeprazole and esomeprazole (acid reflux disease) have off-target inhibitory activity triggering increase of endosomal pH, which inhibits late entry events during a filovirus infection (Long et al., 2015). Also, several studies described heparin as a potent inhibitor of early attachment and co-reception via glycosaminoglycans (Salvador et al., 2013; Cheng et al., 2015; O'Hearn et al., 2015).

However, further studies are needed to unravel the antiviral mechanism of those FDA-approved compounds. Although these compounds have already passed safety tests, it is important to underline that dosing and treatment window will likely be a critical issue before use.

7. Conclusion

Antibodies such as ZMapp[®] have demonstrated that inhibitors targeting filovirus entry might provide a potent antiviral strategy. However the development of small molecules inhibiting this crucial viral step is still hampered by both the insufficient molecular characterization of the attachment/uptake mechanism, and the multiple molecular events occurring in endosomal compartments. To date, the identification of the main characters is clear, but the fine-tuning of each step has to be clarified. The inhibitors presented herein appear to function at several entry steps, but the most susceptible one has yet to be identified. Targeting pre-fusion GP_{1,2} or the GP₂/NPC1 interface might represent the best strategy to develop filovirus-specific treatments. With a more integrative vision in mind, antiviral therapies might advantageously complement vaccines and serological therapies during filovirus outbreaks with their simple logistics especially in West and Central Africa.

Acknowledgements

This work was supported by a DGA-MRIS and Aix-Marseille Université scholarship and by the ANR grant ANR-ST14-ASTR-0026 for the project VMTaseln.

References

- Akpovwa, H., 2016. Chloroquine could be used for the treatment of filoviral infections and other viral infections that emerge or emerged from viruses requiring an acidic pH for infectivity. *Cell Biochem. Funct.* 34, 191–196. <http://dx.doi.org/10.1002/cbf.3182>.
- Aleksandrowicz, P., Marzi, A., Biedenkopf, N., Beimforde, N., Becker, S., Hoenen, T., Feldmann, H., Schnittler, H.-J., 2011. Ebola virus enters host cells by macropinocytosis and clathrin-mediated endocytosis. *J. Infect. Dis.* 204 (Suppl. 3), S957–S967. <http://dx.doi.org/10.1093/infdis/jir326>.
- Alvarez, C., Lasala, F., Carrillo, J., Muñoz, O., Corbí, A., Delgado, R., 2002. C-type Lectins DC-SIGN and L-SIGN Mediate Cellular Entry by Ebola Virus in Cis and in Trans [WWW Document]. URL <http://www.ncbi.nlm.nih.gov/pubmed/?term=12050398> (accessed 03.01.16).
- Ascenzi, P., Bocedi, A., Heptonstall, J., Capobianchi, M.R., Di Caro, A., Mastrangelo, E., Bolognesi, M., Ippolito, G., 2008. Ebolavirus and marburgvirus: insight the Filoviridae family. *Mol. Asp. Med.* 29, 151–185. <http://dx.doi.org/10.1016/j.jam.2007.09.005>.
- Balzarini, J., Hatse, S., Vermeire, K., Princen, K., Aquaro, S., Perno, C.-F., De Clercq, E., Egberink, H., Vanden Mooter, G., Peumans, W., Van Damme, E., Schols, D., 2004. Mannose-specific plant lectins from the Amariyllidaceae family qualify as efficient microbicides for prevention of human immunodeficiency virus infection. *Antimicrob. Agents Chemother.* 48, 3858–3870. <http://dx.doi.org/10.1128/AAC.48.10.3858-3870.2004>.
- Barrientos, L.G., Rollin, P.E., 2007. Release of cellular proteases into the acidic extracellular milieu exacerbates Ebola virus-induced cell damage. *Virology* 358, 1–9. <http://dx.doi.org/10.1016/j.virol.2006.08.018>.
- Barton, C., Kouokam, J.C., Lasnik, A.B., Foreman, O., Cambon, A., Brock, G., Montefiori, D.C., Vojdani, F., McCormick, A.A., O'Keefe, B.R., Palmer, K.E., 2014. Activity of and effect of subcutaneous treatment with the broad-spectrum antiviral lectin griffithsin in two laboratory rodent models. *Antimicrob. Agents Chemother.* 58, 120–127. <http://dx.doi.org/10.1128/AAC.01407-13>.
- Basu, A., Li, B., Mills, D.M., Panchal, R.G., Cardinale, S.C., Butler, M.M., Peet, N.P., Majgier-Baranowska, H., Williams, J.D., Patel, I., Moir, D.T., Bavari, S., Ray, R., Farzan, M.R., Rong, L., Bowlin, T.L., 2011. Identification of a small-molecule entry inhibitor for filoviruses. *J. Virol.* 85, 3106–3119. <http://dx.doi.org/10.1128/JVI.01456-10>.
- Basu, A., Mills, D.M., Mitchell, D., Ndungo, E., Williams, J.D., Herbert, A.S., Dye, J.M., Moir, D.T., Chandran, K., Patterson, J.L., Rong, L., Bowlin, T.L., 2015. Novel small molecule entry inhibitors of ebola virus. *J. Infect. Dis.* 212 (Suppl. 2), S425–S434. <http://dx.doi.org/10.1093/infdis/jiv223>.
- Bavari, S., Bosio, C.M., Wiegand, E., Ruthel, G., Will, A.B., Geisbert, T.W., Hevey, M., Schmaljohn, C., Schmaljohn, A., Aman, M.J., 2002. Lipid raft microdomains: a gateway for compartmentalized trafficking of Ebola and Marburg viruses. *J. Exp. Med.* 195, 593–602.
- Beck, S., Henß, L., Weidner, T., Herrmann, J., Müller, R., Chao, Y.K., Grimm, C., Weber, C., Sliva, K., Schnerle, B.S., 2016. Identification of entry inhibitors of Ebola virus pseudotyped vectors from a myxobacterial compound library. *Antivir. Res.* 132, 85–91. <http://dx.doi.org/10.1016/j.antiviral.2016.05.017>.
- Becker, S., Spiess, M., Klenk, H.D., 1995. The asialoglycoprotein receptor is a potential liver-specific receptor for Marburg virus. *J. Gen. Virol.* 76, 393–399. <http://dx.doi.org/10.1099/0022-1317-76-2-393>.
- Bhattacharyya, S., Hope, T.J., Young, J.A.T., 2011. Differential requirements for clathrin endocytic pathway components in cellular entry by Ebola and Marburg glycoprotein pseudovirions. *Virology* 419, 1–9. <http://dx.doi.org/10.1016/j.virol.2011.07.018>.
- Bhattacharyya, S., Warfield, K.L., Ruthel, G., Bavari, S., Aman, M.J., Hope, T.J., 2010. Ebola virus uses clathrin-mediated endocytosis as an entry pathway. *Virology* 401, 18–28. <http://dx.doi.org/10.1016/j.virol.2010.02.015>.
- Bornholdt, Z.A., Ndungo, E., Fusco, M.L., Bale, S., Flyak, A.I., Crowe, J.E., Chandran, K., Saphire, E.O., 2016. Host-primed ebola virus GP exposes a hydrophobic NPC1 receptor-binding pocket, revealing a target for broadly neutralizing antibodies. *mBio* 7, e02154–2115. <http://dx.doi.org/10.1128/mBio.02154-15>.
- Brecher, M., Schornberg, K.L., Delos, S.E., Fusco, M.L., Saphire, E.O., White, J.M., 2012. Cathepsin cleavage potentiates the Ebola virus glycoprotein to undergo a subsequent fusion-relevant conformational change. *J. Virol.* 86, 364–372. <http://dx.doi.org/10.1128/JVI.05708-11>.
- Brindley, M.A., Hunt, C.L., Kondratowicz, A.S., Bowman, J., Sinn, P.L., McCray, P.B., Quinn, K., Weller, M.L., Chiorini, J.A., Maury, W., 2011. Tyrosine kinase receptor Axl enhances entry of Zaire ebolavirus without direct interactions with the viral glycoprotein. *Virology* 415, 83–94. <http://dx.doi.org/10.1016/j.virol.2011.04.002>.
- Carette, J.E., Raaben, M., Wong, A.C., Herbert, A.S., Obernosterer, G., Mulherkar, N., Kuehne, A.I., Kranzusch, P.J., Griffin, A.M., Ruthel, G., Dal Cin, P., Dye, J.M., Whelan, S.P., Chandran, K., Brummelkamp, T.R., 2011. Ebola virus entry requires the cholesterol transporter Niemann-Pick C1. *Nature* 477, 340–343. <http://dx.doi.org/10.1038/nature10348>.
- Cenedella, R.J., 2009. Cholesterol synthesis inhibitor U18666A and the role of sterol metabolism and trafficking in numerous pathophysiological processes. *Lipids* 44, 477–487. <http://dx.doi.org/10.1007/s11745-009-3305-7>.
- Chandran, K., Sullivan, N.J., Felbor, U., Whelan, S.P., Cunningham, J.M., 2005. Endosomal proteolysis of the Ebola virus glycoprotein is necessary for infection. *Science* 308, 1643–1645. <http://dx.doi.org/10.1126/science.1110656>.
- Chan, S.Y., Speck, R.F., Ma, M.C., Goldsmith, M.A., 2000. Distinct mechanisms of entry by envelope glycoproteins of Marburg and Ebola (Zaire) viruses. *J. Virol.* 74, 4933–4937.
- Cheng, H., Lear-Rooney, C.M., Johansen, L., Varhegyi, E., Chen, Z.W., Olinger, G.G., Rong, L., 2015. Inhibition of ebola and Marburg virus Entry by G Protein-coupled receptor antagonists. *J. Virol.* 89, 9932–9938. <http://dx.doi.org/10.1128/JVI.01337-15>.
- Choi, J.H., Croyle, M.A., 2013. Emerging targets and novel approaches to Ebola virus prophylaxis and treatment. *BioDrugs Clin. Immunother. Biopharm. Gene Ther.* 27, 565–583. <http://dx.doi.org/10.1007/s40259-013-0046-1>.
- Côté, M., Misasi, J., Ren, T., Bruchez, A., Lee, K., Filone, C.M., Hensley, L., Li, Q., Ory, D., Chandran, K., Cunningham, J., 2011. Small molecule inhibitors reveal Niemann-Pick C1 is essential for Ebola virus infection. *Nature* 477, 344–348. <http://dx.doi.org/10.1038/nature10380>.
- de La Vega, M.-A., Wong, G., Kobinger, G.P., Qiu, X., 2015. The multiple roles of sGP in Ebola pathogenesis. *Viral Immunol.* 28, 3–9. <http://dx.doi.org/10.1089/vim.2014.0068>.
- Dias, J.M., Kuehne, A.I., Abelson, D.M., Bale, S., Wong, A.C., Halfmann, P., Muhammad, M.A., Fusco, M.L., Zak, S.E., Kang, E., Kawaoka, Y., Chandran, K., Dye, J.M., Saphire, E.O., 2011. A shared structural solution for neutralizing ebolaviruses. *Nat. Struct. Mol. Biol.* 18 (12), 1424–1427. <http://dx.doi.org/10.1038/nsmb.2150>.
- Dolnik, O., Volchkova, V.A., Escudero-Perez, B., Lawrence, P., Klenk, H.-D., Volchkov, V.E., 2015. Shedding of ebola virus surface glycoprotein is a mechanism of self-regulation of cellular cytotoxicity and has a direct effect on virus infectivity. *J. Infect. Dis.* 212 (Suppl. 2), S322–S328. <http://dx.doi.org/10.1093/infdis/jiv268>.
- Dolnik, O., Volchkova, V., Garten, W., Carbonnelle, C., Becker, S., Kahnt, J., Ströher, U., Klenk, H.-D., Volchkov, V., 2004. Ectodomain shedding of the glycoprotein GP of Ebola virus. *EMBO J.* 23, 2175–2184. <http://dx.doi.org/10.1038/sj.emboj.7600219>.
- Dominguez-Soto, A., Aragoneses-Fenoll, L., Martin-Gayo, E., Martinez-Prats, L., Colmenares, M., Naranjo-Gomez, M., Borrás, F.E., Munoz, P., Zubiar, M., Toribio, M.L., Delgado, R., Corbi, A.L., 2007. The DC-SIGN-related lectin LSECtin mediates antigen capture and pathogen binding by human myeloid cells. *Blood* 109, 5337–5345. <http://dx.doi.org/10.1182/blood-2006-09-048058>.
- Dowall, S.D., Bosworth, A., Watson, R., Bewley, K., Taylor, I., Rayner, E., Hunter, L., Pearson, G., Easterbrook, L., Pitman, J., Hewson, R., Carroll, M.W., 2015. Chloroquine inhibited Ebola virus replication in vitro but failed to protect against infection and disease in the in vivo Guinea pig model. *J. Gen. Virol.* 96, 3484–3492. <http://dx.doi.org/10.1099/jgv.0.000309>.
- Elshabrawy, H.A., Fan, J., Haddad, C.S., Ratiá, K., Broder, C.C., Caffrey, M., Prabhakar, B.S., 2014. Identification of a broad-spectrum antiviral small molecule against severe acute respiratory syndrome coronavirus and Ebola, Hendra, and Nipah viruses by using a novel high-throughput screening assay. *J. Virol.* 88, 4353–4365. <http://dx.doi.org/10.1128/JVI.03050-13>.
- Empig, C.J., Goldsmith, M.A., 2002. Association of the caveola vesicular system with cellular entry by filoviruses. *J. Virol.* 76, 5266–5270.
- Escudero-Pérez, B., Volchkova, V.A., Dolnik, O., Lawrence, P., Volchkov, V.E., 2014. Shed GP of Ebola virus triggers immune activation and increased vascular permeability. *PLoS Pathog.* 10, e1004509. <http://dx.doi.org/10.1371/journal.ppat.1004509>.
- Falzarano, D., Krokhn, O., Wahl-Jensen, V., Seebach, J., Wolf, K., Schnittler, H.-J., Feldmann, H., 2006. Structure-function analysis of the soluble glycoprotein, sGP, of Ebola virus. *Chembiochem Eur. J. Chem. Biol.* 7, 1605–1611. <http://dx.doi.org/10.1002/cbic.200600223>.
- Falzarano, D., Saffronetz, D., Prescott, J., Marzi, A., Feldmann, F., Feldmann, H., 2015. Lack of protection against ebola virus from chloroquine in mice and hamsters. *Emerg. Infect. Dis.* 21, 1065–1067. <http://dx.doi.org/10.3201/eid2106.150176>.
- Feldmann, H., Geisbert, T.W., 2011. Ebola haemorrhagic fever. *Lancet Lond. Engl.* 377, 849–862. [http://dx.doi.org/10.1016/S0140-6736\(10\)60667-8](http://dx.doi.org/10.1016/S0140-6736(10)60667-8).
- Feldmann, H., Volchkov, V.E., Volchkova, V.A., Klenk, H.D., 1999. The glycoproteins of Marburg and Ebola virus and their potential roles in pathogenesis. *Arch. Virol.* (Suppl. 15), 159–169.
- Francica, J.R., Matukonis, M.K., Bates, P., 2009. Requirements for cell rounding and surface protein down-regulation by Ebola virus glycoprotein. *Virology* 383, 237–247. <http://dx.doi.org/10.1016/j.virol.2008.10.029>.
- Francica, J.R., Varela-Rohena, A., Medvec, A., Plesa, G., Riley, J.L., Bates, P., 2010. Steric shielding of surface epitopes and impaired immune recognition induced by the ebola virus glycoprotein. *PLoS Pathog.* 6, e1001098. <http://dx.doi.org/10.1371/journal.ppat.1001098>.
- Gallagher, W.R., Garry, R.F., 2015. Modeling of the Ebola virus delta peptide reveals a potential lytic sequence motif. *Viruses* 7, 285–305. <http://dx.doi.org/10.3390/v7100285>.
- Gehring, G., Rohrmann, K., Atenchong, N., Mittler, E., Becker, S., Dahlmann, F., Pöhlmann, S., Vondran, F.W.R., David, S., Manns, M.P., Ciesek, S., von Hahn, T., 2014. The clinically approved drugs amiodarone, dronedarone and verapamil inhibit filovirus cell entry. *J. Antimicrob. Chemother.* 69, 2123–2131. <http://dx.doi.org/10.1093/jac/dku091>.

- Gnirss, K., Kühl, A., Karsten, C., Glowacka, I., Bertram, S., Kaup, F., Hofmann, H., Pöhlmann, S., 2012. Cathepsins B and L activate Ebola but not Marburg virus glycoproteins for efficient entry into cell lines and macrophages independent of TMPRSS2 expression. *Virology* 424, 3–10. <http://dx.doi.org/10.1016/j.virol.2011.11.031>.
- Gong, X., Qian, H., Zhou, X., Wu, J., Wan, T., Cao, P., Huang, W., Zhao, X., Wang, X., Wang, P., Shi, Y., Gao, G.F., Zhou, Q., Yan, N., 2016. Structural insights into the Niemann-Pick C1 (NPC1)-Mediated cholesterol transfer and ebola infection. *Cell* 165, 1467–1478. <http://dx.doi.org/10.1016/j.cell.2016.05.022>.
- Gramberg, T., Hofmann, H., Möller, P., Lalor, P.F., Marzi, A., Geier, M., Krumbiegel, M., Winkler, T., Kirchhoff, F., Adams, D.H., Becker, S., Münch, J., Pöhlmann, S., 2005. LSECtin interacts with filovirus glycoproteins and the spike protein of SARS coronavirus. *Virology* 340, 224–236. <http://dx.doi.org/10.1016/j.virol.2005.06.026>.
- Gramberg, T., Soilleux, E., Fisch, T., Lalor, P.F., Hofmann, H., Wheelodon, S., Cotterill, A., Wegele, A., Winkler, T., Adams, D.H., Pöhlmann, S., 2008. Interactions of LSECtin and DC-SIGN/DC-SIGNR with viral ligands: differential pH dependence, internalization and virion binding. *Virology* 373, 189–201. <http://dx.doi.org/10.1016/j.virol.2007.11.001>.
- Harcourt, B.H., Sanchez, A., Offermann, M.K., 1999. Ebola virus selectively inhibits responses to interferons, but not to interleukin-1beta, in endothelial cells. *J. Virol.* 73, 3491–3496.
- Hartlieb, B., Weissenhorn, W., 2006. Filovirus assembly and budding. *Virology* 344, 64–70. <http://dx.doi.org/10.1016/j.virol.2005.09.018>.
- Hashiguchi, T., Fusco, M.L., Bornholdt, Z.A., Lee, J.E., Flyak, A.I., Matsuoka, R., Kohda, D., Yanagi, Y., Hammel, M., Crowe, J.E., Saphire, E.O., 2015. Structural basis for Marburg virus neutralization by a cross-reactive human antibody. *Cell* 160, 904–912. <http://dx.doi.org/10.1016/j.cell.2015.01.041>.
- Higgins, C.D., Koellhoffer, J.F., Chandran, K., Lai, J.R., 2013. C-peptide inhibitors of Ebola virus glycoprotein-mediated cell entry: effects of conjugation to cholesterol and side chain-side chain crosslinking. *Bioorg. Med. Chem. Lett.* 23, 5356–5360. <http://dx.doi.org/10.1016/j.bmcl.2013.07.056>.
- Hoenen, T., Marzi, A., Scott, D.P., Feldmann, F., Callison, J., Safronetz, D., Ebihara, H., Feldmann, H., 2015. Soluble glycoprotein is not required for ebola virus virulence in Guinea pigs. *J. Infect. Dis.* 212 (Suppl. 2), S242–S246. <http://dx.doi.org/10.1093/infdis/jiv111>.
- Hoffmann, M., González Hernández, M., Berger, E., Marzi, A., Pöhlmann, S., 2016. The glycoproteins of all filovirus species use the same host factors for entry into bat and human cells but entry efficiency is species dependent. *PLoS One* 11, e0149651. <http://dx.doi.org/10.1371/journal.pone.0149651>.
- Hunt, C.L., Kolokoltsov, A.A., Davey, R.A., Maury, W., 2011. The Tyro3 receptor kinase Axl enhances macropinocytosis of Zaire ebolavirus. *J. Virol.* 85, 334–347. <http://dx.doi.org/10.1128/JVI.01278-09>.
- Ito, H., Watanabe, S., Takada, A., Kawaoka, Y., 2001. Ebola virus glycoprotein: proteolytic processing, acylation, cell tropism, and detection of neutralizing antibodies. *J. Virol.* 75, 1576–1580. <http://dx.doi.org/10.1128/JVI.75.3.1576-1580.2001>.
- Iwasa, A., Shimajima, M., Kawaoka, Y., 2011. sGP serves as a structural protein in Ebola virus infection. *J. Infect. Dis.* 204 (Suppl. 3), S897–S903. <http://dx.doi.org/10.1093/infdis/jir313>.
- Johansen, L., Brannan, J., Delos, S., Shoemaker, C., Stossel, A., Lear, C., Hoffstrom, B., Dewald, L., Schornberg, K., Scully, C., Lehár, J., Hensley, J., White, J., Olinger, G., 2013. FDA-Approved Selective Estrogen Receptor Modulators Inhibit Ebola Virus Infection [WWW Document]. URL: <http://www.ncbi.nlm.nih.gov/pubmed/?term=23785035> (accessed 03.02.16).
- Kampmann, T., Mueller, D.S., Mark, A.E., Young, P.R., Kobe, B., 2006. The Role of histidine residues in low-pH-mediated viral membrane fusion, 1993 *Struct. Lond. Engl.* 14, 1481–1487. <http://dx.doi.org/10.1016/j.str.2006.07.011>.
- Keyaerts, E., Li, S., Vijgen, L., Rysman, E., Verbeeck, J., Van Ranst, M., Maes, P., 2009. Antiviral activity of chloroquine against human coronavirus OC43 infection in newborn mice. *Antimicrob. Agents Chemother.* 53, 3416–3421. <http://dx.doi.org/10.1128/AAC.01509-08>.
- Keyaerts, E., Vijgen, L., Maes, P., Neyts, J., Van Ranst, M., 2004. In vitro inhibition of severe acute respiratory syndrome coronavirus by chloroquine. *Biochem. Biophys. Res. Commun.* 323, 264–268. <http://dx.doi.org/10.1016/j.bbrc.2004.08.085>.
- Kondratowicz, A.S., Lennemann, N.J., Sinn, P.L., Davey, R.A., Hunt, C.L., Moller-Tank, S., Meyerholz, D.K., Rennett, P., Mullins, R.F., Brindley, M., Sandersfeld, L.M., Quinn, K., Weller, M., McCray, P.B., Chiorini, J., Maury, W., 2011. T-cell immunoglobulin and mucin domain 1 (TIM-1) is a receptor for Zaire Ebolavirus and Lake Victoria Marburgvirus. *Proc. Natl. Acad. Sci. U. S. A.* 108, 8426–8431. <http://dx.doi.org/10.1073/pnas.1019030108>.
- Kouznetsova, J., Sun, W., Martínez-Romero, C., Tawa, G., Shinn, P., Chen, C.Z., Schimmer, A., Sanderson, P., McKew, J.C., Zheng, W., García-Sastre, A., 2014. Identification of 53 compounds that block Ebola virus-like particle entry via a repurposing screen of approved drugs. *Emerg. Microbes Infect.* 3, e84. <http://dx.doi.org/10.1038/emi.2014.88>.
- Kugelman, J.R., Lee, M.S., Rossi, C.A., McCarthy, S.E., Radoshitzky, S.R., Dye, J.M., Hensley, L.E., Honko, A., Kuhn, J.H., Jahrling, P.B., Warren, T.K., Whitehouse, C.A., Bavari, S., Palacios, G., 2012. Ebola virus genome plasticity as a marker of its assaying history: a comparison of in vitro passaging to non-human primate infection. *PLoS one* 7, e50316. <http://dx.doi.org/10.1371/journal.pone.0050316>.
- Kuhn, J.H., Becker, S., Ebihara, H., Geisbert, T.W., Johnson, K.M., Kawaoka, Y., Lipkin, W.I., Negredo, A.I., Netesov, S.V., Nichol, S.T., Palacios, G., Peters, C.J., Tenorio, A., Volchkov, V.E., Jahrling, P.B., 2010. Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virus abbreviations. *Arch. Virol.* 155, 2083–2103. <http://dx.doi.org/10.1007/s00705-010-0814-x>.
- Kuroda, M., Fujikura, D., Nanbo, A., Marzi, A., Noyori, O., Kajihara, M., Maruyama, J., Matsuno, K., Miyamoto, H., Yoshida, R., Feldmann, H., Takada, A., 2015. Interaction between TIM-1 and NPC1 is important for cellular entry of ebola virus. *J. Virol.* 89, 6481–6493. <http://dx.doi.org/10.1128/JVI.03156-14>.
- Lee, J.E., Fusco, M.L., Hessel, A.J., Oswald, W.B., Burton, D.R., Saphire, E.O., 2008. Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature* 454, 177–182. <http://dx.doi.org/10.1038/nature07082>.
- Lee, J.E., Saphire, E.O., 2009. Ebolavirus glycoprotein structure and mechanism of entry. *Future Virol.* 4, 621–635. <http://dx.doi.org/10.2217/fvl.09.56>.
- Lee, K., Ren, T., Côté, M., Gholamreza, B., Misasi, J., Bruchez, A., Cunningham, J., 2013. Inhibition of ebola virus infection: identification of Niemann-Pick C1 as the target by optimization of a chemical probe. *ACS Med. Chem. Lett.* 4, 239–243. <http://dx.doi.org/10.1021/ml300370k>.
- Lee, J., Gregory, S.M., Nelson, E.A., White, J.M., Tamm, L.K., 2016. The roles of histidines and charged residues as potential triggers of a conformational change in the fusion loop of Ebola virus glycoprotein. *PLoS One* 11. <http://dx.doi.org/10.1371/journal.pone.0152527>.
- Liang, B., Li, Z., Jenni, S., Rahme, A.A., Morin, B.M., Grant, T., Grigorieff, N., Harrison, S.C., Whelan, S.P.J., 2015. Structure of the L Protein of vesicular stomatitis virus from electron cryomicroscopy. *Cell* 162, 314–327. <http://dx.doi.org/10.1016/j.cell.2015.06.018>.
- Lin, G., Simmons, G., Pöhlmann, S., Baribaud, F., Ni, H., Leslie, G.J., Haggarty, B.S., Bates, P., Weissman, D., Hoxie, J.A., Doms, R.W., 2003. Differential N-linked glycosylation of human immunodeficiency virus and Ebola virus envelope glycoproteins modulates interactions with DC-SIGN and DC-SIGNR. *J. Virol.* 77, 1337–1346.
- Liu, N., Tao, Y., Brenowitz, M.D., Girvin, M.E., Lai, J.R., 2015. Structural and functional studies on the Marburg virus GP2 fusion loop. *J. Infect. Dis.* 212 (Suppl. 2), S146–S153. <http://dx.doi.org/10.1093/infdis/jiv030>.
- Liu, S.-Y., Aliyari, R., Chikere, K., Li, G., Marsden, M.D., Smith, J.K., Pernet, O., Guo, H., Nusbaum, R., Zack, J.A., Freiberg, A.N., Su, L., Lee, B., Cheng, G., 2013. Interferon-inducible cholesterol-25-hydroxylase broadly inhibits viral entry by production of 25-hydroxycholesterol. *Immunity* 38, 92–105. <http://dx.doi.org/10.1016/j.immuni.2012.11.005>.
- Long, J., Wright, E., Molesti, E., Temperton, N., Barclay, W., 2015. Antiviral therapies against Ebola and other emerging viral diseases using existing medicines that block virus entry. *F1000Research* 4, 30. <http://dx.doi.org/10.12688/f1000research.6085.2>.
- Madrid, P.B., Chopra, S., Manger, I.D., Gilfillan, L., Keepers, T.R., Shurtleff, A.C., Green, C.E., Iyer, L.V., Dilks, H.H., Davey, R.A., Kolokoltsov, A.A., Carrion, R., Patterson, J.L., Bavari, S., Panchal, R.G., Warren, T.K., Wells, J.B., Moos, W.H., Burke, R.L., Tanga, M.J., 2013. A systematic screen of FDA-approved drugs for inhibitors of biological threat agents. *PLoS One* 8, e60579. <http://dx.doi.org/10.1371/journal.pone.0060579>.
- Manicassamy, B., Wang, J., Rumschlag, E., Tymen, S., Volchkova, V., Volchkov, V., Rong, L., 2007. Characterization of Marburg virus glycoprotein in viral entry. *Virology* 358, 79–88. <http://dx.doi.org/10.1016/j.virol.2006.06.041>.
- Markosyan, R.M., Miao, C., Zheng, Y.M., Melikyan, G.B., Liu, S.L., Cohen, F.S., 2016. Induction of cell-cell fusion by Ebola virus glycoprotein: low pH is not a trigger. *PLoS Pathog.* 12. <http://dx.doi.org/10.1371/journal.ppat.1005373>.
- Maruyama, J., Miyamoto, H., Kajihara, M., Ogawa, H., Maeda, K., Sakoda, Y., Yoshida, R., Takada, A., 2014. Characterization of the envelope glycoprotein of a novel filovirus, llovio virus. *J. Virol.* 88, 99–109. <http://dx.doi.org/10.1128/JVI.02265-13>.
- Marzi, A., Gramberg, T., Simmons, G., Möller, P., Rennekamp, A.J., Krumbiegel, M., Geier, M., Eisemann, J., Turza, N., Saunier, B., Steinkasserer, A., Becker, S., Bates, P., Hofmann, H., Pöhlmann, S., 2004. DC-SIGN and DC-SIGNR interact with the glycoprotein of Marburg virus and the S protein of severe acute respiratory syndrome coronavirus. *J. Virol.* 78, 12090–12095. <http://dx.doi.org/10.1128/JVI.78.21.12090-12095.2004>.
- Marzi, A., Möller, P., Hanna, S.L., Harrer, T., Eisemann, J., Steinkasserer, A., Becker, S., Baribaud, F., Pöhlmann, S., 2007. Analysis of the interaction of Ebola virus glycoprotein with DC-SIGN (dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin) and its homologue DC-SIGNR. *J. Infect. Dis.* 196 (Suppl. 2), S237–S246. <http://dx.doi.org/10.1086/520607>.
- Marzi, A., Reinheckel, T., Feldmann, H., 2012. Cathepsin B & L are not required for ebola virus replication. *PLoS Negl. Trop. Dis.* 6, e1923. <http://dx.doi.org/10.1371/journal.pntd.0001923>.
- Matsuno, K., Kishida, N., Usami, K., Igarashi, M., Yoshida, R., Nakayama, E., Shimajima, M., Feldmann, H., Irimura, T., Kawaoka, Y., Takada, A., 2010a. Different potential of C-type lectin-mediated entry between Marburg virus strains. *J. Virol.* 84, 5140–5147. <http://dx.doi.org/10.1128/JVI.02021-09>.
- Matsuno, K., Nakayama, E., Noyori, O., Marzi, A., Ebihara, H., Irimura, T., Feldmann, H., Takada, A., 2010b. C-type lectins do not act as functional receptors for filovirus entry into cells. *Biochem. Biophys. Res. Commun.* 403, 144–148. <http://dx.doi.org/10.1016/j.bbrc.2010.10.136>.
- Mehedi, M., Falzarano, D., Seebach, J., Hu, X., Carpenter, M.S., Schnittler, H.-J., Feldmann, H., 2011. A new Ebola virus nonstructural glycoprotein expressed through RNA editing. *J. Virol.* 85, 5406–5414. <http://dx.doi.org/10.1128/JVI.02190-10>.
- Mehedi, M., Hoenen, T., Robertson, S., Ricklefs, S., Dolan, M.A., Taylor, T., Falzarano, D., Ebihara, H., Porcella, S.F., Feldmann, H., 2013. Ebola virus RNA

- editing depends on the primary editing site sequence and an upstream secondary structure. *PLoS Pathog.* 9, e1003677. <http://dx.doi.org/10.1371/journal.ppat.1003677>.
- Mendoza, E.J., Qiu, X., Kobinger, G.P. Progression of Ebola therapeutics during the 2014–2015 outbreak. *Trends Mol. Med.* 22, 164–173. doi: <http://dx.doi.org/10.1016/j.molmed.2015.12.005>.
- Michelow, I.C., Lear, C., Scully, C., Prugar, L.I., Longley, C.B., Yantosca, L.M., Ji, X., Karpel, M., Brudner, M., Takahashi, K., Spear, G.T., Ezekowitz, R.A., Schmidt, E.V., Olinger, G.G., 2011. High-dose mannose-binding lectin therapy for Ebola virus infection, 2011. *J. Infect. Dis.* 203 (2), 175–179. <http://dx.doi.org/10.1093/infdis/jiq025>.
- Miller, E.H., Harrison, J.S., Radoshitzky, S.R., Higgins, C.D., Chi, X., Dong, L., Kuhn, J.H., Bavari, S., Lai, J.R., Chandran, K., 2011. Inhibition of Ebola virus entry by a C-peptide targeted to endosomes. *J. Biol. Chem.* 286, 15854–15861. <http://dx.doi.org/10.1074/jbc.M110.207084>.
- Miller, E.H., Obernosterer, G., Raaben, M., Herbert, A.S., Deffieu, M.S., Krishnan, A., Ndungo, E., Sandesara, R.G., Carette, J.E., Kuehne, A.I., Ruthel, G., Pfeffer, S.R., Dye, J.M., Whelan, S.P., Brummelkamp, T.R., Chandran, K., 2012. Ebola virus entry requires the host-programmed recognition of an intracellular receptor. *EMBO J.* 31, 1947–1960. <http://dx.doi.org/10.1038/emboj.2012.53>.
- Miller, M.E., Adhikary, S., Kolokoltsov, A.A., Davey, R.A., 2012. Ebola virus requires acid sphingomyelinase activity and plasma membrane sphingomyelin for infection. *J. Virol.* 86, 7473–7483. <http://dx.doi.org/10.1128/JVI.00136-12>.
- Mohan, G.S., Li, W., Ye, L., Compans, R.W., Yang, C., 2012. Antigenic subversion: a novel mechanism of host immune evasion by Ebola virus. *PLoS Pathog.* 8, e1003065. <http://dx.doi.org/10.1371/journal.ppat.1003065>.
- Mohan, G.S., Ye, L., Li, W., Monteiro, A., Lin, X., Sapkota, B., Pollack, B.P., Compans, R.W., Yang, C., 2015. Less is more: ebola virus surface glycoprotein expression levels regulate virus production and infectivity. *J. Virol.* 89, 1205–1217. <http://dx.doi.org/10.1128/JVI.01810-14>.
- Moller-Tank, S., Kondratowicz, A.S., Davey, R.A., Rennert, P.D., Maury, W., 2013. Role of the phosphatidylserine receptor TIM-1 in enveloped-virus entry, 2013. *Aug J. Virol.* 87 (15), 8327–8341. <http://dx.doi.org/10.1128/JVI.01025-13>.
- Moller-Tank, S., Maury, W., 2014. Phosphatidylserine receptors: enhancers of enveloped virus entry and infection. *Virology* 468–470, 565–580. <http://dx.doi.org/10.1016/j.virol.2014.09.009>.
- Mühlberger, E., 2007. Filovirus replication and transcription. *Future Virol.* 2, 205–215. <http://dx.doi.org/10.2217/17460794.2.2.205>.
- Muñoz, A., Sigwalt, D., Illescas, B.M., Luczkowiak, J., Rodríguez-Pérez, L., Nierengarten, I., Holler, M., Remy, J.-S., Buffet, K., Vincent, S.P., Rojo, J., Delgado, R., Nierengarten, J.-F., Martín, N., 2016. Synthesis of giant globular multivalent glycofullerenes as potent inhibitors in a model of Ebola virus infection. *Nat. Chem.* 8, 50–57. <http://dx.doi.org/10.1038/nchem.2387>.
- Nakayama, E., Tomabechi, D., Matsuno, K., Kishida, N., Yoshida, R., Feldmann, H., Takada, A., 2011. Antibody-dependent enhancement of Marburg virus infection. *J. Infect. Dis.* 204 (Suppl. 3), S978–S985. <http://dx.doi.org/10.1093/infdis/jir334>.
- Nakayama, E., Yokoyama, A., Miyamoto, H., Igarashi, M., Kishida, N., Matsuno, K., Marzi, A., Feldmann, H., Ito, K., Saijo, M., Takada, A., 2010. Enzyme-linked immunosorbent assay for detection of filovirus species-specific antibodies. *Clin. Vaccine Immunol.* CVI 17, 1723–1728. <http://dx.doi.org/10.1128/CI.00170-10>.
- Nanbo, A., Imai, M., Watanabe, S., Noda, T., Takahashi, K., Neumann, G., Halfmann, P., Kawaoka, Y., 2010. Ebola virus is internalized into host cells via macropinosomes in a viral glycoprotein-dependent manner. *PLoS Pathog.* 6, e1001121. <http://dx.doi.org/10.1371/journal.ppat.1001121>.
- Ndungo, E., Herbert, A.S., Raaben, M., Obernosterer, G., Biswas, R., Miller, E.H., Wirchnianski, A.S., Carette, J.E., Brummelkamp, T.R., Whelan, S.P., Dye, J.M., Chandran, K., 2016. A single residue in Ebola virus receptor NPC1 influences cellular host range in reptiles. *mSphere* 1, 07–16. <http://dx.doi.org/10.1128/mSphere.00007-16>.
- Negredo, A., Palacios, G., Vázquez-Morón, S., González, F., Dopazo, H., Molero, F., Juste, J., Quetglas, J., Savji, N., de la Cruz Martínez, M., Herrera, J.E., Pizarro, M., Hutchison, S.K., Echevarría, J.E., Lipkin, W.I., Tenorio, A., 2011. Discovery of an ebolavirus-like filovirus in europe. *PLoS Pathog.* 7, e1002304. <http://dx.doi.org/10.1371/journal.ppat.1002304>.
- Nelson, E.A., Barnes, A.B., Wiehle, R.D., Fontenot, G.K., Hoenen, T., White, J.M., 2016. Clomiphene and its isomers block ebolavirus particle entry and infection with similar potency: potential therapeutic implications. *Viruses* 8 (8). <http://dx.doi.org/10.3390/v8080206>.
- Ng, M., Ndungo, E., Jangra, R.K., Cai, Y., Postnikova, E., Radoshitzky, S.R., Dye, J.M., Ramírez de Arellano, E., Negredo, A., Palacios, G., Kuhn, J.H., Chandran, K., 2014. Cell entry by a novel European filovirus requires host endosomal cysteine proteases and Niemann-Pick C1. *Virology* 468–470, 637–646. <http://dx.doi.org/10.1016/j.virol.2014.08.019>.
- Ng, M., Ndungo, E., Kaczmarek, M.E., Herbert, A.S., Binger, T., Kuehne, A.I., Jangra, R.K., Hawkins, J.A., Gifford, R.J., Biswas, R., Demogines, A., James, R.M., Yu, M., Brummelkamp, T.R., Drosten, C., Wang, L.-F., Kuhn, J.H., Müller, M.A., Dye, J.M., Sawyer, S.L., Chandran, K., 2015. Filovirus receptor NPC1 contributes to species-specific patterns of ebolavirus susceptibility in bats. *eLife* 4. <http://dx.doi.org/10.7554/eLife.11785>.
- Nina, J., 2014. Ebolaviruses: a 2014 review for clinicians. *Acta Médica Port.* 27, 625–633.
- Nishimura, H., Yamaya, M., 2015. A synthetic serine protease inhibitor, nafamostat mesilate, is a drug potentially applicable to the treatment of ebola virus disease. *Tohoku J. Exp. Med.* 237, 45–50. <http://dx.doi.org/10.1620/tjem.237.45>.
- Noda, T., Ebihara, H., Muramoto, Y., Fujii, K., Takada, A., Sagara, H., Kim, J.H., Kida, H., Feldmann, H., Kawaoka, Y., 2006. Assembly and budding of ebolavirus. *PLoS Pathog.* 2, e99. <http://dx.doi.org/10.1371/journal.ppat.0020099>.
- Nyakatura, E., Frei, J., Lai, J., 2015. Chemical and Structural Aspects of Ebola Virus Entry Inhibitors [WWW Document]. URL: <http://www.ncbi.nlm.nih.gov/pubmed/?term=25984565> (accessed 03.02.16).
- O'Hearn, A., Wang, M., Cheng, H., Lear-Rooney, C.M., Koning, K., Rumschlag-Booms, E., Varhegyi, E., Olinger, G., Rong, L., 2015. Role of EXT1 and glycosaminoglycans in the early stage of filovirus entry. *J. Virol.* 89, 5441–5449. <http://dx.doi.org/10.1128/JVI.03689-14>.
- Olejnik, J., Ryabchikova, E., Corley, R.B., Mühlberger, E., 2011. Intracellular events and cell fate in filovirus infection. *Viruses* 3, 1501–1531. <http://dx.doi.org/10.3390/v3081501>.
- Olinger, G.G., Pettitt, J., Kim, D., Working, C., Bohorov, O., Bratcher, B., Hiatt, E., Hume, S.D., Johnson, A.K., Morton, J., Pauly, M., Whaley, K.J., Lear, C.M., Biggins, J.E., Scully, C., Hensley, L., Zeitlin, L., 2012. Delayed treatment of Ebola virus infection with plant-derived monoclonal antibodies provides protection in rhesus macaques. *Proc. Natl. Acad. Sci. U. S. A.* 109, 18030–18035. <http://dx.doi.org/10.1073/pnas.1213709109>.
- Pécheur, E.-I., Borisevich, V., Halfmann, P., Morrey, J.D., Smeets, D.F., Prichard, M., Mire, C.E., Kawaoka, Y., Geisbert, T.W., Polyak, S.J., 2016. The synthetic antiviral drug arbidol inhibits globally prevalent pathogenic viruses. *J. Virol.* 90, 3086–3092. <http://dx.doi.org/10.1128/JVI.02077-15>.
- Powlesland, A.S., Fisch, T., Taylor, M.E., Smith, D.F., Tissot, B., Dell, A., Pöhlmann, S., Drickamer, K., 2008. A novel mechanism for LSECtin binding to Ebola virus surface glycoprotein through truncated glycans. *J. Biol. Chem.* 283, 593–602. <http://dx.doi.org/10.1074/jbc.M706292200>.
- Qiu, X., Audet, J., Wong, G., Fernando, L., Bello, A., Pillet, S., Alimonti, J.B., Kobinger, G.P., 2013a. Sustained protection against Ebola virus infection following treatment of infected nonhuman primates with ZMab. *Sci. Rep.* 3, 3365. <http://dx.doi.org/10.1038/srep03365>.
- Qiu, X., Audet, J., Wong, G., Pillet, S., Bello, A., Cabral, T., Strong, J.E., Plummer, F., Corbett, C.R., Alimonti, J.B., Kobinger, G.P., 2012. Successful treatment of ebola virus-infected cynomolgus macaques with monoclonal antibodies. *Sci. Transl. Med.* 4, 138ra81. <http://dx.doi.org/10.1126/scitranslmed.3003876>.
- Qiu, X., Wong, G., Audet, J., Bello, A., Fernando, L., Alimonti, J.B., Fausther-Bovendo, H., Wei, H., Aviles, J., Hiatt, E., Johnson, A., Morton, J., Swope, K., Bohorov, O., Bohorova, N., Goodman, C., Kim, D., Pauly, M.H., Velasco, J., Pettitt, J., Olinger, G.G., Whaley, K., Xu, B., Strong, J.E., Zeitlin, L., Kobinger, G.P., 2014. Reversion of advanced Ebola virus disease in nonhuman primates with ZMapp. *Nature* 514, 47–53. <http://dx.doi.org/10.1038/nature13777>.
- Qiu, X., Wong, G., Fernando, L., Audet, J., Bello, A., Strong, J., Alimonti, J.B., Kobinger, G.P., 2013b. mAbs and Ad-vectored IFN- α therapy rescue Ebola-infected nonhuman primates when administered after the detection of viremia and symptoms. *Sci. Transl. Med.* 5, 207ra143. <http://dx.doi.org/10.1126/scitranslmed.3006605>.
- Radoshitzky, S.R., Warfield, K.L., Chi, X., Dong, L., Kota, K., Bradfute, S.B., Gearhart, J.D., Retterer, C., Kranzusch, P.J., Misasi, J.N., Hogenbirk, M.A., Wahl-Jensen, V., Volchkov, V.E., Cunningham, J.M., Jahrling, P.B., Aman, M.J., Bavari, S., Farzan, M., Kuhn, J.H., 2011. Ebolavirus delta-peptide immunoadhesins inhibit marburgvirus and ebolavirus cell entry. *J. Virol.* 85, 8502–8513. <http://dx.doi.org/10.1128/JVI.02600-10>.
- Ramanan, P., Shabman, R.S., Brown, C.S., Amarasinghe, G.K., Basler, C.F., Leung, D.W., 2011. Filoviral immune evasion mechanisms. *Viruses* 3, 1634–1649. <http://dx.doi.org/10.3390/v3091634>.
- Ray, R.B., Basu, A., Steele, R., Beyene, A., McHowat, J., Meyer, K., Ghosh, A.K., Ray, R. Ebola virus glycoprotein-mediated anoikis of primary human cardiac microvascular endothelial cells. *Virology* 321(2):181–188.
- Reynard, O., Borowiak, M., Volchkova, V.A., Delpout, S., Mateo, M., Volchkov, V.E., 2009. Ebolavirus glycoprotein GP masks both its own epitopes and the presence of cellular surface proteins. *J. Virol.* 83, 9596–9601. <http://dx.doi.org/10.1128/JVI.00784-09>.
- Rhein, B.A., Maury, W.J., 2015. Ebola virus entry into host cells: identifying therapeutic strategies. *Curr. Clin. Microbiol. Rep.* 2, 115–124. <http://dx.doi.org/10.1007/s40588-015-0021-3>.
- Rhein, B.A., Brouillette, R.B., Schaack, G.A., Chiorini, J.A., Maury, W., 2016. Characterization of human and murine T-cell Immunoglobulin Mucin domain 4 (TIM-4) IgV domain residues critical for Ebola virus entry. *J. Virol.* 90, 6097–6111. <http://dx.doi.org/10.1128/JVI.00100-16>.
- Rodríguez-Lafresse, C., Rousson, R., Bonnet, J., Pentchev, P.G., Louisot, P., Vanier, M.T., 1990. Abnormal cholesterol metabolism in imipramine-treated fibroblast cultures. Similarities with Niemann-Pick type C disease. *Biochim. Biophys. Acta* 1043, 123–128.
- Saeed, M.F., Kolokoltsov, A.A., Albrecht, T., Davey, R.A., 2010. Cellular entry of ebola virus involves uptake by a macropinosome-like mechanism and subsequent trafficking through early and late endosomes. *PLoS Pathog.* 6, e1001110. <http://dx.doi.org/10.1371/journal.ppat.1001110>.
- Sakurai, Y., Kolokoltsov, A.A., Chen, C.-C., Tidwell, M.W., Bauta, W.E., Klugbauer, N., Grimm, C., Wahl-Schott, C., Biel, M., Davey, R.A., 2015. Ebola virus. Two-pore channels control Ebola virus host cell entry and are drug targets for disease treatment. *Science* 347, 995–998. <http://dx.doi.org/10.1126/science.1258758>.
- Salvador, B., Sexton, N.R., Carrion, R., Nunneley, J., Patterson, J.L., Steffen, I., Lu, K., Muench, M.O., Lembo, D., Simmons, G., 2013. Filoviruses utilize glycosaminoglycans for their attachment to target cells. *J. Virol.* 87, 3295–3304. <http://dx.doi.org/10.1128/JVI.01621-12>.
- Sanchez, A., 2007. Analysis of filovirus entry into vero e6 cells, using inhibitors of

- endocytosis, endosomal acidification, structural integrity, and cathepsin (B and L) activity. *J. Infect. Dis.* 196 (Suppl. 2), S251–S258. <http://dx.doi.org/10.1086/520597>.
- Sanchez, A., Kiley, M.P., Holloway, B.P., 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.
- Sanchez, A., Trappier, S.G., Mahy, B.W., Peters, C.J., Nichol, S.T., 1996. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc. Natl. Acad. Sci. U. S. A.* 93, 3602–3607.
- Savarino, A., Boelaert, J.R., Cassone, A., Majori, G., Cauda, R., 2003. Effects of chloroquine on viral infections: an old drug against today's diseases? *Lancet Infect. Dis.* 3, 722–727.
- Savarino, A., Gennero, L., Sperber, K., Boelaert, J.R., 2001. The anti-HIV-1 activity of chloroquine. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* 20, 131–135.
- Schoggins, J.W., Randall, G., 2013. Lipids in innate antiviral defense. *Cell Host Microbe* 14, 379–385. <http://dx.doi.org/10.1016/j.chom.2013.09.010>.
- Schornberg, K.L., Shoemaker, C.J., Dube, D., Abshire, M.Y., Delos, S.E., Bouton, A.H., White, J.M., 2009. Alpha5beta1-integrin controls ebolavirus entry by regulating endosomal cathepsins. *Proc. Natl. Acad. Sci. U. S. A.* 106, 8003–8008. <http://dx.doi.org/10.1073/pnas.0807578106>.
- Schornberg, K., Matsuyama, S., Kabsch, K., Delos, S., Bouton, A., White, J., 2006. Role of endosomal cathepsins in entry mediated by the Ebola virus glycoprotein. *J. Virol.* 80, 4174–4178. <http://dx.doi.org/10.1128/JVI.80.8.4174-4178.2006>.
- Shah, P.P., Wang, T., Kaletsky, R.L., Myers, M.C., Purvis, J.E., Jing, H., Hury, D.M., Greenbaum, D.C., Smith, A.B., Bates, P., Diamond, S.L., 2010. A small-molecule oxocarbazate inhibitor of human cathepsin L blocks severe acute respiratory syndrome and ebola pseudotyped virus infection into human embryonic kidney 293T cells. *Mol. Pharmacol.* 78, 319–324. <http://dx.doi.org/10.1124/mol.110.064261>.
- Shimajima, M., Takada, A., Ebihara, H., Neumann, G., Fujioka, K., Irimura, T., Jones, S., Feldmann, H., Kawaoka, Y., 2006. Tyro3 family-mediated cell entry of Ebola and Marburg viruses. *J. Virol.* 80, 10109–10116. <http://dx.doi.org/10.1128/JVI.01157-06>.
- Shoemaker, C.J., Schornberg, K.L., Delos, S.E., Scully, C., Pajouhesh, H., Olinger, G.G., Johansen, L.M., White, J.M., 2013. Multiple cationic amphiphiles induce a Niemann-Pick C phenotype and inhibit Ebola virus entry and infection. *PLoS One* 8, e56265. <http://dx.doi.org/10.1371/journal.pone.0056265>.
- Siegert, R., Shu, H.L., Slenczka, W., Peters, D., Müller, G., 1967. [On the etiology of an unknown human infection originating from monkeys]. *Dtsch. Med. Wochenschr* 1946 (92), 2341–2343. <http://dx.doi.org/10.1055/s-0028-1106144>.
- Simmons, G., 2013. Filovirus entry. *Adv. Exp. Med. Biol.* 790, 83–94. <http://dx.doi.org/10.1007/978-1-4614-7651-1-5>.
- Simmons, G., Reeves, J.D., Grogan, C.C., Vandenberghe, L.H., Baribaud, F., Whitbeck, J.C., Burke, E., Buchmeier, M.J., Soilleux, E.J., Riley, J.L., Doms, R.W., Bates, P., Pöhlmann, S., 2003. DC-SIGN and DC-SIGNR bind ebola glycoproteins and enhance infection of macrophages and endothelial cells. *Virology* 305, 115–123.
- Simmons, G., Wool-Lewis, R.J., Baribaud, F., Netter, R.C., Bates, P., 2002. Ebola virus glycoproteins induce global surface protein down-modulation and loss of cell adherence. *J. Virol.* 76, 2518–2528.
- St Vincent, M.R., Colpitts, C.C., Ustinov, A.V., Muqadas, M., Joyce, M.A., Barsby, N.L., Epand, R.F., Epand, R.M., Khramyshev, S.A., Valueva, O.A., Korshun, V.A., Tyrrell, D.L.J., Schang, L.M., 2010. Rigid amphipathic fusion inhibitors, small molecule antiviral compounds against enveloped viruses. *Proc. Natl. Acad. Sci. U. S. A.* 107, 17339–17344. <http://dx.doi.org/10.1073/pnas.1010026107>.
- Takada, A., 2012. Filovirus tropism: cellular molecules for viral entry. *Front. Microbiol.* 3, 34. <http://dx.doi.org/10.3389/fmicb.2012.00034>.
- Takada, A., Ebihara, H., Feldmann, H., Geisbert, T.W., Kawaoka, Y., 2007. Epitopes required for antibody-dependent enhancement of Ebola virus infection. *J. Infect. Dis.* 196 (Suppl. 2), S347–S356. <http://dx.doi.org/10.1086/520581>.
- Takada, A., Feldmann, H., Ksiazek, T.G., Kawaoka, Y., 2003. Antibody-dependent enhancement of Ebola virus infection. *J. Virol.* 77, 7539–7544.
- Takada, A., Fujioka, K., Tsuiji, M., Morikawa, A., Higashi, N., Ebihara, H., Kobasa, D., Feldmann, H., Irimura, T., Kawaoka, Y., 2004. Human macrophage C-type lectin specific for galactose and N-acetylgalactosamine promotes filovirus entry. *J. Virol.* 78, 2943–2947.
- Takada, A., Kawaoka, Y., 2001. The pathogenesis of Ebola hemorrhagic fever. *Trends Microbiol.* 9, 506–511.
- Takada, A., Watanabe, S., Ito, H., Okazaki, K., Kida, H., Kawaoka, Y., 2000. Down-regulation of beta1 integrins by Ebola virus glycoprotein: implication for virus entry. *Virology* 278, 20–26. <http://dx.doi.org/10.1006/viro.2000.0601>.
- Tsuda, Y., Hoenen, T., Banadyga, L., Weisend, C., Ricklefs, S.M., Porcella, S.F., Ebihara, H., 2015. An improved reverse genetics system to overcome cell-type-dependent ebola virus genome plasticity. *J. Infect. Dis.* 212 (Suppl. 2), S129–S137. <http://dx.doi.org/10.1093/infdis/jiu681>.
- van der Linden, W.A., Schulze, C.J., Herbert, A.S., Krause, T.B., Wirchnianski, A.A., Dye, J.M., Chandran, K., Bogoy, M., 2016. Cysteine cathepsin inhibitors as anti-Ebola agents. *ACS Infect. Dis.* 2, 173–179. <http://dx.doi.org/10.1021/acsinfecdis.5b00130>.
- Volchkova, V.A., Dolnik, O., Martinez, M.J., Reynard, O., Volchkov, V.E., 2011. Genomic RNA editing and its impact on Ebola virus adaptation during serial passages in cell culture and infection of Guinea pigs. *J. Infect. Dis.* 204 (Suppl. 3), S941–S946. <http://dx.doi.org/10.1093/infdis/jir321>.
- Volchkova, V.A., Klenk, H.D., Volchkov, V.E., 1999. Delta-peptide is the carboxy-terminal cleavage fragment of the nonstructural small glycoprotein sGP of Ebola virus. *Virology* 265, 164–171. <http://dx.doi.org/10.1006/viro.1999.0034>.
- Volchkova, V.A., Dolnik, O., Martinez, M.J., Reynard, O., Volchkov, V.E., 2015. RNA editing of the GP gene of ebola virus is an important pathogenicity factor. *J. Infect. Dis.* 212 (Suppl. 2), S226–S233. <http://dx.doi.org/10.1093/infdis/jiv309>.
- Volchkov, V.E., Becker, S., Volchkova, V.A., Ternovoj, V.A., Kotov, A.N., Netesov, S.V., Klenk, H.D., 1995. GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. *Virology* 214, 421–430.
- Volchkov, V.E., Volchkova, V.A., Ströher, U., Becker, S., Dolnik, O., Cieplik, M., Garten, W., Klenk, H.D., Feldmann, H. Proteolytic processing of Marburg virus glycoprotein. *Virology* 268(1):1–6. doi:<http://dx.doi.org/10.1006/viro.1999.0110>.
- Volchkov, V.E., Volchkova, V.A., Muhlberger, E., Kolesnikova, L.V., Weik, M., Dolnik, O., Klenk, H.D., 2001. Recovery of infectious Ebola virus from complementary DNA: RNA editing of the GP gene and viral cytotoxicity. *Science* 291, 1965–1969. <http://dx.doi.org/10.1126/science.1057269>.
- Wahl-Jensen, V.M., Afanasieva, T.A., Seebach, J., Ströher, U., Feldmann, H., Schnittler, H.-J., 2005. Effects of Ebola virus glycoproteins on endothelial cell activation and barrier function. *J. Virol.* 79, 10442–10450. <http://dx.doi.org/10.1128/JVI.79.16.10442-10450.2005>.
- Wang, H., Shi, Y., Song, J., Qi, J., Lu, G., Yan, J., Gao, G.F., 2016. Ebola viral glycoprotein binds to its endosomal receptor Niemann-Pick C1. *Cell* 164, 258–268. <http://dx.doi.org/10.1016/j.cell.2015.12.044>.
- Wang, Y., Cui, R., Li, G., Gao, Q., Yuan, S., Altmeyer, R., Zou, G., 2016. Teicoplanin inhibits Ebola pseudovirus infection in cell culture. *Antivir. Res.* 125, 1–7. <http://dx.doi.org/10.1016/j.antiviral.2015.11.003>.
- White, J.M., Schornberg, K.L., 2012. A new player in the puzzle of filovirus entry. *Nat. Rev. Microbiol.* 10, 317–322. <http://dx.doi.org/10.1038/nrmicro2764>.
- Wolf, M.C., Freiberg, A.N., Zhang, T., Akyol-Ataman, Z., Grock, A., Hong, P.W., Li, J., Watson, N.F., Fang, A.Q., Aguilar, H.C., Porotto, M., Honko, A.N., Damoiseaux, R., Miller, J.P., Woodson, S.E., Chantansirivisal, S., Fontanes, V., Negrete, O.A., Krogstad, P., Dasgupta, A., Moscona, A., Hensley, L.E., Whelan, S.P., Faull, K.F., Holbrook, M.R., Jung, M.E., Lee, B., 2010. A broad-spectrum antiviral targeting entry of enveloped viruses. *Proc. Natl. Acad. Sci. U. S. A.* 107, 3157–3162. <http://dx.doi.org/10.1073/pnas.0909587107>.
- Xia, S., Eom, S.H., Konigsberg, W.H., Wang, J., 2012. Bidentate and tridentate metal-ion coordination states within ternary complexes of RB69 DNA polymerase. *Protein Sci. Publ. Protein Soc.* 21, 447–451. <http://dx.doi.org/10.1002/pro.2026>.
- Yuan, S., Cao, L., Ling, H., Dang, M., Sun, Y., Zhang, X., Chen, Y., Zhang, L., Su, D., Wang, X., Rao, Z., 2015. TIM-1 acts a dual-attachment receptor for Ebolavirus by interacting directly with viral GP and the PS on the viral envelope. *Protein Cell* 6, 814–824. <http://dx.doi.org/10.1007/s12328-015-0220-y>.
- Zeitlin, L., Whaley, K.J., Olinger, G.G., Jacobs, M., Gopal, R., Qiu, X., Kobinger, G.P., 2016. Antibody therapeutics for Ebola virus disease. *Curr. Opin. Virol.* 17, 45–49. <http://dx.doi.org/10.1016/j.coviro.2016.01.006>.
- Zhao, Y., Ren, J., Harlos, K., Stuart, D.I., 2016. Structure of glycosylated NPC1 luminal domain C reveals insights into NPC2 and Ebola virus interactions. *FEBS Lett.* 590, 605–612. <http://dx.doi.org/10.1002/1873-3468.12089>.
- Zhou, Y., Vedantham, P., Lu, K., Agudelo, J., Carrión, R., Nunneley, J.W., Barnard, D., Pöhlmann, S., McKerrow, J.H., Renslo, A.R., Simmons, G., 2015. Protease inhibitors targeting coronavirus and filovirus entry. *Antivir. Res.* 116, 76–84. <http://dx.doi.org/10.1016/j.antiviral.2015.01.011>.

CHAPITRE 3

Protéines des filovirus et drug discovery : Bases structurales et fonctionnelles du cycle réplicatif

Résumé de la Revue 2

Titre : Filovirus proteins for antiviral drug discovery: Structure/function bases of the replication cycle

Auteurs : Martin B, Canard B, Decroly E

Référence : Antiviral Res. 2017 May; 141:48-61. doi: 10.1016/j.antiviral.2017.02.004. Epub 2017 Feb 10.

Cette seconde revue résume les connaissances structurales des protéines constituant la nucléocapside, les différents mécanismes moléculaires et enzymes virales impliqués dans le cycle de transcription/réplication du génome viral. L'ensemble de ces mécanismes constitue une cible d'intérêt pour le développement de futurs antiviraux. Après un bref rappel de la classification et de l'organisation générale du cycle des filovirus (**Figure 1**), l'organisation structurale et fonctionnelle de la nucléocapside a été décrite de façon générale (**Figure 2**). Elle est constituée de l'ARN génomique viral, de la nucléoprotéine NP, de la polymérase L et des protéines virales VP30, VP35 et VP24. Chacune de ces protéines a fait l'objet d'une analyse structure/fonction, notamment pour la protéine L (**Figure 3**), pour dégager les principales caractéristiques de ces protéines. La protéine VP24 n'a pas été étudiée dans cette revue du fait qu'elle est impliquée dans des événements plus tardifs et sans rôle direct dans le cycle répliatif. Ensuite, les mécanismes de transcription et de réplication ont été décrits en mettant l'accent sur la modulation de l'activité des différentes protéines virales, en particulier VP30, pour réguler ce cycle (**Figure 4**). Pour finir, les différentes stratégies antivirales ont été récapitulées (**Tableau 1**) en partant des données *in silico* générées par des études de «docking» jusqu'à la caractérisation pré-clinique des composés les plus prometteurs, tels que certains analogues nucléosidiques. D'autres stratégies antivirales plus originales ont enfin été décrites mais elles restent encore préliminaires.

Revue 2



Review

Filovirus proteins for antiviral drug discovery: Structure/function bases of the replication cycle



Baptiste Martin, Bruno Canard, Etienne Decroly*

Laboratoire Architecture et Fonction des Macromolécules Biologiques (AFMB), UMR7257 CNRS, Parc Scientifique de Luminy, Aix-Marseille Université, Marseille, France

ARTICLE INFO

Article history:

Received 19 November 2016

Received in revised form

12 January 2017

Accepted 7 February 2017

Available online 10 February 2017

Keywords:

Ebola virus

Marburg virus

Filovirus

Antiviral therapy

Mononegavirales

ABSTRACT

Filoviruses are important pathogens that cause severe and often fatal hemorrhagic fever in humans for which no approved vaccines and antiviral treatments are yet available. In an earlier article (Martin et al., *Antiviral Research*, 2016), we reviewed the role of the filovirus surface glycoprotein in replication and as a target for drugs and vaccines. In this review, we focus on recent findings on the filovirus replication machinery and how they could be used for the identification of new therapeutic targets and the development of new antiviral compounds. First, we summarize the recent structural and functional advances on the molecules involved in filovirus replication/transcription cycle, particularly the NP, VP30, VP35 proteins, and the “large” protein L, which harbors the RNA-dependent RNA polymerase (RdRp) and mRNA capping activities. These proteins are essential for viral mRNA synthesis and genome replication, and consequently they constitute attractive targets for drug design. We then describe how these insights into filovirus replication mechanisms and the structure/function characterization of the involved proteins have led to the development of new and innovative antiviral strategies that may help reduce the filovirus disease case fatality rate through post-exposure or prophylactic treatments.

© 2017 Elsevier B.V. All rights reserved.

Contents

1. Introduction	49
2. Nucleocapsid structural and functional organization	49
2.1. Nucleocapsid structure	50
2.2. The nucleoprotein NP	50
2.3. The viral proteins VP30 and VP35	51
2.4. The large polymerase L	51
3. The transcription/replication cycle	53
3.1. Filovirus RNA transcription	53
3.2. Genome replication	53
3.3. Regulation of the transcription/replication cycle	53
4. Inhibitors of the polymerase complex	54
4.1. <i>In silico</i> models and docking screenings	55
4.2. Nucleoside analogs	55
4.3. Targeting host proteins	56
4.4. Protein-protein interaction inhibitors	56
4.5. Other strategies	56
5. Conclusion	57
Acknowledgements	57
References	57

* Corresponding author.

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

1. Introduction

Filoviruses are filamentous, enveloped, non-segmented, negative-sense RNA viruses (NNS) that belong to the order *Mononegavirales*. The family *Filoviridae* includes the genera *Ebolavirus* and *Marburgvirus* that comprise highly pathogenic viruses, responsible for hemorrhagic fevers for which there are no approved therapies. In addition, in 2002, Lloviu virus (LLOV), a new filovirus that so far has only been identified in bats, was discovered and proposed to be classified in the novel genus *Cuevavirus* (Ascenzi et al., 2008; Negrodo et al., 2011). The filovirus classification and nomenclature have recently been overhauled (Bukreyev et al., 2014). Accordingly, the genera *Marburgvirus* and *Cuevavirus* each include a single species: *Marburg marburgvirus* (Marburg virus – MARV) and LLOV, respectively. In contrast, the genus *Ebolavirus* includes five virus species: *Zaire ebolavirus* (Ebola virus – EBOV), *Sudan ebolavirus* (Sudan virus – SUDV), *Tai Forest ebolavirus* (Tai Forest virus - TAFV), *Bundibugyo ebolavirus* (Bundibugyo virus - BDBV) and *Reston ebolavirus* (Reston virus - RESTV).

Filoviruses share a common genome organization with seven open reading frames that encode all viral proteins (Fig. 1): a nucleoprotein (NP), the viral proteins VP35 and VP40, the surface glycoprotein (GP), the viral proteins VP30 and VP24, and the “large” polymerase (L) (Ascenzi et al., 2008; Negrodo et al., 2011). GP is the protein that mediates filovirus entry into host cells. This step is a key point in the viral life cycle and an attractive target for antiviral strategies (Martin et al., 2016). After entry by macropinocytosis, the viral nucleocapsid (i.e., the viral genome encapsidated by NP and associated with other viral proteins) is released into the host cell

cytoplasm, allowing the synthesis of new viral proteins and genomes. Then, nucleocapsid and progeny virion assembly and budding are orchestrated by NP, VP24 and VP40, and enhanced by GP (Hartlieb and Weissenhorn, 2006; Noda et al., 2006). VP24 condenses the ribonucleoprotein (RNP) complex (e.g. the NP-bound RNA genome) that may fix L at the 3' end of genomes, thus blocking the transcription/replication cycle (Watt et al., 2014). Whole viral particle assembly is then driven by VP40 (assisted by NP in the case of MARV) that recruits host proteins important for efficient virus egress and spread (Hartlieb and Weissenhorn, 2006; Noda et al., 2006; Silva et al., 2012; Mittler et al., 2013).

In an earlier article, we provided a structure/function analysis of filovirus surface glycoproteins and viral entry into cells (Martin et al., 2016). In this review, we summarize the current knowledge about the structure/function relationships of the viral proteins that drive and regulate filovirus genome transcription/replication. We detail the complexes in which viral proteins are involved and their structural features. We then describe their role in the viral cycle and compare their functionalities with those of homolog proteins from prototype viruses within the order *Mononegavirales*. Finally, we show how these insights can help developing new and innovative antiviral strategies, with a specific focus on small organic compounds.

2. Nucleocapsid structural and functional organization

Filoviruses share the replication/transcription strategy of NNS RNA viruses of the order *Mononegavirales* (Fig. 1) (Whelan et al., 2004). Briefly, upon host cell entry through virus

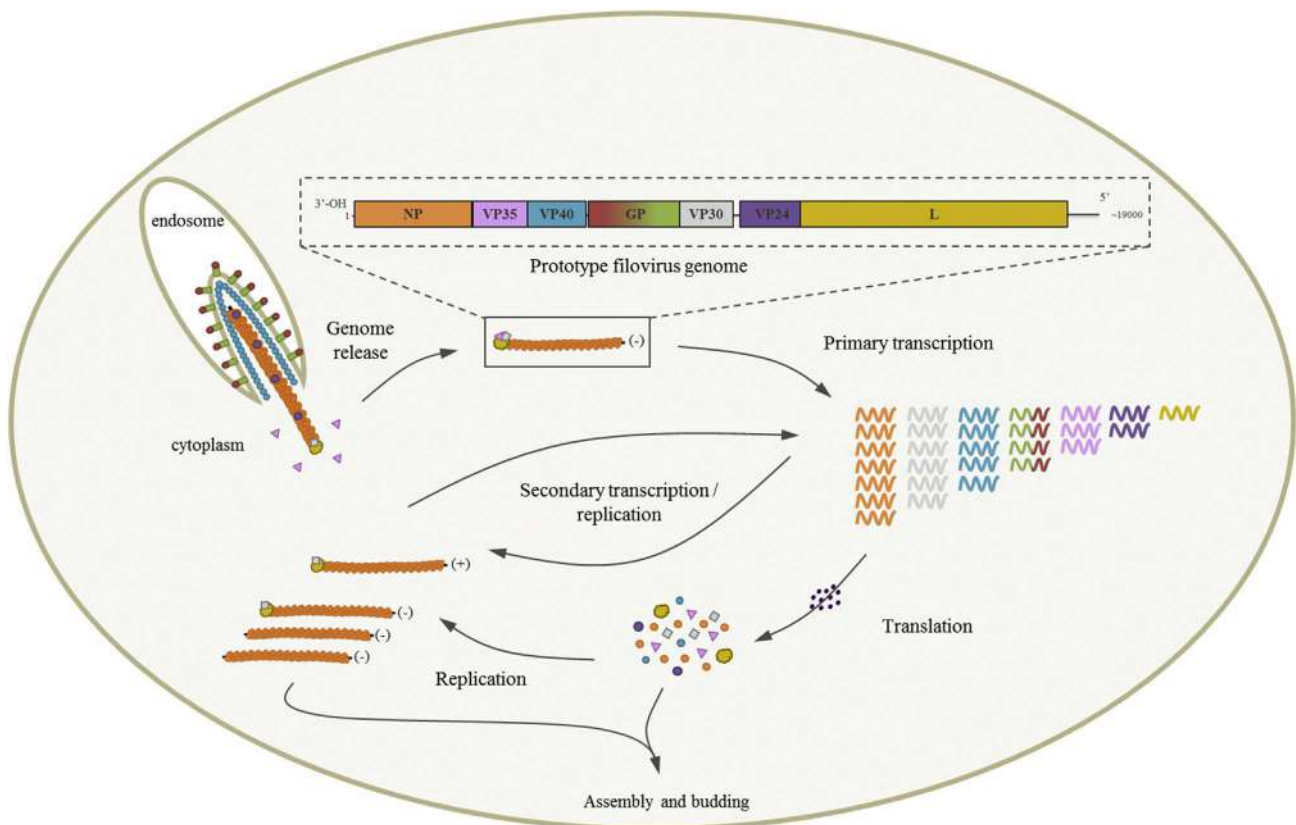


Fig. 1. Filovirus replication cycle. After entry into the host cell (reviewed in Martin et al., 2016), the filovirus genome is released in the cytoplasm. The negative-sense, single-stranded RNA genome is first sequentially transcribed (primary transcription) and then mRNAs are translated into viral proteins. With the increase of viral protein concentration and host cell factor regulation, the polymerase complex switches to replication mode to generate anti-genomes and new genomes. Finally, a balance between secondary transcription and replication leads to the assembly and budding of progeny virions.

macropinocytosis and endosomal processing, the virus RNA genome, which is embedded in the RNP complex, serves as template for discontinuous transcription catalyzed by the polymerase L (Mühlberger et al., 1999). This protein also harbors the catalytic activities required for mRNA polyadenylation and capping (Poch et al., 1990). Once the nucleocapsid has entered the cytoplasm, the polymerase, which is closely associated with the RNP complex, transcribes the viral RNA, producing the seven monocistronic capped viral mRNAs that are translated into viral proteins. Then, when the concentration of neo-synthesized proteins increases in the cytosol, the polymerase switches to a replicative mode to generate positive-sense RNA intermediates (anti-genomes) (Noton and Fearn, 2015). These anti-genome RNAs are used as templates for the synthesis of progeny negative RNA genomes.

Both genomes and anti-genomes are encapsidated by NP. NP intimately binds to the viral RNA and form a complex with the polymerase L and viral proteins VP24, VP30 and VP35. Although this replication/transcription strategy is shared by all viruses of the order *Mononegavirales*, some striking differences in the structural and functional properties essential for virus replication have been highlighted in filoviruses. These include the mechanism that drives RNA embedding in the nucleocapsid, the structure of the genomic replication promoter, the use of RNA editing mechanisms for GP production (Martin et al., 2016), the proteins that form the active RNP complex, and regulation of the switch between transcription and replication.

2.1. Nucleocapsid structure

In filoviruses, the RNP complex is composed of five proteins (NP, L, VP24, VP30 and VP35) that, together, package the 19kb RNA genome to form the nucleocapsid, a characteristic filamentous flexible structure (Fig. 2). Structural analysis of EBOV nucleocapsid at 19 Å resolution (Bharat et al., 2012) and of MARV nucleocapsid at about 26 Å resolution (Bharat et al., 2011) by cryo-electron microscopy and cryo-electron tomography with sub-tomogram averaging showed that the RNP complex is a left-handed helix with about 11–12 subunits per turn, with a pitch of about 7 nm and an outer diameter of about 41 nm. The symmetry of EBOV and MARV RNPs differs. EBOV nucleocapsid has fewer NP subunits per helix turn, but more turns in the helix per virion than MARV nucleocapsid. Consequently, EBOV virions are longer than MARV virions.

NP and VP30 show an essential RNA-binding activity (John et al., 2007; Noda et al., 2010); however, NP should be the main player in RNP complex formation. NP binds to the RNA genome and each NP covers a six-base RNA fragment. Therefore the replication promoter obeys the “rule of six”, although the genome length is not divisible by six (Bharat et al., 2011, 2012). VP24 and VP35 also are essential for the formation of the double-layered filovirus RNP complex (Huang et al., 2002; Watanabe et al., 2006) (see 5.2.). This is the most complicated type of nucleocapsid within the order *Mononegavirales* where, generally, the RNP complex is composed of only three proteins that coat the genome. With the rapid improvement of new tools, such as cryo-electron microscopy and cryo-electron tomography, it should soon be possible to obtain the high-resolution structure of this complex structure.

2.2. The nucleoprotein NP

NP shows the highest expression level in infected cells and is the main nucleocapsid component (roughly 3200 NP molecules per virion) (Bharat et al., 2012). Together with VP30, VP35 and L, it forms the minimal RNP complex responsible for virus transcription and replication (Mühlberger et al., 1998). Biochemical and structural studies demonstrated that NP drives viral RNA encapsidation.

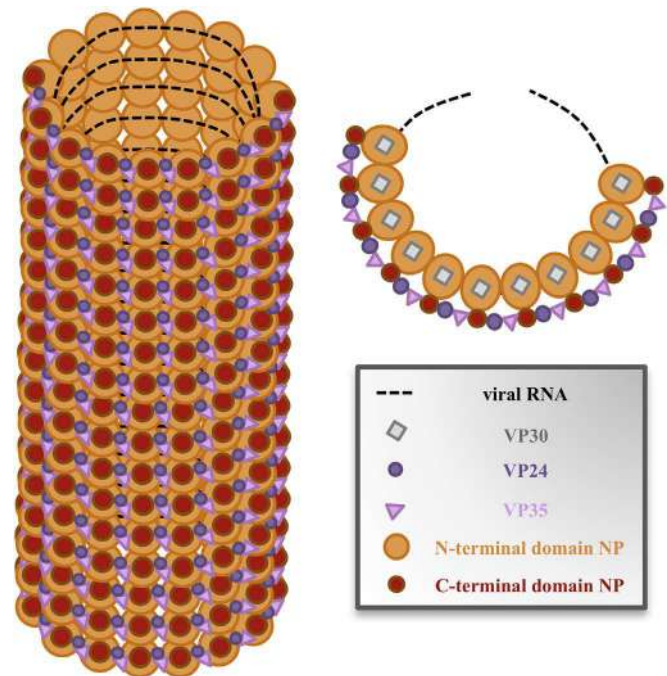


Fig. 2. Filovirus nucleocapsid ultrastructure. The nucleocapsid is a nucleoprotein complex in which the viral genome RNA is coated by five nucleoproteins: the nucleoprotein NP (its N-terminal domain is involved in RNA binding and nucleocapsid assembly, and the C-terminal is necessary for protein-protein interactions), the viral proteins VP30, VP35 and VP24 and the polymerase L. The whole complex forms a flexible left-handed helix with a pitch of about 7 nm and an outer diameter of about 41 nm. The nucleocapsid ultrastructure has been solved by cryo-electron microscopy and cryo-electron tomography. The L protein is not represented here because it has not yet been structurally described in the nucleocapsid complex.

NP includes an N-terminal and a C-terminal domain that harbor specific functions. The EBOV NP core structure was recently solved (Dziubańska et al., 2014; Dong et al., 2015).

The N-terminal domain carries the oligomerization and packaging functions that are usually found in nucleoproteins of other viruses of the order *Mononegavirales* (DiCarlo et al., 2007; Dong et al., 2015; Peng et al., 2016). This domain contains two separate lobes (N and C), mostly constituted of α -helices. Similarly to other mononegaviruses, the N- and C-lobes clamp a highly positively charged groove, driving RNA binding and encapsidation. The involvement of these domains in RNA binding/encapsidation was confirmed by mutagenesis (Mühlberger et al., 1999; Watanabe et al., 2006; Noda et al., 2010; Leung et al., 2015). Additionally, structural homology with other mononegavirus NPs that bind to RNA in the outer side of the nucleocapsid complex (Tawar et al., 2009; Peng et al., 2016) suggests that filovirus NP binds to RNA in a similar way but this still remains to be demonstrated. Another hydrophobic coiled-coil motif, which is highly conserved in all filoviruses, has been described in the N-terminal domain of NP (DiCarlo et al., 2007; Dong et al., 2015). This motif forms a pocket that might accommodate VP35 (Leung et al., 2015) and constitutes an attractive antiviral target (see 4.4.) (Kao et al., 2010; Gerritz et al., 2011).

In contrast to the N-terminal domain, the C-terminal domain has no homology with other known mononegavirus NPs and its sequence is poorly conserved among filoviruses (Dziubańska et al., 2014). Its X-ray structure revealed a globular domain in which β -strands alternate with α -helices (Dziubańska et al., 2014). This domain could serve as a hub for protein-protein interactions within the nucleocapsid that are different from what has been observed in

any other virus of the order *Mononegavirales* (Beniac et al., 2012). The C-terminal domain is also an important antigenic determinant that could be used for virus identification and species-specific diagnostic tests (Saijo et al., 2001; Changula et al., 2013; Sherwood and Hayhurst, 2013). Indeed, the recently solved core structures of BDBV and TAFV NP show a similar conformation to that of EBOV NP (Baker et al., 2016), but differences in the amino-acid sequence could pave the way to the development of sequence-specific anti-NP antibodies for diagnostic purposes. No MARV NP structure is available yet, but NMR data suggest that its folding might be different from that of ebolavirus NPs (Baker et al., 2016).

Filovirus NPs are post-translationally modified proteins. Specifically, ebolavirus NPs exhibit O-glycosylation and sialylation that are required for nucleocapsid assembly (Huang et al., 2002; Watanabe et al., 2006). Both ebolavirus and MARV NPs can also be phosphorylated. The role of these phosphorylations has been studied in MARV. In infected cells, 92kDa and 94kDa NP forms can be detected and the 94kDa form, which can be converted into the 92kDa form by dephosphorylation, is the isoform incorporated into virions (Becker et al., 1994). In contrast to NPs of other viruses of the order *Mononegavirales* that have only a few phosphorylation sites, MARV NP has seven phosphorylatable regions that are distributed in the C-terminal domain (Lötfering et al., 1999). These phosphorylations might modulate NP affinity for RNA to regulate RNA synthesis, as has been proposed for other mononegaviruses (Yang et al., 1999; DiCarlo et al., 2011). Indeed, it is thought that RNA-bound NP inhibits the access of the polymerase L to its RNA template for nucleocapsid formation. Upon phosphorylation, NP affinity for the viral RNA is reduced, probably because of repulsion of negatively charged NP residues toward phosphate groups. Hence, RNA becomes accessible to the polymerase complex for transcription and replication (DiCarlo et al., 2011). As EBOV and MARV NP phosphorylation profiles are similar (Elliott et al., 1985; Peyrol et al., 2013), these phosphorylations might have a comparable regulatory role in both genera.

Although structure/function analysis indicates that NP carries the RNA binding properties, RNA encapsidation in EBOV also requires VP35 and VP24 (Huang et al., 2002; Noda et al., 2005). These proteins induce a conformational change in NP, resulting in the formation of nucleocapsid-like structures that are competent for RNA encapsidation (Watanabe et al., 2006).

2.3. The viral proteins VP30 and VP35

The structure of the transcription activator VP30 has been partially solved and several functional domains have been identified (Hartlieb et al., 2007; Clifton et al., 2014). In ebolaviruses, its N-terminal domain harbors a leucine-zipper motif involved in homooligomerization for transcription activation. This stretch of hydrophobic amino acids leads to VP30 hexamerization that represents the active conformation to support viral mRNA transcription (Hartlieb et al., 2003, 2007). The N-terminal domain also carries a zinc finger that plays a critical role in transcription activation (Modrof et al., 2003; Enterlein et al., 2006). Although VP30 is thought to contribute to RNA binding, the molecular mechanism underlying its role in transcription regulation remains elusive (Modrof et al., 2003; John et al., 2007). Its C-terminal structure has been solved by X-ray crystallography. It forms a globular structure composed of α -helices that can homodimerize. Moreover, it interacts with NP via a basic amino-acid cluster that plays a role also in VP30 transcriptional control activity (Hartlieb et al., 2007; Kirchdoerfer et al., 2016). This suggests that NP and transcription regulation are tightly linked.

Like NP, VP30 can be phosphorylated (Elliott et al., 1985; Becker and Mühlberger, 1999). EBOV VP30 phosphorylation sites occur in

two conserved clusters of three serine residues (Modrof et al., 2002) and two threonine residues (Ilinykh et al., 2014). Interestingly, these sites overlap with an arginine-rich region that participates in RNA binding (Biedenkopf et al., 2016b). Similarly, a stretch of seven serine residues represents MARV VP30 main phosphorylation site, although phosphorylation of still unidentified threonine residues has been also suggested (Modrof et al., 2001). VP30 phosphorylation contributes to the transcription/replication switch (see 3.1.) and nucleocapsid assembly, and is, therefore, an interesting target for antiviral drugs (see 4.3.) (Ilinykh et al., 2014). Differently from ebolaviruses, MARV VP30 does not seem to have a central role in minigenome transcription systems because NP, L and VP35 are sufficient for transcription and replication (Mühlberger et al., 1999). Nonetheless, in the absence of MARV VP30, the intracellular level of all other viral proteins is reduced (Fowler et al., 2005). Also, it might be different in the context of viral infection, especially since MARV VP30 is required for virus rescue (Enterlein et al., 2006). In conclusion, VP30 acts as a transcription activator in EBOV but not in MARV, where it could be a structural component probably involved in proper nucleocapsid assembly.

The L cofactor VP35 is one of the best studied filoviral proteins, which is similar to the P protein of NNS viruses. VP35 is a multifunctional protein important for nucleocapsid assembly, replication and transcription as well as for suppressing the host innate immunity in all filoviruses (Leung et al., 2010; Feagins and Basler, 2015). It has a flexible N-terminal region with a central oligomerization domain with a coiled-coil motif that is thought to participate in protein-protein interactions, including in VP35 self-interaction (Möller et al., 2005; Reid et al., 2005) and association with the polymerase L (Trunschke et al., 2013). Interestingly, N-terminus structural analyses of ebolavirus VP35 revealed that this protein should oligomerize as a tetramer (Luthra et al., 2015; Bruhn et al., 2017) while the recent structure of the MARV VP35 shows a trimeric arrangement (Bruhn et al., 2017).

As supported by biochemical data, MARV and EBOV VP35 may indeed oligomerize differently, partially explaining the specificities of their related replication strategies (Mühlberger et al., 1998, 1999). VP35 also contains a flexible linker region that interacts with NP and regulates NP function and RNA synthesis (Leung et al., 2015). This region is called NP-binding peptide (NPBP) and is necessary to maintain NP in a non-oligomeric and RNA-free state, probably by regulating the balance between NP oligomeric/monomeric forms, thus facilitating the transfer of monomeric NP to the viral template RNA (Kirchdoerfer et al., 2015; Leung et al., 2015). VP35 C-terminal domain binds double stranded-RNA binding (Cárdenas et al., 2006; Basler and Amarasinghe, 2009; Kimberlin et al., 2010) and is responsible for suppressing the host immune response (reviewed in Basler, 2015). Interestingly, this C-terminal domain also contains a basic patch which is not critical for interferon inhibition, but is important for viral RNA synthesis. Indeed, many of these basic residues are involved in NP interaction, supporting the role of VP35 as a central regulator bridging NP and L, which makes it a filovirus-specific target for antiviral development (see 4.4.) (Prins et al., 2010). These regulatory proteins are key factors for the virus life cycle, especially for the transcription/replication switch. However, their structural assembly and regulation need to be better deciphered to thoroughly understand their involvement in filovirus life cycle and to target them with specific compounds.

2.4. The large polymerase L

While several studies have brought insights into the structure/function of NP, VP35 and VP30, much less is known about the essential L protein. This large protein of about 2200 amino acids has

an estimated molecular mass of 250 kDa (Volchkov et al., 1999). By comparative analyses, six conserved regions (I to VI) have been identified in all viruses of the order *Mononegavirales* (Fig. 3). They include three enzymatic domains that are required for replication/transcription (RNA-dependent RNA polymerase, RdRp), polyadenylation (RdRp) and mRNA capping (poly-ribonucleotidyltransferase, PRNTase and methyltransferase, MTase) (Poch et al., 1990; Volchkov et al., 1999; Ferron et al., 2002; Ayub and Waheed, 2016). No crystal structure of the L protein of viruses belonging to the family *Filoviridae* is available yet. However, the structure of the L protein of vesicular stomatitis virus (VSV), which also belongs to the order *Mononegavirales*, was recently determined by cryo-electron microscopy at a resolution of 3.8 Å (Liang et al., 2015). Specifically, the structure of the RdRp domain confirms that NNS virus polymerases share common features with

other viral RNA polymerases, such as the right-hand shape with fingers, palm and thumb subdomains. Moreover, the RdRp and PRNTase domains are intimately linked through a loop in the PRNTase domain that projects into the RdRp catalytic site, in a position that suggests a role as priming loop.

This finding also suggests that RNA synthesis and capping are highly coordinated in mononegaviruses. This was consistent with earlier biochemical studies showing that cap addition is a co-transcriptional process essential for RNA elongation catalyzed by the polymerase L (Rose et al., 1977). Moreover, RNA capping in VSV is ensured by a PRNTase rather than a guanylyltransferase (GTase) (Ogino and Banerjee, 2007). Indeed, the capping reaction requires the formation of a covalent linkage between the 5' end of the RNA and a histidine residue in a conserved HR motif (in conserved region V) in L (Fig. 3). The PRNTase also binds to GDP via a GXXT motif

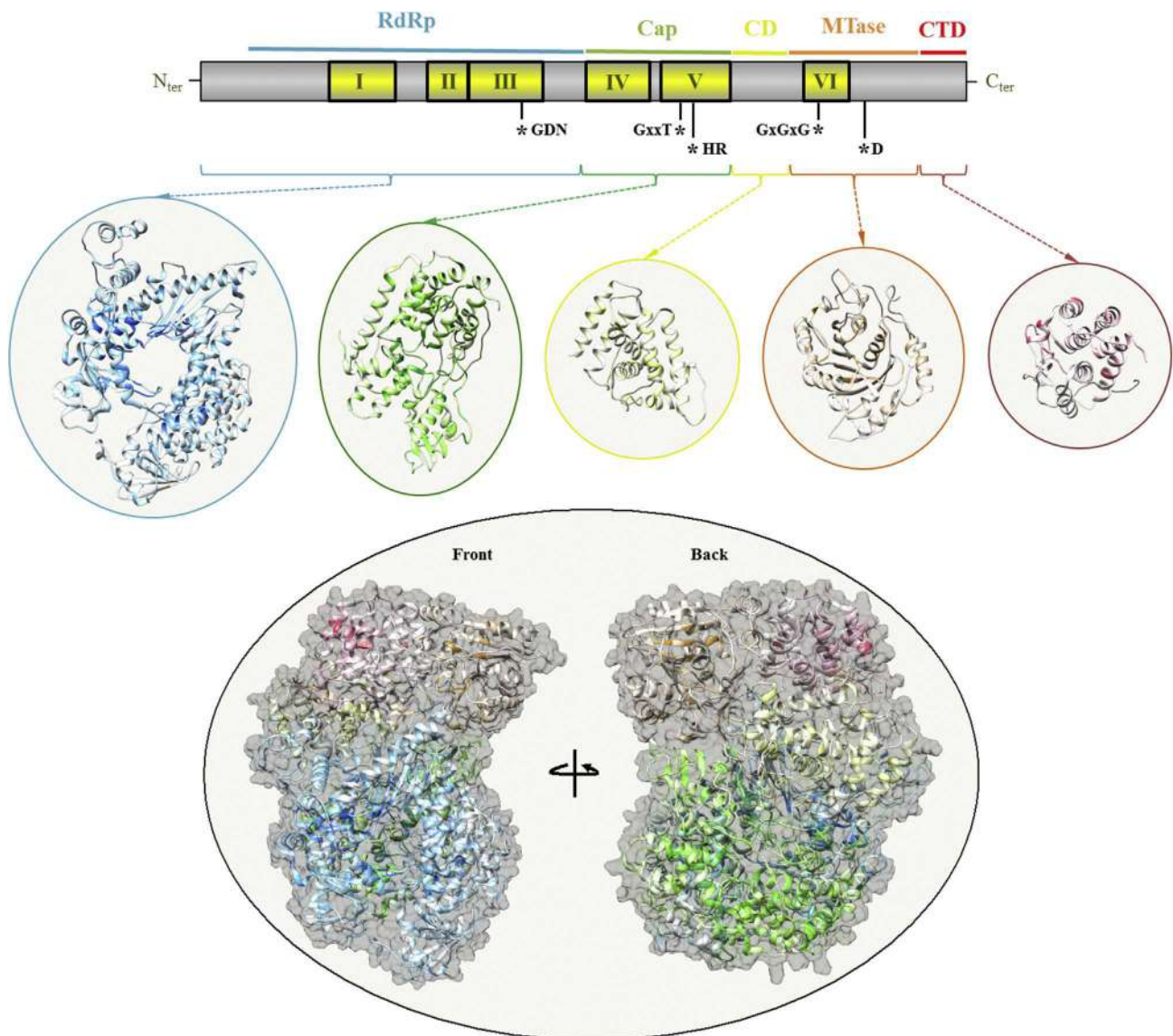


Fig. 3. VSV L protein structure and conservation throughout the order *Mononegavirales*. The activities of the L protein have been characterized for vesicular stomatitis virus (VSV) and have been correlated with its recently solved structure presented here (PDB 5A22). The complete protein includes domains with different activities: the RNA-dependent RNA polymerase (RdRp, blue), the poly-ribonucleotidyltransferase (Cap or PRNTase, green), a highly variable domain (CD, yellow), the cap-methyltransferase (MTase, orange) and the C-terminal domain (CTD, red). Based on sequencing data and VSV structure, conserved regions (I to VI; yellow boxes) and residues in the 3D model at the bottom have been identified as not conserved (white) to highly conserved (brightly colored) in all viruses of the order *Mononegavirales*. While most L residues are poorly conserved, the core residues of the different activities are highly conserved (highlighted with an asterisk in the schematic representation of the L protein, upper panel), suggesting high functional homology among mononegaviruses.

(also in conserved region V) that is located about 70 residues downstream of the HR motif and transfers a GMP moiety to the 5' phosphate of the nascent RNA covalently bound to a histidine residue, allowing attachment of the cap.

After capping, the cap structure is converted into a cap-1 structure (^mGpppN_m). Unlike in standard eukaryotic RNA capping pathways, the 2'O methylation of the N1 ribose of RNA precedes the N7 methylation of the cap guanosine by the MTase domain (conserved region VI) present at the C-terminus of the L protein (Fig. 3). These cap-MTase activities were described using several full-length L proteins (Li et al., 2006) and more recently using the MTase with the C-terminal domain of human metapneumovirus, the structure of which was solved by X-ray crystallography at a resolution of 2.2 Å (Paesen et al., 2015). Although most of the enzymatic activities embedded in the filovirus L protein have not been identified yet, the strong homology with other mononegavirus L proteins might provide a working surrogate model. Nevertheless, L structural and functional features need to be determined because this protein represents an attractive antiviral target due to its essential role in the replication cycle (Ferron et al., 2012), adding to the fact that it is essential for RNA stability and stimulates RNA translation into viral proteins via mRNA recognition by the host translation initiation factor eIF4E. Moreover, 2'O-methylation of cap structures is marker of “self” (Züst et al., 2011; Hyde and Diamond, 2015; Schubert-Wagner et al., 2015; Devarkar et al., 2016). Therefore, inhibition of the viral 2'O MTase leads to recognition of viral RNA by host cell sensors, such as RIG-I and MDA5, and consequently to interferon production (Daffis et al., 2010; Kumar et al., 2014). In turn, interferon-stimulated genes (ISGs), can detect mis-capped RNAs and restrict their translation into proteins (Bray et al., 2002) (Daffis et al., 2010; Kumar et al., 2014). Interestingly, particular ISGs, such as IFIT1, do not restrict infection of EBOV in mice, which may be another trick of filoviruses to escape from innate immunity oversight (Pinto et al., 2015).

3. The transcription/replication cycle

3.1. Filovirus RNA transcription

Transcription of NNS viruses requires a minimal set of proteins: NP, the phosphoprotein P and L (Conzelmann, 2004). In filoviruses, NP, VP35 and L are sufficient for transcription of MARV mini-genomes (Mühlberger et al., 1998), whereas ebolaviruses also require VP30 (Mühlberger et al., 1999; Boehmann et al., 2005). VP30 might regulate the early step of transcription initiation and reinitiation (see 3.3.) (Fig. 4) (Modrof et al., 2002; Martínez et al., 2008). VP30 is actually required to initiate transcription in the presence of a hairpin loop, whereas disruption of the hairpin enables VP30-independent transcription initiation at the first gene start (Weik et al., 2002).

Filovirus transcription is a discontinuous process that generates seven monocistronic capped and poly-adenylated RNAs in ebolaviruses and MARV (Mühlberger et al., 1996; Weik et al., 2002) (Fig. 1), and one bicistronic mRNA in LLOV (Negredo et al., 2011). The polymerase recognizes gene-start (GS) and gene-end (GE) sequences in intergenic sequences (from 57 to 684 nucleotides in size) of the genomic RNA to synthesize the different viral mRNAs. The current model postulates that the polymerase complex recognizes a single polymerase-binding site at the 3' end of the genome to initiate RNA synthesis. Once bound, the polymerase complex transcribes sequentially the genes from the 3' to the 5' end of the genome by stopping and reinitiating at GE and GS junctions, respectively (Whelan et al., 2004). During this process, the polymerase complex might fall from the RNA template in intergenic sequences, leading to the progressive decrease of mRNA expression

level observed in infected cells (Fig. 1). The filovirus GS sequence includes a stretch of 12 nucleotides, which varies slightly between MARV (3' CUUCUUA/GUAAUU) and EBOV (3' CUC/ACUUCUAAUU) (Mühlberger et al., 1996), and the 3' UAAUUC motif followed by a stretch of five to six uridine residues that are required for the poly-A tail addition by a transcriptional stuttering mechanism (Whelan et al., 2004). Interestingly, these regulatory elements show a high degree of sequence homology between marburgviruses and ebolaviruses (Sanchez et al., 1993; Mühlberger et al., 1996), although the polymerase complexes remain virus-specific (Mühlberger et al., 1999; Boehmann et al., 2005). This suggests that the structure and/or composition of these complexes might be crucial for their specificity (Mühlberger et al., 1999).

3.2. Genome replication

In NNS RNA viruses, the polymerase L is not sufficient to catalyze RNA replication and the synthesis of the positive-strand anti-genome that is subsequently used to synthesize progeny genomes. The L protein is associated with P to form the active viral polymerase complex that interacts preferentially with NP-bound RNA rather than uncoated RNA (Whelan et al., 2004). Mini-replicon systems have been used to investigate in detail MARV and EBOV genome replication (Mühlberger et al., 1998, 1999; Groseth et al., 2005). In such systems (reviewed in Hoenen et al., 2011), a genomic viral RNA template, called a mini-genome or mini-replicon, is modified to remove some or all of the virus open reading frames and replace them with a reporter gene (GFP or luciferase) that contains non-coding regions of the virus of interest harboring the minimal signals for replication and transcription. This minigenome can be used as a template to monitor the activity of the virus polymerase by reporter gene expression. These systems made it possible to show that NP, VP35 and L constitute the minimal set of proteins necessary for genome replication (Mühlberger et al., 1999). This suggests that VP35, in filoviruses, and P, in NNS RNA viruses, have a similar role and that the active polymerase complex includes VP35 and L (Mühlberger et al., 1999). Moreover, VP35 forms homo-oligomers via its amino-terminal coiled-coil motif that interact with L and NP (Becker et al., 1998; Reid et al., 2005; Zinzula et al., 2009). This oligomerization could be a prerequisite for interaction with the N-terminal region of L (Möller et al., 2005; Prins et al., 2010; Trunschke et al., 2013).

VP35, L, NP and RNA form a huge complex thought to contain all the catalytic functions required for genome replication, starting at the RNA 3' end. The 3' end of filovirus genomes and anti-genomes contains 155- and 176-nucleotide-long promoters, respectively. The 176-nucleotide-long promoter harbors conserved hairpin structures that may be used for replication initiation (Calain et al., 1999). The EBOV replication promoter belongs to the class of bipartite promoters, such as those identified in *Paramixoviridae*, rather than to the class of monopartite promoters of *Rhabdoviridae* and *Pneumoviridae*. The first promoter element involves the leader region (nucleotides 1–55), and the second promoter element spans nucleotides 81–128 within the untranslated region of the first gene that encodes NP. A spacer region, in which the transcription start signal of the NP gene is present, separates the two promoter elements. The structure of the second promoter element is unique among NNS RNA viruses and consists of a stretch of eight consecutive (UN)₅ hexamers, where N can be any nucleotide (Weik et al., 2005).

3.3. Regulation of the transcription/replication cycle

Recent studies have identified different functional elements in VP30, such as homodimerization and hexamerization (Hartlieb

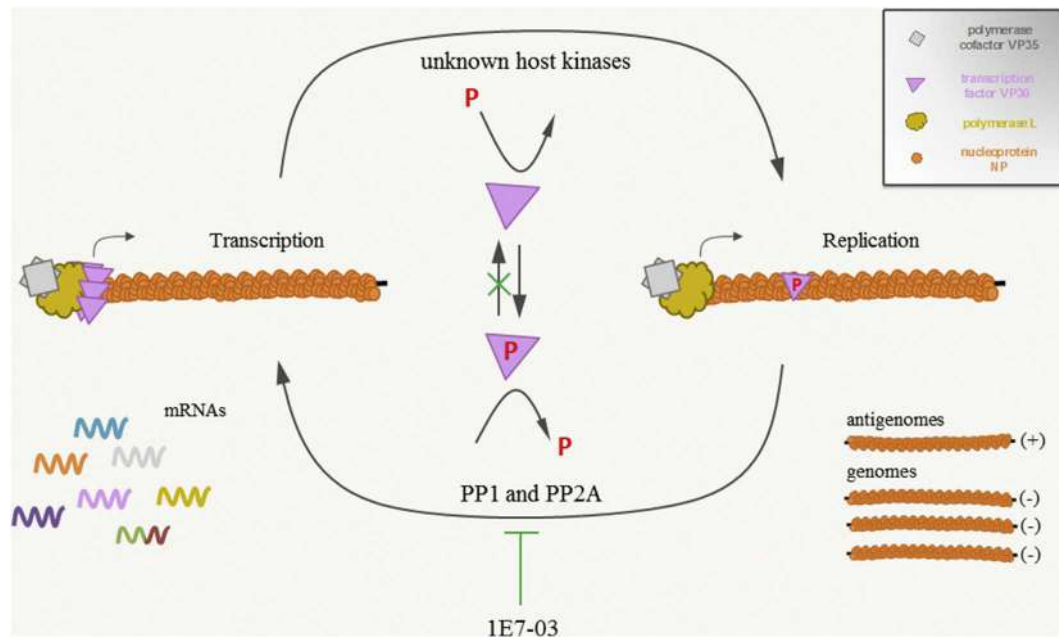


Fig. 4. Ebola virus VP30 regulatory role in transcription/replication cycle. The viral protein VP30 is unquestionably implicated in the regulation of the Ebola virus transcription/replication cycle (Mühlberger et al., 1999) through its phosphorylation state. Following nucleocapsid release into the cytoplasm, NP-associated and phosphorylated VP30 is dephosphorylated by protein phosphatase 1 and 2A (PP1 and PP2A). Dephosphorylated VP30 hexamerizes and interacts with the polymerase complex to trigger primary transcription. mRNAs are translated and viral protein concentrations increase in the cytoplasm. The balance between phosphorylated, dephosphorylated and rephosphorylated VP30 regulates the switch from secondary VP30-dependent transcription to VP30-free replication, leading to the generation of anti-genomes and then new genomes. This regulation is an attractive anti-viral therapeutic target because compounds like 1E7-03 can inhibit phosphatases and consequently RNA transcription.

et al., 2003, 2007; Modrof et al., 2003), phosphorylation (Biedenkopf et al., 2013, 2016a; Ilinykh et al., 2014), and binding to NP (Hartlieb et al., 2007; Biedenkopf et al., 2013; Kirchdoerfer et al., 2016), L (Groseth et al., 2009) and RNA (John et al., 2007; Schlereth et al., 2016). These elements allowed building a VP30-regulated transcription activation model, albeit incomplete, in ebolaviruses (Fig. 4).

Briefly, after release into the cytoplasm, the nucleocapsid is transported to sites of viral RNA synthesis. VP30 phosphorylation is essential to ensure its association and co-transport with the nucleocapsid because it increases VP30 affinity for NP (Biedenkopf et al., 2013). Indeed, non-phosphorylated VP30 can easily leave the nucleocapsid during transport and consequently might not be available for primary viral transcription. For the initiation of viral transcription, VP30 must be dephosphorylated by the host protein phosphatase 1 and 2A (PP1 and PP2A), which could represent anti-filovirus targets (see 4.3.) (Ilinykh et al., 2014). Dephosphorylated VP30 becomes fully active and associates with VP35 and L at the transcription start site (Groseth et al., 2009; Biedenkopf et al., 2013). This triggers transcription initiation and reinitiation (Weik et al., 2002; Martínez et al., 2008). Non-phosphorylated VP30 assembles into a hexamer that forms a ternary complex with viral RNA and VP35. It has been hypothesized that VP30 clamps together the RNA template and VP35/L, thus increasing the polymerase affinity for negative-sense RNA to boost the rate of transcription initiation (Biedenkopf et al., 2016b; Schlereth et al., 2016). After primary transcription, new VP30 proteins can be phosphorylated, dephosphorylated or re-phosphorylated by the host phosphatases and not-yet-identified kinases, leading to a balance between secondary transcription and VP30-free replication (Biedenkopf et al., 2016a). In MARV, VP30 does not seem to be involved in transcription regulation, but is necessary for protein synthesis (Mühlberger et al., 1999; Fowler et al., 2005). MARV VP30 phosphorylation is crucial for NP binding and thus for VP30 localization in inclusion

bodies where transcription and replication take place (Modrof et al., 2001); however, its role in transcription regulation remains unclear.

VP30 is unique to filoviruses because no homolog protein has been found in other viruses of the order *Mononegavirales*. Indeed, in other mononegaviruses, only the highly phosphorylated phosphoprotein P, which is considered to be similar to the non-phosphorylated or only weakly phosphorylated filovirus VP35, coordinates transcription and replication, according to its phosphorylation status (Saikia et al., 2008; Fuentes et al., 2010; Sun et al., 2011; Sugai et al., 2012; Mondal et al., 2014; Pickar et al., 2014). The only comparable regulation loop has been partly characterized in pneumoviruses, notably in respiratory syncytial virus (RSV) that encodes the additional transcriptional support factor M2-1 (Fearn and Collins, 1999). In contrast to VP30, M2-1 supports transcription elongation and anti-termination when phosphorylated, and acts as an essential phosphorylation-dependent viral transcription factor (Fearn and Collins, 1999; Mason et al., 2003). Indeed, its dynamic and reversible phosphorylation is required for its function in transcription, presumably by regulating the binding of other unknown factors (Tanner et al., 2014).

In conclusion, the mechanism regulating the balance between transcription and replication is still elusive. The identification of VP30 kinases and a complete characterization of L/VP35/VP30/NP coordination might help elucidating the mechanism by which filoviruses regulate their transcription/replication cycle. As VP30 is a crucial regulatory element for virus multiplication, it constitutes a key target for filovirus-specific antiviral strategies.

4. Inhibitors of the polymerase complex

As the viral replication/transcription machinery has no equivalent in the host cell, it is an ideal target for developing potent antiviral strategies. Ribavirin, a broad-spectrum antiviral molecule

active against other hemorrhagic viruses, such as arena- and bunyaviruses (Huggins, 1989; Ignat'ev et al., 1996), was the first drug tested against filoviruses. However, it did not protect animals from a lethal EBOV challenge. Therefore, other strategies have been and are currently being developed to treat filovirus infections (summarized in Table 1).

4.1. *In silico* models and docking screenings

Several groups have taken advantage of the L protein structural homologies across *Mononegavirales* to develop a three-dimensional model of the EBOV polymerase to perform molecular docking studies with nucleotide analogs, including a variety of approved antiviral drugs. For instance, docking experiments were conducted using nucleoside analogs that had been originally selected as inhibitors of other types of RNA polymerases and could potentially block EBOV polymerase also (Jácome et al., 2015; van Hemert et al., 2015). However, in the absence of any crystal structure of the RdRp, not to mention appropriate RdRp:enzyme:nucleotide ternary complexes, these studies have limited value.

Several screens have been performed to identify known nucleoside analogs that can interfere with EBOV replication. Multiple putative inhibitory compounds were identified, such as the cytidine analog β -d-N4-hydroxycytidine (Reynard et al., 2015), and CMLDBU3402 that inhibits also VSV (Filone et al., 2013), and favipiravir (T705). Only the latter seems to bear some reasonable potential as an anti-EBOV drug (see below). It is converted to favipiravir-ribofuranosyl-5'-triphosphate in host cells and is active against a broad spectrum of RNA viruses (Oestereich et al., 2014). The L methyltransferase domain also has been modeled for virtual screening of FDA-approved drugs and nucleotide/nucleoside analogs (Zhao et al., 2016). Several interesting hits have been identified with binding predictions higher than or close to that of S-adenosyl methionine (SAM), the natural methyl donor, which, with respect to cellular MTases, does not automatically qualify these hits as selective candidates. It is also necessary to demonstrate that these compounds have an actual antiviral effect on the EBOV MTase since cap methylation inhibition may restrict viral mRNA translation and induces a strong antiviral response mediated by the host innate immunity (see 2.4.).

Virtual screening has also been performed with plant extracts used in Indian medicine, leading to the selection of molecules that target different pockets of the viral proteins VP30, VP35, VP24 and VP40 (Brown et al., 2014; Setlur et al., 2016).

These *in silico* models led to the identification of several potential ligands of viral proteins among already known molecules, but did not allow identifying new antiviral compounds. Nevertheless, such strategies are useful for selecting interesting compounds from large chemical libraries in order to limit the number of compounds to test in infected cells or in mini-replicon systems.

4.2. Nucleoside analogs

Among the different molecules identified by *in silico* studies, few have been tested *in vitro* and *in vivo* with significant effects. The pyrazine carboxamide derivative favipiravir (T-705) is one of the most interesting drug candidates because of its efficacy against EBOV *in vitro* and in animal models, and its good tolerance profile in adult patients. Moreover, favipiravir could be quickly available because it has already been approved in Japan to treat influenza virus infections (Westover et al., 2016). Favipiravir suppresses EBOV replication in cell culture (IC₅₀ of 110 μ M) and its administration at day 6 post-infection in a mouse model induces rapid EBOV clearance, reduces the concentration of biochemical markers of disease severity, and prevents death in 100% of infected animals (Oestereich et al., 2014). In humans, favipiravir has a post-exposure prophylactic effect on patients recently infected by EBOV and a pre-exposure prophylactic effect on exposed contacts (Bouazza et al., 2015; Madelain et al., 2016; Mora-Rillo et al., 2015; Palich et al., 2016; Schibler et al., 2015). However, the JIKI trial, a historically controlled and single-arm test of feasibility and acceptability of an emergency trial in the context of the large EBOV outbreak of 2013–2015, concluded that favipiravir is unlikely to be efficient in patients with high viremia (Sissoko et al., 2016). Additional studies are needed to improve favipiravir dosing and better determine the therapeutic window.

Other nucleotide analogs have been investigated, notably the adenosine analog BCX4430. This molecule is converted to the 5'-triphosphate derivative BCX6870 and acts as a RNA chain terminator after its incorporation in hepatitis C virus (HCV) RNA (Warren et al., 2014; Taylor et al., 2016), which has a significantly different active-site structure than that of NNS RdRps. In the papers cited above, it was not clear if these compounds had some effects on HCV replicons, raising questions about their MoA, toxicity, and finally, potential use. BCX6870 also is a potent broad-spectrum compound against negative-sense RNA viruses with no obvious mutagenic effect. It has the same pharmacokinetics as a nucleotide, and is currently tested in non-human primates and in healthy humans

Table 1
Summary of filovirus protein structure/function characterization and related antiviral strategies.

Viral protein	Structure	Functions	Inhibition strategies	References
NP	N-terminal and C-terminal domains (EBOV)	RNA encapsidation	protein-protein interaction inhibitors (NP/VP35, NP/NP)	Kirchdoerfer et al., 2015; Leung et al., 2015; Fu et al., 2016
VP35	interferon inhibitory domain (EBOV and MARV)	replication/transcription cofactor	protein-protein interaction inhibitors (VP35/NP and VP35/L)	Trunschke et al., 2013; Kirchdoerfer et al., 2015; Leung et al., 2015; Fu et al., 2016
VP30	C-terminal domain (EBOV)	interferon antagonism transcription activator	not reviewed phosphatase regulation inhibitors	not reviewed Illyin et al., 2014
L	not available	RNA-dependent RNA polymerase	anti-oligomerization peptides nucleoside analogs (favipiravir, BCX4430, GS-5734)	Hartlieb et al., 2003 Oestereich et al., 2014; Warren et al., 2014; Reynard et al., 2015; Warren et al., 2016 Jacobs et al., 2016; Dörnemann et al., 2017 Trunschke et al., 2013
		RNA modifications (polyribonucleotidyltransferase, cap-methyltransferase, and polyadenylation)	protein-protein interaction SAM/SAH balance disruptors	Huggins et al., 1999; Hyde and Diamond, 2015

In conclusion, this table aims to summarize advances in terms of the structure and function of filovirus replication proteins. Progress in our understanding of filovirus biology, together with the urgent need to design new efficient inhibitors, allow the development of innovative antiviral strategies.

(phase I clinical trial). However, *in vitro* data are confusing because this compound shows variable IC₅₀ values when tested with various negative-sense RNA virus polymerases as well as cytotoxicity. No *in vitro* characterization data are available on the molecular mechanism of its antiviral activity in filoviruses. Therefore, fundamental research on its mode of action is required before starting a phase I trial in humans.

More recently, the compound GS-5734, which is a mono-phosphoramidate prodrug of an adenosine analog, has been described with similar interesting capacities but a deeper investigation of its *in vitro* activity is still needed (Warren et al., 2016). Punctual reports of administration in humans as experimental antiviral therapy support that GS-5734 should be a good anti-filovirus candidate (Jacobs et al., 2016; Dörnemann et al., 2017).

Another non-toxic cytosine analog, β -D-N⁴-hydroxycytidine or NHC, has been identified as a replication inhibitor *in vitro* (Reynard et al., 2015). The molecular mechanism of its antiviral activity is still unknown. In the same study, lamivudine, zidovudine and tenofovir failed to inhibit EBOV replication, although they were tested at the same concentration and using the same reporter EBOV replicon model. This raises questions about the reproducibility and reliability of the different replicon systems and protocols used to screen antiviral compounds.

4.3. Targeting host proteins

Several decades ago, it was reported that 3-deazaadenosine and 3-deazaneplanocin, two adenosine analogs, inhibit EBOV by blocking the cellular enzyme S-adenosyl-homocysteine (SAH) hydrolase. SAH is the product of methylation reactions that use SAM as the methyl donor (Wolfe and Borhardt, 1991; De Clercq, 1998). As SAH hydrolase down-regulates the SAH pool in the cell, its inhibition leads to increased intracellular SAH concentration and, consequently, to inhibition of the viral SAM-dependent methyltransferases that are involved in viral RNA cap methylation. Accordingly, 3-deazaadenosine and 3-deazaneplanocin have a broad antiviral profile (de Clercq and Montgomery, 1983) and can inhibit EBOV in infected cells and *in vivo* (Huggins et al., 1999). During EBOV infection, 3-deazaneplanocin A inhibits EBOV replication and also boosts the innate antiviral responses and increases the interferon response *in vitro* (Bray et al., 2000, 2002; Bray, 2001; Hyde and Diamond, 2015). This last effect might be related to the reduction of 2'O methylation of viral mRNA 5' caps that leads to the recognition of these mRNAs by the innate immunity sensors RIG-I and MDA5 (Züst et al., 2011; Schubert-Wagner et al., 2015; Devarkar et al., 2016).

Other molecules, such as the tetrahydrobenzothiazole compound 1, a derivative of the hit compound CID 847035, can generate a similar antiviral response in cultured cells, independently from type I interferons (Reichelt et al., 1989). These molecules represent new promising antiviral tools to expose the virus and to induce a strong antiviral response mediated by the host innate immunity that limits virus replication.

PP1, a cellular protein involved in the regulation of VP30 activity in the transcription/replication cycle (see 3.3.), also is a potential target for anti-filovirus strategies (Ilinykh et al., 2014). Indeed, the PP1 holoenzyme consists of a constant catalytic subunit (PP1 α , PP1 β/δ or PP1 γ) associated with a variable regulatory subunit that determines the phosphatase localization, activity and substrate specificity (Bollen et al., 2010). PP1 inhibition is effective against HIV-1 replication (Amosova et al., 2003, 2005, 2011). This phosphatase is also involved in EBOV replication (Fig. 4), and screening of a small library of molecules led to the identification of the hit compound 1E7-03. Its inhibitory effect on PP1 induces an increase of phosphorylated VP30, thus inhibiting viral transcription.

Importantly, 1E7-03 is non-toxic, probably because it prevents PP1-VP30 interaction. Additionally, host-cell PP1 regulatory subunits bind properly to PP1 through several interacting sites with a much higher affinity (Bollen and Beullens, 2002). Moreover, the similarity in the phosphorylation sites at the N-terminus of VP30 from various filoviruses suggests that PP1 inhibition is a feasible approach for developing a pan-filoviral drug.

4.4. Protein-protein interaction inhibitors

The filovirus replication complex also could be an interesting target, because L, VP30, VP35 and NP must assemble for efficient transcription and replication. Blocking the nucleocapsid formation by using protein-protein interaction inhibitors could lead to the specific inhibition of filoviruses.

The NP-NP and NP-VP35 interactions are essential for filovirus replication complex formation because they anchor the polymerase L to the RNA genome. NP interactions have been targeted successfully in influenza virus (Kao et al., 2010; Gerritz et al., 2011). In EBOV, the first identified NP ligand was a VP35-derived peptide (Kirchdoerfer et al., 2015; Leung et al., 2015). This peptide binds to NP with high affinity and specificity *in vitro*, blocks NP oligomerization, and induces RNA release from NP-RNA complexes. More recently, two compounds, 18 β -glycyrrhetic acid and licochalcone A, have been isolated from crude extracts of licorice roots, used in traditional Chinese medicine (Fu et al., 2016). These compounds also target NP with good affinity. Moreover, 18 β -glycyrrhetic acid binding to NP impairs its stability, triggers protein oligomerization and disrupts the association of viral RNA with oligomeric NP complexes. It would be interesting to test these molecules in replicon systems or in infected cells. Others studies identified aptamers and pyrrolidinone compounds, respectively, from *in silico* screening to minireplicon assays, potentially leading to efficient VP35-NP interaction inhibitors (Binning et al., 2013; Brown et al., 2014).

VP35-L interaction also is crucial for the virus life cycle because VP35 is necessary for the transcription and replication activities of L (Mühlberger et al., 1998). This interaction involves the N-terminal regions of VP35 and L (Trunschke et al., 2013). EBOV VP35 oligomerizes and interacts with L to form functional polymerase complexes and this could be exploited to develop antiviral drugs. Indeed, binding of L fragments to VP35 inhibits EBOV mini-genome replication in replicon systems, demonstrating the potential of this anti-filovirus strategy.

Finally, protein-protein interactions within the polymerase complex could also be targeted also by inhibiting VP30 oligomerization, a crucial step for viral transcription (Hartlieb et al., 2003). The efficacy of this approach has been evaluated *in vitro* using a synthetic peptide designed to mimic the VP30 oligomerization domain. This peptide inhibits viral replication in infected cells, suggesting that inhibition of VP30 oligomerization represents a valuable target for EBOV antiviral drugs.

4.5. Other strategies

An innovative strategy proposes to directly target structured RNA elements (Hermann, 2016). Indeed, replication of RNA viruses relies on structured RNA elements, notably the replication promoter with its large hairpin at the 5' end of the viral genome (see 3.2.) and the intergenic region between VP30 and VP24 that folds into a complex RNA structure. Compounds that promote misfolding of such RNA elements, or that induce specifically filovirus genome fragmentation might inhibit filovirus replication. A related strategy uses ranpirnase (Deshpande et al., 2016), a low molecular weight type III endoribonuclease isolated from leopard frog oocytes

(Hodge et al., 2016). Ranpirinase specifically targets dsRNA intermediates and induces their degradation, thus blocking the viral cycle. First, ranpirinase was used as a pro-apoptotic anti-cancer drug and large phase II/III trials showed its safety and effectiveness in different cancer types (Costanzi et al., 2005; Porta et al., 2008). More recent *in vitro* and *in vivo* studies suggested that it could work as an antiviral agent, particularly against EBOV (Deshpande et al., 2016). Additional studies, particularly in non-human primate EBOV models, are required.

The identification of broad-spectrum antiviral molecules that inhibit viral RNA synthesis could also be an effective strategy. Indeed, NNS viruses share a conserved replication mechanism that involves the protein L. For instance, the indole alkaloid-derivative CMLDBU3402 is effective against VSV and EBOV (Filone et al., 2013). Other indole alkaloid-type molecules also are active against HCV and malaria. This suggests that the indole alkaloid scaffold could be easily improved through medicinal chemistry derivatization. Another compound, the triazolotriazine-derivative triazavirin, has been described as a broad-spectrum antiviral drug against hemorrhagic fever viruses (Kiselev et al., 2015 RF Patent no. 2516936), but reference studies are still missing to support its efficacy in filoviruses.

Finally, other strategies are currently under development, although proof-of-concept studies are still missing or incomplete. For instance, the host innate immunity adaptation could be targeted with microRNA-based vaccines (Golkar et al., 2016), RNA interference is effective in infected cells (Fowler et al., 2005; Geisbert et al., 2006) and developed towards clinical trials (e.g. AVI morpholino-oligomers that target gene-specific viral RNA translation (Heald et al., 2014)) or the TKM small interfering RNAs targeting L and VP35 encoding sequences (Geisbert et al., 2006; Thi et al., 2015; Dunning et al., 2016) and the zinc-finger antiviral protein ZAP (Müller et al., 2007) decreases viral activity in EBOV- and MARV-infected cells. While it is a legitimate anti-filovirus therapy, development of siRNA against filoviruses, such as the TKM strategy, is unfortunately no longer relevant because of the occurrence of fever in recipients in Phase I trial (evidently because the RNA triggered innate responses).

5. Conclusion

Advances in filovirus protein structure/function characterization have allowed the development of antiviral compounds that block enzymes involved in transcription/replication or virus replication *in vitro* and/or *in vivo*. The development of pharmaceutical treatments, starting from these molecules, will be essential to treat infected patients and to restrict virus spread, as witnessed in the recent EBOV outbreak in West Africa. The broadening of our understanding of filovirus biology is severely limited by the fact that work on these viruses is restricted to maximum-containment laboratories (BSL-4). Nevertheless, the progressive accumulation of structural and functional data, as well as the availability of reverse genetic approaches and replicon systems should pave the way to better understanding of the filovirus replication cycle and the development of novel therapeutic compounds.

Acknowledgements

This work was supported by a DGA-MRIS and Aix-Marseille Université scholarship and by the ANR grant ANR-ST14-ASTR-0026 for the project VMTaseln.

References

Ammosova, T., Jerebtsova, M., Beullens, M., Lesage, B., Jackson, A., Kashanchi, F.,

- Southerland, W., Gordeuk, V.R., Bollen, M., Nekhai, S., 2005. Nuclear targeting of protein phosphatase-1 by HIV-1 Tat protein. *J. Biol. Chem.* 280, 36364–36371. <http://dx.doi.org/10.1074/jbc.M503673200>.
- Ammosova, T., Jerebtsova, M., Beullens, M., Voloshin, Y., Ray, P.E., Kumar, A., Bollen, M., Nekhai, S., 2003. Nuclear protein phosphatase-1 regulates HIV-1 transcription. *J. Biol. Chem.* 278, 32189–32194. <http://dx.doi.org/10.1074/jbc.M300521200>.
- Ammosova, T., Yedavalli, V.R.K., Niu, X., Jerebtsova, M., Van Eynde, A., Beullens, M., Bollen, M., Jeang, K.-T., Nekhai, S., 2011. Expression of a protein phosphatase 1 inhibitor, cdNIPP1, increases CDK9 threonine 186 phosphorylation and inhibits HIV-1 transcription. *J. Biol. Chem.* 286, 3798–3804. <http://dx.doi.org/10.1074/jbc.M110.196493>.
- Ascenzi, P., Bocedi, A., Heptonstall, J., Capobianchi, M.R., Di Caro, A., Mastrangelo, E., Bolognesi, M., Ippolito, G., 2008. Ebolavirus and marburgvirus: insight the Filoviridae family. *Mol. Asp. Med.* 29, 151–185. <http://dx.doi.org/10.1016/j.mam.2007.09.005>.
- Ayub, G., Waheed, Y., 2016. Sequence analysis of the L protein of the Ebola 2014 outbreak: insight into conserved regions and mutations. *Mol. Med. Rep.* 13, 4821–4826. <http://dx.doi.org/10.3892/mmr.2016.5145>.
- Baker, L.E., Ellena, J.F., Handing, K.B., Derewenda, U., Utepbergenov, D., Engel, D.A., Derewenda, Z.S., 2016. Molecular architecture of the nucleoprotein C-terminal domain from the Ebola and Marburg viruses. *Acta Crystallogr. Sect. Struct. Biol.* 72, 49–58. <http://dx.doi.org/10.1107/S2059798315021439>.
- Basler, C.F., 2015. Innate immune evasion by filoviruses. *Virology* 479–480, 122–130. <http://dx.doi.org/10.1016/j.virol.2015.03.030>.
- Basler, C.F., Amarasinghe, G.K., 2009. Evasion of interferon responses by Ebola and Marburg viruses. *J. Interferon Cytokine Res. Off. J. Int. Soc. Interferon Cytokine Res.* 29, 511–520. <http://dx.doi.org/10.1089/jir.2009.0076>.
- Becker, S., Huppertz, S., Klenk, H.D., Feldmann, H., 1994. The nucleoprotein of Marburg virus is phosphorylated. *J. Gen. Virol.* 75 (Pt 4), 809–818. <http://dx.doi.org/10.1099/0022-1317-75-4-809>.
- Becker, S., Mühlberger, E., 1999. Co- and posttranslational modifications and functions of Marburg virus proteins. *Curr. Top. Microbiol. Immunol.* 235, 23–34.
- Becker, S., Rinne, C., Hofsäss, U., Klenk, H.D., Mühlberger, E., 1998. Interactions of Marburg virus nucleocapsid proteins. *Virology* 249, 406–417. <http://dx.doi.org/10.1006/viro.1998.9328>.
- Beniac, D.R., Melito, P.L., Devarenes, S.L., Hiebert, S.L., Rabb, M.J., Lamboo, L.L., Jones, S.M., Booth, T.F., 2012. The organisation of Ebola virus reveals a capacity for extensive, modular polyploidy. *PLoS One* 7, e29608. <http://dx.doi.org/10.1371/journal.pone.0029608>.
- Bharat, T.A.M., Noda, T., Riches, J.D., Kraehling, V., Kolesnikova, L., Becker, S., Kawaoaka, Y., Briggs, J.A.G., 2012. Structural dissection of Ebola virus and its assembly determinants using cryo-electron tomography. *Proc. Natl. Acad. Sci. U. S. A.* 109, 4275–4280. <http://dx.doi.org/10.1073/pnas.1120453109>.
- Bharat, T.A.M., Riches, J.D., Kolesnikova, L., Welsch, S., Krähling, V., Davey, N., Parsy, M.-L., Becker, S., Briggs, J.A.G., 2011. Cryo-electron tomography of Marburg virus particles and their morphogenesis within infected cells. *PLoS Biol.* 9, e1001196. <http://dx.doi.org/10.1371/journal.pbio.1001196>.
- Biedenkopf, N., Hartlieb, B., Hoenen, T., Becker, S., 2013. Phosphorylation of Ebola virus VP30 influences the composition of the viral nucleocapsid complex: impact on viral transcription and replication. *J. Biol. Chem.* 288, 11165–11174. <http://dx.doi.org/10.1074/jbc.M113.461285>.
- Biedenkopf, N., Lier, C., Becker, S., 2016a. Dynamic phosphorylation of VP30 is essential for ebola virus life cycle. *J. Virol.* 90, 4914–4925. <http://dx.doi.org/10.1128/JVI.03257-15>.
- Biedenkopf, N., Schlereth, J., Grünweller, A., Becker, S., Hartmann, R.K., 2016b. RNA binding of ebola virus VP30 is essential for activating viral transcription. *J. Virol.* 90, 7481–7496. <http://dx.doi.org/10.1128/JVI.00271-16>.
- Binning, J.M., Wang, T., Luthra, P., Shabman, R.S., Borek, D.M., Liu, G., Xu, W., Leung, D.W., Basler, C.F., Amarasinghe, G.K., 2013. Development of RNA aptamers targeting Ebola virus VP35. *Biochem. Mosc.* 52, 8406–8419. <http://dx.doi.org/10.1021/bi400704d>.
- Boehmann, Y., Enterlein, S., Randolph, A., Mühlberger, E., 2005. A reconstituted replication and transcription system for Ebola virus Reston and comparison with Ebola virus Zaire. *Virology* 332, 406–417. <http://dx.doi.org/10.1016/j.virol.2004.11.018>.
- Bollen, M., Beullens, M., 2002. Signaling by protein phosphatases in the nucleus. *Trends Cell Biol.* 12, 138–145.
- Bollen, M., Peti, W., Ragusa, M.J., Beullens, M., 2010. The extended PP1 toolkit: designed to create specificity. *Trends Biochem. Sci.* 35, 450–458. <http://dx.doi.org/10.1016/j.tibs.2010.03.002>.
- Bouazza, N., Treluyer, J.-M., Foissac, F., Mentré, F., Taburet, A.-M., Guedj, J., Anglaret, X., de Lamballerie, X., Keita, S., Malvy, D., Frange, P., 2015. Favipiravir for children with ebola. *Lancet Lond. Engl.* 385, 603–604. [http://dx.doi.org/10.1016/S0140-6736\(15\)60232-X](http://dx.doi.org/10.1016/S0140-6736(15)60232-X).
- Bray, M., 2001. The role of the Type I interferon response in the resistance of mice to filovirus infection. *J. Gen. Virol.* 82, 1365–1373. <http://dx.doi.org/10.1099/0022-1317-82-6-1365>.
- Bray, M., Driscoll, J., Huggins, J.W., 2000. Treatment of lethal Ebola virus infection in mice with a single dose of an S-adenosyl-L-homocysteine hydrolase inhibitor. *Antivir. Res.* 45, 135–147.
- Bray, M., Raymond, J.L., Geisbert, T., Baker, R.O., 2002. 3-deazaneplanocin A induces massively increased interferon-alpha production in Ebola virus-infected mice. *Antivir. Res.* 55, 151–159.
- Brown, C.S., Lee, M.S., Leung, D.W., Wang, T., Xu, W., Luthra, P., Anantpadma, M.,

- Shabman, R.S., Melito, L.M., MacMillan, K.S., Borek, D.M., Otwonowski, Z., Ramanan, P., Stubbs, A.J., Peterson, D.S., Binning, J.M., Tonelli, M., Olson, M.A., Davey, R.A., Ready, J.M., Basler, C.F., Amarasinghe, G.K., 2014. *In silico* derived small molecules bind the filovirus VP35 protein and inhibit its polymerase cofactor activity. *J. Mol. Biol.* 426, 2045–2058. <http://dx.doi.org/10.1016/j.jmb.2014.01.010>.
- Bruhn, J.F., Kirchdoerfer, R.N., Urata, S.M., Li, S., Tickle, I.J., Bricogne, G., Saphire, E.O., 2017. Crystal structure of the marburg virus VP35 oligomerization domain. *J. Virol.* 91 <http://dx.doi.org/10.1128/JVI.01085-16>.
- Bukreyev, A.A., Chandran, K., Dolnik, O., Dye, J.M., Ebihara, H., Leroy, E.M., Mühlberger, E., Netesov, S.V., Patterson, J.L., Paweska, J.T., Saphire, E.O., Smither, S.J., Takada, A., Towner, J.S., Volchkov, V.E., Warren, T.K., Kuhn, J.H., 2014. Discussions and decisions of the 2012–2014 international committee on taxonomy of viruses (ICTV) Filoviridae study group, January 2012–June 2013. *Arch. Virol.* 159, 821–830. <http://dx.doi.org/10.1007/s00705-013-1846-9>.
- Calain, P., Monroe, M.C., Nichol, S.T., 1999. Ebola virus defective interfering particles and persistent infection. *Virology* 262, 114–128. <http://dx.doi.org/10.1006/viro.1999.9915>.
- Cárdenas, W.B., Loo, Y.-M., Gale, M., Hartman, A.L., Kimberlin, C.R., Martínez-Sobrido, L., Saphire, E.O., Basler, C.F., 2006. Ebola virus VP35 protein binds double-stranded RNA and inhibits alpha/beta interferon production induced by RIG-I signaling. *J. Virol.* 80, 5168–5178. <http://dx.doi.org/10.1128/JVI.02199-05>.
- Changula, K., Yoshida, R., Noyori, O., Marzi, A., Miyamoto, H., Ishijima, M., Yokoyama, A., Kajihara, M., Feldmann, H., Mweene, A.S., Takada, A., 2013. Mapping of conserved and species-specific antibody epitopes on the Ebola virus nucleoprotein. *Virus Res.* 176, 83–90. <http://dx.doi.org/10.1016/j.virusres.2013.05.004>.
- Clifton, M.C., Kirchdoerfer, R.N., Atkins, K., Abendroth, J., Raymond, A., Grice, R., Barnes, S., Moen, S., Lorimer, D., Edwards, T.E., Myler, P.J., Saphire, E.O., 2014. Structure of the Reston ebolavirus VP30 C-terminal domain. *Acta Crystallogr. Sect. F. Struct. Biol. Commun.* 70, 457–460. <http://dx.doi.org/10.1107/S2053230X14003811>.
- Conzelmann, K.K., 2004. Reverse genetics of mononegavirales. *Curr. Top. Microbiol. Immunol.* 283, 1–41.
- Costanzi, J., Sidransky, D., Navon, A., Goldsweig, H., 2005. Ribonucleases as a novel pro-apoptotic anticancer strategy: review of the preclinical and clinical data for ranpirnase. *Cancer Investig.* 23, 643–650. <http://dx.doi.org/10.1080/07357900500283143>.
- Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.-Y., Schneller, S., Züst, R., Dong, H., Thiel, V., Sen, G.C., Fensterl, V., Klimstra, W.B., Pierson, T.C., Buller, R.M., Gale, M., Shi, P.-Y., Diamond, M.S., 2010. 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* 468, 452–456. <http://dx.doi.org/10.1038/nature09489>.
- De Clercq, E., 1998. Carbocyclic adenosine analogues as S-adenosylhomocysteine hydrolase inhibitors and antiviral agents: recent advances. *Nucleosides Nucleotides* 17, 625–634.
- de Clercq, E., Montgomery, J.A., 1983. Broad-spectrum antiviral activity of the carbocyclic analog of 3-deazaadenosine. *Antivir. Res.* 3, 17–24.
- Deshpande, A., McMahon, B., Daughton, A.R., Abeyta, E.L., Hodge, D., Anderson, K., Pillai, S., 2016. Surveillance for emerging diseases with multiplexed point-of-care diagnostics. *Health Secur* 14, 111–121. <http://dx.doi.org/10.1089/hs.2016.0005>.
- Devarkar, S.C., Wang, C., Miller, M.T., Ramanathan, A., Jiang, F., Khan, A.G., Patel, S.S., 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. U. S. A.* 113, 596–601. <http://dx.doi.org/10.1073/pnas.1515152113>.
- DiCarlo, A., Biedenkopf, N., Hartlieb, B., Klussmeier, A., Becker, S., 2011. Phosphorylation of Marburg virus NP region II modulates viral RNA synthesis. *J. Infect. Dis.* 204 (Suppl. 13), S927–S933. <http://dx.doi.org/10.1093/infdis/jir319>.
- DiCarlo, A., Möller, P., Lander, A., Kolesnikova, L., Becker, S., 2007. Nucleocapsid formation and RNA synthesis of Marburg virus is dependent on two coiled coil motifs in the nucleoprotein. *Virol. J.* 4, 105. <http://dx.doi.org/10.1186/1743-422X-4-105>.
- Dong, S., Yang, P., Li, G., Liu, B., Wang, W., Liu, X., Xia, B., Yang, C., Lou, Z., Guo, Y., Rao, Z., 2015. Insight into the Ebola virus nucleocapsid assembly mechanism: crystal structure of Ebola virus nucleoprotein core domain at 1.8 Å resolution. *Protein Cell* 6, 351–362. <http://dx.doi.org/10.1007/s13238-015-0163-3>.
- Dörnemann, J., Burzio, C., Ronsse, A., Sprecher, A., De Clercq, H., Van Herp, M., Kolić, M.-C., Yosifiva, V., Caluwaerts, S., McElroy, A.K., Antierens, A., 2017. First newborn baby to receive experimental therapies survives ebola virus disease. *J. Infect. Dis.* <http://dx.doi.org/10.1093/infdis/jiw493>.
- Dunning, J., Sahr, F., Rojek, A., Gannon, F., Carson, G., Idriss, B., Massaquoi, T., Gandhi, R., Joseph, S., Osman, H.K., Brooks, T.J.G., Simpson, A.J.H., Goodfellow, I., Thorne, L., Arias, A., Merson, L., Castle, L., Howell-Jones, R., Pardinaz-Solis, R., Hope-Gill, B., Ferri, M., Grove, J., Kowalski, M., Stepniewska, K., Lang, T., Whitehead, J., Olliaro, P., Samai, M., Horby, P.W., RAPIDE-TKM trial team, 2016. Experimental Treatment of Ebola Virus Disease with TKM-130803: A Single-Arm Phase 2 Clinical Trial. *PLoS Med.* 13, e1001997. <http://dx.doi.org/10.1371/journal.pmed.1001997>.
- Dziubańska, P.J., Derewenda, U., Ellena, J.F., Engel, D.A., Derewenda, Z.S., 2014. The structure of the C-terminal domain of the Zaire ebolavirus nucleoprotein. *Acta Crystallogr. D. Biol. Crystallogr.* 70, 2420–2429. <http://dx.doi.org/10.1107/S1399004714014710>.
- Elliott, L.H., Kiley, M.P., McCormick, J.B., 1985. Descriptive analysis of Ebola virus proteins. *Virology* 147, 169–176.
- Enterlein, S., Volchkov, V., Weik, M., Kolesnikova, L., Volchkova, V., Klenk, H.-D., Mühlberger, E., 2006. Rescue of recombinant Marburg virus from cDNA is dependent on nucleocapsid protein VP30. *J. Virol.* 80, 1038–1043. <http://dx.doi.org/10.1128/JVI.80.2.1038-1043.2006>.
- Feagins, A.R., Basler, C.F., 2015. Lloviu virus VP24 and VP35 proteins function as innate immune antagonists in human and bat cells. *Virology* 485, 145–152. <http://dx.doi.org/10.1016/j.viro.2015.07.010>.
- Fearn, R., Collins, P.L., 1999. Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. *J. Virol.* 73, 5852–5864.
- Ferron, F., Decroly, E., Selisko, B., Canard, B., 2012. The viral RNA capping machinery as a target for antiviral drugs. *Antivir. Res.* 96, 21–31. <http://dx.doi.org/10.1016/j.antiviral.2012.07.007>.
- Ferron, F., Longhi, S., Henrissat, B., Canard, B., 2002. Viral RNA-polymerases – a predicted 2'-O-ribose methyltransferase domain shared by all Mononegavirales. *Trends Biochem. Sci.* 27, 222–224.
- Filone, C.M., Hodges, E.N., Honeyman, B., Bushkin, G.G., Boyd, K., Platt, A., Ni, F., Strom, K., Hensley, L., Snyder, J.K., Connor, J.H., 2013. Identification of a broad-spectrum inhibitor of viral RNA synthesis: validation of a prototype virus-based approach. *Chem. Biol.* 20, 424–433. <http://dx.doi.org/10.1016/j.chembiol.2013.02.011>.
- Fowler, T., Bamberg, S., Möller, P., Klenk, H.-D., Meyer, T.F., Becker, S., Rudel, T., 2005. Inhibition of Marburg virus protein expression and viral release by RNA interference. *J. Gen. Virol.* 86, 1181–1188. <http://dx.doi.org/10.1099/vir.0.80622-0>.
- Fu, X., Wang, Z., Li, L., Dong, S., Li, Z., Jiang, Z., Wang, Y., Shui, W., 2016. Novel chemical ligands to ebola virus and marburg virus nucleoproteins identified by combining affinity mass spectrometry and metabolomics approaches. *Sci. Rep.* 6, 29680. <http://dx.doi.org/10.1038/srep29680>.
- Fuentes, S.M., Sun, D., Schmitt, A.P., He, B., 2010. Phosphorylation of paramyxovirus phosphoprotein and its role in viral gene expression. *Future Microbiol.* 5, 9–13. <http://dx.doi.org/10.2217/fmb.09.93>.
- Geisbert, T.W., Hensley, L.E., Kagan, E., Yu, E.Z., Geisbert, J.B., Daddario-DiCaprio, K., Fritz, E.A., Jahrling, P.B., McClintock, K., Phelps, J.R., Lee, A.C.H., Judge, A., Jeffs, L.B., MacLachlan, I., 2006. Postexposure protection of Guinea pigs against a lethal ebola virus challenge is conferred by RNA interference. *J. Infect. Dis.* 193, 1650–1657. <http://dx.doi.org/10.1086/504267>.
- Gerritz, S.W., Cianci, C., Kim, S., Pearce, B.C., Deminie, C., Discotto, L., McAuliffe, B., Minassian, B.F., Shi, S., Zhu, S., Zhai, W., Pendri, A., Li, G., Poss, M.A., Edavettal, S., McDonnell, P.A., Lewis, H.A., Maskos, K., Mörtl, M., Kiefersauer, R., Steinbacher, S., Baldwin, E.T., Metzler, W., Bryson, J., Healy, M.D., Philip, T., Zockler, M., Schartman, R., Sinz, M., Leyva-Grado, V.H., Hoffmann, H.-H., Langley, D.R., Meanwell, N.A., Krystal, M., 2011. Inhibition of influenza virus replication via small molecules that induce the formation of higher-order nucleoprotein oligomers. *Proc. Natl. Acad. Sci. U. S. A.* 108, 15366–15371. <http://dx.doi.org/10.1073/pnas.1107906108>.
- Golkar, Z., Bhattar, R., Pace, D.G., Bagasra, O., 2016. Inhibition of ebola virus by anti-ebola miRNAs *in silico*. *J. Infect. Dev. Ctries.* 10, 626–634.
- Groseth, A., Charton, J.E., Sauerborn, M., Feldmann, F., Jones, S.M., Hoenen, T., Feldmann, H., 2009. The Ebola virus ribonucleoprotein complex: a novel VP30-L interaction identified. *Virus Res.* 140, 8–14. <http://dx.doi.org/10.1016/j.virusres.2008.10.017>.
- Groseth, A., Feldmann, H., Theriault, S., Mehmetoglu, G., Flick, R., 2005. RNA polymerase I-driven minigenome system for Ebola viruses. *J. Virol.* 79, 4425–4433. <http://dx.doi.org/10.1128/JVI.79.7.4425-4433.2005>.
- Hartlieb, B., Modrof, J., Mühlberger, E., Klenk, H.-D., Becker, S., 2003. Oligomerization of Ebola virus VP30 is essential for viral transcription and can be inhibited by a synthetic peptide. *J. Biol. Chem.* 278, 41830–41836. <http://dx.doi.org/10.1074/jbc.M307036200>.
- Hartlieb, B., Muziol, T., Weissenhorn, W., Becker, S., 2007. Crystal structure of the C-terminal domain of Ebola virus VP30 reveals a role in transcription and nucleocapsid association. *Proc. Natl. Acad. Sci. U. S. A.* 104, 624–629. <http://dx.doi.org/10.1073/pnas.0606730104>.
- Hartlieb, B., Weissenhorn, W., 2006. Filovirus assembly and budding. *Virology* 344, 64–70. <http://dx.doi.org/10.1016/j.viro.2005.09.018>.
- Heald, A.E., Iversen, P.L., Saoud, J.B., Sazani, P., Charleston, J.S., Axtelle, T., Wong, M., Smith, W.B., Vutikullird, A., Kaye, E., 2014. Safety and pharmacokinetic profiles of phosphorodiamidate morpholino oligomers with activity against ebola virus and marburg virus: results of two single-ascending-dose studies. *Antimicrob. Agents Chemother.* 58, 6639–6647. <http://dx.doi.org/10.1128/AAC.03442-14>.
- Hermann, T., 2016. Small molecules targeting viral RNA. *Wiley Interdiscip. Rev. RNA* 7, 726–743. <http://dx.doi.org/10.1002/wrna.1373>.
- Hodge, T., Draper, K., Basler, C.F., Freiberg, A., Squiquera, L., Sidransky, D., Sulley, J., Taxman, D.J., 2016. Antiviral effect of ranpirnase against Ebola virus. *Antivir. Res.* 132, 210–218. <http://dx.doi.org/10.1016/j.antiviral.2016.06.009>.
- Hoenen, T., Groseth, A., de Kok-Mercado, F., Kuhn, J.H., Wahl-Jensen, V., 2011. Minigenomes, transcription and replication competent virus-like particles and beyond: reverse genetics systems for filoviruses and other negative stranded hemorrhagic fever viruses. *Antivir. Res.* 91, 195–208. <http://dx.doi.org/10.1016/j.antiviral.2011.06.003>.
- Huang, Y., Xu, L., Sun, Y., Nabel, G.J., 2002. The assembly of Ebola virus nucleocapsid requires virion-associated proteins 35 and 24 and posttranslational modification of nucleoprotein. *Mol. Cell* 10, 307–316.
- Huggins, J., Zhang, Z.X., Bray, M., 1999. Antiviral drug therapy of filovirus infections: S-adenosylhomocysteine hydrolase inhibitors inhibit Ebola virus *in vitro* and in a lethal mouse model. *J. Infect. Dis.* 179 (Suppl. 1), S240–S247. <http://dx.doi.org/>

- 10.1086/514316.
- Huggins, J.W., 1989. Prospects for treatment of viral hemorrhagic fevers with ribavirin, a broad-spectrum antiviral drug. *Rev. Infect. Dis.* 11 (Suppl. 4), S750–S761.
- Hyde, J.L., Diamond, M.S., 2015. Innate immune restriction and antagonism of viral RNA lacking 2'-O methylation. *Virology* 479–480, 66–74. <http://dx.doi.org/10.1016/j.virol.2015.01.019>.
- Ignat'ev, G.M., Strel'tsova, M.A., Agafonov, A.P., Kashentseva, E.A., Prozorovskii, N.S., 1996. Experimental study of possible treatment of Marburg hemorrhagic fever with desferal, ribavirin, and homologous interferon. *Vopr. Virusol.* 41, 206–209.
- Ilinykh, P.A., Tigabu, B., Ivanov, A., Ammosova, T., Obukhov, Y., Garron, T., Kumari, N., Kovalsky, D., Platonov, M.O., Naumchik, V.S., Freiberg, A.N., Nekhai, S., Bukreyev, A., 2014. Role of protein phosphatase 1 in dephosphorylation of Ebola virus VP30 protein and its targeting for the inhibition of viral transcription. *J. Biol. Chem.* 289, 22723–22738. <http://dx.doi.org/10.1074/jbc.M114.575050>.
- Jacobs, M., Rodger, A., Bell, D.J., Bhagani, S., Cropley, I., Filipe, A., Gifford, R.J., Hopkins, S., Hughes, J., Jabeen, F., Johannessen, I., Karageorgopoulos, D., Lackenby, A., Lester, R., Liu, R.S.N., MacConnachie, A., Mahungu, T., Martin, D., Marshall, N., Mephram, S., Orton, R., Palmari, M., Patel, M., Perry, C., Peters, S.E., Porter, D., Ritchie, D., Ritchie, N.D., Seaton, R.A., Sreenu, V.B., Templeton, K., Warren, S., Wilkie, G.S., Zambon, M., Gopal, R., Thomson, E.C., 2016. Late Ebola virus relapse causing meningoencephalitis: a case report. *Lancet Lond. Engl.* 388, 498–503. [http://dx.doi.org/10.1016/S0140-6736\(16\)30386-5](http://dx.doi.org/10.1016/S0140-6736(16)30386-5).
- Jácome, R., Becerra, A., Ponce de León, S., Lazzano, A., 2015. Structural analysis of monomeric RNA-dependent polymerases: evolutionary and therapeutic implications. *PLoS One* 10, e0139001. <http://dx.doi.org/10.1371/journal.pone.0139001>.
- John, S.P., Wang, T., Steffen, S., Longhi, S., Schmaljohn, C.S., Jonsson, C.B., 2007. Ebola virus VP30 is an RNA binding protein. *J. Virol.* 81, 8967–8976. <http://dx.doi.org/10.1128/JVI.02523-06>.
- Kao, R.Y., Yang, D., Lau, L.-S., Tsui, W.H.W., Hu, L., Dai, J., Chan, M.-P., Chan, C.-M., Wang, P., Zheng, B.-J., Sun, J., Huang, J.-D., Madar, J., Chen, G., Chen, H., Guan, Y., Yuen, K.-Y., 2010. Identification of influenza A nucleoprotein as an antiviral target. *Nat. Biotechnol.* 28, 600–605. <http://dx.doi.org/10.1038/nbt.1638>.
- Kimberlin, C.R., Bornholdt, Z.A., Li, S., Woods, V.L., MacRae, I.J., Saphire, E.O., 2010. Ebola virus VP35 uses a bimodal strategy to bind dsRNA for innate immune suppression. *Proc. Natl. Acad. Sci. U. S. A.* 107, 314–319. <http://dx.doi.org/10.1073/pnas.0910547107>.
- Kirchdoerfer, R.N., Abelson, D.M., Li, S., Wood, M.R., Saphire, E.O., 2015. Assembly of the ebola virus nucleoprotein from a chaperoned VP35 complex. *Cell Rep.* 12, 140–149. <http://dx.doi.org/10.1016/j.celrep.2015.06.003>.
- Kirchdoerfer, R.N., Moyer, C.L., Abelson, D.M., Saphire, E.O., 2016. The ebola virus VP30-NP interaction is a regulator of viral RNA synthesis. *PLoS Pathog.* 12, e1005937. <http://dx.doi.org/10.1371/journal.ppat.1005937>.
- Kiselev, O.I., Vasin, A.V., Shevryyova, M.P., Deeva, E.G., Sivak, K.V., Egorov, V.V., Tsvetkov, V.B., Yu Egorov, A., Romanovskaya-Romanko, E.A., Stepanova, L.A., Komissarov, A.B., Tsybalova, L.M., Ignatjev, G.M., 2015. Ebola hemorrhagic fever: properties of the pathogen and development of vaccines and chemotherapeutic agents. *Mol. Biol.* 480–493. <http://dx.doi.org/10.1134/S002689331504007X>.
- Kumar, P., Sweeney, T.R., Skabkin, M.A., Skabkina, O.V., Hellen, C.U.T., Pestova, T.V., 2014. 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.* 42, 3228–3245. <http://dx.doi.org/10.1093/nar/gkt1321>.
- Leung, D.W., Borek, D., Luthra, P., Binning, J.M., Anantpadma, M., Liu, G., Harvey, I.B., Su, Z., Endlich-Frazier, A., Pan, J., Shabman, R.S., Chiu, W., Davey, R.A., Otwinowski, Z., Basler, C.F., Amarasinghe, G.K., 2015. An intrinsically disordered peptide from ebola virus VP35 controls viral RNA synthesis by modulating nucleoprotein-RNA interactions. *Cell Rep.* 11, 376–389. <http://dx.doi.org/10.1016/j.celrep.2015.03.034>.
- Leung, D.W., Prins, K.C., Basler, C.F., Amarasinghe, G.K., 2010. Ebola virus VP35 is a multifunctional virulence factor. *Virulence* 1, 526–531.
- Li, J., Wang, J.T., Whelan, S.P.J., 2006. A unique strategy for mRNA cap methylation used by vesicular stomatitis virus. *Proc. Natl. Acad. Sci. U. S. A.* 103, 8493–8498. <http://dx.doi.org/10.1073/pnas.0509821103>.
- Liang, B., Li, Z., Jenni, S., Rahmeh, A.A., Morin, B.M., Grant, T., Grigorieff, N., Harrison, S.C., Whelan, S.P.J., 2015. Structure of the L Protein of vesicular stomatitis virus from electron cryomicroscopy. *Cell* 162, 314–327. <http://dx.doi.org/10.1016/j.cell.2015.06.018>.
- Lötfering, B., Mühlberger, E., Tamura, T., Klenk, H.D., Becker, S., 1999. The nucleoprotein of Marburg virus is target for multiple cellular kinases. *Virology* 255, 50–62. <http://dx.doi.org/10.1006/viro.1998.9577>.
- Luthra, P., Jordan, D.S., Leung, D.W., Amarasinghe, G.K., Basler, C.F., 2015. Ebola virus VP35 interaction with dynein LC8 regulates viral RNA synthesis. *J. Virol.* 89, 5148–5153. <http://dx.doi.org/10.1128/JVI.03652-14>.
- Madelain, V., Nguyen, T.H.T., Olivo, A., de Lamballerie, X., Guedj, J., Taburet, A.-M., Mentré, F., 2016. Ebola virus infection: review of the pharmacokinetic and pharmacodynamic properties of drugs considered for testing in human efficacy trials. *Clin. Pharmacokinet.* 55, 907–923. <http://dx.doi.org/10.1007/s40262-015-0364-1>.
- Martin, B., Hoenen, T., Canard, B., Decroly, E., 2016. Filovirus proteins for antiviral drug discovery: a structure/function analysis of surface glycoproteins and virus entry. *Antivir. Res.* 135, 1–14. <http://dx.doi.org/10.1016/j.antiviral.2016.09.001>.
- Martínez, M.J., Biedenkopf, N., Volchkov, V., Hartlieb, B., Alazard-Dany, N., Reynard, O., Becker, S., Volchkov, V., 2008. Role of Ebola virus VP30 in transcription reinitiation. *J. Virol.* 82, 12569–12573. <http://dx.doi.org/10.1128/JVI.01395-08>.
- Mason, S.W., Aberg, E., Lawetz, C., DeLong, R., Whitehead, P., Liuzzi, M., 2003. Interaction between human respiratory syncytial virus (RSV) M2-1 and P proteins is required for reconstitution of M2-1-dependent RSV minigenome activity. *J. Virol.* 77, 10670–10676.
- Mittler, E., Kolesnikova, L., Herwig, A., Dolnik, O., Becker, S., 2013. Assembly of the Marburg virus envelope. *Cell. Microbiol.* 15, 270–284. <http://dx.doi.org/10.1111/cmi.12076>.
- Modrof, J., Becker, S., Mühlberger, E., 2003. Ebola virus transcription activator VP30 is a zinc-binding protein. *J. Virol.* 77, 3334–3338.
- Modrof, J., Möritz, C., Kolesnikova, L., Konakova, T., Hartlieb, B., Randolph, A., Mühlberger, E., Becker, S., 2001. Phosphorylation of Marburg virus VP30 at serines 40 and 42 is critical for its interaction with NP inclusions. *Virology* 287, 171–182. <http://dx.doi.org/10.1006/viro.2001.1027>.
- Modrof, J., Mühlberger, E., Klenk, H.-D., Becker, S., 2002. Phosphorylation of VP30 impairs ebola virus transcription. *J. Biol. Chem.* 277, 33099–33104. <http://dx.doi.org/10.1074/jbc.M203775200>.
- Möller, P., Pariente, N., Klenk, H.-D., Becker, S., 2005. Homo-oligomerization of Marburgvirus VP35 is essential for its function in replication and transcription. *J. Virol.* 79, 14876–14886. <http://dx.doi.org/10.1128/JVI.79.23.14876-14886.2005>.
- Mondal, A., Victor, K.G., Pudupakam, R.S., Lyons, C.E., Wertz, G.W., 2014. Newly identified phosphorylation site in the vesicular stomatitis virus P protein is required for viral RNA synthesis. *J. Virol.* 88, 1461–1472. <http://dx.doi.org/10.1128/JVI.02384-13>.
- Mora-Rillo, M., Arsuaga, M., Ramírez-Olivencia, G., de la Calle, F., Borobia, A.M., Sánchez-Seco, P., Lago, M., Figueira, J.C., Fernández-Puntero, B., Viejo, A., Negro, A., Nuñez, C., Flores, E., Carcas, A.J., Jiménez-Yuste, V., Lasala, F., García-de-Lorenzo, A., Arnalich, F., Arribas, J.R., La Paz-Carlos III University Hospital Isolation Unit, 2015. Acute respiratory distress syndrome after convalescent plasma use: treatment of a patient with Ebola virus disease contracted in Madrid, Spain. *Lancet Respir. Med.* 3, 554–562. [http://dx.doi.org/10.1016/S2213-2600\(15\)00180-0](http://dx.doi.org/10.1016/S2213-2600(15)00180-0).
- Mühlberger, E., Lötfering, B., Klenk, H.D., Becker, S., 1998. 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.* 72, 8756–8764.
- Mühlberger, E., Trommer, S., Funke, C., Volchkov, V., Klenk, H.D., Becker, S., 1996. Termini of all mRNA species of Marburg virus: sequence and secondary structure. *Virology* 223, 376–380. <http://dx.doi.org/10.1006/viro.1996.0490>.
- Mühlberger, E., Weik, M., Volchkov, V.E., Klenk, H.D., Becker, S., 1999. Comparison of the transcription and replication strategies of marburg virus and Ebola virus by using artificial replication systems. *J. Virol.* 73, 2333–2342.
- Müller, S., Möller, P., Bick, M.J., Wurr, S., Becker, S., Günther, S., Kümmerer, B.M., 2007. Inhibition of filovirus replication by the zinc finger antiviral protein. *J. Virol.* 81, 2391–2400. <http://dx.doi.org/10.1128/JVI.01601-06>.
- Negredo, A., Palacios, G., Vázquez-Morón, S., González, F., Dopazo, H., Molero, J., Juste, J., Quetglas, J., Savji, N., de la Cruz Martínez, M., Herrera, J.E., Pizarro, M., Hutchison, S.K., Echevarría, J.E., Lipkin, W.I., Tenorio, A., 2011. Discovery of an ebolavirus-like filovirus in Europe. *PLoS Pathog.* 7, e1002304. <http://dx.doi.org/10.1371/journal.ppat.1002304>.
- Noda, T., Aoyama, K., Sagara, H., Kida, H., Kawaoka, Y., 2005. Nucleocapsid-like structures of Ebola virus reconstructed using electron tomography. *J. Vet. Med. Sci.* 67, 325–328.
- Noda, T., Ebihara, H., Muramoto, Y., Fujii, K., Takada, A., Sagara, H., Kim, J.H., Kida, H., Feldmann, H., Kawaoka, Y., 2006. Assembly and budding of ebolavirus. *PLoS Pathog.* 2, e99. <http://dx.doi.org/10.1371/journal.ppat.0020099>.
- Noda, T., Hagiwara, K., Sagara, H., Kawaoka, Y., 2010. Characterization of the Ebola virus nucleoprotein-RNA complex. *J. Gen. Virol.* 91, 1478–1483. <http://dx.doi.org/10.1099/vir.0.019794-0>.
- Noton, S.L., Fearn, R., 2015. Initiation and regulation of paramyxovirus transcription and replication. *Virology* 479–480, 545–554. <http://dx.doi.org/10.1016/j.virol.2015.01.014>.
- Oestereich, L., Lüdtke, A., Wurr, S., Rieger, T., Muñoz-Fontela, C., Günther, S., 2014. Successful treatment of advanced Ebola virus infection with T-705 (favipiravir) in a small animal model. *Antivir. Res.* 105, 17–21. <http://dx.doi.org/10.1016/j.antiviral.2014.02.014>.
- Ogino, T., Banerjee, A.K., 2007. Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Mol. Cell* 25, 85–97. <http://dx.doi.org/10.1016/j.molcel.2006.11.013>.
- Paesen, G.C., Collet, A., Sallamand, C., Debart, F., Vasseur, J.-J., Canard, B., Decroly, E., Grimes, J.M., 2015. X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat. Commun.* 6, 8749. <http://dx.doi.org/10.1038/ncomms9749>.
- Palich, R., Gala, J.-L., Petitjean, F., Shepherd, S., Peyrouset, O., Abdoul, B.M., Kinda, M., Danel, C., Augier, A., Anglaret, X., Malvy, D., Blackwell, N., ALIMA N'zérékoré Ebola Treatment Center medical group, 2016. A 6-year-old child with severe ebola virus disease: laboratory-guided clinical care in an ebola treatment center in Guinea. *PLoS Negl. Trop. Dis.* 10, e0004393. <http://dx.doi.org/10.1371/journal.pntd.0004393>.
- Peng, R., Zhu, T., Oladejo, B.O., Musyoki, A.M., Cui, Y., Shi, Y., Wang, P., Gao, G.F., 2016. In vitro assembly of Ebola virus nucleocapsid-like complex expressed in *E. coli*. *Protein Cell.* <http://dx.doi.org/10.1007/s12328-016-0314-1>.
- Peyrol, J., Thizon, C., Gaillard, J.-C., Marchetti, C., Armengaud, J., Rollin-Genet, F.,

2013. Multiple phosphorylatable sites in the Zaire Ebola virus nucleoprotein evidenced by high resolution tandem mass spectrometry. *J. Virol. Methods* 187, 159–165. <http://dx.doi.org/10.1016/j.jviromet.2012.10.003>.
- Pickar, A., Xu, P., Elson, A., Li, Z., Zengel, J., He, B., 2014. Roles of serine and threonine residues of mumps virus P protein in viral transcription and replication. *J. Virol.* 88, 4414–4422. <http://dx.doi.org/10.1128/JVI.03673-13>.
- Pinto, A.K., Williams, G.D., Szretter, K.J., White, J.P., Proença-Módena, J.L., Liu, G., Olejnik, J., Brien, J.D., Ebihara, H., Mühlberger, E., Amarasinghe, G., Diamond, M.S., Boon, A.C.M., 2015. Human and murine IFIT1 proteins do not restrict infection of negative-sense RNA viruses of the orthomyxoviridae, bunyaviridae, and Filoviridae families. *J. Virol.* 89, 9465–9476. <http://dx.doi.org/10.1128/JVI.00996-15>.
- Poch, O., Blumberg, B.M., Bougueleret, L., 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. <http://dx.doi.org/10.1099/0022-1317-71-5-1153>.
- Porta, C., Paglino, C., Mutti, L., 2008. Ranpirinase and its potential for the treatment of unresectable malignant mesothelioma. *Biol. Targets Ther.* 2, 601–609.
- Prins, K.C., Binning, J.M., Shabman, R.S., Leung, D.W., Amarasinghe, G.K., Basler, C.F., 2010. Basic residues within the ebolavirus VP35 protein are required for its viral polymerase cofactor function. *J. Virol.* 84, 10581–10591. <http://dx.doi.org/10.1128/JVI.00925-10>.
- Reichelt, H., Hoffmann, G., Lederer, S., Kretschmar, K.H., 1989. Changes in the pathobiochemical reactions in traumatic-hemorrhagic shock following therapy with blood substitutes based on perfluorocarbon. *Z. Med. Lab. Diagn.* 30, 53–55.
- Reid, S.P., Cárdenas, W.B., Basler, C.F., 2005. Homo-oligomerization facilitates the interferon-antagonist activity of the ebolavirus VP35 protein. *Virology* 341, 179–189. <http://dx.doi.org/10.1016/j.virol.2005.06.044>.
- Reynard, O., Nguyen, X.-N., Alazard-Dany, N., Barateau, V., Cimarelli, A., Volchkov, V.E., 2015. Identification of a new ribonucleoside inhibitor of ebola virus replication. *Viruses* 7, 6233–6240. <http://dx.doi.org/10.3390/v7122934>.
- Rose, J.K., Lodish, H.F., Brock, M.L., 1977. Giant heterogeneous polyadenylic acid on vesicular stomatitis virus mRNA synthesized in vitro in the presence of S-adenosylhomocysteine. *J. Virol.* 21, 683–693.
- Saijo, M., Niikura, M., Morikawa, S., Ksiazek, T.G., Meyer, R.F., Peters, C.J., Kurane, I., 2001. Enzyme-linked immunosorbent assays for detection of antibodies to Ebola and Marburg viruses using recombinant nucleoproteins. *J. Clin. Microbiol.* 39, 1–7. <http://dx.doi.org/10.1128/JCM.39.1.1-7.2001>.
- Saikia, P., Gopinath, M., Shaïla, M.S., 2008. Phosphorylation status of the phosphoprotein P of rinderpest virus modulates transcription and replication of the genome. *Arch. Virol.* 153, 615–626. <http://dx.doi.org/10.1007/s00705-008-0034-9>.
- Sanchez, A., Kiley, M.P., Holloway, B.P., 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.
- Schibler, M., Vetter, P., Cherpillod, P., Petty, T.J., Cordey, S., Vieille, G., Yerly, S., Siegrist, C.-A., Samii, K., Dayer, J.-A., Docquier, M., Zdobnov, E.M., Simpson, A.J.H., Rees, P.S.C., Sarria, F.B., Gasche, Y., Chappuis, F., Iten, A., Pittet, D., Pugin, J., Kaiser, L., 2015. Clinical features and viral kinetics in a rapidly cured patient with Ebola virus disease: a case report. *Lancet Infect. Dis.* 15, 1034–1040. [http://dx.doi.org/10.1016/S1473-3099\(15\)00229-7](http://dx.doi.org/10.1016/S1473-3099(15)00229-7).
- Schlereth, J., Grünweller, A., Biedenkopf, N., Becker, S., Hartmann, R.K., 2016. RNA binding specificity of Ebola virus transcription factor VP30. *RNA Biol.* 13, 783–798. <http://dx.doi.org/10.1080/15476286.2016.1194160>.
- Schubert-Wagner, C., Ludwig, J., Bruder, A.K., Herzner, A.-M., Zillinger, T., Godeck, M., Schmidt, T., Schmid-Burgk, J.L., Kerber, R., Wolter, S., Stümpel, J.-P., Roth, A., Bartok, E., Drosten, C., Coch, C., Hornung, V., Barchel, V., Kümmerer, B.M., Hartmann, G., Schlee, M., 2015. A conserved histidine in the RNA sensor RIG-I controls immune tolerance to N1-2'-O-Methylated self RNA. *Immunity* 43, 41–51. <http://dx.doi.org/10.1016/j.immuni.2015.06.015>.
- Setlur, A.S., Naik, S.Y., Skariyachan, S., 2016. Herbal lead as ideal bioactive compounds against probable drug targets of ebola virus in comparison with known chemical analogue: a computational drug discovery perspective. *Interdiscip. Sci. Comput. Life Sci.* <http://dx.doi.org/10.1007/s12539-016-0149-8>.
- Sherwood, L.J., Hayhurst, A., 2013. Ebolavirus nucleoprotein C-termini potentially attract single domain antibodies enabling monoclonal affinity reagent sandwich assay (MARSa) formulation. *PLoS One* 8, e61232. <http://dx.doi.org/10.1371/journal.pone.0061232>.
- Silva, L.P., Vanzile, M., Bavari, S., Aman, J.M.J., Schriemer, D.C., 2012. Assembly of Ebola virus matrix protein VP40 is regulated by latch-like properties of N and C terminal tails. *PLoS One* 7, e39978. <http://dx.doi.org/10.1371/journal.pone.0039978>.
- Sissoko, D., Laouenan, C., Folkesson, E., M'Lebing, A.-B., Beavogui, A.-H., Baize, S., Camara, A.-M., Maes, P., Shepherd, S., Danel, C., Carazo, S., Conde, M.N., Gala, J.-L., Colin, G., Savini, H., Bore, J.A., Le Marais, F., Koundouno, F.R., Petitjean, F., Lamah, M.-C., Diederich, S., Tounkara, A., Poelart, G., Berbain, E., Dindart, J.-M., Duraffour, S., Lefevre, A., Leno, T., Peyrouset, O., Ireng, L., Bangoura, N., Palich, R., Hinzmann, J., Kraus, A., Barry, T.S., Berette, S., Bongono, A., Camara, M.S., Chanfreau Munoz, V., Doumbouya, L., Souley Harouna, Kighoma, P.M., Koundouno, F.R., René Lomou, Loua, C.M., Massala, V., Mousmouni, K., Provost, C., Samake, N., Sekou, C., Soumah, A., Arnould, I., Komano, M.S., Gustin, L., Berutto, C., Camara, D., Camara, F.S., Colpaert, J., Delamou, L., Jansson, L., Kourouma, E., Loua, M., Malme, K., Manfrin, E., Maomou, A., Milinouno, A., Ombelet, S., Sidiboun, A.Y., Verreck, I., Yombouno, P., Bocquain, A., Carbonnelle, C., Carmoi, T., Frange, P., Mely, S., Nguyen, V.-K., Pannetier, D., Taburet, A.-M., Treluyer, J.-M., Kolie, J., Moh, R., Gonzalez, M.C., Kuisma, E., Liedjck, B., Ngabo, D., Rudolf, M., Thom, R., Kerber, R., Gabriel, M., Di Caro, A., Wölfel, R., Badir, J., Bentahir, M., Deccache, Y., Dumont, C., Durant, J.-F., El Bakkouri, K., Gasasira Umamahoro, M., Smits, B., Toufik, N., Van Cauwenbergh, S., Ezzedine, K., D'Ortenzio, E., D'Ortenzio, E., Pizarro, L., Etienne, A., Guedj, J., Fizez, A., Barte de Sainte Fare, E., Murgue, B., Tran-Minh, T., Rapp, C., Pignat, P., Poncin, M., Draguez, B., Allaford Duverger, T., Barbe, S., Baretr, G., Defourny, I., Carroll, M., Raoul, H., Augier, A., Eholie, S.P., Yazdanpanah, Y., Levy-Marchal, C., Antierrens, A., Van Herp, M., Günther, S., de Lamballerie, X., Keïta, S., Mentre, F., Anglaret, X., Malvy, D., JIKI Study Group, 2016. Experimental treatment with favipiravir for ebola virus disease (the JIKI trial): a historically controlled, single-arm proof-of-concept trial in Guinea. *PLoS Med.* 13, e1001967. <http://dx.doi.org/10.1371/journal.pmed.1001967>.
- Sugai, A., Sato, H., Yoneda, M., Kai, C., 2012. Phosphorylation of measles virus phosphoprotein at S86 and/or S151 downregulates viral transcriptional activity. *FEBS Lett.* 586, 3900–3907. <http://dx.doi.org/10.1016/j.febslet.2012.09.021>.
- Sun, D., Luthra, P., Xu, P., Yoon, H., He, B., 2011. Identification of a phosphorylation site within the P protein important for mRNA transcription and growth of parainfluenza virus 5. *J. Virol.* 85, 8376–8385. <http://dx.doi.org/10.1128/JVI.00618-11>.
- Tanner, S.J., Ariza, A., Richard, C.-A., Kyle, H.F., Dods, R.L., Blondot, M.-L., Wu, W., Trincão, J., Trinh, C.H., Hiscox, J.A., Carroll, M.W., Silman, N.J., Eléouët, J.-F., Edwards, T.A., Barr, J.N., 2014. Crystal structure of the essential transcription antiterminator M2-1 protein of human respiratory syncytial virus and implications of its phosphorylation. *Proc. Natl. Acad. Sci. U. S. A.* 111, 1580–1585. <http://dx.doi.org/10.1073/pnas.1317262111>.
- Tawar, R.G., Duquerry, S., Vornrhein, C., Varela, P.F., Damier-Piolle, L., Castagné, N., MacLellan, K., Bedouelle, H., Bricogne, G., Bhella, D., Eléouët, J.-F., Rey, F.A., 2009. Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science* 326, 1279–1283. <http://dx.doi.org/10.1126/science.1177634>.
- Taylor, R., Kotian, P., Warren, T., Panchal, R., Bavari, S., Julander, J., Dobo, S., Rose, A., El-Kattan, Y., Taubenheim, B., Babu, Y., Sheridan, W.P., 2016. BCX4430-A broad-spectrum antiviral adenosine nucleoside analog under development for the treatment of Ebola virus disease. *J. Infect. Public Health* 9, 220–226. <http://dx.doi.org/10.1016/j.jiph.2016.04.002>.
- Thi, E.P., Mire, C.E., Lee, A.C.H., Geisbert, J.B., Zhou, J.Z., Agans, K.N., Snead, N.M., Deer, D.J., Barnard, T.R., Fenton, K.A., MacLachlan, I., Geisbert, T.W., 2015. Lipid nanoparticle siRNA treatment of Ebola-virus-Makona-infected nonhuman primates. *Nature* 521, 362–365. <http://dx.doi.org/10.1038/nature14442>.
- Trunschke, M., Conrad, D., Enterlein, S., Olejnik, J., Brauburger, K., Mühlberger, E., 2013. 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* 441, 135–145. <http://dx.doi.org/10.1016/j.virol.2013.03.013>.
- van Hemert, F.J., Zaaier, H.L., Berkhout, B., 2015. *In silico* prediction of ebolavirus RNA polymerase inhibition by specific combinations of approved nucleotide analogues. *J. Clin. Virol. Off. Publ. Pan Am. Soc. Clin. Virol.* 73, 89–94. <http://dx.doi.org/10.1016/j.jcv.2015.10.020>.
- Volchkov, V.E., Volchkova, V.A., Cherpurnov, A.A., Blinov, V.M., Dolnik, O., Netesov, S.V., Feldmann, H., 1999. Characterization of the L gene and 5' trailer region of Ebola virus. *J. Gen. Virol.* 80 (Pt 2), 355–362. <http://dx.doi.org/10.1099/0022-1317-80-2-355>.
- Warren, T.K., Jordan, R., Lo, M.K., Ray, A.S., Mackman, R.L., Soloveva, V., Siegel, D., Perron, M., Bannister, R., Hui, H.C., Larson, N., Strickley, R., Wells, J., Stuthman, K.S., Van Tongeren, S.A., Garza, N.L., Donnelly, G., Shurtleff, A.C., Retterer, C.J., Gharaibeh, D., Zamani, R., Kenny, T., Eaton, B.P., Grimes, E., Welch, L.S., Gomba, L., Wilhelmens, C.L., Nichols, D.K., Nuss, J.E., Nagle, E.R., Kugelman, J.R., Palacios, G., Doerffler, E., Neville, S., Carra, E., Clarke, M.O., Zhang, L., Lew, W., Ross, B., Wang, Q., Chun, K., Wolfe, L., Babusis, D., Park, Y., Stray, K.M., Trancheva, I., Feng, J.Y., Barauskas, O., Xu, Y., Wong, P., Braun, M.R., Flint, M., McMullan, L.K., Chen, S.-S., Fearn, R., Swaminathan, S., Mayers, D.L., Spiropoulos, C.F., Lee, W.A., Nichol, S.T., Chihlar, T., Bavari, S., 2016. Therapeutic efficacy of the small molecule GS-5734 against Ebola virus in rhesus monkeys. *Nature* 531, 381–385. <http://dx.doi.org/10.1038/nature17180>.
- Warren, T.K., Wells, J., Panchal, R.G., Stuthman, K.S., Garza, N.L., Van Tongeren, S.A., Dong, L., Retterer, C.J., Eaton, B.P., Pegoraro, G., Honnold, S., Bantia, S., Kotian, P., Chen, X., Taubenheim, B.R., Welch, L.S., Minning, D.M., Babu, Y.S., Sheridan, W.P., Bavari, S., 2014. Protection against filovirus diseases by a novel broad-spectrum nucleoside analogue BCX4430. *Nature* 508, 402–405. <http://dx.doi.org/10.1038/nature13027>.
- Watanabe, S., Noda, T., Kawaoka, Y., 2006. Functional mapping of the nucleoprotein of Ebola virus. *J. Virol.* 80, 3743–3751. <http://dx.doi.org/10.1128/JVI.80.8.3743-3751.2006>.
- Watt, A., Moukambi, F., Banadyga, L., Groseth, A., Callison, J., Herwig, H., Ebihara, H., Feldmann, H., Hoenen, T., 2014. A novel life cycle modeling system for Ebola virus shows a genome length-dependent role of VP24 in virus infectivity. *J. Virol.* 88, 10511–10524. <http://dx.doi.org/10.1128/JVI.01272-14>.
- Weik, M., Enterlein, S., Schlenz, K., Mühlberger, E., 2005. The Ebola virus genomic replication promoter is bipartite and follows the rule of six. *J. Virol.* 79, 10660–10671. <http://dx.doi.org/10.1128/JVI.79.16.10660-10671.2005>.
- Weik, M., Modrof, J., Klenk, H.-D., Becker, S., Mühlberger, E., 2002. Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *J. Virol.* 76, 8532–8539.
- Westover, J.B., Sefing, E.J., Bailey, K.W., Van Wettere, A.J., Jung, K.-H., Dagley, A., Wandersee, L., Downs, B., Smees, D.F., Furuta, Y., Bray, M., Gowen, B.B., 2016.

- Low-dose ribavirin potentiates the antiviral activity of favipiravir against hemorrhagic fever viruses. *Antivir. Res.* 126, 62–68. <http://dx.doi.org/10.1016/j.antiviral.2015.12.006>.
- Whelan, S.P.J., Barr, J.N., Wertz, G.W., 2004. Transcription and replication of non-segmented negative-strand RNA viruses. *Curr. Top. Microbiol. Immunol.* 283, 61–119.
- Wolfe, M.S., Borchardt, R.T., 1991. S-adenosyl-L-homocysteine hydrolase as a target for antiviral chemotherapy. *J. Med. Chem.* 34, 1521–1530.
- Yang, J., Koprowski, H., Dietzschold, B., Fu, Z.F., 1999. Phosphorylation of rabies virus nucleoprotein regulates viral RNA transcription and replication by modulating leader RNA encapsidation. *J. Virol.* 73, 1661–1664.
- Zhao, Z., Martin, C., Fan, R., Bourne, P.E., Xie, L., 2016. Drug repurposing to target Ebola virus replication and virulence using structural systems pharmacology. *BMC Bioinforma.* 17, 90. <http://dx.doi.org/10.1186/s12859-016-0941-9>.
- Zinzula, L., Esposito, F., Mühlberger, E., Trunschke, M., Conrad, D., Piano, D., Tramontano, E., 2009. Purification and functional characterization of the full length recombinant Ebola virus VP35 protein expressed in *E. coli*. *Protein Expr. Purif.* 66, 113–119. <http://dx.doi.org/10.1016/j.pep.2009.02.008>.
- Züst, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B.W., Ziebuhr, J., Szretter, K.J., Baker, S.C., Barchet, W., Diamond, M.S., Siddell, S.G., Ludwig, B., Thiel, V., 2011. Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat. Immunol.* 12, 137–143. <http://dx.doi.org/10.1038/ni.1979>.

CHAPITRE 4

**Protéines des filovirus et drug discovery :
Bases structurales et fonctionnelles de
l'assemblage et du bourgeonnement**

Résumé de la Revue 3

Titre : Filovirus proteins for antiviral drug discovery: Structure/function report of assembly and budding

Auteurs : Martin B, Volchkov V, Decroly E

Référence : submitted to Antiviral Research.

Cette troisième revue a pour but d'introduire les aspects structure/fonction des protéines virales impliquées dans l'assemblage et le bourgeonnement des nouveaux virions ainsi que les potentielles stratégies antivirales visant ces étapes tardives du cycle viral. Après une brève introduction replaçant les filovirus dans la classification des virus et leurs caractéristiques, les données structurales des protéines virales impliquées dans l'assemblage, VP40 et VP24, ont été détaillées ainsi que les propriétés fonctionnelles des différents domaines de ces protéines (**Figure 1**). Par la suite, nous avons décrit comment les protéines virales interagissent ensemble et recrutent des protéines cellulaires afin d'assurer la formation de nouvelles particules virales et permettre un nouveau cycle infectieux (**Figure 2 et 3**). Enfin, les avancées concernant le développement de nouvelles stratégies antivirales visant les étapes tardives d'assemblage ont été résumées (**Tableau 1**). Ces nouvelles stratégies antivirales sont, à ce jour, peu nombreuses et à des étapes précoces de développement car les mécanismes moléculaires de ces étapes tardives et les interactions mises en jeu pour le cycle viral des filovirus restent partiellement compris.

Revue 3



Filovirus proteins for antiviral drug discovery: Structure/function of proteins involved in assembly and budding

Baptiste Martin^a, Olivier Reynard^b, Viktor Volchkov^{b,**}, Etienne Decroly^{a,*}

^a Laboratoire Architecture et Fonction des Macromolécules Biologiques (AFMB), UMR7257 CNRS, Parc Scientifique de Luminy, Aix-Marseille Université, Marseille, France

^b Molecular Basis of Viral Pathogenicity, International Centre for Research in Infectiology (CIRI), INSERM U1111, CNRS UMR5308, Université Lyon 1, Ecole Normale Supérieure de Lyon, Lyon, France



ARTICLE INFO

Keywords:

Ebola virus
Filovirus
Antiviral therapy
Mononegavirales
Marburg virus
Virion assembly

ABSTRACT

There are no approved medications for the treatment of Marburg or Ebola virus infection. In two previous articles (Martin *et al.*, 2016, Martin *et al.*, 2017), we reviewed surface glycoprotein and replication proteins structure/function relationship to decipher the molecular mechanisms of filovirus life cycle and identify antiviral strategies. In the present article, we recapitulate knowledge about the viral proteins involved in filovirus assembly and budding. First we describe the structural data available for viral proteins associated with virus assembly and virion egress and then, we integrate the structural features of these proteins in the functional context of the viral replication cycle. Finally, we summarize recent advances in the development of innovative antiviral strategies to target filovirus assembly and egress. The development of such prophylactic or post-exposure treatments could help controlling future filovirus outbreaks.

1. Introduction

This review is the third of a series of three reviews that aim to recapitulate the structure/function relationship of filovirus proteins. We have previously published two papers reviewing aspects of the filovirus entry (Martin *et al.*, 2016) and replication cycle (Martin *et al.*, 2017), with the intent of identifying potential targets for antiviral intervention. This third article is now dealing with filovirus assembly and budding. After a short outline on taxonomy and genomic organization, we first described structures of proteins involved in filovirus assembly and budding. Then, we depicted an overview of molecular mechanisms driving nucleocapsid formation, virion assembly and egress. We finally proposed different strategies that might be interesting for the development of filovirus antiviral therapies.

2. Background: taxonomic classification and genomic organization

This viral family belongs to the order *Mononegavirales*, which includes enveloped, single strand, negative polarity RNA-containing viruses. Filoviruses are classified in three genera: *Marburgvirus*, *Ebolavirus* and *Cuevavirus* (Bukreyev *et al.*, 2014). The genus *Ebolavirus* includes five species: *Zaire ebolavirus* (Zaire virus – EBOV), *Sudan*

ebolavirus (Sudan virus – SUDV), *Tai Forest ebolavirus* (Tai Forest virus – TAFV), *Bundibugyo ebolavirus* (Bundibugyo virus – BDBV) and *Reston ebolavirus* (Reston virus – RESTV). The genus *Marburgvirus*, represented by the single species, *Marburg marburgvirus*, contains two separate viruses, Marburg virus (MARV) and Ravn virus (RAVV). The genus *Cuevavirus* also contains a single species, *Lloviu cuevavirus*, represented by Lloviu virus (LLOV). Noteworthy, sequencing and partial characterization of LLOV have been performed in the absence of replicating virus isolates (Negredo *et al.*, 2011). Unlike RESTV and most likely LLOV, other four ebolavirus species and marburgviruses are highly pathogenic for humans.

All filoviruses share a common genomic organization. Their genomes encode seven structural proteins: the nucleoprotein NP, the viral proteins VP35, VP40, VP30 and VP24, the surface glycoprotein GP, and the polymerase L (Feldmann *et al.*, 1992; Sanchez *et al.*, 1993; Negredo *et al.*, 2011). GP is responsible for virus attachment to the target cells and drives the membrane fusion process, leading to the release of the ribonucleoprotein (RNP) complexes in the cytosol (reviewed in Martin *et al.*, 2016). Viral RNP contains the RNA genome, which is tightly encapsidated by NP, and is associated with polymerase L and its co-factors VP35 and VP30. Recently, it has been shown that VP24 also constitutes viral RNP. Genome replication and transcription occurs in the host cell cytoplasm where inclusion bodies are formed later post

* Corresponding author.

** Corresponding author.

E-mail addresses: viktor.volchkov@inserm.fr (V. Volchkov), etienne.decroly@afmb.univ-mrs.fr (E. Decroly).

infection (reviewed in Martin et al., 2017). VP24 is required for viral nucleocapsids maturation and silencing of VP24 expression prevents the release of the virions (Bamberg et al., 2005; Mateo et al., 2011). VP40 orchestrates the virus assembly process by recruiting newly formed nucleocapsids to the membrane, where viral particles acquire their envelope and are released by budding at the cell surface.

Assembly and budding are the final steps in filovirus intracellular cycle and, therefore, represent good targets for the development of innovative antiviral therapies. In this third review, we focus on these last steps of filovirus life cycle. First, we resume the knowledge about the different viral proteins involved in assembly and budding. We then describe how they regulate each other and how viruses interact with the cellular machinery to drive newly formed nucleocapsids from perinuclear inclusion bodies to the plasma membrane where budding occurs. Finally, we summarize recent advances in the development of innovative antiviral strategies to target filovirus assembly and budding.

3. Structure/function of filovirus assembly proteins

Assembly and release of filoviral particles is a complex process requiring interactions between different viral and host cell proteins, and which altogether divert the host cell trafficking machinery to trigger virion budding (see 4.).

3.1. Features of the nucleocapsid-associated protein VP24

Among viruses of the order *Mononegavirales*, VP24 is unique to viruses of the family *Filoviridae* (Sanchez et al., 1993). Currently, it is accepted that VP24, previously thought to be a minor matrix protein, indeed functions as a nucleocapsid maturation factor and a transcription/replication modulator (Han et al., 2003; Noda et al., 2007; Watanabe et al., 2007; Mateo et al., 2011; Zhang et al., 2012; Watt et al., 2014). In the case of ebolaviruses, VP24 is also involved in an interferon signaling antagonism and plays a role in increasing virus pathogenicity during virus passaging in guinea pigs (Reid et al., 2006). In case of marburgviruses, VP24 hijacks antioxidative stress pathway by targeting nrf2-negative regulator keap1 (Page et al., 2014). The structures of MARV, SUDV and RESTV VP24 have been solved by X-ray crystallography (Zhang et al., 2012, 2014) (Fig. 1). Although VP24 amino acid sequences are barely conserved (30%), it is likely that its overall structural organization is conserved among filoviruses (Zhang et al., 2014). Its structure reveals a α/β single domain that is organized in a pyramidal shape with two neighboring concave pockets at the bottom of the pyramid that contains highly conserved residues (green and pink in Fig. 1). A first difference between MARV and SUDV concerns residues 201 to 217 that form a great β -sheet in MARV and a central small α -helix within a flexible loop encased between small β -strands in SUDV (Zhang et al., 2012, 2014). Additionally, it is thought that the N-terminal region (1–5) of VP24 regulates filovirus nucleocapsid assembly and oligomerization (Han et al., 2003; Bamberg et al., 2005; Noda et al., 2007) (see 4.1.). It is interesting to note that structural differences have also been detected within this N-terminal region. Indeed, in MARV, VP24 N-terminus is arranged as an extended flexible strand. Conversely, in ebolaviruses, it forms a rigid α -helix that extends from the apex of the VP24 pyramid shape. This apex binds to a conserved hydrophobic pocket along one of the pyramid faces. Finally, it should be noted that the proposed immunosuppressive domain of ebolavirus VP24, which is held in a short α -helix at the top of the pyramid (in blue in Fig. 1), is also found in the non-immunosuppressive MARV VP24 (not detailed in this review).

Additional investigations on structure comparison and structure/function characterization are needed to better understand how VP24 drives the filovirus assembly and budding. It would also be interesting to elucidate the structure/function bases of ebolavirus VP24 interferon activity inhibition to design potent innate immunity restorative molecules.

3.2. Structural insights on the matrix protein VP40

The viral protein VP40 is one of the most abundantly expressed viral proteins. This protein is considered as the “matrix” protein of filoviruses because it shares analogous features of other so called proteins in other RNA viruses (oligomerization and self-assembly, hydrophobicity and membrane binding, hijacking of cellular traffic) (Timmins et al., 2004). Most of the matrix proteins of other mononegaviruses are shorter and they do not show any significant homology with filovirus VP40, indicating major structural differences and thus a folding that is unique to filoviruses (Dessen et al., 2000). VP40 intracellular distribution varies during the viral life cycle (Kolesnikova et al., 2004a, b; Nanbo et al., 2013) because it orchestrates the distribution of other viral components and viral assembly (Mittler et al., 2013; Nanbo et al., 2013) (see 4.2.). Additionally, MARV VP40 carries an interferon antagonist domain, forming a ring-like structure (Timmins et al., 2003; Oda et al., 2015) that is absent in ebolaviruses.

Moreover, despite VP40 C-terminal domain is poorly conserved (16% identical) between ebolaviruses and marburgviruses, their VP40 show many structural and functional similarities (Bukreyev et al., 1993a, b; Licata et al., 2003; Timmins et al., 2003; Jasenosky and Kawaoka, 2004; Yamayoshi and Kawaoka, 2007; Liu et al., 2010). Indeed, the structural analysis of VP40 indicated that the protein is composed of two distinct β -sandwich domains (Dessen et al., 2000; Clifton et al., 2015; Oda et al., 2015) (Fig. 1). Conversely, the VP40 N-terminal domain shows a higher percentage of similarity (42% of sequence identity) with conserved fold, and is essential for oligomerization (Hoenen et al., 2010) and budding.

It is likely that VP40 can adopt different quaternary structures that confer its various roles during the viral cycle (Gomis-Rüth et al., 2003; Hoenen et al., 2010; Bornholdt et al., 2013; Radzimanowski et al., 2014). The predominantly dimeric form acts as a building block via the N-terminal residues 52–65 and 108–117 (Bornholdt et al., 2013). This dimer form allows contacts with VP40 filaments, which are linear oligomers that interact with the plasma membrane (see 4.2.1.). VP40 can also form octameric rings structures that can bind to RNA (Gomis-Rüth et al., 2003). This structural arrangement is crucial for transcription regulation (Hoenen et al., 2005; Bornholdt et al., 2013). The N-terminal domain of ebolavirus VP40 contains also conserved overlapping tetrapeptide motifs (PTAP and PPXY) that are called late domains (Irie et al., 2005). These domains drive the assembly and budding mechanism by recruiting host cell components that are involved in multi-vesicular body biogenesis (MVB) (Harty et al., 2000; Licata et al., 2003) (see 4.2.1.). In the case of MARV, active late domains are split between NP and VP40 (Dolnik et al., 2010).

VP40 C-terminal domain, which is less conserved, contains domains that regulate its intracellular trafficking and membrane binding (Scianimanico et al., 2000; Adu-Gyamfi et al., 2013; Soni et al., 2013) (Fig. 1). It carries hydrophobic residues for anchoring to the inner leaflet monolayer of the plasma membrane. Upon membrane binding, VP40 undergoes a conformational change that favors its polymerization through interaction between the N-terminus of different VP40 molecules (see 3.3.) (Ruigrok et al., 2000). Also, in VP40 C-terminal domain, there is a patch of cationic residues that should interact with the anionic polar head of phospholipids present at the plasma membrane. This interaction is strengthened by an N-terminal loop adjacent to the C-terminal anchor for efficient VP40-mediated filovirus assembly, budding and egress (Adu-Gyamfi et al., 2014).

VP40 is a phosphorylated protein, like NP and VP30. Several phosphorylated tyrosine residues have been identified in its N-terminal domain (Kolesnikova et al., 2012) and in EBOV late domain motif PPXY (Kolesnikova et al., 2012). Therefore, VP40 phosphorylation could play a critical regulatory role in viral budding; especially because it significantly affects RNP complex recruitment to the plasma membrane (Kolesnikova et al., 2012) (see 4.2.4.).

The multi-functionality could be regulated by VP40 C- and N-tails

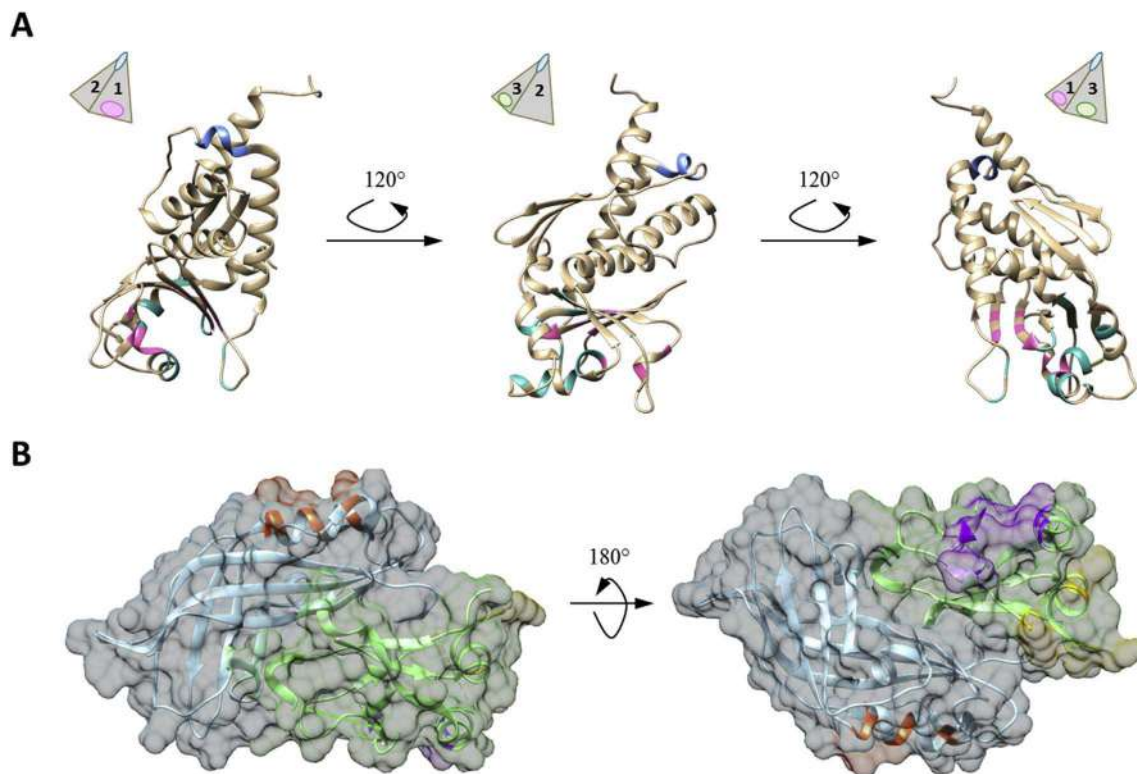


Fig. 1. Structural insights into the viral proteins VP24 and VP40: (A) Based on EBOV structure obtained by X-ray crystallography (PDB: 4M0Q), VP24 structural features were mapped and schematized according to its pyramidal shape. Two conserved pockets (a hydrophobic cavity, in green, and a polar cavity, in pink) are located in two different faces of the pyramid base, but no role has been clearly assigned to these regions. The small α -helix at the pyramid apex (in blue) is involved in immune suppression in ebolaviruses, but not in MARV. The oligomerization and nucleocapsid formation domains are included in the N- and C-terminal domains that have not been resolved in the available structures. (B) The Sudan virus VP40 structure (PDB: 4LD8) revealed an N-terminal domain (in light blue) and a C-terminal domain (in green). Functional domains also have been mapped: the N-terminal dimerization domain (in orange), the C-terminal oligomerization domain (in yellow) and the basic patch for membrane interaction (in purple). Late domains for hijacking the host cell transport system and the basic loop that strengthens membrane interaction are located in the unresolved parts of VP40 structure.

that appear to lock VP40 in a dimer state until the recognition of still unidentified molecular triggers (Silva et al., 2012; Wijesinghe et al., 2017). This suggests a mechanism by which late stages of viral assembly might be coordinated with viral packaging and release.

3.3. Other assembly-related proteins

Whereas VP40 and VP24 coordinate assembly and budding by recruiting the host multivesicular body machinery to drive viral particle formation, GP and NP also play important roles during filovirus assembly.

GP structure and function have been reviewed previously (Martin et al., 2016). Briefly, GP, the only surface protein of filoviruses, is a homotrimeric assemblage of GP_{1,2} heterodimeric complexes (Lee et al., 2008). Surface GP_{1,2} consist of two disulfide-linked furin cleavage products GP₁ and GP₂ (Volchikov et al., 1998b). GP₁ is composed of two domains: an ectodomain that carries a highly glycosylated mucin-like domain, and a core that bears the receptor-binding site. GP₂ is the membrane-anchoring moiety with an extracellular domain containing fusion loop and a short cytoplasmic tail. This surface protein is responsible for virus entry, but also participate in virus egress (Gustin et al., 2015). GP is indeed the second budding force of particles as it helps budding of VP40 VLPs and is also able to bud alone (Noda et al., 2002).

NP structure/function aspects have also been reviewed recently (Martin et al., 2017). It is the major component of the RNP complex and drives viral RNA encapsidation. Two separate domains can be distinguished within NP: the N-terminal domain harbors the oligomerization and packaging functions, essential for RNA encapsidation and replication regulation (DiCarlo et al., 2007; Dong et al., 2015; Peng

et al., 2016); whereas the C-terminal domain is considered to be a hub for protein-protein interactions, notably with VP40 (Dziubańska et al., 2014), and may be phosphorylated for modulating its RNA affinity (Becker et al., 1994; Lötfering et al., 1999).

4. Filovirus assembly and budding

4.1. Nucleocapsid formation

The nucleocapsid is a RNP complex made of five viral proteins (NP, VP35, VP30, VP24 and L) and the encapsidated RNA genome. As soon as it is synthesized, RNA is encapsidated by NP. NP affinity for RNA could be regulated by a phosphorylation/dephosphorylation cycle (Becker et al., 1994; DiCarlo et al., 2011; Peyrol et al., 2013). VP35, VP30 and L are also associated to the RNP complex to initiate primary transcription when the complex is released in the cytoplasm of newly infected cells after entry. NP, VP35, and VP24 are necessary and sufficient for the formation of nucleocapsid-like structures in the presence of RNA (Huang et al., 2002). VP24 appears to play a key role for RNP complex assembly, ready to be active for new infection (Watanabe et al., 2004; Mateo et al., 2011). Indeed, it was demonstrated that the increase of intracellular VP24 concentration is a key factor regulating the virus assembly because it favors the formation of condensed RNP complexes (Fig. 2). In this RNP complex type, the polymerase is locked at the 3' end of the genome, thus inhibiting replication and transcription, but allowing the right arrangement to trigger transcription after nucleocapsid uncoating (Watt et al., 2014). VP24 is definitely an original and central protein that orchestrates nucleocapsid formation, thus conditioning the infectivity of future viral particles. Therefore, this viral protein is a very important target for the development of anti-filovirus

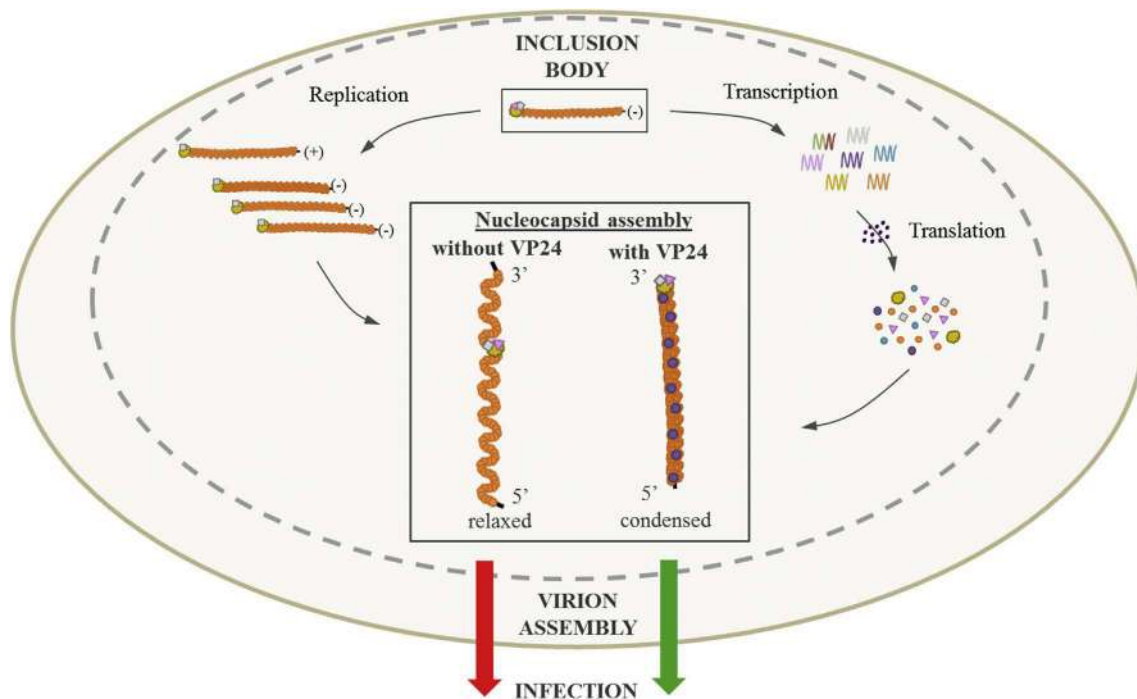


Fig. 2. VP24 contribution to nucleocapsid assembly: Relaxed ribonucleoprotein complexes are used as templates for filovirus replication and transcription/translation within inclusion bodies. When enough genomic copies and proteins have been generated, new nucleocapsids are formed. This requires VP24 (in blue) that leads to nucleocapsid condensation and locks the polymerase at the 3' end of the genome, thus blocking replication and transcription. Condensed nucleocapsids are efficiently assembled and allow new infections with competent primary transcription. If VP24 is absent, nucleocapsid stays in a relaxed arrangement, making assembly and infection less efficient, particularly because the L protein is not locked in the right conformation.

molecules.

4.2. Virion assembly

4.2.1. VP40 orchestration

VP40 play an important role in filovirus assembly and egress but is also involved in the regulation of viral transcription (Bornholdt et al., 2013) thus representing a unique target to control virus replication cycle.

First, VP40 hijacks the host Endosomal Sorting Complex Required for Transport (ESCRT) pathway, which is normally involved in protein trafficking, cytokinesis regulation, and endosomes recycling (McDonald and Martin-Serrano, 2009; Li et al., 2012; Votteler and Sundquist, 2013) through interactions mediated by its N-terminal late domains (Fig. 3). In EBOV, after oligomerization, the N-terminal PPEY motif recruits the NEDD4 ubiquitin ligase by interacting with its WW domains (Licata et al., 2003; Yasuda et al., 2003). This interaction induces VP40 ubiquitination that is essential for viral assembly and egress (Irie et al., 2005; Neumann et al., 2005). Then, ubiquitinated VP40 interacts with TSG101 via its other late domain PTAP that recruits the ESCRT-I complex and probably downstream partners, such as ESCRT II and III with an ATP-dependent release via CHMP/VPS4 proteins, similarly to what observed for other enveloped viruses (e.g., HIV) (Licata et al., 2003; Martin-Serrano et al., 2003; Strack et al., 2003; von Schwedler et al., 2003). However, VP40 PTAP late domain mutants can still recruit proteins involved downstream of the ESCRT pathway (VPS4, VPS28 and VPS37B) (Silvestri et al., 2007) and late domain mutations in VP40 (or NP for MARV) do not completely inhibit viral particle budding (Urata et al., 2007; Dolnik et al., 2010). Therefore, other interactions possibly play key roles in assembly and budding like the COPII system in which Sec24C interacts with VP40 residues 303–307 to trigger its intracellular traffic to the membrane (Yamayoshi et al., 2008).

Furthermore, VP40 phosphorylation is critical for filovirus assembly (Fig. 3). Mutations of MARV VP40 phosphorylatable residues impair VP40 recruitment of nucleocapsids to the plasma membrane, but do not

alter other VP40 functions (e.g., membrane association, oligomerization, virus-like particle formation and budding) (Kolesnikova et al., 2012). Additionally, non-phosphorylated MARV VP40 is localized within perinuclear inclusion bodies, while phosphorylated VP40 is membrane-associated (Mittler et al., 2013). Therefore, VP40 phosphorylation could regulate nucleocapsid transport. To date, it has been demonstrated that the kinase ABL1 catalyzes this phosphorylation for ebolaviruses (García et al., 2012) (see 5.); however, it is not known where VP40 phosphorylation occurs. While MARV VP40 phosphorylation seems to be better established than for EBOV, its actual role still has to be clarified. Interestingly, phosphorylation of the matrix protein is not conserved in all mononegaviruses. Indeed, matrix protein phosphorylation has been observed in vesicular stomatitis virus, Sendai virus and respiratory syncytial virus (Lamb and Choppin, 1977; Lambert et al., 1988; Kaptur et al., 1992), but not in mumps virus and Newcastle disease virus (Rima et al., 1980; Chambers and Samson, 1980). Therefore, the functional significance of this phosphorylation remains unclear even for most viruses of the order *Mononegavirales* (Kaptur et al., 1992, 1995).

4.2.2. GP processing

Synthesis and maturation of filoviral surface glycoprotein GP involve a complex sequence of co- and post-translational events. Two precursor molecules and proteolytically processed mature GP_{1,2} have been identified in GP expressing cells (Volchkov et al., 1998a, 2000). First GP precursor, preGPer, represents an ER-associated form, which forms trimers and contains oligomannosidic N-glycans. The second precursor, preGP, represents a Golgi-associated form, which acquires mature carbohydrates. Shortly after appearance, preGP is processed into mature GP_{1,2} consisting of two disulfide-linked cleavage products, the amino-terminal fragment GP₁ and the carboxyl-terminal fragment GP₂ (Volchkov et al., 1998a, 2000). All filoviral surface glycoproteins are cleaved during processing at the site R-X-K/R-R that is recognized by cellular proprotein convertases, in particular by serine proteases with subtilisin/kexin catalytic domains. The site is located in close

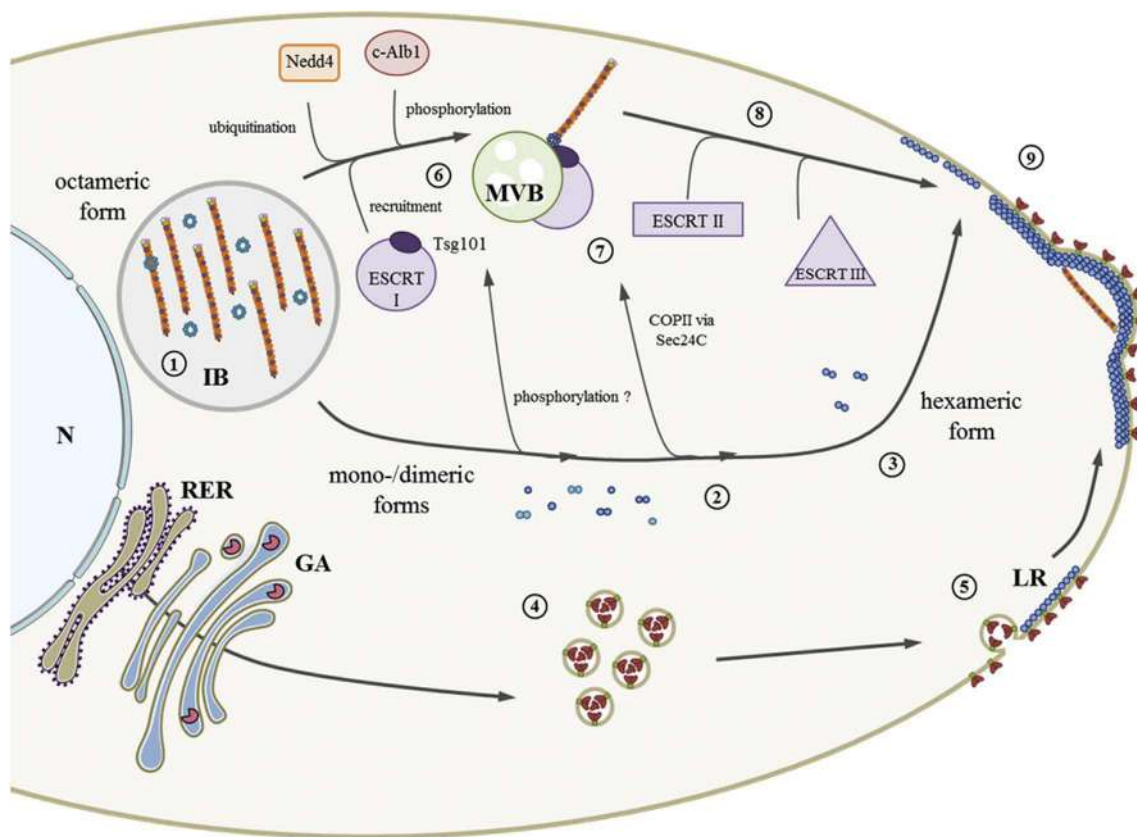


Fig. 3. VP40 orchestration of filovirus assembly: After transcription/replication in inclusion bodies (IB) (1), VP40 is synthesized (light blue) and recruited to the plasma membrane as phosphorylated dimers (dark blue) that assemble into hexamers (2) and longer oligomers (3). Concomitantly, GP is processed through the rough endoplasmic reticulum (RER) and the Golgi apparatus (GA) where it is cleaved in GP_{1,2} by furin-like proteases (pink). VP40 octamers interact with nucleocapsids. Other factors modify VP40 (6): NEDD4 ubiquitinates and ABL1 phosphorylates VP40. VP40 also interacts with the ESCRT I export system via TSG101 to allow nucleocapsid transport using the multivesicular body trafficking (MVB) mechanism (7). Nucleocapsids are then transferred to the membrane by ESCRT II and III (8). VP40 polymers, GP_{1,2} and nucleocapsids are finally assembled together to form new viral particles that bud from the plasma membrane (9).

proximity to a domain involved in fusion and cleavage is believed to facilitate processes requiring exposure of fusion loop during virus entry. Finally, GP_{1,2} is transported to the plasma membrane where it localizes at cholesterol-enriched microdomains, enriched in VP40 (Bavari et al., 2002; Noda et al., 2002; Mittler et al., 2007). GP₁ is responsible for the virus attachment and receptor recognition (Mittler et al., 2007, 2011; Lee and Saphire, 2009; Mittler et al., 2011), whereas GP₂ ensures membrane anchoring and catalyzes membrane fusion and viral entry. Noteworthy, GP₂ subunit bears structural and sequence similarity with the retroviral fusion proteins (Volchkov et al., 1992; A. Bukreyev et al., 1993a, b). Remarkably, short cytoplasmic tail of GP₂ contains conserved among filoviruses two C-terminal acylated cysteine residues (Funke et al., 1995; Ito et al., 2001). Their role in virus assembly remains largely unknown. They thought to be involved in the association of GP with lipid-raft microdomains where budding of virus particles occurs. Mutation of these cysteine residues in EBOV GP resulted in failure of GP localization to the rafts (Bavari et al., 2002). However, it also appears that removal of these cysteine residues in the transmembrane region of GP showed no discernible effects on virus infectivity (Ito et al., 2001). When GP_{1,2} is expressed in the absence of VP40, it will be transported through the cellular secretory pathway and localizes at the plasma membrane (Volchkov et al., 1998; Sanger et al., 2001; Kolesnikova et al., 2004b). Remarkably co-expression of EBOV GP and VP40 resulted in dramatic increase in release of virus-like particles which indeed contain the raft marker GM1 and display the morphological characteristics of EBOV virions (Bavari et al., 2002; Noda et al., 2002; Mittler et al., 2007). They are thought to be involved in a domain that putatively interacts with VP40. Indeed, Thus, it is likely that GP_{1,2} could accumulate in VP40-enriched microdomains thus facilitating its

integration in the viral envelope. Nevertheless, the mechanism governing the interaction between the GP_{1,2} transmembrane domain and VP40 is still enigmatic (Reynard et al., 2011).

4.2.3. NP role

NP encapsidates genomic RNA and recruits VP30, VP35 and L to inclusion bodies to form new nucleocapsids, together with VP24. NP promotion of budding activity suggests that it also interacts with VP40 (Licata et al., 2004; Noda et al., 2006). Actually, it has been hypothesized that an octameric form of VP40 takes up RNP complexes (Kolesnikova et al., 2004a, 2004b) and carries them from inclusion bodies formed by accumulation of new RNP complexes to assembly and budding sites (Fig. 3). These interactions are regulated via NP and VP40 phosphorylation. In addition, it is likely that VP40 interaction with the cytoskeleton also plays a role in nucleocapsid transport (see 4.2.4.).

4.2.4. RNP complex trafficking

In the host cell, protein trafficking is mediated by the cytoskeleton. Upon a trigger by VP40, transport-competent nucleocapsids are transported towards the plasma membrane along the cytoskeleton. This indicates that VP40 requires active transport through the retrograde late endosomal pathway (Kolesnikova et al., 2002, 2004a, 2004b). It is not clear whether microtubules, actin or both are involved, but filoviruses can hijack the multi-vesicular body trafficking mechanism, notably through the use of the cytoskeleton network.

To date, there is a debate on how filovirus RNP complexes are brought to the membrane. Evidence tends to support microtubules or actin-dependent RNP trafficking. Microtubules have been involved in filovirus budding by electron microscopy studies in presence of

inhibitors of microtubule assembly (Noda et al., 2006). Furthermore, VP40 directly binds to microtubules via a tubulin binding-like domain and this stimulates tubulin polymerization and stabilization (Ruthel et al., 2005). On the other hand, recent live cell imaging studies showed that nucleocapsids are transported to the plasma membrane along actin fibers (Adu-Gyamfi et al., 2012; Schudt et al., 2013, 2015). In MARV, Tsg101 is co-transported with nucleocapsids and recruits IQGAP1, an actin cytoskeleton regulator that impairs MARV release when down-regulated (Lu et al., 2013; Dolnik et al., 2014). Even these results suggest a link between microtubules and actin fibers and filovirus RNP complex the MVB trafficking machinery, the molecular basis of this interaction remains to be clarified. Nonetheless, it is likely that VP40 phosphorylation state is related to its interactions with the cytoskeleton because it is linked to nucleocapsid transport to budding sites (see 4.2.1.).

4.3. Budding

Budding sites are organized by VP40 and GP_{1,2} at the plasma membrane, specifically at the basolateral side of polarized cells (Sänger et al., 2001) but for endothelial cells where budding occurs at the apical membrane (Schnittler et al., 1993) and in filopodia of mobile cells (Schudt et al., 2013, 2015).

VP40 drives the formation of these specific sites but, in its absence, GP_{1,2} still appears to be concentrated in lipid rafts (Bavari et al., 2002; Noda et al., 2002; Mittler et al., 2007). As mentioned in 3.2., VP40 C-terminal domain bears sequence features for membrane binding that are critical for virus release (Fig. 3). EBOV VP40 dimers are anchored thanks to hydrophobic interactions more than halfway through the phosphatidyl serine-enriched inner leaflet monolayer of the plasma membrane (Adu-Gyamfi et al., 2013; Soni et al., 2013). Dimers can then assemble into hexamers. Rearrangements in the hexamers might displace the C-terminal domains from the N-terminal domains, revealing the oligomerization sites. Therefore, these rearranged hexamers assemble in long filamentous structures that form the matrix layer and facilitate budding (Bornholdt et al., 2013). Differently from EBOV, MARV VP40 dimers associate with phosphatidyl serine-containing membranes through its C-terminal basic patch (Oda et al., 2015). Following this association, MARV VP40 oligomerizes through an N-terminal interface, but possibly also via C-terminal end-to-end contacts, like EBOV (Bornholdt et al., 2013; Wijesinghe et al., 2017). This emphasizes that modest VP40 sequence identity (mentioned in 3.2.) can lead to different functions, such as lipid binding properties.

Recently, it has been shown that VP40 lipid binding properties are essential for virus assembly and egress. In EBOV, VP40 interacts with the hydrocarbon core of lipid membranes that contains phosphatidyl serine residues (Adu-Gyamfi et al., 2013; Soni et al., 2013) required for normal VP40 plasma membrane localization and oligomerization and also for virus egress (Adu-Gyamfi et al., 2015). VP40 interaction with other lipids, such as phosphatidylinositol-4,5-bisphosphate, phosphatidylinositol-3-phosphate and phosphatidylinositol-3,4,5-trisphosphate (Johnson et al., 2016), allows the formation of oligomers. Indeed, phosphatidyl serine residues are necessary for VP40 dimer to 12-mer oligomer formation and phosphatidylinositol derivatives for 12-mer to larger oligomer assemblage. Recently, it was reported that VP40 filaments are structurally similar to Bin Amphiphysin Rvs (BAR) domains (Mim and Unger, 2012). These domains are present in proteins involved in membrane dynamics that mediate membrane curvature changes.

Finally, a model for virus release suggests that new viral particles assemble through lateral association of nucleocapsids with the plasma membrane (Noda et al., 2006; Welsch et al., 2010). Then, nucleocapsids wrapped by the plasma membrane are pushed out in protrusions containing one or multiple nucleocapsid, which form single or multi-genome particles at the cell surface (Beniac et al., 2012).

5. Antiviral approaches to target assembly and budding

The incomplete knowledge on filovirus assembly and budding has impaired the design of small inhibitors. For instance, only the proof of concept of VP24 interest as therapeutic target has been published using phosphorodiamidate morpholino oligomers *in vivo* that specifically inhibit VP24 synthesis (Warren et al., 2015). However, several computational studies and few *in vitro* investigations on VP40 have paved the way to the development of anti-VP40 molecules (Stahelin, 2014).

Broad-spectrum strategies could potentially be developed to target the key cellular mechanisms involved in filovirus maturation, assembly and budding. Potential targets are ER α -glucosidases and the cytoskeleton. ER α -glucosidases are essential for the maturation of glycosylated envelop proteins, like GP. The imino sugar 1-deoxynojirimycin and its derivatives, which mimic glucose with a nitrogen atom replacing the sugar ring oxygen, competitively inhibit ER α -glucosidases (Dwek et al., 2002). These compounds have shown some potency against MARV, EBOV and other enveloped viruses (Chang et al., 2011, 2013). However, as they target cellular processes, high toxicity issues could be expected that may slow down their therapeutic development. Other potent broad-spectrum strategies are currently under study, such as targeting the cellular membrane bilayer. Inhibition of dihydroceramide desaturase should rigidify lipid-ordered membrane domains, thus limiting viral budding, as observed *in vitro* for HIV (Vieira et al., 2010). Another promising approach is the inhibition of cellular kinases. Nilotinib and imatinib (inhibitor antibodies against the tyrosine kinase ABL1) prevent VP40 phosphorylation and consequently late infection events (García et al., 2012). However, such strategies target cellular mechanisms and the potential side effects during therapy have not been evaluated.

Specific strategies to target viral components, such as the matrix protein VP40, also have been proposed. Potential druggable sites have been identified on VP40 by computational and *in vitro* studies (summarized in Table 1). For instance, VP40/host protein interactions could be disrupted by targeting the PTAP-dependent interactions with host ESCRT proteins in filoviruses and other viruses (Huang et al., 1995; Licata et al., 2003; Irie et al., 2005; Silvestri et al., 2007; Lu et al., 2014). Targeting the matrix protein PPTY-dependent interactions with NEDD4 could also prove a broad-spectrum strategy because this motif is largely present in filoviruses and enveloped RNA virus matrix proteins (Licata et al., 2003; Yasuda et al., 2003; Bornholdt et al., 2013; Han et al., 2014). Small organic compounds that target the virus-host interactions might represent a good compromise between targeting host factors and viral proteins, with virus-specificity and low cytotoxicity (Lu et al., 2014; Han et al., 2014; Loughran et al., 2016). Other strategies propose to target VP40 self-assembly at the plasma membrane, specifically the N-terminal dimerization interface, the loop for plasma membrane interaction, the C-terminal basic patch, and the hydrophobic loop for plasma membrane insertion (Timmins et al., 2003; Adu-Gyamfi et al., 2013, 2014; Bornholdt et al., 2013; Soni et al., 2013; Adu-Gyamfi et al., 2014; Mirza and Ikram, 2016). However, N-terminal dimerization or C-terminal assembly inhibition could be overcome by increased protein synthesis (Stahelin, 2014). Likewise, drug leads against VP40-mediated RNA binding have been described and should be further evaluated *in vitro* (Gomis-Rüth et al., 2003; Hoenen et al., 2005; Tamilvanan and Hopper, 2013; Bornholdt et al., 2013; Abazari et al., 2015; Karthick et al., 2016). Finally, other ligand clusters have been identified using standard bioinformatics techniques and these domains could be used for the development of other antiviral drug against filoviruses (Joob and Wiwanitkit, 2016).

A significant number of research avenues are open to target filovirus assembly and budding. Additional investigations on the latest infection event mechanisms might help to refine the proof of concept of such inhibition strategies, especially those focusing on VP40 plasticity and multi-functionality in these late events.

Table 1

EBOV VP40 regions that could be targeted for antiviral drug development. As a conclusion, this table summarizes VP40 structure/function characterization and related antiviral strategies. It highlights the central role of VP40 in late events of the filovirus life cycle and the multiple related inhibition strategies.

Region	Function	Inhibition strategy	Antiviral compounds	References
PTAP (late domain)	Interaction with the ESCRT-I complex via TSG101	Assembly and egress disruption	Proof of concept	Huang et al., 1995, Licata et al., 2003, Irie et al., 2005, Silvestri et al., 2007, Lu et al., 2014
PPEY (late domain)	Interaction with NEDD4	Assembly and egress disruption	quinoxalines, compounds 4 & 5	Licata et al., 2003, Yasuda et al., 2003, Bornholdt et al., 2013, Han et al., 2014, Lu et al., 2014, Loughran et al., 2016
Residues 52–65 and 108–117	VP40 dimerization	Self-assembly inhibition	Proof of concept	Timmins et al., 2003, Bornholdt et al., 2013, Soni et al., 2013, Mirza and Ikram, 2016
N-terminal loop	Plasma membrane recruitment	Assembly and egress disruption	Proof of concept	Bornholdt et al., 2013, Adu-Gyamfi et al., 2014, Mirza and Ikram, 2016
Octamerization domain	RNA binding	Viral replication block	Proof of concept	Gomis-Rüth et al., 2003, Timmins et al., 2003, Hoenen et al., 2005, Tamilvanan and Hopper, 2013, Bornholdt et al., 2013, Abazari et al., 2015, Karthick et al., 2016
C-terminal basic patch	Electrostatic interaction with the plasma membrane	Assembly and egress disruption	Proof of concept	Timmins et al., 2003, Adu-Gyamfi et al., 2014, Soni et al., 2013, Bornholdt et al., 2013, Mirza and Ikram, 2016
C-terminal hydrophobic loop	Electrostatic interaction with the plasma membrane	Assembly and egress disruption	Proof of concept	Timmins et al., 2003, Adu-Gyamfi et al., 2014, Soni et al., 2013, Bornholdt et al., 2013, Mirza and Ikram, 2016
VP40 phosphorylation	Nucleocapsid recruitment	Assembly and egress disruption	nilotinib and imatinib	García et al., 2012

6. Conclusion

It is likely that assembly and budding present a number of common features among enveloped virus. Indeed, most enveloped viruses hijack the host cell budding machinery by interacting with cellular factors of multivesicular bodies (Schmitt and Lamb, 2004; Morita and Sundquist, 2004). As this machinery is essential for many biological processes, its inhibition is speculative. The identification of specific inhibitors requires a better understanding of some specificities of VP40 self-assembly and VP24 roles in virus assembling. Disruption of VP40 late domain interactions with cellular factors, inhibition of endosomal pathway components that are essential for particle formation, or interference with viral membrane protein self-assembly and interaction with other viral partners are all feasible targets for the development of antiviral molecules. Knowledge of the complex processes shared by different viruses should expedite the development of efficacious therapeutics that could be used to treat different types of infections, including by filoviruses.

Acknowledgements

This work was supported by a DGA-MRIS and Aix-Marseille Université scholarship, by the ANR grants ANR-ST14-ASTR-0026 for the project VMTaseIn and ANR-14-EBOL-002-01.

References

Abazari, D., Moghtadaei, M., Behvarmanesh, A., Ghannadi, B., Aghaei, M., Behruznia, M., Rigi, G., 2015. Molecular docking based screening of predicted potential inhibitors for VP40 from Ebola virus. *Bioinformatics* 11, 243–247. <http://dx.doi.org/10.6026/97320630011243>.

Adu-Gyamfi, E., Digman, M.A., Gratton, E., Stahelin, R.V., 2012. Single-particle tracking demonstrates that actin coordinates the movement of the Ebola virus matrix protein. *Biophys. J.* 103, L41–L43. <http://dx.doi.org/10.1016/j.bpj.2012.09.026>.

Adu-Gyamfi, E., Johnson, K.A., Fraser, M.E., Scott, J.L., Soni, S.P., Jones, K.R., Digman, M.A., Gratton, E., Tessier, C.R., Stahelin, R.V., 2015. Host cell plasma membrane phosphatidylserine regulates the assembly and budding of Ebola virus. *J. Virol.* 89, 9440–9453. <http://dx.doi.org/10.1128/JVI.01087-15>.

Adu-Gyamfi, E., Soni, S.P., Jee, C.S., Digman, M.A., Gratton, E., Stahelin, R.V., 2014. A loop region in the N-terminal domain of Ebola virus VP40 is important in viral assembly, budding, and egress. *Viruses* 6, 3837–3854. <http://dx.doi.org/10.3390/v6103837>.

Adu-Gyamfi, E., Soni, S.P., Xue, Y., Digman, M.A., Gratton, E., Stahelin, R.V., 2013. The

Ebola virus matrix protein penetrates into the plasma membrane: a key step in viral protein 40 (VP40) oligomerization and viral egress. *J. Biol. Chem.* 288, 5779–5789. <http://dx.doi.org/10.1074/jbc.M112.443960>.

Bamberg, S., Kolesnikova, L., Möller, P., Klenk, H.-D., Becker, S., 2005. VP24 of Marburg virus influences formation of infectious particles. *J. Virol.* 79, 13421–13433. <http://dx.doi.org/10.1128/JVI.79.21.13421-13433.2005>.

Bavari, S., Bosio, C.M., Wiegand, E., Ruthel, G., Will, A.B., Geisbert, T.W., Hevey, M., Schmaljohn, C., Schmaljohn, A., Aman, M.J., 2002. Lipid raft microdomains: a gateway for compartmentalized trafficking of Ebola and Marburg viruses. *J. Exp. Med.* 195, 593–602.

Becker, S., Huppertz, S., Klenk, H.D., Feldmann, H., 1994. The nucleoprotein of Marburg virus is phosphorylated. *J. Gen. Virol.* 75 (Pt 4), 809–818. <http://dx.doi.org/10.1099/0022-1317-75-4-809>.

Beniac, D.R., Melito, P.L., Devarenes, S.L., Hiebert, S.L., Rabb, M.J., Lamboo, L.L., Jones, S.M., Booth, T.F., 2012. The organisation of Ebola virus reveals a capacity for extensive, modular polyploidy. *PLoS One* 7, e29608. <http://dx.doi.org/10.1371/journal.pone.0029608>.

Bornholdt, Z.A., Noda, T., Abelson, D.M., Halfmann, P., Wood, M.R., Kawaoka, Y., Saphire, E.O., 2013. Structural rearrangement of ebola virus VP40 begets multiple functions in the virus life cycle. *Cell* 154, 763–774. <http://dx.doi.org/10.1016/j.cell.2013.07.015>.

Bukreyev, A., Volchkov, V.E., Blinov, V.M., Netesov, S.V., 1993a. The GP-protein of Marburg virus contains the region similar to the “immunosuppressive domain” of oncogenic retrovirus P15E proteins. *FEBS Lett.* 323, 183–187.

Bukreyev, A.A., Volchkov, V.E., Blinov, V.M., Netesov, S.V., 1993b. The VP35 and VP40 proteins of filoviruses. Homology between Marburg and Ebola viruses. *FEBS Lett.* 322, 41–46.

Bukreyev, A.A., Chandran, K., Dolnik, O., Dye, J.M., Ebihara, H., Leroy, E.M., Mühlberger, E., Netesov, S.V., Patterson, J.L., Paweska, J.T., Saphire, E.O., Smither, S.J., Takada, A., Towner, J.S., Volchkov, V.E., Warren, T.K., Kuhn, J.H., 2014. Discussions and decisions of the 2012–2014 International Committee on taxonomy of viruses (ICTV) Filoviridae study group, January 2012–June 2013. *Arch. Virol.* 159, 821–830. <http://dx.doi.org/10.1007/s00705-013-1846-9>.

Chambers, P., Samson, A.C., 1980. A new structural protein for Newcastle disease virus. *J. Gen. Virol.* 50, 155–166. <http://dx.doi.org/10.1099/0022-1317-50-1-155>.

Chang, J., Schul, W., Yip, A., Xu, X., Guo, J.-T., Block, T.M., 2011. Competitive inhibitor of cellular α -glucosidases protects mice from lethal dengue virus infection. *Antivir. Res.* 92, 369–371. <http://dx.doi.org/10.1016/j.antiviral.2011.08.003>.

Chang, J., Warren, T.K., Zhao, X., Gill, T., Guo, F., Wang, L., Comunale, M.A., Du, Y., Alonzi, D.S., Yu, W., Ye, H., Liu, F., Guo, J.-T., Mehta, A., Cuconati, A., Butters, T.D., Bavari, S., Xu, X., Block, T.M., 2013. Small molecule inhibitors of ER α -glucosidases are active against multiple hemorrhagic fever viruses. *Antivir. Res.* 98, 432–440. <http://dx.doi.org/10.1016/j.antiviral.2013.03.023>.

Clifton, M.C., Bruhn, J.F., Atkins, K., Webb, T.L., Baydo, R.O., Raymond, A., Lorimer, D.D., Edwards, T.E., Myler, P.J., Saphire, E.O., 2015. High-resolution crystal structure of dimeric VP40 from Sudan ebolavirus. *J. Infect. Dis.* 212 (Suppl. 2), S167–S171. <http://dx.doi.org/10.1093/infdis/jiv090>.

Dessen, A., Volchkov, V., Dolnik, O., Klenk, H.D., Weissenhorn, W., 2000. Crystal structure of the matrix protein VP40 from Ebola virus. *EMBO J.* 19, 4228–4236. <http://dx.doi.org/10.1093/emboj/19.16.4228>.

DiCarlo, A., Biedenkopf, N., Hartlieb, B., Klussmeier, A., Becker, S., 2011.

- Phosphorylation of Marburg virus NP region II modulates viral RNA synthesis. *J. Infect. Dis.* 204 (Suppl. 3), S927–S933. <http://dx.doi.org/10.1093/infdis/jir319>.
- DiCarlo, A., Möller, P., Lander, A., Kolesnikova, L., Becker, S., 2007. Nucleocapsid formation and RNA synthesis of Marburg virus is dependent on two coiled coil motifs in the nucleoprotein. *Virology* 357, 105. <http://dx.doi.org/10.1016/j.virol.2006.11.015>.
- Dolnik, O., Kolesnikova, L., Stevermann, L., Becker, S., 2010. Tsg101 is recruited by a late domain of the nucleocapsid protein to support budding of Marburg virus-like particles. *J. Virol.* 84, 7847–7856. <http://dx.doi.org/10.1128/JVI.00476-10>.
- Dolnik, O., Kolesnikova, L., Welsch, S., Strecker, T., Schudt, G., Becker, S., 2014. Interaction with Tsg101 is necessary for the efficient transport and release of nucleocapsids in marburg virus-infected cells. *PLoS Pathog.* 10, e1004463. <http://dx.doi.org/10.1371/journal.ppat.1004463>.
- Dong, S., Yang, P., Li, G., Liu, B., Wang, W., Liu, X., Xia, B., Yang, C., Lou, Z., Guo, Y., Rao, Z., 2015. Insight into the Ebola virus nucleocapsid assembly mechanism: crystal structure of Ebola virus nucleoprotein core domain at 1.8 Å resolution. *Protein Cell* 6, 351–362. <http://dx.doi.org/10.1007/s13238-015-0163-3>.
- Dwek, R.A., Butters, T.D., Platt, F.M., Zitzmann, N., 2002. Targeting glycosylation as a therapeutic approach. *Nat. Rev. Drug Discov.* 1, 65–75. <http://dx.doi.org/10.1038/nrd708>.
- Dziubańska, P.J., Derewenda, U., Ellena, J.F., Engel, D.A., Derewenda, Z.S., 2014. The structure of the C-terminal domain of the Zaire ebolavirus nucleoprotein. *Acta Crystallogr. D Biol. Crystallogr.* 70, 2420–2429. <http://dx.doi.org/10.1107/S1399004714014710>.
- Feldmann, H., Mühlberger, E., Randolph, A., Will, C., Kiley, M.P., Sanchez, A., Klenk, H.D., 1992. Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle. *Virus Res.* 24, 1–19.
- Funke, C., Becker, S., Dartsch, H., Klenk, H.D., Mühlberger, E., 1995. Acylation of the Marburg virus glycoprotein. *Virology* 208, 289–297. <http://dx.doi.org/10.1006/viro.1995.1151>.
- García, M., Cooper, A., Shi, W., Bornmann, W., Carrion, R., Kalman, D., Nabel, G.J., 2012. Productive replication of Ebola virus is regulated by the c-Abl1 tyrosine kinase. *Sci. Transl. Med.* 4, 123ra24. <http://dx.doi.org/10.1126/scitranslmed.3003500>.
- Gomis-Rüth, F.X., Dessen, A., Timmins, J., Bracher, A., Kolesnikova, L., Becker, S., Klenk, H.D., Weissenhorn, W., 2003. The matrix protein VP40 from Ebola virus octamerizes into pore-like structures with specific RNA binding properties. *Struct. Lond. Engl.* 11, 423–433.
- Gustin, J.K., Bai, Y., Moses, A.V., Douglas, J.L., 2015. Ebola virus glycoprotein promotes enhanced viral egress by preventing Ebola VP40 from associating with the host restriction factor BST2/Tetherin. *J. Infect. Dis.* 212 (Suppl. 2), S181–S190. <http://dx.doi.org/10.1093/infdis/jiv125>.
- Han, Z., Boshra, H., Sunyer, J.O., Zwiers, S.H., Paragas, J., Hartly, R.N., 2003. Biochemical and functional characterization of the Ebola virus VP24 protein: implications for a role in virus assembly and budding. *J. Virol.* 77, 1793–1800.
- Han, Z., Lu, J., Liu, Y., Davis, B., Lee, M.S., Olson, M.A., Ruthel, G., Freedman, B.D., Schnell, M.J., Wrobel, J.E., Reitz, A.B., Hartly, R.N., 2014. Small-molecule probes targeting the viral PPxY-host Nedd4 interface block egress of a broad range of RNA viruses. *J. Virol.* 88, 7294–7306. <http://dx.doi.org/10.1128/JVI.00591-14>.
- Harty, R.N., Brown, M.E., Wang, G., Huijbreghse, J., Hayes, F.P., 2000. A PPxY motif within the VP40 protein of Ebola virus interacts physically and functionally with a ubiquitin ligase: implications for filovirus budding. *Proc. Natl. Acad. Sci. U. S. A.* 97, 13871–13876. <http://dx.doi.org/10.1073/pnas.250277297>.
- Hoenen, T., Biedenkopf, N., Ziebeck, F., Jung, S., Groseth, A., Feldmann, H., Becker, S., 2010. Oligomerization of Ebola virus VP40 is essential for particle morphogenesis and regulation of viral transcription. *J. Virol.* 84, 7053–7063. <http://dx.doi.org/10.1128/JVI.00737-10>.
- Hoenen, T., Volchkov, V., Kolesnikova, L., Mittler, E., Timmins, J., Ottmann, M., Reynard, O., Becker, S., Weissenhorn, W., 2005. VP40 octamers are essential for Ebola virus replication. *J. Virol.* 79, 1898–1905. <http://dx.doi.org/10.1128/JVI.79.3.1898-1905.2005>.
- Huang, M., Orenstein, J.M., Martin, M.A., Freed, E.O., 1995. p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. *J. Virol.* 69, 6810–6818.
- Huang, Y., Xu, L., Sun, Y., Nabel, G.J., 2002. The assembly of Ebola virus nucleocapsid requires virion-associated proteins 35 and 24 and posttranslational modification of nucleoprotein. *Mol. Cell* 10, 307–316.
- Irie, T., Licata, J.M., Hartly, R.N., 2005. Functional characterization of Ebola virus L-domains using VSV recombinants. *Virology* 336, 291–298. <http://dx.doi.org/10.1016/j.virol.2005.03.027>.
- Ito, H., Watanabe, S., Takada, A., Kawaoka, Y., 2001. Ebola virus glycoprotein: proteolytic processing, acylation, cell tropism, and detection of neutralizing antibodies. *J. Virol.* 75, 1576–1580. <http://dx.doi.org/10.1128/JVI.75.3.1576-1580.2001>.
- Jasenosky, L.D., Kawaoka, Y., 2004. Filovirus budding. *Virus Res.* 106, 181–188. <http://dx.doi.org/10.1016/j.virusres.2004.08.014>.
- Johnson, K.A., Taghon, G.J.F., Scott, J.L., Stahelin, R.V., 2016. The Ebola Virus matrix protein, VP40, requires phosphatidylinositol 4,5-bisphosphate (PI(4,5)P2) for extensive oligomerization at the plasma membrane and viral egress. *Sci. Rep.* 6, 19125. <http://dx.doi.org/10.1038/srep19125>.
- Joob, B., Wiwanitkit, V., 2016. Determination of ligand cluster and binding site within VP40 of Ebola virus: Clue for drug development. *Asian Pac. J. Trop. Med.* 9, 410–411. <http://dx.doi.org/10.1016/j.apjtm.2016.03.002>.
- Kaptur, P.E., McCreedy, B.J., Lyles, D.S., 1992. Sites of in vivo phosphorylation of vesicular stomatitis virus matrix protein. *J. Virol.* 66, 5384–5392.
- Kaptur, P.E., McKenzie, M.O., Wertz, G.W., Lyles, D.S., 1995. Assembly functions of vesicular stomatitis virus matrix protein are not disrupted by mutations at major sites of phosphorylation. *Virology* 206, 894–903. <http://dx.doi.org/10.1006/viro.1995.1012>.
- Karthick, V., Nagasundaram, N., Doss, C.G.P., Chakraborty, C., Siva, R., Lu, A., Zhang, G., Zhu, H., 2016. Virtual screening of the inhibitors targeting at the viral protein 40 of Ebola virus. *Infect. Dis. Poverty* 5, 12. <http://dx.doi.org/10.1186/s40249-016-0105-1>.
- Kolesnikova, L., Bamberg, S., Berghöfer, B., Becker, S., 2004a. The matrix protein of Marburg virus is transported to the plasma membrane along cellular membranes: exploiting the retrograde late endosomal pathway. *J. Virol.* 78, 2382–2393.
- Kolesnikova, L., Berghöfer, B., Bamberg, S., Becker, S., 2004b. Multivesicular bodies as a platform for formation of the Marburg virus envelope. *J. Virol.* 78, 12277–12287. <http://dx.doi.org/10.1128/JVI.78.22.12277-12287.2004>.
- Kolesnikova, L., Bugany, H., Klenk, H.-D., Becker, S., 2002. VP40, the matrix protein of Marburg virus, is associated with membranes of the late endosomal compartment. *J. Virol.* 76, 1825–1838.
- Kolesnikova, L., Mittler, E., Schudt, G., Shams-Eldin, H., Becker, S., 2012. Phosphorylation of Marburg virus matrix protein VP40 triggers assembly of nucleocapsids with the viral envelope at the plasma membrane. *Cell Microbiol.* 14, 182–197. <http://dx.doi.org/10.1111/j.1462-5822.2011.01709.x>.
- Lamb, R.A., Choppin, P.W., 1977. The synthesis of Sendai virus polypeptides in infected cells. III. Phosphorylation of polypeptides. *Virology* 81, 382–397.
- Lambert, D.M., Hambor, J., Diebold, M., Galinski, B., 1988. Kinetics of synthesis and phosphorylation of respiratory syncytial virus polypeptides. *J. Gen. Virol.* 69 (Pt 2), 313–323. <http://dx.doi.org/10.1099/0022-1317-69-2-313>.
- Lee, J.E., Fusco, M.L., Hessel, A.J., Oswald, W.B., Burton, D.R., Saphire, E.O., 2008. Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature* 454, 177–182. <http://dx.doi.org/10.1038/nature07082>.
- Lee, J.E., Saphire, E.O., 2009. Ebola virus glycoprotein structure and mechanism of entry. *Future Virol.* 4, 621–635. <http://dx.doi.org/10.2217/fvl.09.56>.
- Li, Z., Tian, H., Liu, T., 2012. [Cellular ESCRT complex and its roles in enveloped viruses budding]. *Sheng Wu Gong Cheng Xue Bao Chin. J. Biotechnol.* 28, 1031–1037.
- Licata, J.M., Johnson, R.F., Han, Z., Hartly, R.N., 2004. Contribution of ebola virus glycoprotein, nucleoprotein, and VP24 to budding of VP40 virus-like particles. *J. Virol.* 78, 7344–7351. <http://dx.doi.org/10.1128/JVI.78.14.7344-7351.2004>.
- Licata, J.M., Simpson-Holley, M., Wright, N.T., Han, Z., Paragas, J., Hartly, R.N., 2003. Overlapping motifs (PTAP and PPEY) within the Ebola virus VP40 protein function independently as late budding domains: involvement of host proteins TSG101 and VPS-4. *J. Virol.* 77, 1812–1819.
- Liu, Y., Cocka, L., Okumura, A., Zhang, Y.-A., Sunyer, J.O., Hartly, R.N., 2010. Conserved motifs within Ebola and Marburg virus VP40 proteins are important for stability, localization, and subsequent budding of virus-like particles. *J. Virol.* 84, 2294–2303. <http://dx.doi.org/10.1128/JVI.02034-09>.
- Löttering, B., Mühlberger, E., Tamura, T., Klenk, H.D., Becker, S., 1999. The nucleoprotein of Marburg virus is target for multiple cellular kinases. *Virology* 255, 50–62. <http://dx.doi.org/10.1006/viro.1998.9577>.
- Loughran, H.M., Han, Z., Wrobel, J.E., Decker, S.E., Ruthel, G., Freedman, B.D., Hartly, R.N., Reitz, A.B., 2016. Quinoxaline-based inhibitors of Ebola and Marburg VP40 egress. *Bioorg. Med. Chem. Lett* 26, 3429–3435. <http://dx.doi.org/10.1016/j.bmcl.2016.06.053>.
- Lu, J., Han, Z., Liu, Y., Liu, W., Lee, M.S., Olson, M.A., Ruthel, G., Freedman, B.D., Hartly, R.N., 2014. A host-oriented inhibitor of Junin Argentine hemorrhagic fever virus egress. *J. Virol.* 88, 4736–4743. <http://dx.doi.org/10.1128/JVI.03757-13>.
- Lu, J., Qu, Y., Liu, Y., Jambusaria, R., Han, Z., Ruthel, G., Freedman, B.D., Hartly, R.N., 2013. Host IQGAP1 and Ebola virus VP40 interactions facilitate virus-like particle egress. *J. Virol.* 87, 7777–7780. <http://dx.doi.org/10.1128/JVI.00470-13>.
- Martin, B., Canard, B., Decroly, E., 2017. Filovirus proteins for antiviral drug discovery: structure/function bases of the replication cycle. *Antivir. Res.* 141, 48–61. <http://dx.doi.org/10.1016/j.antiviral.2017.02.004>.
- Martin, B., Hoenen, T., Canard, B., Decroly, E., 2016. Filovirus proteins for antiviral drug discovery: a structure/function analysis of surface glycoproteins and virus entry. *Antivir. Res.* 135, 1–14. <http://dx.doi.org/10.1016/j.antiviral.2016.09.001>.
- Martin-Serrano, J., Yarovoy, A., Perez-Caballero, D., Bieniasz, P.D., Yarovoy, A., 2003. Divergent retroviral late-budding domains recruit vacuolar protein sorting factors by using alternative adaptor proteins. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12414–12419. <http://dx.doi.org/10.1073/pnas.2133846100>.
- Mateo, M., Carbonnelle, C., Martinez, M.J., Reynard, O., Page, A., Volchkov, V.A., Volchkov, V.E., 2011. Knockdown of Ebola virus VP24 impairs viral nucleocapsid assembly and prevents virus replication. *J. Infect. Dis.* 204 (Suppl. 3), S892–S896. <http://dx.doi.org/10.1093/infdis/jir311>.
- McDonald, B., Martin-Serrano, J., 2009. No strings attached: the ESCRT machinery in viral budding and cytokinesis. *J. Cell Sci.* 122, 2167–2177. <http://dx.doi.org/10.1242/jcs.028308>.
- Mim, C., Unger, V.M., 2012. Membrane curvature and its generation by BAR proteins. *Trends Biochem. Sci.* 37, 526–533. <http://dx.doi.org/10.1016/j.tics.2012.09.001>.
- Mirza, M.U., Ikram, N., 2016. Integrated computational approach for virtual hit identification against Ebola viral proteins VP35 and VP40. *Int. J. Mol. Sci.* 17. <http://dx.doi.org/10.3390/ijms17111748>.
- Mittler, E., Kolesnikova, L., Hartlieb, B., Davey, R., Becker, S., 2011. The cytoplasmic domain of Marburg virus GP modulates early steps of viral infection. *J. Virol.* 85, 8188–8196. <http://dx.doi.org/10.1128/JVI.00453-11>.

- Mittler, E., Kolesnikova, L., Herwig, A., Dolnik, O., Becker, S., 2013. Assembly of the Marburg virus envelope. *Cell Microbiol.* 15, 270–284. <http://dx.doi.org/10.1111/cmi.12076>.
- Mittler, E., Kolesnikova, L., Strecker, T., Garten, W., Becker, S., 2007. Role of the transmembrane domain of marburg virus surface protein GP in assembly of the viral envelope. *J. Virol.* 81, 3942–3948. <http://dx.doi.org/10.1128/JVI.02263-06>.
- Morita, E., Sundquist, W.I., 2004. Retrovirus budding. *Annu. Rev. Cell Dev. Biol.* 20, 395–425. <http://dx.doi.org/10.1146/annurev.cellbio.20.010403.102350>.
- Nanbo, A., Watanabe, S., Halfmann, P., Kawaoka, Y., 2013. The spatio-temporal distribution dynamics of Ebola virus proteins and RNA in infected cells. *Sci. Rep.* 3, 1206. <http://dx.doi.org/10.1038/srep01206>.
- Negredo, A., Palacios, G., Vázquez-Morón, S., González, F., Dopazo, H., Molero, F., Juste, J., Quetglas, J., Savji, N., de la Cruz Martínez, M., Herrera, J.E., Pizarro, M., Hutchison, S.K., Echevarría, J.E., Lipkin, W.I., Tenorio, A., 2011. Discovery of an ebolavirus-like filovirus in europe. *PLoS Pathog.* 7, e1002304. <http://dx.doi.org/10.1371/journal.ppat.1002304>.
- Neumann, G., Ebihara, H., Takada, A., Noda, T., Kobasa, D., Jasenosky, L.D., Watanabe, S., Kim, J.H., Feldmann, H., Kawaoka, Y., 2005. Ebola virus VP40 late domains are not essential for viral replication in cell culture. *J. Virol.* 79, 10300–10307. <http://dx.doi.org/10.1128/JVI.79.16.10300-10307.2005>.
- Noda, T., Ebihara, H., Muramoto, Y., Fujii, K., Takada, A., Sagara, H., Kim, J.H., Kida, H., Feldmann, H., Kawaoka, Y., 2006. Assembly and budding of Ebolavirus. *PLoS Pathog.* 2, e99. <http://dx.doi.org/10.1371/journal.ppat.0020099>.
- Noda, T., Halfmann, P., Sagara, H., Kawaoka, Y., 2007. Regions in Ebola virus VP24 that are important for nucleocapsid formation. *J. Infect. Dis.* 196 (Suppl. 2), S247–S250. <http://dx.doi.org/10.1086/520596>.
- Noda, T., Sagara, H., Suzuki, E., Takada, A., Kida, H., Kawaoka, Y., 2002. Ebola virus VP40 drives the formation of virus-like filamentous particles along with GP. *J. Virol.* 76, 4855–4865.
- Oda, S.-I., Noda, T., Wijesinghe, K.J., Halfmann, P., Bornholdt, Z.A., Abelson, D.M., Armbrust, T., Stahelin, R.V., Kawaoka, Y., Saphire, E.O., 2015. Crystal structure of Marburg virus VP40 reveals a broad, basic patch for matrix assembly and a requirement of the N-Terminal domain for immunosuppression. *J. Virol.* 90, 1839–1848. <http://dx.doi.org/10.1128/JVI.01597-15>.
- Page, A., Volchikova, V.A., Reid, S.P., Mateo, M., Bagnaud-Baule, A., Nemirov, K., Shurtleff, A.C., Lawrence, P., Reynard, O., Ottmann, M., Lotteau, V., Biswal, S.S., Thimmulappa, R.K., Bavari, S., Volchok, V.E., 2014. Marburgvirus hijacks nrf2-dependent pathway by targeting nrf2-negative regulator keap1. *Cell Rep.* 6, 1026–1036. <http://dx.doi.org/10.1016/j.celrep.2014.02.027>.
- Peng, R., Zhu, T., Oladejo, B.O., Musyoki, A.M., Cui, Y., Shi, Y., Wang, P., Gao, G.F., 2016. In vitro assembly of Ebola virus nucleocapsid-like complex expressed in *E. coli*. *Protein Cell.* <http://dx.doi.org/10.1007/s13238-016-0314-1>.
- Peyrol, J., Thizon, C., Gaillard, J.-C., Marchetti, C., Armengaud, J., Rollin-Genetet, F., 2013. Multiple phosphorylatable sites in the Zaire Ebolavirus nucleoprotein evidenced by high resolution tandem mass spectrometry. *J. Virol Meth.* 187, 159–165. <http://dx.doi.org/10.1016/j.jviromet.2012.10.003>.
- Radzimanowski, J., Effantin, G., Weissenhorn, W., 2014. Conformational plasticity of the Ebola virus matrix protein. *Protein Sci. Publ. Protein Soc.* 23, 1519–1527. <http://dx.doi.org/10.1002/pro.2541>.
- Reid, S.P., Leung, L.W., Hartman, A.L., Martinez, O., Shaw, M.L., Carbonnelle, C., Volchok, V.E., Nichol, S.T., Basler, C.F., 2006. Ebola virus VP24 binds karyopherin alpha1 and blocks STAT1 nuclear accumulation. *J. Virol.* 80, 5156–5167. <http://dx.doi.org/10.1128/JVI.02349-05>.
- Reynard, O., Nemirov, K., Page, A., Mateo, M., Raoul, H., Weissenhorn, W., Volchok, V.E., 2011. Conserved proline-rich region of Ebola virus matrix protein VP40 is essential for plasma membrane targeting and virus-like particle release. *J. Infect. Dis.* 204 (Suppl. 3), S884–S891. <http://dx.doi.org/10.1093/infdis/jir359>.
- Rima, B.K., Roberts, M.W., McAdam, W.D., Martin, S.J., 1980. Polypeptide synthesis in mumps virus-infected cells. *J. Gen. Virol.* 46, 501–505. <http://dx.doi.org/10.1099/0022-1317-46-2-501>.
- Ruigrok, R.W., Schoehn, G., Dessen, A., Forest, E., Volchok, V., Dolnik, O., Klenk, H.D., Weissenhorn, W., 2000. Structural characterization and membrane binding properties of the matrix protein VP40 of Ebola virus. *J. Mol. Biol.* 300, 103–112. <http://dx.doi.org/10.1006/jmbi.2000.3822>.
- Ruthel, G., Demmin, G.L., Kallstrom, G., Javid, M.P., Badie, S.S., Will, A.B., Nelle, T., Schokman, R., Nguyen, T.L., Carra, J.H., Bavari, S., Aman, M.J., 2005. Association of ebola virus matrix protein VP40 with microtubules. *J. Virol.* 79, 4709–4719. <http://dx.doi.org/10.1128/JVI.79.8.4709-4719.2005>.
- Sanchez, A., Kiley, M.P., Holloway, B.P., 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.
- Sänger, C., Mühlberger, E., Ryabchikova, E., Kolesnikova, L., Klenk, H.D., Becker, S., 2001. Sorting of Marburg virus surface protein and virus release take place at opposite surfaces of infected polarized epithelial cells. *J. Virol.* 75, 1274–1283. <http://dx.doi.org/10.1128/JVI.75.3.1274-1283.2001>.
- Schmitt, A.P., Lamb, R.A., 2004. Escaping from the cell: assembly and budding of negative-strand RNA viruses. *Curr. Top. Microbiol. Immunol.* 283, 145–196.
- Schnittler, H.J., Mahner, F., Drenckhahn, D., Klenk, H.D., Feldmann, H., 1993. Replication of Marburg virus in human endothelial cells. A possible mechanism for the development of viral hemorrhagic disease. *J. Clin. Invest.* 91, 1301–1309. <http://dx.doi.org/10.1172/JCI116329>.
- Schudt, G., Dolnik, O., Kolesnikova, L., Biedenkopf, N., Herwig, A., Becker, S., 2015. Transport of Ebolavirus nucleocapsids is dependent on actin polymerization: live-cell imaging analysis of Ebolavirus-infected cells. *J. Infect. Dis.* 212 (Suppl. 2), S160–S166. <http://dx.doi.org/10.1093/infdis/jiv083>.
- Schudt, G., Kolesnikova, L., Dolnik, O., Sodeik, B., Becker, S., 2013. Live-cell imaging of Marburg virus-infected cells uncovers actin-dependent transport of nucleocapsids over long distances. *Proc. Natl. Acad. Sci. U. S. A.* 110, 14402–14407. <http://dx.doi.org/10.1073/pnas.1307681110>.
- Scianimanco, S., Schoehn, G., Timmins, J., Ruigrok, R.H., Klenk, H.D., Weissenhorn, W., 2000. Membrane association induces a conformational change in the Ebola virus matrix protein. *EMBO J.* 19, 6732–6741. <http://dx.doi.org/10.1093/emboj/19.24.6732>.
- Silva, L.P., Vanzile, M., Bavari, S., Aman, J.M.J., Schriemer, D.C., 2012. Assembly of Ebola virus matrix protein VP40 is regulated by latch-like properties of N and C terminal tails. *PLoS One* 7, e39978. <http://dx.doi.org/10.1371/journal.pone.0039978>.
- Silvestri, L.S., Ruthel, G., Kallstrom, G., Warfield, K.L., Swenson, D.L., Nelle, T., Iversen, P.L., Bavari, S., Aman, M.J., 2007. Involvement of vacuolar protein sorting pathway in Ebola virus release independent of TSG101 interaction. *J. Infect. Dis.* 196 (Suppl. 2), S264–S270. <http://dx.doi.org/10.1086/520610>.
- Soni, S.P., Adu-Gyamfi, E., Yong, S.S., Jee, C.S., Stahelin, R.V., 2013. The Ebola virus matrix protein deeply penetrates the plasma membrane: an important step in viral egress. *Biophys. J.* 104, 1940–1949. <http://dx.doi.org/10.1016/j.bpj.2013.03.021>.
- Stahelin, R.V., 2014. Could the Ebola virus matrix protein VP40 be a drug target? *Expert Opin. Ther. Targets* 18, 115–120. <http://dx.doi.org/10.1517/14728222.2014.863877>.
- Strack, B., Calistri, A., Craig, S., Popova, E., Göttlinger, H.G., 2003. AIP1/ALIX is a binding partner for HIV-1 p6 and ELAV p9 functioning in virus budding. *Cell* 114, 689–699.
- Tamilvanan, T., Hopper, W., 2013. High-throughput virtual screening and docking studies of matrix protein vp40 of ebola virus. *Bioinformation* 9, 286–292. <http://dx.doi.org/10.6026/97320630009286>.
- Timmins, J., Ruigrok, R.W.H., Weissenhorn, W., 2004. Structural studies on the Ebola virus matrix protein VP40 indicate that matrix proteins of enveloped RNA viruses are analogues but not homologues. *FEMS Microbiol. Lett.* 233, 179–186. <http://dx.doi.org/10.1016/j.femsle.2004.03.002>.
- Timmins, J., Schoehn, G., Kohlhaas, C., Klenk, H.-D., Ruigrok, R.W.H., Weissenhorn, W., 2003. Oligomerization and polymerization of the filovirus matrix protein VP40. *Virology* 312, 359–368.
- Urata, S., Noda, T., Kawaoka, Y., Morikawa, S., Yokosawa, H., Yasuda, J., 2007. Interaction of Tsg101 with Marburg virus VP40 depends on the PPPY motif, but not the PT/SAP motif as in the case of Ebola virus, and Tsg101 plays a critical role in the budding of Marburg virus-like particles induced by VP40, NP, and GP. *J. Virol.* 81, 4895–4899. <http://dx.doi.org/10.1128/JVI.02829-06>.
- Vieira, C.R., Munoz-Olaya, J.M., Sot, J., Jiménez-Baranda, S., Izquierdo-Useros, N., Abad, J.L., Apellániz, B., Delgado, R., Martínez-Picado, J., Alonso, A., Casas, J., Nieva, J.L., Fabriás, G., Mañes, S., Goñi, F.M., 2010. Dihydrospingomyelin impairs HIV-1 infection by rigidifying liquid-ordered membrane domains. *Chem. Biol.* 17, 766–775. <http://dx.doi.org/10.1016/j.chembiol.2010.05.023>.
- Volchok, V.E., Blinov, V.M., Netesov, S.V., 1992. The envelope glycoprotein of Ebola virus contains an immunosuppressive-like domain similar to oncogenic retroviruses. *FEBS Lett.* 305, 181–184.
- Volchok, V.E., Feldmann, H., Volchikova, V.A., Klenk, H.D., 1998a. Processing of the Ebola virus glycoprotein by the proprotein convertase furin. *Proc. Natl. Acad. Sci. U. S. A.* 95, 5762–5767.
- Volchok, V.E., Volchikova, V.A., Slenczka, W., Klenk, H.D., Feldmann, H., 1998b. Release of viral glycoproteins during Ebola virus infection. *Virology* 245, 110–119. <http://dx.doi.org/10.1006/viro.1998.9143>.
- Volchok, V.E., Volchikova, V.A., Ströher, U., Becker, S., Dolnik, O., Cieplik, M., Garten, W., Klenk, H.D., Feldmann, H., 2000. Proteolytic processing of Marburg virus glycoprotein. *Virology* 268, 1–6. <http://dx.doi.org/10.1006/viro.1999.0110>.
- von Schwedler, U.K., Stuchell, M., Müller, B., Ward, D.M., Chung, H.-Y., Morita, E., Wang, H.E., Davis, T., He, G.-P., Cimbara, D.M., Scott, A., Kräusslich, H.-G., Kaplan, J., Morham, S.G., Sundquist, W.I., 2003. The protein network of HIV budding. *Cell* 114, 701–713.
- Votteler, J., Sundquist, W.I., 2013. Virus budding and the ESCRT pathway. *Cell Host Microbe* 14, 232–241. <http://dx.doi.org/10.1016/j.chom.2013.08.012>.
- Warren, T.K., Whitehouse, C.A., Wells, J., Welch, L., Heald, A.E., Charleston, J.S., Szani, P., Reid, S.P., Iversen, P.L., Bavari, S., 2015. A single phosphoramidate morpholino oligomer targeting VP24 protects rhesus monkeys against lethal Ebola virus infection. *mBio* 6. <http://dx.doi.org/10.1128/mBio.02344-14>.
- Watanabe, S., Noda, T., Halfmann, P., Jasenosky, L., Kawaoka, Y., 2007. Ebola virus (EBOV) VP24 inhibits transcription and replication of the EBOV genome. *J. Infect. Dis.* 196 (Suppl. 2), S284–S290. <http://dx.doi.org/10.1086/520582>.
- Watanabe, S., Watanabe, T., Noda, T., Takada, A., Feldmann, H., Jasenosky, L.D., Kawaoka, Y., 2004. Production of novel ebola virus-like particles from cDNAs: an alternative to ebola virus generation by reverse genetics. *J. Virol.* 78, 999–1005.
- Watt, A., Moukambi, F., Banadyga, L., Groseth, A., Callison, J., Herwig, A., Ebihara, H., Feldmann, H., Hoenen, T., 2014. A novel life cycle modeling system for Ebola virus shows a genome length-dependent role of VP24 in virus infectivity. *J. Virol.* 88, 10511–10524. <http://dx.doi.org/10.1128/JVI.01272-14>.
- Welsch, S., Kolesnikova, L., Krähling, V., Riches, J.D., Becker, S., Briggs, J.A.G., 2010. Electron tomography reveals the steps in filovirus budding. *PLoS Pathog.* 6, e1000875. <http://dx.doi.org/10.1371/journal.ppat.1000875>.
- Wijesinghe, K.J., Urata, S., Bhattarai, N., Kooijman, E.E., Gerstman, B.S., Chapagain, P.P., Li, S., Stahelin, R.V., 2017. Detection of lipid-induced structural changes of the Marburg virus matrix protein VP40 using hydrogen/deuterium exchange-mass spectrometry. *J. Biol. Chem.* 292, 6108–6122. <http://dx.doi.org/10.1074/jbc.M116.758300>.
- Yamayoshi, S., Kawaoka, Y., 2007. Mapping of a region of Ebola virus VP40 that is important in the production of virus-like particles. *J. Infect. Dis.* 196 (Suppl. 2),

- S291–S295. <http://dx.doi.org/10.1086/520595>.
- Yamayoshi, S., Noda, T., Ebihara, H., Goto, H., Morikawa, Y., Lukashevich, I.S., Neumann, G., Feldmann, H., Kawaoka, Y., 2008. Ebola virus matrix protein VP40 uses the COPII transport system for its intracellular transport. *Cell Host Microbe* 3, 168–177. <http://dx.doi.org/10.1016/j.chom.2008.02.001>.
- Yasuda, J., Nakao, M., Kawaoka, Y., Shida, H., 2003. Nedd4 regulates egress of Ebola virus-like particles from host cells. *J. Virol.* 77, 9987–9992.
- Zhang, A.P.P., Bornholdt, Z.A., Abelson, D.M., Saphire, E.O., 2014. Crystal structure of Marburg virus VP24. *J. Virol.* 88, 5859–5863. <http://dx.doi.org/10.1128/JVI.03565-13>.
- Zhang, A.P.P., Bornholdt, Z.A., Liu, T., Abelson, D.M., Lee, D.E., Li, S., Woods, V.L., Saphire, E.O., 2012. The ebola virus interferon antagonist VP24 directly binds STAT1 and has a novel, pyramidal fold. *PLoS Pathog.* 8, e1002550. <http://dx.doi.org/10.1371/journal.ppat.1002550>.

CHAPITRE 5

**Immunité innée, filovirus et structure
coiffe: bases moléculaires et drug design**

Les filovirus entraînent des infections aiguës et sévères, notamment en raison de leur réplication efficace médiée par les enzymes virales présentées dans les chapitres précédents et associée à une subversion de la réponse inflammatoire. Ces éléments conduisent aux manifestations cliniques telles que des fièvres hémorragiques à filovirus, incluant une forte fièvre, des fuites vasculaires et une coagulopathie. La réplication incontrôlée des filovirus cache en réalité de puissants mécanismes, portés par les protéines virales, permettant de supprimer la réponse immunitaire innée antivirale. Dans ce chapitre, l'objectif est de présenter les bases moléculaires sur lesquelles reposent ces mécanismes d'échappement des filovirus à l'immunité innée. Dans un premier temps, les principales voies de l'immunité innée antivirale ciblées par les filovirus sont présentées. Ensuite, nous avons mis en lien ces différentes voies avec les mécanismes d'échappement portés par différentes protéines des filovirus. La relation structure/fonction de VP35 avec ses différents mécanismes d'échappement ainsi que les mécanismes de subversion de la réponse immunitaire innée portés par VP24 chez EBOV et LLOV et VP40 chez MARV sont ensuite résumés. Enfin, la coiffe des ARNm viraux jouant également un rôle dans l'échappement à l'immunité innée, son intérêt en tant que cible thérapeutique est discuté en s'appuyant sur des travaux prometteurs.

1. L'immunité innée antivirale

Les cellules eucaryotes ont acquis des mécanismes de défense limitant leur invasion par des pathogènes. Lorsqu'une cellule est infectée par un organisme exogène tel qu'un virus, la première ligne de défense est constituée de mécanismes immunitaires précoces regroupés sous le nom d'immunité innée [1]. L'activation des senseurs de l'immunité innée va d'une part déclencher une cascade de signalisation permettant la synthèse d'effecteurs cellulaires antiviraux et, d'autre part, recruter les effecteurs de l'immunité dite adaptative qui engendrera une réponse de plus haute spécificité et confèrera une protection durable avec un effet mémoire.

1.1. Détection des acides nucléiques viraux

L'immunité innée est activée en réponse à la reconnaissance de motifs moléculaires associés aux pathogènes (MAMPs) par des récepteurs cellulaires appelés « pattern recognition receptors » (PRRs) [2]. Différents types de PRRs permettent de déceler une infection de la cellule par un virus à ARN. En effet, deux classes de PRRs peuvent détecter les ARNs viraux grâce à des mécanismes de détection du « non-soi » [3]: des senseurs cytosoliques ubiquitaires de la famille du « Retinoic acid Inducible Gene 1 » (RIG-I) (RLRs) [4] et certains récepteurs membranaires de la famille des « Toll-like receptors » (TLRs) exprimées plus spécifiquement dans les cellules immunitaires [5].

1.1.1. Le senseur cytosolique RIG-I

Le senseur RIG-I est une protéine cytosolique constituée de trois domaines distincts (**Figure 3**): une région N-terminale composée d'un tandem de domaines destiné au recrutement et à l'activation de caspases (CARDs), un domaine central portant une activité hélicase à ARN avec un motif caractéristique DEX(D/H) permettant de dérouler l'ARN en hydrolysant de l'ATP grâce à une activité ATPase et un domaine C-terminal (CTD) dans lequel est intégré un domaine répresseur (RD). Lorsque RIG-I n'est pas activé, les domaines CARDs sont séquestrés par le RD [6]. Quand RIG-I détecte un ARN exogène, le RD libère les domaines CARDs qui sont ubiquitinylés entraînant la tétramérisation

de RIG-I et l'induction de la cascade de signalisation pour la synthèse d'interférons (INFs) (voir 1.2.) [7-9].

Le récepteur RIG-I a, dans un premier temps, été décrit comme un facteur de reconnaissance de l'extrémité 5' triphosphate (5' p_γp_βp_α-ARN) des ARNs double brin (db), notamment les ARNs génomiques des virus à ARN négatif [10-12]. En effet, l'analyse structurale du domaine CTD de RIG-I a révélé que ce domaine interagit principalement avec les charges négatives portées par les phosphates α et β. Plus récemment, son implication a été étendue à la reconnaissance des structures coiffe [13,14] permettant ainsi de discriminer les ARNs du « soi » qui sont 2'O-méthylés des ARNs du « non-soi » qui ne sont pas 2'O-méthylés (voir 3.).



Figure 3 – Représentation structurale des « RIG-I-like receptors »

Les domaines clés impliqués dans les mécanismes de reconnaissance de l'ARN et de transduction du signal ont été représentés sous forme de boîtes. RIG-I est composé de deux CARDS (« caspase activation and recruitment domain ») à l'extrémité N-terminale, d'un domaine central portant une fonction hélicase DEX(D/H) et d'un domaine C-terminal qui contient un domaine répresseur (RD) impliqué dans l'autorégulation de RIG-I. MDA5 contient les mêmes domaines que RIG-I mais ne possède pas de RD. LGP2 est également semblable à RIG-I mais ne contient aucun CARD.

Adapté de Loo *et al.* (2011).

1.1.2. Le senseur cytosolique MDA5

MDA5, pour « Melanoma differentiation associated gene 5 », est également un facteur cytosolique qui possède une organisation en domaines semblable à celle de RIG-I mais ne contient pas de domaine d'autorégulation suggérant qu'il adopte une forme ouverte sans ligand (**Figure 3**) [4,15]. Ce senseur est activé par la présence de longs ARNs db (> 300 paires de base) et peut également discriminer les ARNs du « soi » et du « non-soi » [16]. Toutefois, les mécanismes d'interaction MDA5/ARN restent incomplètement documentés (voir 3.). Il semble que MDA5 se fixe sous forme de monomère sur des ARNs db puis s'oligomérisent en structure hélicoïdale [17] ce qui provoque un changement de conformation exposant les CARDS qui initient le processus de signalisation induisant la production d'INFs de type I (voir 1.2.) [12,18].

1.1.3. Le senseur cytosolique LGP2

Le facteur « Laboratory of Genetics and Physiology 2 » (LGP2) diffère des deux autres senseurs dans le sens où il ne possède pas de domaine CARD (**Figure 3**) et donc ne transmet pas un signal suite à la détection d'un acide nucléique exogène [4]. Cependant, les dernières données structurales montrent que son domaine hélicase lui permet d'interagir avec des ARNs exogènes au milieu de la séquence (comme MDA5) et son domaine CTD lui permet de discriminer les différentes extrémités de l'ARN (comme RIG-I) ce qui ferait de LGP2 une « chimère » entre RIG-I et MDA5 [19]. D'après la littérature récente, LGP2 favorise l'interaction MDA5-ARN db et l'oligomérisation de MDA5 à faible

concentration [20] mais a un effet antagoniste à forte concentration [21,22]. Il a donc été proposé que LGP2 soit un régulateur de MDA5 [23]. Au début de l'infection, LGP2 permettrait de favoriser la détection d'ARN viraux par MDA5 puis il servirait de rétrocontrôle négatif sur MDA5 suite à la stimulation des cellules par les INFs et à l'augmentation de l'expression de LGP2.

1.1.4. Les senseurs membranaires TLRs

Les TLRs sont des glycoprotéines transmembranaires exprimées principalement à la surface des cellules immunitaires (macrophages et cellules dendritiques notamment). Ces TLRs sont composés d'un domaine extracellulaire (ou luminal) contenant des régions riches en leucine (LRRs) impliquées dans la reconnaissance des MAMPs, un domaine transmembranaire et un domaine cytoplasmique Toll/IL-1 (TIR) responsable de la transduction du signal pour permettre la production d'INFs (voir 1.2.) [24].

Les TLRs 1, 2, 4, 5, 6 et 11 sont présents à la surface des cellules et dans l'endosome et sont responsables de la reconnaissance des composants des parois cellulaires bactériennes et fongiques [5]. Les TLRs 3, 7, 8 et 9 sont localisés dans le réseau endosomal pour détecter les acides nucléiques exogènes [25]. Ils sont impliqués dans le cadre de la réponse antivirale (**Figure 3**). En effet, le TLR3 détecte l'ARN db par interactions électrostatiques non spécifiques entre les résidus phosphates de l'ARN et les cycles imidazoles portés par 4 résidus histidines sur TLR3. Les TLRs 7 et 8 reconnaissent de l'ARN viral sb riche en GU. Le TLR 9, quant à lui, permet de déceler les motifs CpG non méthylés contenus dans l'ADN de pathogènes viraux.

Contrairement aux RLRs, la compartimentalisation des TLRs dans l'endosome fait que l'enjeu n'est pas de discriminer le « soi » du « non soi », mais de détecter des molécules qui ne devraient pas être présentes à la surface des cellules ou dans l'endosome.

1.2. Activation de la synthèse d'interféron de type I

L'activation des RLRs, par un ARN viral notamment, entraîne une cascade de signalisation permettant la production des INFs de type I (α et β) (**Figure 4**) [4, 26]. Les CARDS de RIG-I et MDA5 déclenchent la cascade de signalisation via une interaction homotypique avec le CARD du facteur de stimulation du promoteur de l'INF- β (IPS-1, également appelé MAVS, VISA ou Cardif). L'activation des CARDS d'IPS-1 entraîne leur accumulation à la surface des mitochondries ce qui active les kinases « tank binding kinase 1 » (TBK1) et « I- κ -B kinase ϵ » (IKK ϵ). Ces kinases activées IKK ϵ et TBK1 vont respectivement phosphoryler les facteurs de transcription « Interferon regulatory factors 3 et 7 » (IRF3 et IRF7) et le facteur nucléaire κ -B (NF- κ B). Ces facteurs migrent alors vers le noyau pour interagir avec les promoteurs des gènes cibles induisant la production des INFs α et β et de cytokines pro-inflammatoires.

Les TLRs ont des mécanismes de détection différents mais la transduction du signal passe par des facteurs communs aux RLRs induisant également la synthèse d'INFs (**Figure 4**) [2,27,28]. Lorsque les TLRs détectent de l'ARN viral, ceux-ci vont s'homodimériser permettant à leurs domaines TIRs de recruter des adaptateurs cytosoliques nécessaire à la transduction du signal. Pour TLR 3, l'adaptateur est le facteur « Toll-like receptor adaptor molecule 1 » (TICAM-1) qui va permettre l'activation des facteurs de transcription IRF3 et NF- κ B. Les TLRs 7 et 9 recrutent le facteur « myeloid differentiation primary response gene 88 » (MyD88) qui va induire l'activation de IRF7 et NF- κ B. Ces facteurs

intermédiaires vont par la suite induire la synthèse d'INFs de type I et de cytokines pro-inflammatoires comme les RLRs.

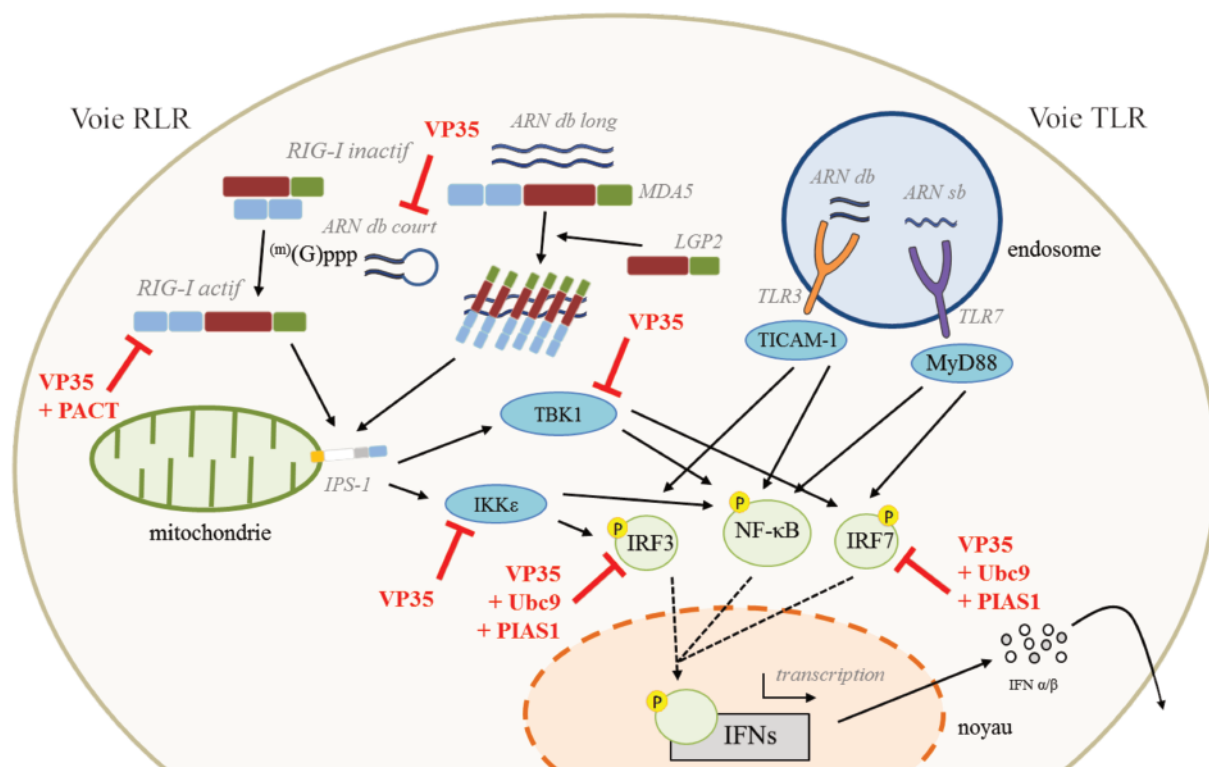


Figure 4 – Détection des acides nucléiques exogènes et stimulation de la synthèse d'interférons

L'activation de l'immunité innée se fait pas la reconnaissance de motifs moléculaires exogènes. Deux voies sont responsables de la détection des virus via la reconnaissance de leurs acides nucléiques: la voie des « RIG-I-like receptors » (RLRs) et la voie des « Toll-like receptors » (TLRs). La présence d'ARN double brin (db) court avec une extrémité 5' triphosphate active RIG-I tandis que l'ARN double brin long active MDA5. LGP2 participe à l'oligomérisation de MDA5 le long de son ARN cible. Ces deux RLRs activent la molécule adaptatrice IPS-1 située à la surface des mitochondries qui, lui, active deux kinases, IKKε et TBK1, qui activent par phosphorylation les facteurs de transcription IRF3 et IRF7, respectivement, ainsi que NFJκB. Ces facteurs migrent dans le noyau pour interagir avec le promoteur de nombreux gènes codant pour des facteurs antiviraux tels que les interférons. La présence d'ARN db ou simple brin déclenche certains récepteurs endosomaux, TLR3 et TLR7 respectivement. TLR3 active TICAM-1 qui phosphoryle IRF3 et NFJκB. TLR7 active MyD88 qui phosphoryle IRF7 et NFJκB. Ces facteurs, une fois phosphorylés, migrent dans le noyau afin d'activer la transcription des INFs. Les filovirus codent la protéine VP35 qui est capable de bloquer différentes étapes dans ses voies de signalisation en interagissant ou non avec d'autres partenaires cellulaires (PACT, protéines de SUMOylation Ubc9 et PIAS1) des protéines impliquées.

1.3. Réponse à l'interféron

Les INFs sont des glycoprotéines de la famille des cytokines qui regroupe des facteurs de régulation du système immunitaire. On distingue trois types d'INFs selon leurs propriétés et leur origine [26]. Les INFs de type I (α, β, ω, κ et ε) sont synthétisés par la plupart des cellules et ont une action antimicrobienne. L'INF de type II ou INF-γ est principalement sécrété par certaines cellules immunitaires (« natural killer cells », lymphocytes auxiliaires Th-1 et lymphocytes cytotoxiques Tc) afin d'activer et d'amplifier la réponse immunitaire cellulaire. Les INFs de type III ou INF-λ ont été décrits plus récemment et ont une action antivirale complémentaire aux INFs de type I dans les tissus les plus exposés aux infections (tube digestif, système respiratoire, etc.).

Les INFs de type I α et β sont sécrétés suite à la détection d'antigènes dans la cellule. Ils agissent de manière autocrine et paracrine grâce à leur récepteur dimérique IFNAR1/2 (Figure 5) [29,30]. La

reconnaissance ligand/récepteur conduit au recrutement de la « janus kinase 1 » (JAK1) et de la « tyrosine kinase 2 » (TYR2) qui vont phosphoryler les facteurs de transcription « signal transducers and activators of transcription 1 et 2 » (STAT1 et STAT2). Ces deux protéines forment alors un complexe hétérotrimérique appelé ISGF3 avec le facteur IRF9. Ce complexe migre au noyau pour stimuler la transcription d'« interféron stimulated genes » (ISGs). Parmi ces ISGs, les gènes codant pour différentes classes de protéines sont stimulés comme les RLRs et les TLRs, la « double-stranded RNA dependent protein kinase » (PKR), la 2'-5'-oligoadénylate synthétase (OAS), des déaminases spécifiques des adénosines d'ARN db (ADARs) [31] ou encore d'autres « interféron-induced proteins » comme les IFITs [32] qui ont un rôle dans l'inhibition précoce de la réplication du pathogène comme IFIT1 qui inhibe la traduction des ARNs non ou mal coiffés [33,34] (voir 3.).

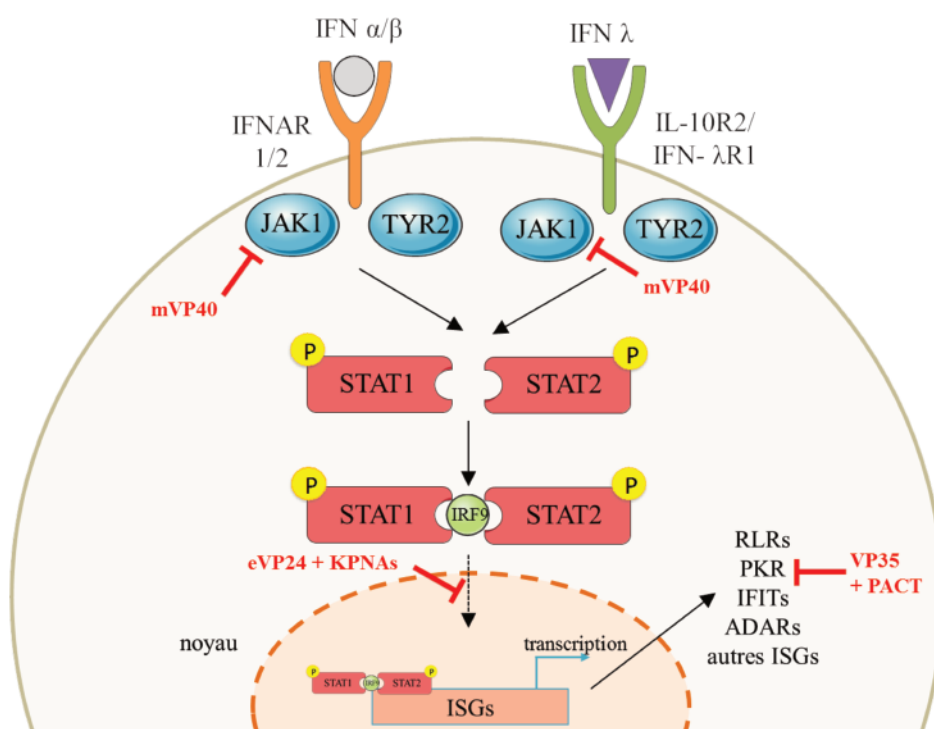


Figure 5 – Voie de synthèse d'effecteurs en réponse à une stimulation par des interférons

La liaison des interférons (INFs) de type I (α ou β) à leur récepteur IFNAR1/2 permet le recrutement et l'activation des kinases JAK1 et TYR2 qui vont ensuite phosphoryler les protéines STAT1 et STAT2. L'activation des protéines STATs conduit au recrutement du facteur IRF9 qui forme alors un complexe hétérotrimériques qui est transféré au noyau pour stimuler la transcription de gènes en réponse aux INFs (ISGs). Cette voie de signalisation est également activée par les interférons de type 3 (INFs λ) via le récepteur IL10R2/INF- λ R1. Dans ce processus, les filovirus ont développé des mécanismes d'échappement. VP24 du virus Ebola (eVP24) permet de bloquer, par compétition avec l'agent de transfert, le transfert de STAT1/2+IRF9 au noyau. VP40 du virus Marburg (mVP40) permet de bloquer le recrutement et l'activation de la protéine JAK1. La protéine VP35 de l'ensemble des filovirus est également capable de bloquer l'effecteur PKR impliquée dans la dégradation des ARNs.

Les INFs λ ont un fonctionnement très similaires aux INFs de type 1 α et β [35,36]. La différence majeure entre les deux types d'INF est le récepteur qui, dans le cas des INFs λ , est le dimère IL-10R2/IFN- λ R1 qui n'est exprimé que dans les tissus les plus exposés.

2. Les mécanismes d'échappement des filovirus à l'immunité innée

Les virus ont mis en place différentes stratégies afin de limiter leur détection par l'immunité innée et ainsi de pouvoir se répliquer. Ces stratégies sont aussi diverses qu'originales et touchent la cascade de signalisation de la voie des INFs à différents niveaux [37,38]. Chez les filovirus, plusieurs interférences avec la réponse antivirale ont été mises en évidence et reliées à certaines protéines virales permettant une réplication virale robuste (récapitulée dans les **Figures 4 et 5**).

2.1. VP35 et inhibition de la voie des senseurs cytosoliques

VP35 est une protéine multifonctionnelle servant de cofacteur de la polymérase L [39] et de facteur de virulence inhibant l'immunité innée [40]. En effet, chez EBOV et MARV, VP35 est capable d'interférer avec les processus en amont et en aval de la détection des ARNs par les RLRs. VP35 bloque en effet la voie de signalisation médiée par les RLRs et prévient la phosphorylation des facteurs IRF3 et IRF7, court-circuitant ainsi la voie de synthèse des INFs de type I [41].

2.1.1. Interactions VP35 et immunité innée

VP35 interagit avec les kinases IKK ϵ et TBK1 comme un « substrat-leurre » ce qui empêche donc le recrutement et l'activation de IRF3 en aval et entraîne la phosphorylation de VP35 [42]. VP35 est également capable d'interagir avec IRF3 ou IRF7, l'enzyme E2 Ubc9 SUMO-conjuguée, et la ligase SUMO-E3 PIAS1 pour favoriser la SUMOylation des IRFs [43,44]. Cette modification altère la cascade de signalisation en aval et donc la synthèse d'INFs.

L'inhibition de la détection des ARNs viraux par les RLRs est médiée par un domaine de fixation des ARNs db (appelé domaine d'inhibition des INFs, IID) porté en CTD de VP35 [45]. Ce domaine permettrait de recruter les ARNs db limitant ainsi leur détection par les RLRs [46-48] tels que RIG-I [49] (voir **2.1.2.**). En outre, VP35 peut interagir avec la protéine PACT. Cette protéine multifonctionnelle est impliquée dans la régulation de RIG-I et de la PKR, une protéine codée par un « interferon stimulated gene » (ISG) qui porte une activité kinase des ARNs db [50,51]. VP35 bloque l'interaction entre PACT et RIG-I *in vitro* prévenant donc l'activation de RIG-I. Cette inhibition fait intervenir le même site basique permettant l'interaction de VP35 avec les ARNs db [49], suggérant un mécanisme de régulation dépendant des concentrations de VP35, de PACT et des ARNs db.

Le domaine IID de la protéine VP35 joue donc un rôle clé pour filovirus puisqu'il permet d'inhiber différents acteurs de la détection des ARNs viraux et d'assurer les différentes fonctions de VP35 dans le cycle répliatif de ces virus. L'étude de la structure d'IID est donc un point clé pour le développement d'antiviraux.

2.1.2. Déterminants structuraux de VP35 pour la reconnaissance des ARNs db

VP35 est constitué de deux domaines : un domaine N-terminal permettant l'oligomérisation de VP35 et le domaine CTD IID. La région comprise entre les résidus 220 et 340 du domaine CTD a été définie comme la région minimale permettant le recrutement d'ARNs db [45,52]. Chez EBOV, ce domaine est composé d'un sous-domaine en hélices α et un sous-domaine en feuillets β [45]. En comparant les structures de ce domaine chez différents filovirus, une conservation structurale a été mise en évidence en lien avec certaines surfaces moléculaires parfaitement conservées [48,53-55]. Deux régions riches en acides aminés basiques ont été mises en évidence dans l'IID : le « first basic patch » (FBP) et le « central basic patch » (CBP). Ces patches sont cruciaux pour les fonctions de

VP35. En effet, le CBP, très conservé au sein des filovirus, est impliqué dans la reconnaissance d'ARN db et dans l'inhibition des voies d'activation des INFs [45,48,55]. Le FBP est, quant à lui, conservé uniquement dans les virus du genre *Ebolavirus* et semble plutôt impliqué dans la réplication virale [56].

Malgré la conservation globale de l'IID de VP35 chez les filovirus, quelques différences structurales et biochimiques sont à relever. En effet, les IIDs ne reconnaissent pas l'ARN db avec la même affinité. En effet, des résultats *in vitro* ont montré que VP35 d'EBOV peut interagir avec des ARNs db > 8 paires de base (pb) tandis que VP35 de RESTV et MARV peuvent interagir avec des ARNs db > 12 pb [53]. L'affinité de l'ARN db est plus grande pour le IID d'EBOV que celle du IID de MARV ce qui pourrait s'expliquer par un mécanisme de recrutement de l'ARN différent. En effet, il semble que VP35 d'EBOV puisse accommoder à la fois les extrémités et la chaîne d'ARN db tandis que VP35 de MARV ne reconnaît que la chaîne d'ARN db [57]. VP35 d'EBOV pourrait donc masquer les extrémités des ARNs db vis-à-vis des senseurs cytosoliques.

2.2. Subversion de la réponse à l'INF

En plus de la prévention de l'activation de l'immunité innée par VP35, les filovirus peuvent également bloquer la réponse aux INFs induisant une forte répression de la synthèse des effecteurs antiviraux cellulaires [58,59] (voir 1.3.). En effet, différentes protéines virales permettent d'interférer avec la voie JAK/STAT permettant la transduction du signal en réponse aux INFs.

Chez les virus du genre *Ebolavirus*, VP24 (noté eVP24) est impliquée dans les événements précoces d'assemblage mais également dans l'inhibition de la réponse aux INFs. En effet, eVP24 peut interagir avec 3 protéines de la famille des karyophérines α (KPNAs), KPNA 1, 5 et 6. Ces protéines se lient normalement à STAT1 afin de transporter le complexe activé STAT1/2+IRF9 au noyau. Or, eVP24 se lie sur le même site d'interaction ce qui, par compétition, bloque l'importation du complexe activé ce qui entraîne une inhibition de la synthèse des ISGs [60,61].

Contrairement à la protéine VP24 du virus Ebola, VP24 de MARV ne semble pas interagir avec les KPNAs ni affecter l'expression d'ISGs [62]. Cependant, VP40 de ce virus (mVP40) bloque spécifiquement l'activation de JAK1 [62,63]. Le mécanisme d'interférence de VP40 avec l'activation JAK1 n'a pas encore été décrit et aucune interaction directe entre mVP40 et JAK1 n'a été démontrée. Ainsi, contrairement à EBOV, MARV empêche la phosphorylation de STAT1 et STAT2 et donc la transduction du signal vers la synthèse des ISGs.

Enfin, les filovirus ont également développés des mécanismes de défense contre les ISGs produites suite à la réponse aux INFs. Deux cibles ont été identifiées à ce jour : la PKR, dont le mécanisme d'inhibition par VP35 a été succinctement présenté précédemment (voir 2.1.1.) [63,64] et dont le mécanisme moléculaire reste à démontrer, et la téthérine. Cette dernière est un facteur de restriction encodée par un ISG qui bloque l'assemblage et le bourgeonnement de certains virus en ancrant les particules virales à la surface des cellules [65,66]. Alors que le relargage de particules virales vides par surexpression de la VP40 des filovirus semble altéré par la téthérine [67,68], le relargage de particules virales infectieuses dans le contexte du virus entier montre un échappement à la téthérine, vraisemblablement grâce à la glycoprotéine de surface GP [69-71]. Les mécanismes moléculaires permettant à la GP de contrecarrer les effets de la téthérine sont encore peu documentés mais la GP ne

semble influencer ni sur l'expression de la téthérine, ni sur son association au sein des radeaux lipidiques [71-73].

Pour finir, le virus LLOV, qui a récemment été identifié chez des chauves-souris [74], présente les mêmes modes d'action contre l'immunité innée cellulaire [75]. En effet, la protéine VP24 de LLOV peut inhiber la translocation des protéines STATs grâce à son interaction avec KPNA5. En outre, VP35 inhibe la phosphorylation d'IRF3 et l'activation de la PKR. Elle semble également pouvoir lier de l'ARN db. Par contre, VP40 de LLOV ne semble pas inhiber JAK1. Ces observations suggèrent donc que les mécanismes d'échappement à l'immunité innée de LLOV se rapprochent plutôt de ceux d'EBOV ce qui confirme leur proximité génétique.

2.3. Subversion d'autres mécanismes cellulaires

Les filovirus ne bloquent pas seulement les mécanismes de l'immunité innée. En effet, il semble également que ces virus puissent s'adapter à leur hôte afin de pouvoir se répliquer et de détourner la machinerie de traduction. Les filovirus présentent certaines particularités par rapport aux autres mononégavirus. En effet, ces virus produisent des ARNm possédant de longues régions non traduites aux extrémités 5' et 3' (5'UTRs et 3'UTRs). Les 5'UTRs des ARNm de VP35, VP30, VP24 et L d'EBOV contiennent un motif AUG (uAUG) qui, pour les ARNm de VP30, VP24 et L marque le début d'une petite phase ouverte de lecture (uORF) [76]. Ces uORFs semblent réprimer la traduction des ORFs codant pour les protéines virales. Suite à un stress, le facteur d'initiation de la traduction eIF-2 α est phosphorylé entraînant une diminution globale de la traduction. Ce mécanisme permet à la fois à la cellule de se remettre du stress et de limiter la réplication virale dans le cadre d'une infection. Cependant, certains ARNm cellulaires qui contiennent des uORFs voient leur traduction amplifiée dans ces conditions [77]. De fait, un rôle similaire a été proposé pour les uORFs mises en évidence chez les filovirus. En effet, il a été montré dans un système réplicon que la traduction de la L était observée dans des conditions de stress où le facteur eIF-2 α est phosphorylé. En outre, la mutation de cette uORF de la L conduit à une diminution significative de la réplication d'EBOV *in vitro*. Ces résultats suggèrent donc que l'uORFs de l'ARNm de la L aide EBOV à maintenir la production de sa polymérase L lorsque la cellule est en état de stress comme c'est le cas lors de l'activation de l'immunité innée par les INFs [76].

Un autre exemple de mécanisme de subversion des processus cellulaires est l'interaction et la modulation de la protéine cellulaire « kelch-like ech-associated protein 1 » (Keap1) par mVP24 [78-80]. Keap1 dirige normalement l'ubiquitine ligase Cul3 sur différentes protéines tel que le facteur nucléaire « erythroid-derived 2)-like 2 » (Nrf2), un facteur de transcription responsable de l'activation des promoteurs contenant un « antioxidant response element » (ARE). Les gènes sous contrôle de ce type de promoteur permettent aux cellules de récupérer suite à un stress. En conditions physiologiques, Nrf2 est continuellement dégradée à cause de son ubiquitinylation par Keap1. Or, en situation de stress, l'interaction Nrf2-Keap1 est perdue entraînant la transcription des gènes sous promoteur ARE [81]. mVP24 interagit avec Keap1 sur le même site d'interaction que Nrf2 ce qui entraîne l'expression des gènes de stress [78]. L'intérêt de ce mécanisme dans la survie du virus n'a pas été démontré mais il a été proposé qu'il augmente la survie des cellules infectées.

En conclusion, il est donc clair que les filovirus ont développé de nombreuses stratégies leur permettant de se prémunir de leur détection par l'immunité innée et d'adapter leurs mécanismes de

réplication à leur hôte. Ces mécanismes de protection touchent tous les niveaux de la cascade de signalisation menant à la synthèse d'INfs puis d'ISGs. Les ARNs viraux portent également des modifications telles que la coiffe des ARNm qui sont impliquées dans la discrimination du « soi » et du « non-soi ».

3. Coiffe des ARNs: implication dans l'échappement de l'immunité innée et cible thérapeutique potentielle

La coiffe des ARNm est une modification co-transcriptionnelle portée à l'extrémité 5' de ces ARNs chez les eucaryotes (**Figure 6**). Cette structure chimique est constituée d'une guanosine reliée au premier nucléotide de l'extrémité 5' par une liaison 5'-5' triphosphate. Cette structure peut être méthylée en position N7 de la guanosine (coiffe-0) et aux positions 2'O des riboses des premier (N1, coiffe-1). Dans certains cas, des méthyltions supplémentaires sont observées au second (N2), et quatrième (N4) nucléotides de l'extrémité 5' de l'ARNm (appelées respectivement coiffe 2 et coiffe-4). Le degré de méthylation de la structure coiffe dépend des niveaux de complexité des eucaryotes. D'autres modifications de la coiffe ont également été décrites sur certains ARNs comme la N6 méthylation ou la triméthylguanosine de la coiffe (voir 3.1.3.).

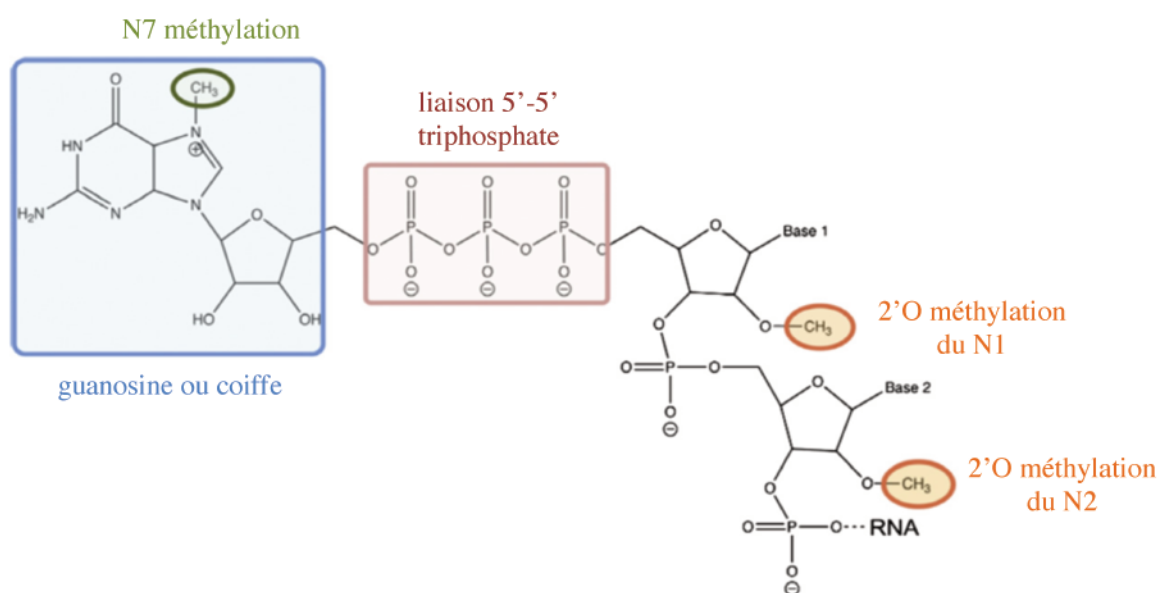


Figure 6 – Structure chimique de la coiffe des ARNs

La coiffe est une structure chimique présente à l'extrémité 5' des ARNm cellulaires et d'ARNs viraux. Cette structure est composée d'une guanosine liée à l'extrémité 5' de l'ARN par une liaison 5'-5' triphosphate. Cette guanosine est méthylée en position N7 et forme une coiffe minimale dite coiffe-0. Cette structure peut également être méthylée sur les positions 2'O du ribose du premier nucléotide (coiffe-1) et 2'O du ribose du second nucléotide de l'extrémité 5' de l'ARN (coiffe-2).

3.1. La coiffe chez les eucaryotes

3.1.1. Synthèse

La voie de synthèse de la coiffe chez les eucaryotes est dite « conventionnelle ». Elle comprend trois réactions enzymatiques chez les protozoaires et quatre chez les métazoaires (**Figure 7**). La synthèse débute par l'hydrolyse du phosphate γ à l'extrémité 5' triphosphate de l'ARN naissant ($5' p_{\gamma}p_{\beta}p_{\alpha}N1$ -ARN où N1 correspond au premier nucléotide de l'ARN) par une ARN 5' triphosphatase

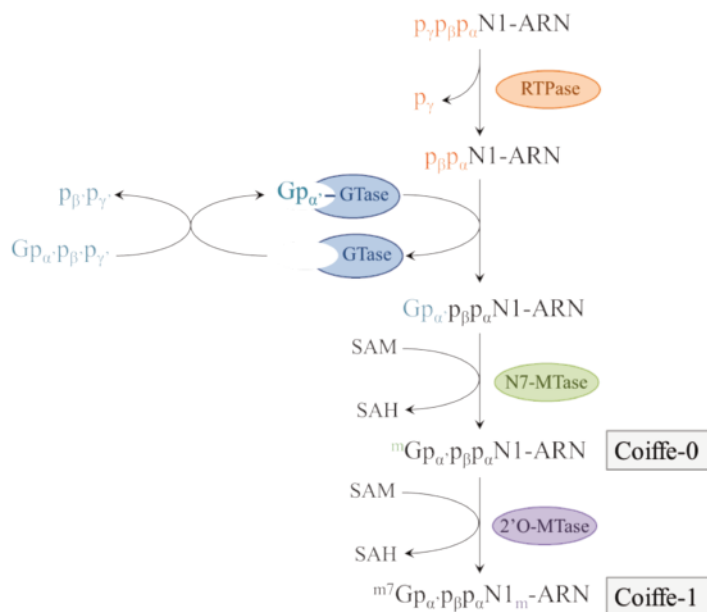


Figure 7 – Voie « conventionnelle » de synthèse de la coiffe chez les eucaryotes

La formation de la coiffe-0 à l'extrémité 5' des ARNm est assurée par trois enzymes. La triphosphatase (RTPase) hydrolyse le phosphate γ de l'ARN triphosphate ($p_\gamma p_\beta p_\alpha N1$ -ARN où N1 correspond au premier nucléotide de l'ARN). Un ARN diphosphate est ainsi généré et un phosphate inorganique est libéré. La guanylyltransférase (GTase) forme ensuite un lien covalent avec le phosphate α d'un GTP (GTase-Gp $_\alpha$) et libère du pyrophosphate. La GTase transfère ensuite le GMP sur l'ARN diphosphate pour donner un ARN dit coiffé Gp $_\alpha$ p $_\beta$ p $_\alpha$ N1-ARN. L'étape finale fait intervenir la N7-méthyltransférase (N7 MTase) qui transfère le groupement méthyle du cofacteur S-adénosyl méthionine (SAM) sur la position N7 de la guanine de la coiffe pour former la coiffe-0. Cette réaction génère le sous-produit le S-adénosyl homocystéine (SAH). Enfin, une méthylation supplémentaire est possible (chez les eucaryotes supérieurs) sur la position 2'O du ribose du premier nucléotide N1 de l'ARN grâce à une 2'O méthyltransférase (2'O MTase) conduisant à la formation d'une coiffe-1.

(RTPase), générant une extrémité 5' diphosphate ($5' p_\beta p_\alpha N1$ -ARN). Ensuite, la guanylyltransférase (GTase) lie une molécule de GMP de manière covalente sur une lysine, formant l'intermédiaire GTase-Gp $_\alpha$. Ce GMP est ensuite transféré sur l'ARN diphosphate pour former l'ARN coiffé (Gp $_\alpha$ p $_\beta$ p $_\alpha$ NGARN). Le résidu guanine de la coiffe est ensuite méthylié en position N7 par une (guanine-N7)-méthyltransférase (N7 MTase) dépendante du cofacteur S-adénosyl méthionine (SAM), générant ainsi une structure coiffe-0 (m Gp $_\alpha$ p $_\beta$ p $_\alpha$ N1-ARN). Les métazoaires catalysent des méthylations supplémentaires. En effet, les premier (N1) et deuxième (N2) nucléotides de l'ARN peuvent également être méthyliés sur la position 2'O de leur résidu ribose par une (nucléoside-2'O)-méthyltransférase (2'O MTase) permettant la synthèse de structures coiffe-1 (m Gp $_\alpha$ p $_\beta$ p $_\alpha$ N1 $_m$ -ARN) ou coiffe-2 (m Gp $_\alpha$ p $_\beta$ p $_\alpha$ N1 $_m$ N2 $_m$ -ARN). La synthèse des coiffe-0 et coiffe-1 est nucléaire tandis que la synthèse de la coiffe-2 dépend d'une 2'O MTase cytoplasmique [82,83].

3.1.2. Rôles

La coiffe joue des rôles biologiques primordiaux puisque l'inhibition des enzymes qui permettent sa synthèse est létale pour les cellules eucaryotes. En effet, cette structure est nécessaire à l'ARNm pour sa transcription complète, son épissage, son export nucléaire, sa traduction et sa stabilité. En outre, la coiffe est maintenant considérée comme un marqueur moléculaire du « soi ».

Chez les eucaryotes, l'ajout de la coiffe s'effectue uniquement sur les transcrits synthétisés par l'ARN polymérase II lorsque leur longueur atteint 20 à 25 nucléotides [84]. A ce stade, des phosphorylations spécifiques du domaine CTD de l'ARN polymérase II induit le recrutement des enzymes de synthèse de la coiffe [85]. La synthèse de la coiffe est donc régulée par un cycle de phosphorylation/déphosphorylation [86,87] qui augmente par la suite la processivité de l'ARN polymérase II.

La présence de la coiffe-0 $m^7GpppN1$ sur les ARNm régule l'efficacité d'épissage. L'épissage des ARNm est médié par le complexe de fixation de la coiffe nucléaire CBC. Le complexe CBC/coiffe favorise l'interaction entre la protéine ribonucléique nucléaire U1, snRNP U1, et le site d'épissage en 5' [88-90]. Cette interaction initie l'assemblage du spliceosome qui initiera l'épissage du premier intron [88]. Le complexe CBC reste associé à l'ARNm durant le processus d'épissage, et également lorsque l'ARNm mature quitte le spliceosome [89]. Les ARNm sont ensuite empaquetés dans des complexes ribonucléoprotéiques (RNP) permettant l'export des ARNm vers le cytoplasme, à travers les pores nucléaires [91]. Après export des ARNm dans le cytoplasme, le CBC se dissocie de l'ARNm coiffé afin de permettre la liaison du facteur d'initiation de la traduction eIF-4E.

Dans le cytoplasme, les ribosomes et les facteurs de traduction sont recrutés afin d'assurer la traduction de l'ARNm en protéine. Le facteur d'initiation de la traduction des eucaryotes eIF-4E reconnaît spécifiquement la coiffe [92]. Après reconnaissance de la coiffe, eIF-4E recrute eIF-4G et l'hélicase eIF-4A pour former le complexe eIF-4F [93] qui recrute à son tour la petite sous-unité 40S du ribosome via le complexe eIF-3 [94]. Le complexe d'initiation de la traduction « balaye » la région non-traduite dans la direction 5' vers 3' jusqu'au codon d'initiation AUG dans le contexte de séquences Kozak. Une fois le codon d'initiation atteint, les eIFs sont libérés et la sous-unité 40S recrute la sous-unité 60S pour former le complexe 80S permettant le début de l'élongation de la traduction [95].

La coiffe contribue également à la stabilité des ARNm en protégeant physiquement les ARNm de leur dégradation par des exoribonucléases 5'-3' [96,97]. La demi-vie des ARNm chez les eucaryotes peut varier de quelques minutes à quelques jours, ce qui dépend de l'efficacité de recrutement de la machinerie de dégradation des ARNm. De manière générale, la déadénylase cellulaire hydrolyse la queue poly(A) en 3' des ARNm qui peuvent ensuite être dégradés par décoiffage et digestion par l'exoribonucléase Xrn1 ou par recrutement de l'exosome [97].

Enfin, la 2'O méthylation du N1 de l'ARNm est un facteur déterminant dans la détection des ARN par les senseurs de l'immunité innée. Les ARN coiffés et dépourvus de cette méthylation sont spécifiquement reconnus par les récepteurs cytosoliques RIG-I et MDA5 conduisant à la production d'IFNs [14,16] (voir 3.3.). En outre, la protéine IFIT1, une protéine produite suite à la stimulation par les IFNs de type I (ISG), se lie spécifiquement aux ARNs non 2'O-méthylés et empêche leur traduction [34,98,99] (voir 3.3.). La 2'O méthylation apparaît donc comme un marqueur du « soi » qui évite aux ARNs cellulaires de déclencher l'immunité innée mais également de se prémunir de ses effets lors de son activation.

3.1.3. Autres types de coiffe

Il existe d'autres types de structures protectrices présentes à l'extrémité 5' de certains ARNs cellulaires. Des coiffes « γ -méthyle » ($H_3CpppARN$) ont été identifiées sur des petits ARNs nucléaires (snARNs) synthétisés par l'ARN polymérase III. Cette structure résulte de la méthylation de l'ARN triphosphorylé par une γ -MTase. Il existe également des coiffes 2,2,7 triméthylguanosine (TMG) qui sont produites après une méthylation d'une coiffe-0 par la triméthylguanosine synthase 1 (TSG1) [100]. Cette structure est notamment présente à l'extrémité 5' de quelques snARNs transcrits par l'ARN polymérase II et semble adresser les snARNs au noyau pour leur recrutement dans l'épissage des ARNm [101]. D'autres modifications supplémentaires de la structure coiffe peuvent également avoir lieu sur une coiffe-1 impliquant un résidu adénosine en N1. Dans ce cas, l'adénosine peut être

également méthylée en position N6 [102]. Le rôle de cette coiffe plus « exotiques » reste partiellement compris. Il semble que la N6 méthylation joue un rôle dans la stabilité des ARNm et par conséquent, dans leur niveau de traduction [103,104]. Les structures coiffes sont présentes sur la plupart des ARNm eucaryotes à l'exception des ARNm polyA isolés des mitochondries [105], des ARNr et des ARNt [106].

3.2. Voies de synthèse de la coiffe chez les virus

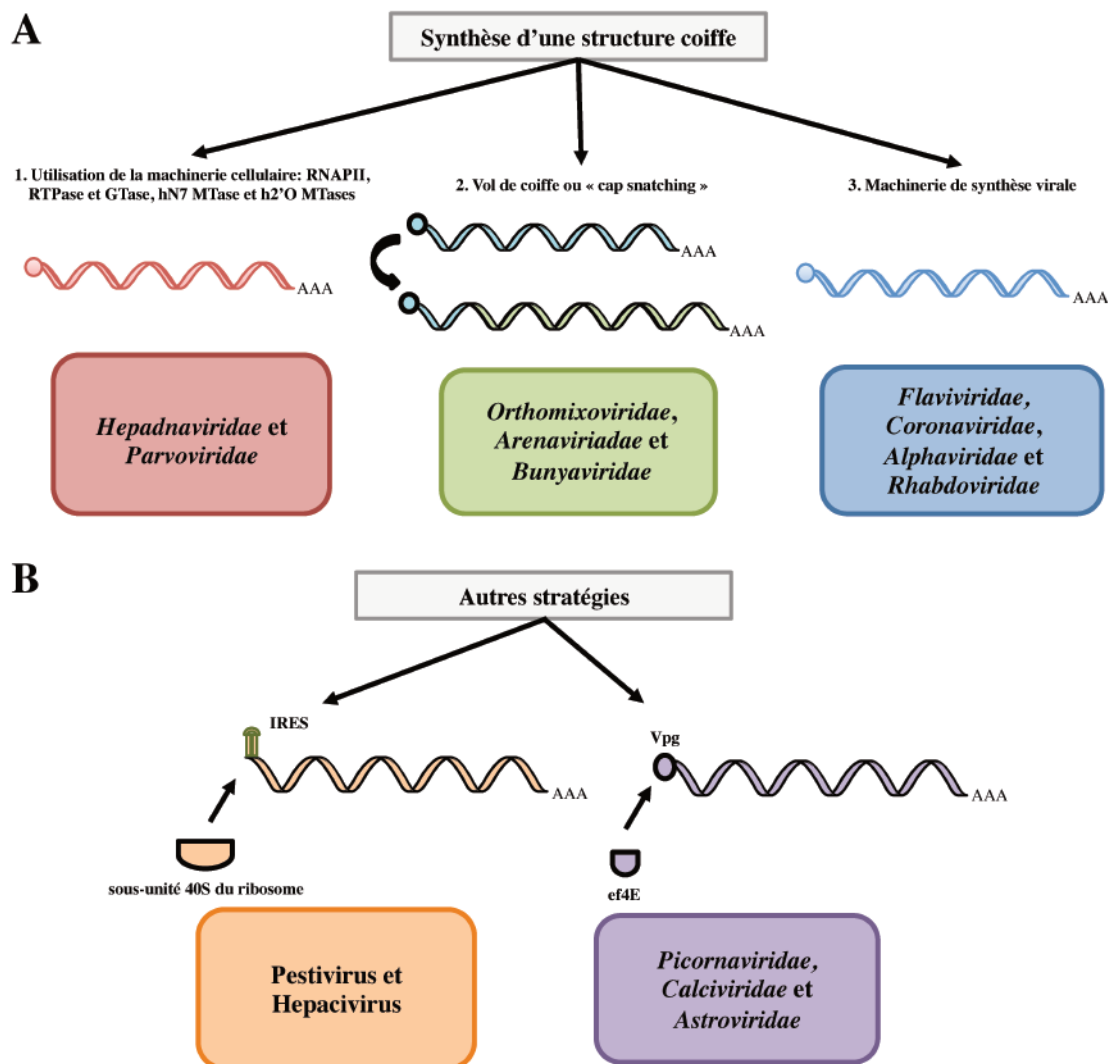


Figure 8 – Stratégies de traduction des ARNs viraux

Les virus ne codent pas pour leur propre machinerie de traduction et doivent donc détourner celle de leur hôte. (A) La voie la plus commune est la synthèse d'une coiffe semblable aux ARNs cellulaires. Différentes stratégies ont été mises en place par les virus pour synthétiser cette coiffe. Ils synthétisent leur structure coiffe en utilisant la machinerie cellulaire (plupart des virus à ADN), ils volent la structure coiffe des transcrits cellulaires ou ils codent leurs propres enzymes virales (plupart des virus à ARN). (B) Certains virus n'ont pas recours à la coiffe pour leur traduction. Par exemple, les hépacivirus et les pestivirus possèdent des « Internal ribosome entry sites » (IRES) à l'extrémité 5' de leurs ARNs, une séquence repliée en une structure spécifique reconnue par la sous-unité 40S du ribosome. Les virus de la famille des *Picornaviridae*, *Calciviridae* et *Astroviridae* ont une traduction qui dépend de la protéine VPg qui se lie de façon covalente à l'extrémité 5' du génome viral afin de recruter le facteur eIF-4E qui permet le recrutement des facteurs d'initiation de la traduction.

De nombreux virus ont développé des stratégies permettant à leur ARNs d'acquérir une structure coiffe mimant les coiffes cellulaires. Outre la protection contre les exoribonucléases, la coiffe virale

assure le détournement de la machinerie traductionnelle cellulaire pour la traduction des protéines virales. Elle permet également aux ARNs viraux d'échapper à la détection par les senseurs de l'immunité innée de l'hôte (voir 3.3.).

Les virus ont développé des voies de synthèse de la coiffe originales (**Figure 8**). A ce jour, trois stratégies d'acquisition de la coiffe par les virus ont été décrites. La première concerne les virus qui synthétisent leurs ARNm par la voie conventionnelle et utilisent l'ARN polymérase II ainsi que la machinerie de formation de la coiffe cellulaire. C'est le cas de la majorité des virus à ADN, des virus à ARN de la famille des *Retroviridae* et des *Bornaviridae*. La seconde stratégie retrouvée chez les *Orthomyxoviridae*, *Arenaviridae* et *Bunyaviridae* consiste à subtiliser les coiffes des ARNm cellulaires dans le cytoplasme. Cette stratégie de vol de coiffe ou « cap snatching » implique une activité endonucléase qui clive les ARNm coiffés cellulaires en aval de la coiffe [107-109]. Ces ARNs sont ensuite utilisés comme amorce par la polymérase virale afin de synthétiser les ARNs viraux coiffés. Enfin, la dernière stratégie consiste à coder des protéines virales dédiées à la synthèse de la coiffe. Il s'agit d'une stratégie exploitée par de nombreux virus à ARN et le virus de la vaccine, virus à ADN, qui se répliquent dans le cytoplasme des cellules infectées.

Cependant, certains virus ne coiffent pas leur ARNm (**Figure 8**). Dans ce cas, la traduction est assurée par un mécanisme indépendant de la structure coiffe. Par exemple, les *Pestivirus* et les *Hépacivirus* possèdent dans les régions 5' non traduites de leurs ARNm un site d'entrée interne au ribosome (IRES) qui recrute directement la sous-unité 40S. Un autre exemple concerne les picornavirus ou les calicivirus qui lient de manière covalente la protéine « viral protein genome-linked » (VPg) à l'extrémité 5' de leurs ARNm pour interagir avec eIF-4E [110].

3.3. Coiffe comme marqueur du « soi »

Le système de l'immunité innée initie les réponses défensives à l'infection virale via la reconnaissance des MAMPs comme des signatures du « non-soi ». Cette reconnaissance est médiée par les PRRs (TLRs et RLRs), précédemment décrits (voir 1.). Différents senseurs et effecteurs de la réponse immunitaire innée médiée par les INFs semblent capables de discriminer la coiffe des ARNs présents dans le cytoplasme et ce, grâce notamment à la 2'O méthylation du premier nucléotide.

3.3.1. Bases moléculaire de la détection des 2'O méthylations par l'immunité innée

Différents facteurs de l'immunité innée ont été montrés comme sensibles aux 2'O méthylations de la coiffe des ARNs. Cette discrimination entre des ARNs 2'O méthylés ou non se fait à la fois au niveau de la détection par les RLRs (RIG-I et MDA5) et au niveau de certains effecteurs de la réponse aux INFs (IFIT1 par exemple).

Le CTD du senseur RIG-I est capable de détecter les ARNs mal coiffés (triphosphate, coiffé ou coiffe-0) mais pas des ARNs avec une coiffe-1 (**Figure 9A**). Différentes études structurales et fonctionnelles ont révélé que ce domaine réalise un contact spécifique avec le groupement OH en position 2'O du ribose du premier nucléotide de l'ARN grâce au résidu très conservé H830 du CTD de RIG-I [10,111-114]. Ces études ont montré que des ARNs avec une coiffe-1 ne sont pas reconnus par RIG-I et n'induisent pas la synthèse d'INFs au contraire de leurs homologues aux extrémités triphosphate ou coiffe-0 [13,14]. En outre, la mutation H830A permet la détection d'ARN avec une coiffe-1. Ces résultats indiquent donc que RIG-I discrimine les ARNs 2'O-méthylés des autres ARNs par un mécanisme de reconnaissance hautement spécifique lié aux différences d'encombrement stérique

des structures coiffe-0 et coiffe-1. MDA5 est un autre senseur cytosolique également sensible aux 2'O méthylations [16]. Bien que la réplication de virus privé de 2'O MTase soit sensible à l'expression de MDA5 (voir 3.3.2.) [115], les mécanismes moléculaires permettant à MDA5 de discriminer les ARNs non 2'O méthylés restent mal compris. Etant donné que MDA a été montré comme étant un senseur de chaînes d'ARN db, une hypothèse serait que ce facteur discrimine les ARNs en détectant des 2'O méthylations internes et non sur la coiffe. En effet plusieurs 2'O MTases virales sont capables de méthyler les 2'OH des adénosines au sein des ARNs viraux [116,117].

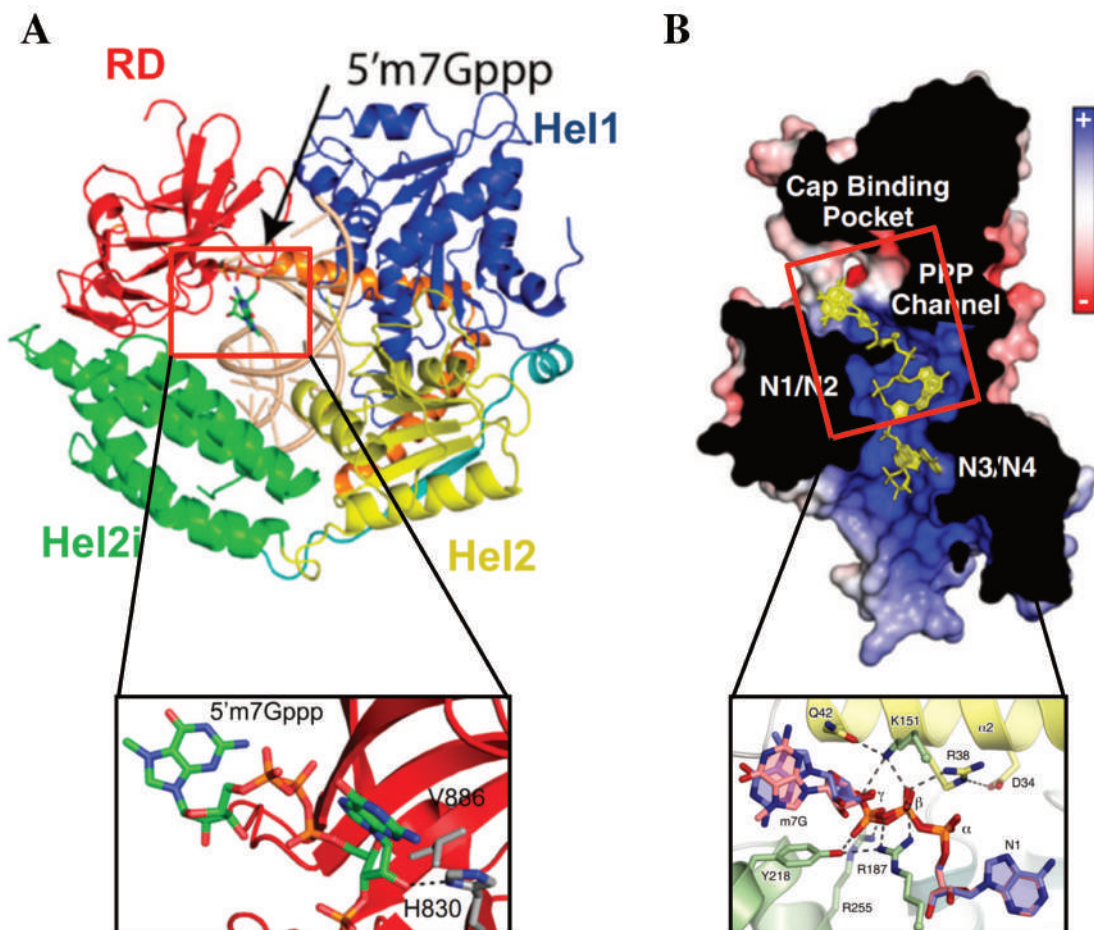


Figure 9 – Bases moléculaires de la discrimination des ARNs par la 2'O méthylation

Les structures de RIG-I et de IFIT1 en accord avec les données biochimiques montrent que ces deux protéines discriminent les ARNs exogènes des ARNs cellulaires grâce à la 2'O méthylation. (A) Le senseur cytosolique RIG-I possède un site de fixation de la coiffe qui interagit via le résidu H830 également avec le groupement 2'OH du ribose du premier nucléotide de l'ARN. La 2'O méthylation altère cette interaction et entraîne un encombrement stérique. (B) L'effecteur IFIT1, induit par une stimulation aux INFs, forme un site de fixation de la coiffe hydrophobe (blanc) et un tunnel chargé positivement (bleu) qui peut accommoder les premiers nucléotides de l'extrémité 5' de l'ARN. Dans ce tunnel, les résidus 2'OH des riboses des premiers résidus sont impliqués dans l'interaction avec IFIT1 qui, lorsqu'ils sont méthylés, diminuent (coiffe-1) voire inhibent (coiffe-2) la capacité d'IFIT1 à reconnaître l'ARN.

Suite à la détection d'ARNs mal coiffés, différents effecteurs comme les IFITs sont synthétisés en réponse aux INFs. Parmi ces IFITs, IFIT1 et son isoforme IFIT1B peuvent se fixer sur les ARNs mal coiffés détectés en amont pour inhiber leur traduction. Cette inhibition résulte de la compétition avec le recrutement d'eIF-4E [34]. Ce mécanisme a également été démontré comme étant dépendant des 2'O méthylations de la coiffe (Figure 9B) [34]. IFIT1 forme un tunnel chargé positivement avec une extension hydrophobe capable de recruter la structure coiffe-0. Plusieurs conformations de l'ARN ont

été décrites suggérant une plasticité du site de fixation de l'ARN. La 2'O méthylation du N1 peut inhiber la reconnaissance de l'ARN par IFIT1 selon sa séquence suggérant que la 2'O méthylation du N1 n'est pas suffisante pour prévenir la reconnaissance de l'ARN en présence de fortes concentrations d'IFIT1. En revanche, la méthylation supplémentaire en N2 abolit complètement cette reconnaissance de l'ARN. IFIT1 est donc un exemple supplémentaire où les 2'O méthylations de la coiffe jouent un rôle dans la discrimination du « soi » et du « non-soi » [118].

3.3.2. La 2'O méthylation de la coiffe comme stratégie d'échappement des virus

Les rôles respectifs de MDA5 et IFIT1 dans la reconnaissance des ARNs viraux portant des structures coiffe-0 ont été confirmés par des expériences de génétique inverse sur différents virus à ARN en mutant leur 2'O MTase. Chez le SARS-CoV, la mutation du site catalytique de l'activité 2'O MTase de la protéine nsp16 produit des virus dont la capacité de réplication est atténuée en culture cellulaire et *in vivo*. En revanche, dans les souris KO MDA5^{-/-} ou IFIT1^{-/-}, le SARS-CoV mutant a une virulence similaire au virus sauvage [115]. Cette étude suggère que l'absence de la 2'O méthylation en N1 des ARNs viraux active la voie de signalisation des RLRs qui, par la suite, induit la production d'IFIT1 qui inhibe la traduction virale [16, 115]. Ces observations sont corroborées par plusieurs études de la réplication des flavivirus et coronavirus (MHV) déficients en activité 2'OMTase [119]. Quant à RIG-I, son rôle de senseur des 2'O méthylation antiviral a été mis en évidence chez les flavivirus [13]. En utilisant un système réplicon du virus de la fièvre jaune (YFV) muté dans le site catalytique de la 2'O MTase, la réplication virale est diminuée dans des cellules A549 mais pas dans des cellules Vero, connues pour une déficience dans la voie de signalisation IFN de type I. En outre, les titres viraux des virus sauvage et muté sont similaires dans les cellules déficientes en RIG-I et l'expression d'IFIT1 est plus fortement induite par le mutant que par la souche sauvage [13].

Enfin il est intéressant de noter que certains virus échappent aux senseurs RLRs alors qu'ils ne sont pas capables de synthétiser des ARNs avec des structures coiffe-1. Par exemple, les alphavirus comme le virus de l'encéphalite équine vénézuélienne (VEEV) échappent à la détection par RIG-I et IFIT1 alors que leurs ARNm portent une structure coiffe-0. Il semble que l'absence de détection des ARNs viraux chez ces virus résulte de la présence d'une structure tige/boucle à l'extrémité 5' de leurs ARNs. Des mutations de l'extrémité 5' altérant cette structure tige/boucle de l'ARN viral restaurent l'induction des INFs et diminuent la virulence du virus [120].

L'ensemble de ces observations indique que les 2'O méthylations des structures coiffes sont une signature du « soi ». De nombreux virus se répliquant dans le cytoplasme des cellules de l'hôte ont évolué pour coder des 2'O MTases afin de masquer leurs ARNs aux senseurs et effecteurs de l'immunité innée antivirale. Etant donné les rôles largement décrits de la coiffe dans la stabilité des ARNm et leur traduction et la récente implication de la coiffe dans les mécanismes de l'immunité innée, il semble qu'inhiber les activités MTases virales permettrait à la fois de limiter la réplication virale et de restaurer la réponse antivirale de l'immunité innée.

3.4. Inhiber la synthèse des coiffes virales, une stratégie pertinente ?

La mise en évidence du rôle des 2'O méthylations dans l'immunité innée a également incité plusieurs groupes à tester l'efficacité de virus défectifs en 2'O MTase comme stratégie vaccinale reposant sur un virus vivant atténué. Cette stratégie a été testée chez plusieurs familles virales. L'immunisation des souris par différents flavivirus atténués en mutant les résidus catalytiques de la

2'O MTase virale protège les souris contre une infection par les souches sauvages [121,122]. Une approche de vaccination similaire a également été décrite chez le métapneumovirus humain (hMPV) qui montre que l'immunisation de souris par une souche de hMPV déficiente en activité 2'O MTase induit la production d'anticorps neutralisants et une protection durable contre le hMPV sauvage [123]. Un autre exemple d'application de cette stratégie a été réalisé chez le SARS-CoV. Les souris vaccinées par SARS-CoV sans 2'O MTase sont protégées face à une injection d'une dose létale de virus sauvage. Cette stratégie vaccinale originale donc est prometteuse et pourrait être applicable à l'ensemble des virus à ARN codant pour une 2'O MTase.

Les voies de synthèse de la coiffe chez les virus étant non conventionnelles, elles impliquent des mécanismes enzymatiques différents de ceux engagés pour la synthèse de la coiffe des ARNm cellulaires laissant penser que des inhibiteurs spécifiques puissent être développés. Récemment, d'autres groupes ont donc tenté de développer des inhibiteurs contre les MTases virales [124,125]. Différentes stratégies ont été envisagées. Des analogues du SAM avec une meilleure affinité que le SAM permettraient par compétition d'inhiber tout type d'activité SAM-dépendante [126]. Cette stratégie est cependant limitée par le fait que le site de recrutement du SAM est extrêmement conservé entre MTases cellulaires et virales. Des inhibiteurs de la SAH hydrolase altèrent les activités MTases par rétrocontrôle négatif et semblent capable d'inhiber la réplication virale [127]. La limitation de cette stratégie réside dans le fait que la SAH hydrolase régule également les MTases cellulaires SAM-dépendantes. Enfin, chez certains virus (coronavirus et virus de la vaccine), les activités MTases virales sont régulées allostériquement par d'autres protéines virales [128]. Des inhibiteurs de ces interactions protéine-protéine pourraient permettre une inhibition spécifique.

En conclusion, les filovirus possèdent un arsenal leur permettant d'échapper aux premières lignes de défense de l'hôte et ainsi de pouvoir assurer leur réplication. Les ARNm de ces virus semblent porter une coiffe-1 [129] mais les 2'O méthylations n'ont jamais été mises en évidence. Chez les filovirus, la protéine L est responsable des activités de réplication et de transcription du génome. Cette protéine semble également porter les différentes activités nécessaires pour la synthèse de la coiffe dont un domaine possède des motifs caractéristiques d'une activité MTase.

Bibliographie

- [1] : Seth RB, Sun L, Chen ZJ. Antiviral innate immunity pathways. *Cell Res.* 2006 Feb;16(2):141-7.
- [2] : Jeong E, Lee JY. Intrinsic and extrinsic regulation of innate immune receptors. *Yonsei Med J.* 2011 May;52(3):379-92.
- [3] : Schlee M, Hartmann G. Discriminating self from non-self in nucleic acid sensing. *Nat Rev Immunol.* 2016 Sep;16(9):566-80.
- [4] : Goubau D, Deddouche S, Reis e Sousa C. Cytosolic sensing of viruses. *Immunity.* 2013 May 23;38(5):855-69.
- [5] : Lester SN, Li K. Toll-like receptors in antiviral innate immunity. *J Mol Biol.* 2014 Mar 20;426(6):1246-64.
- [6] : Kowalinski I, Lunardi T, McCarthy AA, Loubser J, Brunel J, Grigorov B, Gerlier D, Cusack S. Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. *Cell.* 2011 Oct 14;147(2):423-35.
- [7] : Gack MU, Shin YC, Joo CH, Urano T, Liang C, Sun L, Takeuchi O, Akira S, Chen Z, Inoue S, Jung JU. TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature.* 2007 Apr 19;446(7138):916-920.
- [8] : Zeng W, Sun L, Jiang X, Chen X, Hou F, Adhikari A, Xu M, Chen ZJ. Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. *Cell.* 2010 Apr 16;141(2):315-30.
- [9] : Jiang X, Kinch LN, Brautigam CA, Chen X, Du F, Grishin NV, Chen ZJ. Ubiquitin-induced oligomerization of the RNA sensors RIG-I and MDA5 activates antiviral innate immune response. *Immunity.* 2012 Jun 29;36(6):959-73.
- [10] : Wang Y, Ludwig J, Schuberth C, Goldeck M, Schlee M, Li H, Juraneck S, Sheng G, Micura R, Tuschl T, Hartmann G, Patel DJ. Structural and functional insights into 5'-ppp RNA pattern recognition by the innate immune receptor RIG-I. *Nat Struct Mol Biol.* 2010 Jul;17(7):781-7.
- [11] : Weber M, Gawanbacht A, Habjan M, Rang A, Borner C, Schmidt AM, Veitinger S, Jacob R, Devignot S, Kochs G, García-Sastre A, Weber F. Incoming RNA virus nucleocapsids containing a 5'-triphosphorylated genome activate RIG-I and antiviral signaling. *Cell Host Microbe.* 2013 Mar 13;13(3):336-46.
- [12] : Bruns AM, Horvath CM. LGP2 synergy with MDA5 in RLR-mediated RNA recognition and antiviral signaling. *Cytokine.* 2015 Aug;74(2):198-206.
- [13] : Schuberth-Wagner C, Ludwig J, Bruder AK, Herzner AM, Zillinger T, Goldeck M, Schmidt T, Schmid-Burgk JL, Kerber R, Wolter S, Stümpel JP, Roth A, Bartok E, Drost C, Coch C, Hornung V, Barchet W, Kümmerer BM, Hartmann G, Schlee M. A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2'-O-Methylated Self RNA. *Immunity.* 2015 Jul 21;43(1):41-51.
- [14] : Devarkar SC, Wang C, Miller MT, Ramanathan A, Jiang F, Khan AG, Patel SS, Marcotrigiano J. Structural basis for m7G recognition and 2'-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proc Natl Acad Sci U S A.* 2016 Jan 19;113(3):596-601.
- [15] : Berke IC, Modis Y. MDA5 cooperatively forms dimers and ATP-sensitive filaments upon binding double-stranded RNA. *EMBO J.* 2012 Apr 4;31(7):1714-26.
- [16] : Züst R, Cervantes-Barragan L, Habjan M, Maier R, Neuman BW, Ziebuhr J, Szretter KJ, Baker SC, Barchet W, Diamond MS, Siddell SG, Ludewig B, Thiel V. 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 Feb;12(2):137-43.
- [17] : Berke IC, Yu X, Modis Y, Egelman EH. MDA5 assembles into a polar helical filament on dsRNA. *Proc Natl Acad Sci U S A.* 2012 Nov 6;109(45):18437-41.

- [18] : Wu B, Peisley A, Richards C, Yao H, Zeng X, Lin C, Chu F, Walz T, Hur S. Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5. *Cell*. 2013 Jan 17;152(1-2):276-89.
- [19] : 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*. 2016 May 19;62(4):586-602.
- [20] : Bruns AM, Leser GP, Lamb RA, Horvath CM. The innate immune sensor LGP2 activates antiviral signaling by regulating MDA5-RNA interaction and filament assembly. *Mol Cell*. 2014 Sep 4;55(5):771-81.
- [21] : 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*. 2012 Jan;86(2):705-17.
- [22] : 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*. 2013 Jan 11;288(2):938-46.
- [23] : Rodriguez KR, Bruns AM, Horvath CM. MDA5 and LGP2: accomplices and antagonists of antiviral signal transduction. *J Virol*. 2014 Aug;88(15):8194-200.
- [24] : Gao D, Li W. Structures and recognition modes of toll-like receptors. *Proteins*. 2017 Jan;85(1):3-9.
- [25] : Zhang Z, Ohto U, Shimizu T. Toward a structural understanding of nucleic acid-sensing Toll-like receptors in the innate immune system. *FEBS Lett*. 2017 Jul 7.
- [26] : Fensterl V, Chattopadhyay S, Sen GC. No Love Lost Between Viruses and Interferons. *Annu Rev Virol*. 2015 Nov;2(1):549-72.
- [27] : Iwasaki A, Medzhitov R. Control of adaptive immunity by the innate immune system. *Nat Immunol*. 2015 Apr;16(4):343-53.
- [28] : Satoh T, Akira S. Toll-Like Receptor Signaling and Its Inducible Proteins. *Microbiol Spectr*. 2016 Dec;4(6).
- [29] : Loo YM, Gale M Jr. Immune signaling by RIG-I-like receptors. *Immunity*. 2011 May 27;34(5):680-92.
- [30] : Fitzgerald ME, Rawling DC, Vela A, Pyle AM. An evolving arsenal: viral RNA detection by RIG-I-like receptors. *Curr Opin Microbiol*. 2014 Aug;20:76-81.
- [31] : George CX, Gan Z, Liu Y, Samuel CE. Adenosine deaminases acting on RNA, RNA editing, and interferon action. *J Interferon Cytokine Res*. 2011 Jan;31(1):99-117.
- [32] : Leung DW, Amarasinghe GK. When your cap matters: structural insights into self vs non-self recognition of 5' RNA by immunomodulatory host proteins. *Curr Opin Struct Biol*. 2016 Feb;36:133-41.
- [33] : Pichlmair A, Lassnig C, Eberle CA, Górna MW, Baumann CL, Burkard TR, Bürckstümmer T, Stefanovic A, Krieger S, Bennett KL, Rülcke T, Weber F, Colinge J, Müller M, Superti-Furga G. IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. *Nat Immunol*. 2011 Jun 5;12(7):624-30.
- [34] : Kumar P, Sweeney TR, Skabkin MA, Skabkina OV, Hellen CU, Pestova TV. Inhibition of translation by IFIT family members is determined by their ability to interact selectively with the 5'-terminal regions of cap0-, cap1- and 5'ppp- mRNAs. *Nucleic Acids Res*. 2014 Mar;42(5):3228-45.
- [35] : Wack A, Terczyńska-Dyla E, Hartmann R. Guarding the frontiers: the biology of type III interferons. *Nat Immunol*. 2015 Aug;16(8):802-9.
- [36] : Kotenko SV, Durbin JE. Contribution of type III interferons to antiviral immunity: location, location, location. *J Biol Chem*. 2017 May 5;292(18):7295-7303.
- [37] : García-Sastre A. Ten Strategies of Interferon Evasion by Viruses. *Cell Host Microbe*. 2017 Aug 9;22(2):176-184.
- [38] : Gerlier D, Lyles DS. Interplay between innate immunity and negative-strand RNA viruses: towards a rational model. *Microbiol Mol Biol Rev*. 2011 Sep;75(3):468-90.

- [39] : Martin B, Canard B, Decroly E. Filovirus proteins for antiviral drug discovery: Structure/function bases of the replication cycle. *Antiviral Res.* 2017 May;141:48-61.
- [40] : Leung DW, Prins KC, Basler CF, Amarasinghe GK. Ebola virus VP35 is a multifunctional virulence factor. *Virulence.* 2010 Nov-Dec;1(6):526-31. Epub 2010 Nov 1.
- [41] : Basler CF, Mikulasova A, Martinez-Sobrido L, Paragas J, Mühlberger E, Bray M, Klenk HD, Palese P, García-Sastre A. The Ebola virus VP35 protein inhibits activation of interferon regulatory factor 3. *J Virol.* 2003 Jul;77(14):7945-56.
- [42] : Prins KC, Cárdenas WB, Basler CF. Ebola virus protein VP35 impairs the function of interferon regulatory factor-activating kinases IKKepsilon and TBK-1. *J Virol.* 2009 Apr;83(7):3069-77.
- [43] : Chang TH, Kubota T, Matsuoka M, Jones S, Bradfute SB, Bray M, Ozato K. Ebola Zaire virus blocks type I interferon production by exploiting the host SUMO modification machinery. *PLoS Pathog.* 2009 Jun;5(6):e1000493.
- [44] : Kubota T, Matsuoka M, Chang TH, Taylor P, Sasaki T, Tashiro M, Kato A, Ozato K. Virus infection triggers SUMOylation of IRF3 and IRF7, leading to the negative regulation of type I interferon gene expression. *J Biol Chem.* 2008 Sep 12;283(37):25660-70.
- [45] : Leung DW, Ginder ND, Fulton DB, Nix J, Basler CF, Honzatko RB, Amarasinghe GK. Structure of the Ebola VP35 interferon inhibitory domain. *Proc Natl Acad Sci U S A.* 2009 Jan 13;106(2):411-6.
- [46] : Cárdenas WB, Loo YM, Gale M Jr, Hartman AL, Kimberlin CR, Martínez-Sobrido L, Saphire EO, Basler CF. Ebola virus VP35 protein binds double-stranded RNA and inhibits alpha/beta interferon production induced by RIG-I signaling. *J Virol.* 2006 Jun;80(11):5168-78.
- [47] : 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.* 2004 Oct 25;328(2):177-84.
- [48] : Leung DW, Prins KC, Borek DM, Farahbakhsh M, Tufariello JM, Ramanan P, Nix JC, Helgeson LA, Otwinowski Z, Honzatko RB, Basler CF, Amarasinghe GK. Structural basis for dsRNA recognition and interferon antagonism by Ebola VP35. *Nat Struct Mol Biol.* 2010 Feb;17(2):165-72.
- [49] : Luthra P, Ramanan P, Mire CE, Weisend C, Tsuda Y, Yen B, Liu G, Leung DW, Geisbert TW, Ebihara H, Amarasinghe GK, Basler CF. Mutual antagonism between the Ebola virus VP35 protein and the RIG-I activator PACT determines infection outcome. *Cell Host Microbe.* 2013 Jul 17;14(1):74-84.
- [50] : Patel RC, Sen GC. PACT, a protein activator of the interferon-induced protein kinase, PKR. *EMBO J.* 1998 Aug 3;17(15):4379-90.
- [51] : Kok KH, Lui PY, Ng MH, Siu KL, Au SW, Jin DY. The double-stranded RNA-binding protein PACT functions as a cellular activator of RIG-I to facilitate innate antiviral response. *Cell Host Microbe.* 2011 Apr 21;9(4):299-309.
- [52] : Reid SP, Cárdenas WB, Basler CF. Homo-oligomerization facilitates the interferon-antagonist activity of the ebolavirus VP35 protein. *Virology.* 2005 Oct 25;341(2):179-89. Epub 2005 Aug 10.
- [53] : Ramanan P, Edwards MR, Shabman RS, Leung DW, Endlich-Frazier AC, Borek DM, Otwinowski Z, Liu G, Huh J, Basler CF, Amarasinghe GK. Structural basis for Marburg virus VP35-mediated immune evasion mechanisms. *Proc Natl Acad Sci U S A.* 2012 Dec 11;109(50):20661-6.
- [54] : Bale S, Julien JP, Bornholdt ZA, Kimberlin CR, Halfmann P, Zandonatti MA, Kunert J, Kroon GJ, Kawaoka Y, MacRae IJ, Wilson IA, Saphire EO. Marburg virus VP35 can both fully coat the backbone and cap the ends of dsRNA for interferon antagonism. *PLoS Pathog.* 2012 Sep;8(9):e1002916.
- [55] : Kimberlin CR, Bornholdt ZA, Li S, Woods VL Jr, MacRae IJ, Saphire EO. Ebola virus VP35 uses a bimodal strategy to bind dsRNA for innate immune suppression. *Proc Natl Acad Sci U S A.* 2010 Jan 5;107(1):314-9.

- [56] : Prins KC, Binning JM, Shabman RS, Leung DW, Amarasinghe GK, Basler CF. Basic residues within the ebolavirus VP35 protein are required for its viral polymerase cofactor function. *J Virol.* 2010 Oct;84(20):10581-91.
- [57] : Messaoudi I, Amarasinghe GK, Basler CF. Filovirus pathogenesis and immune evasion: insights from Ebola virus and Marburg virus. *Nat Rev Microbiol.* 2015 Nov;13(11):663-76.
- [58] : Harcourt BH, Sanchez A, Offermann MK. Ebola virus selectively inhibits responses to interferons, but not to interleukin-1beta, in endothelial cells. *J Virol.* 1999 Apr;73(4):3491-6.
- [59] : Kash JC, Mühlberger E, Carter V, Grosch M, Perwitasari O, Proll SC, Thomas MJ, Weber F, Klenk HD, Katze MG. Global suppression of the host antiviral response by Ebola- and Marburgviruses: increased antagonism of the type I interferon response is associated with enhanced virulence. *J Virol.* 2006 Mar;80(6):3009-20.
- [60] : Reid SP, Valmas C, Martinez O, Sanchez FM, Basler CF. Ebola virus VP24 proteins inhibit the interaction of NPI-1 subfamily karyopherin alpha proteins with activated STAT1. *J Virol.* 2007 Dec;81(24):13469-77. Epub 2007 Oct 10.
- [61] : Xu W, Edwards MR, Borek DM, Feagins AR, Mittal A, Alinger JB, Berry KN, Yen B, Hamilton J, Brett TJ, Pappu RV, Leung DW, Basler CF, Amarasinghe GK. Ebola virus VP24 targets a unique NLS binding site on karyopherin alpha 5 to selectively compete with nuclear import of phosphorylated STAT1. *Cell Host Microbe.* 2014 Aug 13;16(2):187-200.
- [62] : Valmas C, Grosch MN, Schumann M, Olejnik J, Martinez O, Best SM, Krähling V, Basler CF, Mühlberger E. Marburg virus evades interferon responses by a mechanism distinct from ebola virus. *PLoS Pathog.* 2010 Jan 15;6(1):e1000721.
- [63] : Rodig SJ, Meraz MA, White JM, Lampe PA, Riley JK, Arthur CD, King KL, Sheehan KC, Yin L, Pennica D, Johnson EM Jr, Schreiber RD. Disruption of the Jak1 gene demonstrates obligatory and nonredundant roles of the Jaks in cytokine-induced biologic responses. *Cell.* 1998 May 1;93(3):373-83.
- [64] : Shabman RS, Gulcicek EE, Stone KL, Basler CF. The Ebola virus VP24 protein prevents hnRNP C1/C2 binding to karyopherin $\alpha 1$ and partially alters its nuclear import. *J Infect Dis.* 2011 Nov;204 Suppl 3:S904-10.
- [65] : Neil SJ, Zang T, Bieniasz PD. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. *Nature.* 2008 Jan 24;451(7177):425-30.
- [66] : Van Damme N, Goff D, Katsura C, Jorgenson RL, Mitchell R, Johnson MC, Stephens EB, Guatelli J. The interferon-induced protein BST-2 restricts HIV-1 release and is downregulated from the cell surface by the viral Vpu protein. *Cell Host Microbe.* 2008 Apr 17;3(4):245-52.
- [67] : Jouvenet N, Neil SJ, Zhadina M, Zang T, Kratovac Z, Lee Y, McNatt M, Hatzioannou T, Bieniasz PD. Broad-spectrum inhibition of retroviral and filoviral particle release by tetherin. *J Virol.* 2009 Feb;83(4):1837-44.
- [68] : Neil SJ, Sandrin V, Sundquist WI, Bieniasz PD. An interferon-alpha-induced tethering mechanism inhibits HIV-1 and Ebola virus particle release but is counteracted by the HIV-1 Vpu protein. *Cell Host Microbe.* 2007 Sep 13;2(3):193-203.
- [69] : Radoshitzky SR, Dong L, Chi X, Clester JC, Retterer C, Spurgers K, Kuhn JH, Sandwick S, Ruthel G, Kota K, Boltz D, Warren T, Kranzusch PJ, Whelan SP, Bavari S. Infectious Lassa virus, but not filoviruses, is restricted by BST-2/tetherin. *J Virol.* 2010 Oct;84(20):10569-80.
- [70] : Kaletsky RL, Francica JR, Agrawal-Gamse C, Bates P. Tetherin-mediated restriction of filovirus budding is antagonized by the Ebola glycoprotein. *Proc Natl Acad Sci U S A.* 2009 Feb 24;106(8):2886-91.
- [71] : Kühl A, Banning C, Marzi A, Votteler J, Steffen I, Bertram S, Glowacka I, Konrad A, Stürzl M, Guo JT, Schubert U, Feldmann H, Behrens G, Schindler M, Pöhlmann S. The Ebola virus glycoprotein and HIV-1 Vpu employ different strategies to counteract the antiviral factor tetherin. *J Infect Dis.* 2011 Nov;204 Suppl 3:S850-60.

- [72]: Lopez LA, Yang SJ, Exline CM, Rengarajan S, Haworth KG, Cannon PM. Anti-tetherin activities of HIV-1 Vpu and Ebola virus glycoprotein do not involve removal of tetherin from lipid rafts. *J Virol.* 2012 May;86(10):5467-80.
- [73]: Lopez LA, Yang SJ, Hauser H, Exline CM, Haworth KG, Oldenburg J, Cannon PM. Ebola virus glycoprotein counteracts BST-2/Tetherin restriction in a sequence-independent manner that does not require tetherin surface removal. *J Virol.* 2010 Jul;84(14):7243-55.
- [74]: Negredo A, Palacios G, Vázquez-Morón S, González F, Dopazo H, Molero F, Juste J, Quetglas J, Savji N, de la Cruz Martínez M, Herrera JE, Pizarro M, Hutchison SK, Echevarría JE, Lipkin WI, Tenorio A. Discovery of an ebolavirus-like filovirus in Europe. *PLoS Pathog.* 2011 Oct;7(10):e1002304.
- [75]: Feagins AR, Basler CF. Lloviu virus VP24 and VP35 proteins function as innate immune antagonists in human and bat cells. *Virology.* 2015 Nov;485:145-52.
- [76]: Shabman RS, Hoenen T, Groseth A, Jabado O, Binning JM, Amarasinghe GK, Feldmann H, Basler CF. An upstream open reading frame modulates ebola virus polymerase translation and virus replication. *PLoS Pathog.* 2013 Jan;9(1):e1003147.
- [77]: Wek RC, Jiang HY, Anthony TG. Coping with stress: eIF2 kinases and translational control. *Biochem Soc Trans.* 2006 Feb;34(Pt 1):7-11.
- [78]: Edwards MR, Johnson B, Mire CE, Xu W, Shabman RS, Speller LN, Leung DW, Geisbert TW, Amarasinghe GK, Basler CF. The Marburg virus VP24 protein interacts with Keap1 to activate the cytoprotective antioxidant response pathway. *Cell Rep.* 2014 Mar 27;6(6):1017-25.
- [79]: Pichlmair A, Kandasamy K, Alvisi G, Mulhern O, Sacco R, Habjan M, Binder M, Stefanovic A, Eberle CA, Goncalves A, Bürkstümmer T, Müller AC, Fauster A, Holze C, Lindsten K, Goodbourn S, Kochs G, Weber F, Bartenschlager R, Bowie AG, Bennett KL, Colinge J, Superti-Furga G. Viral immune modulators perturb the human molecular network by common and unique strategies. *Nature.* 2012 Jul 26;487(7408):486-90.
- [80]: Page A, Volchkova VA, Reid SP, Mateo M, Bagnaud-Baule A, Nemirov K, Shurtleff AC, Lawrence P, Reynard O, Ottmann M, Lotteau V, Biswal SS, Thimmulappa RK, Bavari S, Volchkov VE. Marburgvirus hijacks nrf2-dependent pathway by targeting nrf2-negative regulator keap1. *Cell Rep.* 2014 Mar 27;6(6):1026-36.
- [81]: Copple IM. The Keap1-Nrf2 cell defense pathway--a promising therapeutic target? *Adv Pharmacol.* 2012;63:43-79.
- [82]: Furuichi Y, Shatkin AJ. Viral and cellular mRNA capping: past and prospects. *Adv Virus Res.* 2000;55:135-84.
- [83]: Decroly E, Ferron F, Lescar J, Canard B. Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol.* 2011 Dec 5;10(1):51-65.
- [84]: Coppola JA, Field AS, Luse DS. Promoter-proximal pausing by RNA polymerase II in vitro: transcripts shorter than 20 nucleotides are not capped. *Proc Natl Acad Sci U S A.* 1983 Mar;80(5):1251-5.
- [85]: Hirose Y, Manley JL. RNA polymerase II and the integration of nuclear events. *Genes Dev.* 2000 Jun 15;14(12):1415-29.
- [86]: Sims RJ 3rd, Belotserkovskaya R, Reinberg D. Elongation by RNA polymerase II: the short and long of it. *Genes Dev.* 2004 Oct 15;18(20):2437-68.
- [87]: McCracken S, Fong N, Rosonina E, Yankulov K, Brothers G, Siderovski D, Hessel A, Foster S, Shuman S, Bentley DL. 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* 1997 Dec 15;11(24):3306-18.
- [88]: Izaurralde E, Lewis J, Gamberi C, Jarmolowski A, McGuigan C, Mattaj JW. A cap-binding protein complex mediating U snRNA export. *Nature.* 1995 Aug 24;376(6542):709-12.

- [89] : Lewis JD, Izaurralde E, Jarmolowski A, McGuigan C, Mattaj IW. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev.* 1996 Jul 1;10(13):1683-98.
- [90] : Nagai K. RNA-protein complexes. *Curr Opin Struct Biol.* 1996 Feb;6(1):53-61.
- [91] : Izaurralde E, Mattaj IW. RNA export. *Cell.* 1995 Apr 21;81(2):153-9.
- [92] : Fischer PM. Cap in hand: targeting eIF4E. *Cell Cycle.* 2009 Aug 15;8(16):2535-41.
- [93] : Hentze MW. eIF4G: a multipurpose ribosome adapter? *Science.* 1997 Jan 24;275(5299):500-1.
- [94] : Sonenberg N. eIF4E, the mRNA cap-binding protein: from basic discovery to translational research. *Biochem Cell Biol.* 2008 Apr;86(2):178-83.
- [95] : Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol.* 2010 Feb;11(2):113-27.
- [96] : Beelman CA, Parker R. Degradation of mRNA in eukaryotes. *Cell.* 1995 Apr 21;81(2):179-83.
- [97] : Liu H, Kiledjian M. Decapping the message: a beginning or an end. *Biochem Soc Trans.* 2006 Feb;34(Pt 1):35-8.
- [98] : Daffis S, Szretter KJ, Schriewer J, Li J, Youn S, Errett J, Lin TY, Schnell S, Zust R, Dong H, Thiel V, Sen GC, Fensterl V, Klimstra WB, Pierson TC, Buller RM, Gale M Jr, Shi PY, Diamond MS. 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature.* 2010 Nov 18;468(7322):452-6.
- [99] : Hyde JL, Diamond MS. Innate immune restriction and antagonism of viral RNA lacking 2'-O methylation. *Virology.* 2015 May;479-480:66-74.
- [100] : Chang J, Schwer B, Shuman S. Mutational analyses of trimethylguanosine synthase (Tgs1) and Mud2: proteins implicated in pre-mRNA splicing. *RNA.* 2010 May;16(5):1018-31.
- [101] : Gupta S, Busch RK, Singh R, Reddy R. Characterization of U6 small nuclear RNA cap-specific antibodies. Identification of gamma-monomethyl-GTP cap structure in 7SK and several other human small RNAs. *J Biol Chem.* 1990 Nov 5;265(31):19137-42.
- [102] : Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, Linder B, Pickering BF, Vasseur JJ, Chen Q, Gross SS, Elemento O, Debart F, Kiledjian M, Jaffrey SR. Reversible methylation of m6Am in the 5' cap controls mRNA stability. *Nature.* 2017 Jan 19;541(7637):371-375.
- [103] : Lin S, Choe J, Du P, Triboulet R, Gregory RI. The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Mol Cell.* 2016 May 5;62(3):335-45.
- [104] : Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, Linder B, Pickering BF, Vasseur JJ, Chen Q, Gross SS, Elemento O, Debart F, Kiledjian M, Jaffrey SR. Reversible methylation of m6Am in the 5' cap controls mRNA stability. *Nature.* 2017 Jan 19;541(7637):371-375.
- [105] : Grohmann K, Amairic F, Crews S, Attardi G. Failure to detect "cap" structures in mitochondrial DNA-coded poly(A)-containing RNA from HeLa cells. *Nucleic Acids Res.* 1978 Mar;5(3):637-51.
- [106] : Furuichi Y. Discovery of m(7)G-cap in eukaryotic mRNAs. *Proc Jpn Acad Ser B Phys Biol Sci.* 2015;91(8):394-409.
- [107] : Dias A, Bouvier D, Crépin T, McCarthy AA, Hart DJ, Baudin F, Cusack S, Ruigrok RW. The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit. *Nature.* 2009 Apr 16;458(7240):914-8.
- [108] : Morin B, Coutard B, Lelke M, Ferron F, Kerber R, Jamal S, Frangeul A, Baronti C, Charrel R, de Lamballerie X, Vonrhein C, Lescar J, Bricogne G, Günther S, Canard B. The N-terminal domain of the arenavirus L protein is an RNA endonuclease essential in mRNA transcription. *PLoS Pathog.* 2010 Sep 16;6(9):e1001038.

- [109] : Reguera J, Weber F, Cusack S. Bunyaviridae RNA polymerases (L-protein) have an N-terminal, influenza-like endonuclease domain, essential for viral cap-dependent transcription. *PLoS Pathog.* 2010 Sep 16;6(9):e1001101.
- [110] : Goodfellow I, Chaudhry Y, Gioldasi I, Gerondopoulos A, Natoni A, Labrie L, Laliberté JF, Roberts L. Calicivirus translation initiation requires an interaction between VPg and eIF 4 E. *EMBO Rep.* 2005 Oct;6(10):968-72.
- [111] : Luo D, Ding SC, Vela A, Kohlway A, Lindenbach BD, Pyle AM. Structural insights into RNA recognition by RIG-I. *Cell.* 2011 Oct 14;147(2):409-22.
- [112] : Lu C, Xu H, Ranjith-Kumar CT, Brooks MT, Hou TY, Hu F, Herr AB, Strong RK, Kao CC, Li P. The structural basis of 5' triphosphate double-stranded RNA recognition by RIG-I C-terminal domain. *Structure.* 2010 Aug 11;18(8):1032-43.
- [113] : Vela A, Fedorova O, Ding SC, Pyle AM. The thermodynamic basis for viral RNA detection by the RIG-I innate immune sensor. *J Biol Chem.* 2012 Dec 14;287(51):42564-73.
- [114] : Kohlway A, Luo D, Rawling DC, Ding SC, Pyle AM. Defining the functional determinants for RNA surveillance by RIG-I. *EMBO Rep.* 2013 Sep;14(9):772-9.
- [115] : Menachery VD, Yount BL Jr, Josset L, Gralinski LE, Scobey T, Agnihothram S, Katze MG, Baric RS. Attenuation and restoration of severe acute respiratory syndrome coronavirus mutant lacking 2'-O-methyltransferase activity. *J Virol.* 2014 Apr;88(8):4251-64.
- [116] : Dong H, Chang DC, Hua MH, Lim SP, Chionh YH, Hia F, Lee YH, Kukkaro P, Lok SM, Dedon PC, Shi PY. 2'-O methylation of internal adenosine by flavivirus NS5 methyltransferase. *PLoS Pathog.* 2012;8(4):e1002642.
- [117] : Coutard B, Barral K, Lichière J, Selisko B, Martin B, Aouadi W, Lombardia MO, Debart F, Vasseur JJ, Guillemot JC, Canard B, Decroly E. Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. *J Virol.* 2017 Feb 14;91(5). pii: e02202-16.
- [118] : Abbas YM, Laudénbach BT, Martínez-Montero S, Cencic R, Habjan M, Pichlmair A, Damha MJ, Pelletier J, Nagar B. Structure of human IFIT1 with capped RNA reveals adaptable mRNA binding and mechanisms for sensing N1 and N2 ribose 2'-O methylations. *Proc Natl Acad Sci U S A.* 2017 Mar 14;114(11):E2106-E2115.
- [119] : Diamond MS. IFIT1: A dual sensor and effector molecule that detects non-2'-O methylated viral RNA and inhibits its translation. *Cytokine Growth Factor Rev.* 2014 Oct;25(5):543-50.
- [120] : Hyde JL, Gardner CL, Kimura T, White JP, Liu G, Trobaugh DW, Huang C, Tonelli M, Paessler S, Takeda K, Klimstra WB, Amarasinghe GK, Diamond MS. A viral RNA structural element alters host recognition of nonself RNA. *Science.* 2014 Feb 14;343(6172):783-7.
- [121] : Züst R, Dong H, Li XF, Chang DC, Zhang B, Balakrishnan T, Toh YX, Jiang T, Li SH, Deng YQ, Ellis BR, Ellis EM, Poidinger M, Zolezzi F, Qin CF, Shi PY, Fink K. Rational design of a live attenuated dengue vaccine: 2'-O-methyltransferase mutants are highly attenuated and immunogenic in mice and macaques. *PLoS Pathog.* 2013;9(8):e1003521.
- [122] : Li SH, Dong H, Li XF, Xie X, Zhao H, Deng YQ, Wang XY, Ye Q, Zhu SY, Wang HJ, Zhang B, Leng QB, Zuest R, Qin ED, Qin CF, Shi PY. Rational design of a flavivirus vaccine by abolishing viral RNA 2'-O methylation. *J Virol.* 2013 May;87(10):5812-9.
- [123] : Zhang Y, Wei Y, Zhang X, Cai H, Niewiesk S, Li J. Rational design of human metapneumovirus live attenuated vaccine candidates by inhibiting viral mRNA cap methyltransferase. *J Virol.* 2014 Oct;88(19):11411-29.
- [124] : Ferron F, Decroly E, Selisko B, Canard B. The viral RNA capping machinery as a target for antiviral drugs. *Antiviral Res.* 2012 Oct;96(1):21-31.

[125] : Decroly E, Canard B. Biochemical principles and inhibitors to interfere with viral capping pathways. *Curr Opin Virol.* 2017 Jun;24:87-96.

[126] : Aouadi W, Eydoux C, Coutard B, Martin B, Debart F, Vasseur JJ, Contreras JM, Morice C, Quérat G, Jung ML, Canard B, Guillemot JC, Decroly E. Toward the identification of viral cap-methyltransferase inhibitors by fluorescence screening assay. *Antiviral Res.* 2017 Aug;144:330-339.

[127] : De Clercq E, Cools M, Balzarini J, Marquez VE, Borcharding DR, Borchardt RT, Drach JC, Kitaoka S, Konno T. Broad-spectrum antiviral activities of neplanocin A, 3-deazaneplanocin A, and their 5'-nor derivatives. *Antimicrob Agents Chemother.* 1989 Aug;33(8):1291-7.

[128] : Aouadi W, Blanjoie A, Vasseur JJ, Debart F, Canard B, Decroly E. Binding of the Methyl Donor S-Adenosyl-l-Methionine to Middle East Respiratory Syndrome Coronavirus 2'-O-Methyltransferase nsp16 Promotes Recruitment of the Allosteric Activator nsp10. *J Virol.* 2017 Feb 14;91(5).

[129] : Weik M, Modrof J, Klenk HD, Becker S, Mühlberger E. Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *J Virol.* 2002 Sep;76(17):8532-9.

CHAPITRE 6

La protéine L chez les virus de l'ordre des *Mononegavirales*

L'ensemble des revues précédemment présentées aux **Chapitres 2, 3 et 4** ont permis de préciser les connaissances actuelles sur la relation structure/fonction des différentes protéines des filovirus et les stratégies thérapeutiques possibles. Le **Chapitre 5** a permis de montrer que les filovirus possèdent de nombreuses voies d'échappement à l'immunité innée. La coiffe des ARNm viraux constitue une de ces stratégies mais est également impliquée dans la protection de ces ARNm et dans leur traduction. Cibler la synthèse de la coiffe virale est donc une stratégie pertinente pour le développement d'une stratégie antivirale. Chez les virus de l'ordre des *Mononegavirales*, la coiffe est synthétisée par la protéine L. Cette protéine est très conservée chez les mononégavirus mais encore peu caractérisée chez les filovirus. Elle porte des activités essentielles au cycle de réplication/transcription ce qui en fait une cible attractive pour le développement de stratégies antivirales. Au laboratoire, nous avons donc choisi de nous intéresser à cette protéine L qui porterait les activités de synthèse de la coiffe chez les filovirus. Ce chapitre s'attache à faire l'état des lieux des connaissances actuelles des aspects structuraux et fonctionnels de la protéine L de ces virus. La majorité des travaux disponibles à ce jour ont été réalisés sur le virus modèle VSV mais des études plus récentes sur d'autres mononégavirus étayent le modèle proposé chez VSV. Après une brève présentation de l'ordre des *Mononegavirales*, nous aborderons donc les aspects structuraux de la protéine L puis nous détaillerons ses différentes activités enzymatiques.

1. Généralités sur les mononégavirus

Les virus de l'ordre des *Mononegavirales* sont des virus à ARN simple brin (sb) monoségmenté de polarité négative (**Figure 10**). Cet ordre viral est divisé en différentes familles : les rhabdovirus (*Rhabdoviridae*), les pneumovirus (*Pneumoviridae*), les paramyxovirus (*Paramyxoviridae*), les filovirus (*Filoviridae*), les bornavirus (*Bornaviridae*), les nouvellement décrits nyavirus (*Nyaviridae*), mymonavirus (*Mymonaviridae*), et sunvirus (*Sunviridae*), et d'autres virus non assignés. Chaque famille est subdivisée en différents genres puis en différentes « espèces » virales. Plusieurs virus d'intérêt médical font partie de ces familles comme la rougeole (MeV, paramyxovirus), les oreillons (MuV, paramyxovirus), la rage (RABV, rhabdovirus), le métapneumovirus humain (hMPV, pneumovirus), le virus respiratoire syncytial humain (hRSV, pneumovirus) ou les différentes souches du virus Ebola (filovirus). En terme général, une espèce type a été étudiée pour chaque genre viral afin de définir ceux-ci et de dégager des caractéristiques de famille. Dans ce travail, l'intérêt a été porté sur plusieurs de ces virus : SUDV pour les filovirus, RABV pour les rhabdovirus, hRSV et hMPV pour les pneumovirus, et le virus de la stomatite vésiculaire (VSV, rhabdovirus) comme virus de référence.

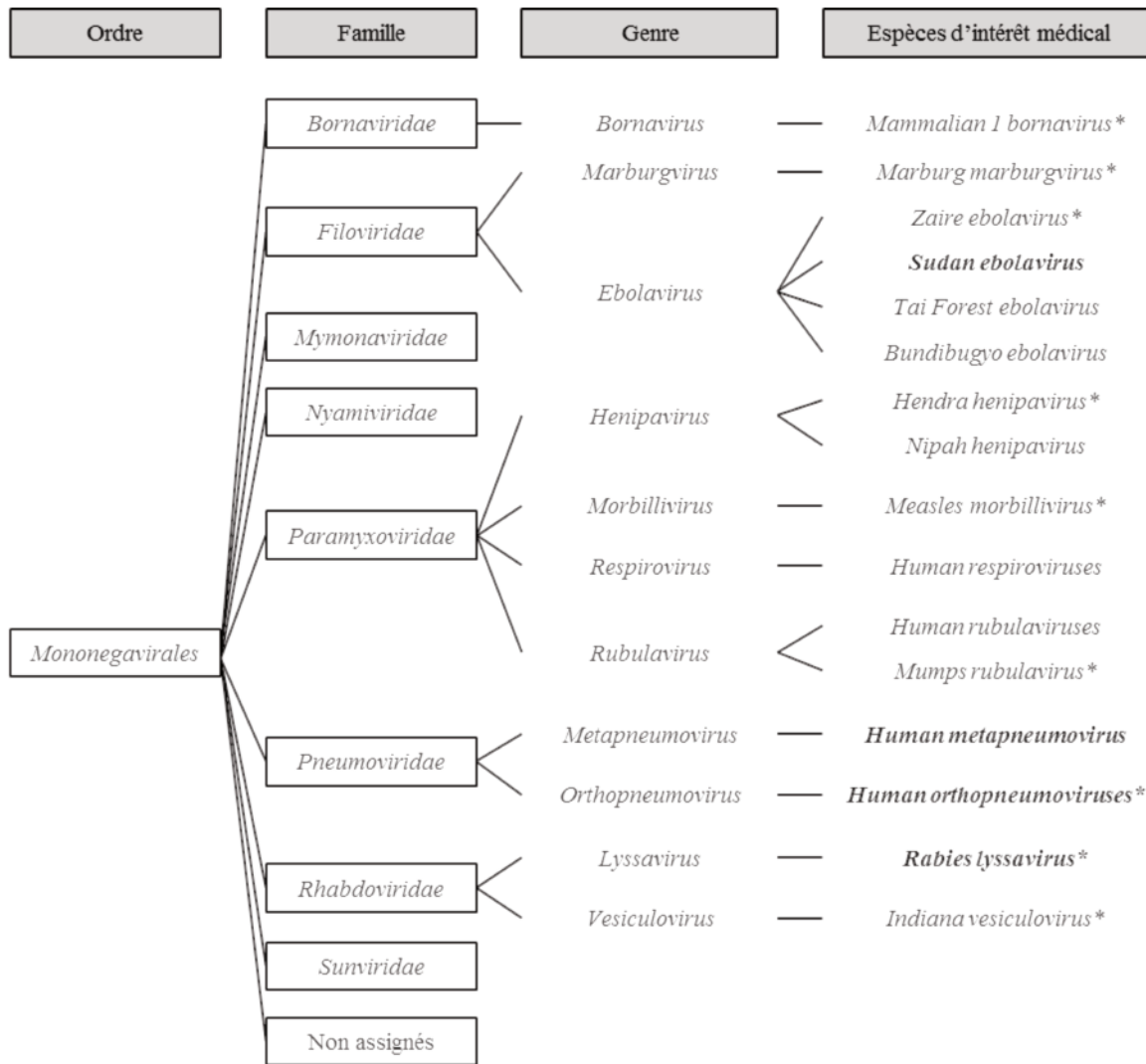


Figure 10 – Classification phylogénétique des virus de l'ordre des *Mononegavirales*

L'ordre des *Mononegavirales* regroupe des virus dont le génome est constitué d'un ARN négatif simple brin monosegmenté. Cet ordre est divisé en différentes familles qui sont elles mêmes subdivisées en genre viraux. Chaque genre peut avoir une ou plusieurs « espèces » virales (non détaillées ici). Pour chaque genre, un virus type est en particulier étudié pour dégager les caractéristiques du genre (*) auquel il appartient. La classification présentée ici n'est pas exhaustive et replace les virus d'intérêt médical majeur dans cette classification. Les virus abordés dans ce manuscrit sont indiqués en gras.

Classification officielle du comité international de taxonomie des virus (<https://talk.ictvonline.org>)

Ces virus ayant tous un génome ARN sb monosegmenté de polarité négatif, leur cycle réplcatif est semblable à celui des filovirus. Brièvement, l'ARN génomique est encapsidé sur toute sa longueur par la nucléoprotéine (N), formant un complexe ribonucléoprotéique (RNP) associé au complexe réplcatif constitué de la protéine L et de son cofacteur, la phosphoprotéine P. Durant la synthèse d'ARN, la N libère vraisemblablement l'ARN afin de permettre au complexe réplcatif d'accéder à celui-ci pour répliquer le génome viral et le transcrire en ARNm. Lors du cycle de réplcation/transcription, les ARNm coiffés, méthylés et polyadénylés produits sont directement disponibles pour leur traduction en protéines via la machinerie cellulaire alors que les ARNs génomiques néosynthétisées sont encapsidés par N au fur et à mesure de leur synthèse.

La protéine L est la protéine la plus conservée chez les virus de l'ordre des *Mononegavirales* [1-3]. Cette conservation résulte probablement du fait que la protéine L possède des activités essentielles à la réplication de ces virus : l'ARN polymérase ARN-dépendante (RdRp) nécessaire à la réplication et à la transcription de l'ARN génomique et les activités de modifications post-transcriptionnelles des ARNs messagers viraux (polyadénylation à l'extrémité 3', et synthèse et méthylations de la structure coiffe à l'extrémité 5') [4-6]. Il est donc important de caractériser les activités enzymatiques de cette protéine L pour mieux comprendre le cycle de réplication et de transcription des mononégavirus et ainsi de développer des stratégies antivirales efficaces contre ces virus.

2. Etude bioinformatique et structurale de la protéine L

La protéine L est une « méga » protéine d'environ 200-250 kDa. Il s'agit certes de la protéine la plus conservée mais elle est aussi la plus faiblement exprimée durant le cycle viral. L'étude des protéines L des virus de l'ordre des *Mononegavirales* est ardue car ce sont des protéines difficiles à surexprimer dans les systèmes d'expression actuels, notamment en raison de leur taille importante. C'est pourquoi la résolution de sa structure et son étude biochimique *in vitro* sont de vrais challenges techniques et scientifiques.

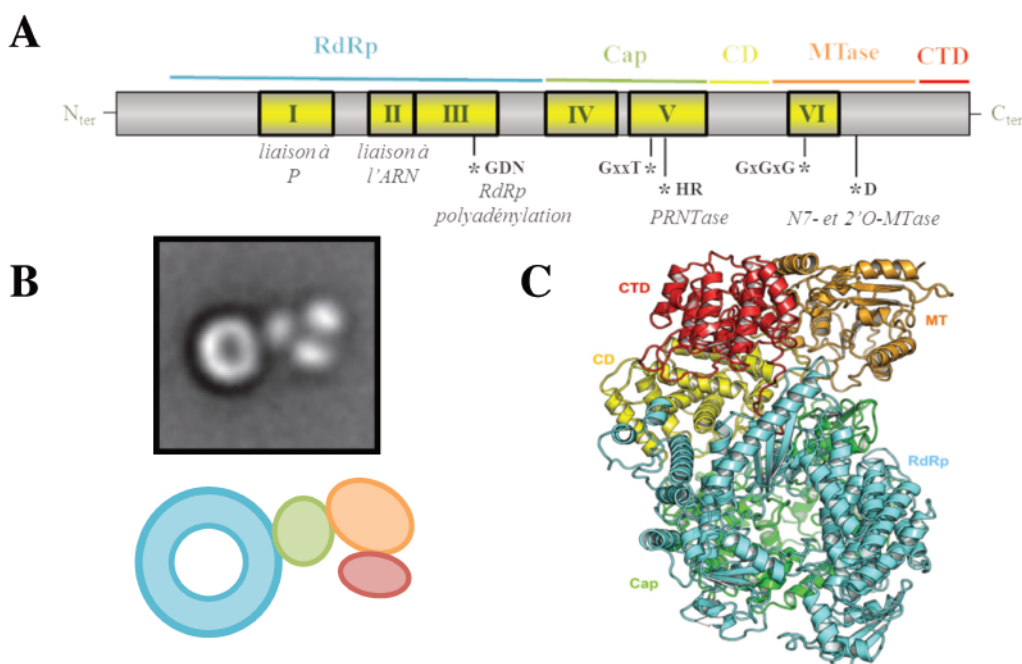


Figure 11 – Structure de la protéine L des mononégavirus

La caractérisation de la structure de la protéine L des mononégavirus a été un processus long et périlleux. (A) En 1990, Poch *et al.* ont mis en évidence différentes régions conservées (I à VI) dans la séquence des protéines L des mononégavirus dont certains motifs (en gras) portent les différentes activités de la protéine L (en italique). (B) En 2010, Rahmeh *et al.* ont, pour la première fois, réussi à exprimer et purifier une L entière (VSV), leur permettant d'avoir les premiers clichés par microscopie électronique à coloration négative des différents domaines de cette protéine. Par protéolyse ménagée, ils ont proposé un modèle selon lequel l'anneau central serait l'ARN polymérase ARN-dépendante (RdRp) et les trois petits domaines porteraient les activités de synthèse de la coiffe. (C) En 2015, grâce aux progrès de la microscopie, Liang *et al.* ont finalement réussi le tour de force de résoudre la structure de la L du VSV par cryomicroscopie électronique (PDB: 5A22). Cette nouvelle structure a permis de reconsidérer les éléments antérieurs et d'apporter un nouveau modèle. L'anneau est en fait constitué de la RdRp et du domaine d'ajout de la coiffe (Cap) et les trois modules portent un domaine connecteur (CD), un domaine méthyltransférase de la coiffe (MTase) et un petit domaine C-terminal (CTD). Bien que les rôles des domaines CD et CTD ne soient pas connus à ce jour, cette structure apporte une vision bien plus claire de la relation structure/fonction des différentes activités portées par la L chez tous les mononégavirus.

En 1990, l'étude d'alignements de séquence de la protéine L de différents mononégavirus a mis en évidence l'existence de six régions conservées (CRs, I à VI) (**Figure 11**) [1, 7]. Ces CRs ont été par la suite associés aux différentes activités portées par L (voir ci-dessous) [8-16]. Ces régions conservées sont séparées les unes des autres par des régions de plus faible homologie de séquence qui constituent des « régions charnières » entre les CRs. Plus récemment, des expériences de transcomplémentation de domaines de la protéine L ont suggéré que cette protéine est organisée en domaines structuraux indépendants reliés par des régions flexibles [17].

Lorsqu'il a été possible de purifier la L entière du VSV, des expériences de microscopie électronique en coloration négative ont confirmé l'organisation de la L en différents domaines structuraux indépendants (**Figure 11**) [18]. Son architecture globale est constituée d'un anneau (« donut ») auquel sont liés trois domaines globulaires flexibles. L'étude de ces différents éléments architecturaux par protéolyse ménagée a permis de proposer un premier modèle postulant que l'anneau porterait l'activité RdRp et les trois globules C-terminaux seraient responsables des activités de synthèse de la coiffe. L'interaction de la L avec son cofacteur P a également été décrite par microscopie électronique. La L adopte de manière surprenante une architecture en « 6 » en présence de P où les trois domaines globulaires forment un coude en prolongement de l'anneau principal. Ce réarrangement suggère que la L existe sous différentes conformations et que celles-ci assurent la coordination entre les activités de polymérisation et les activités enzymatique de synthèse de la coiffe.

En 2015, avec l'avènement de la cryomicroscopie électronique (cryo-EM), la structure de la protéine L du VSV a été résolue à une résolution de 3,8 Å [19]. Cette structure révèle une organisation des domaines fonctionnels différente (**Figure 11**), dans laquelle l'anneau décrit en microscopie électronique à coloration négative est constitué de deux domaines fonctionnels distincts intimement intriqués, la RdRp en forme d'anneau et le domaine « Cap » ou polyribonucleotidyltransférase (PRNTase), responsable de l'ajout de la coiffe. Les domaines modulaires reliés au premier module annulaire sont bien au nombre de trois mais ce qui était suggéré comme le domaine Cap est en fait un domaine connecteur (CD). La fonction du CD reste mal comprise à ce jour. Au contraire, les deux autres modules C-terminaux restent identifiés comme étant le domaine MTase et un petit domaine CTD de séquence et de longueur variable chez les mononégavirus. Alors que les domaines RdRp, Cap et MTase ont fait l'objet d'une caractérisation biochimique permettant de mieux comprendre leurs fonctions (voir ci-dessous), les domaines CD et CTD ont été décrits comme des éléments de stabilisation de la structure de L et de l'interaction L-P [19].

3. Fonctionnalités de la polymérase

Le domaine RdRp forme une structure dite en main droite composée de trois sous-domaines, « doigts », « paume » et « pouce », qui est commune à d'autres ARN ou ADN polymérases (**Figure 12A**) [19]. Ce domaine, intimement lié au domaine Cap, forme une structure en anneau identifiable en microscopie électronique. Cette conformation lui permet d'assurer le passage de l'ARN durant la polymérisation.

La RdRp contient les trois premières CRs. La CRI est composée principalement de résidus hydrophobes qui interviendraient dans l'interaction avec la protéine P [20- 24]. La CRII, qui possède une grande proportion d'acides aminés chargés positivement, serait impliquée dans l'interaction avec l'ARN pour le positionner dans le site actif de la RdRp [25]. L'intervalle régulier entre résidus

basiques et hydrophobes tous les quatre résidus suggère en outre que cette région d'interaction serait conformée en hélice α amphiphile [1]. La CRIII porte le cœur de l'activité polymérase avec l'identification des motifs A, B, C et D, caractéristiques des RdRp (**Figure 12B**) [26]. Le motif C, enfoui dans le tunnel formé par la « paume », contient notamment la triade catalytique GDN responsable de la formation des liaisons phosphodiesteres entre nucléotides chez les virus de l'ordre des *Mononegavirales* [27]. Lors de la transcription, cette CRIII catalyse également l'activité de polyadénylation qui implique un mécanisme de « glissement » de la polymérase au niveau d'une séquence de 7 uridines à la fin de chaque gène [8]. Après la polyadénylation du transcrit amont, la transcription du gène aval peut commencer avec une efficacité d'environ 70% ce qui se conduit à la synthèse une quantité décroissante d'ARNm correspondant à chaque protéine à mesure que l'on s'éloigne de l'extrémité 5' du génome [28]. Au contraire, lors de la réplication, les éléments en *cis* régulant la transcription sont ignorés par la L qui synthétise alors une copie antisens du brin matrice qui est encapsidée par la N au fur et à mesure de son élongation. Les mécanismes de régulation du « switch » entre transcription et réplication restent encore mal compris mais semblent liés à la concentration des différentes protéines virales.

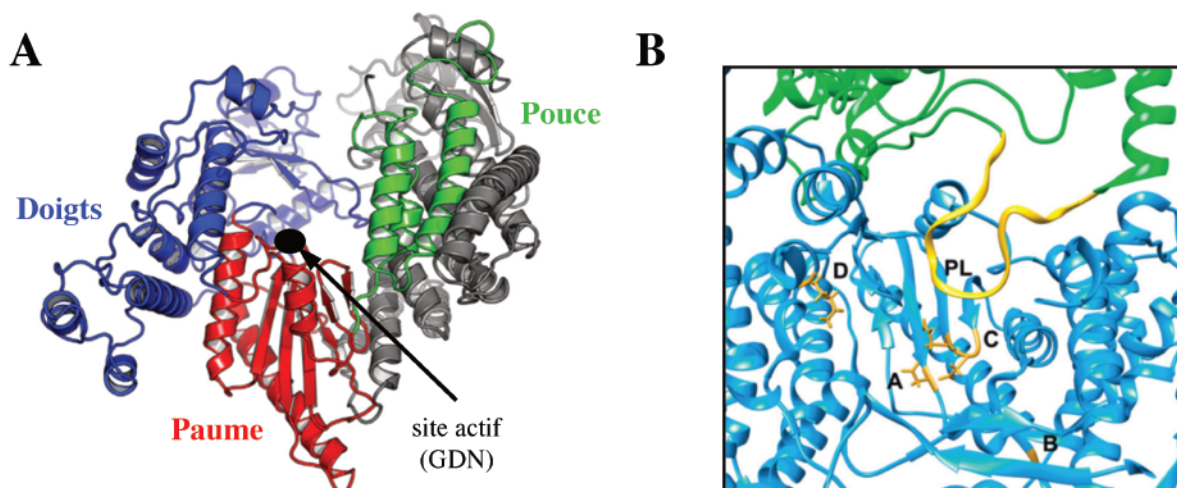


Figure 12 – Structure du domaine ARN polymérase ARN-dépendante du VSV

La protéine L possède les différentes activités permettant la synthèse de nouvelles copies du génome viral et la transcription des ARNm codant pour les différentes protéines virales. (A) L'activité ARN polymérase ARN-dépendante qui permet de polymériser l'ARN lors de la réplication et la transcription du génome est portée par un domaine dit en « main droite » caractérisé par les trois sous-domaines « doigts » (en bleu), « paume » (en rouge) et « pouce » (en vert) (PDB: 5A22). Trois régions de cette polymérase sont très conservées au sein de l'ordre des *Mononegavirales*, les régions CRI, CRII et CRIII. Les CRI et CRII sont des motifs d'interaction avec la phosphoprotéine P et l'ARN, respectivement. La CRIII contient en revanche le site actif de la polymérase constitué du motif GDN, spécifique des mononégavirus. (B) Le domaine polymérase (en bleu) est intimement lié au domaine Cap (en vert) (PDB: 5A22). La CRIII possède 4 motifs conservés A, B, C et D qui permettent la polymérisation de l'ARN. Adjacent à ces motifs, une boucle du domaine Cap (en jaune) se projette dans le site actif de la polymérase. Elle servirait de boucle d'initiation pour la polymérisation. Adapté de Liang B *et al.* (2015) et de Fearn R *et al.* (2017).

4. Ajout de la coiffe par la PRNTase

Le domaine Cap est responsable de l'ajout de la coiffe grâce à son activité PRNTase portée par la CRV (**Figure 11**). Ce domaine contient également la CRIV mais son rôle n'a pas encore été élucidé.

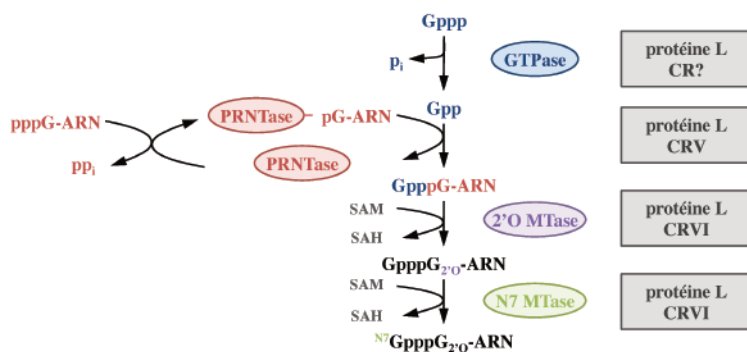


Figure 13 – Voie de synthèse de la coiffe chez les mononégavirus

La protéine L porte l'activité ARN polymérase ARN-dépendante mais aussi les activités nécessaires à la formation de la coiffe des ARNm. La formation de la coiffe chez le VSV n'est pas conventionnelle et semble suivre un consensus au sein des virus de l'ordre des *Mononegavirales*. Pour commencer, du GTP est hydrolysé en GDP par une GTPase encore non identifiée. Dans le même temps, l'ARN, qui est triphosphorylé à l'extrémité 5', se lie de façon covalente à la protéine L au niveau de la région conservée CRV où il est ensuite transféré sur la molécule de GDP qui forme alors la coiffe. L'ARN coiffé est ensuite méthylé grâce au co-substrat donneur de méthyle, la S-adénosyl méthionine (SAM) en position 2'O du ribose du premier nucléotide de l'ARN puis en position N7 de la guanosine de la coiffe par le domaine méthyltransférase de la L qui porte la CRVI. La coiffe est alors complète, permettant une traduction efficace des ARNm, une protection des 5' exonucléases cellulaires et un moyen d'échapper à la détection des senseurs de l'immunité innée.

Le GDP formé se fixe par des interactions de faible énergie sur le motif GXXT et l'extrémité 5' de l'ARN est transférée par attaque nucléophile sur le GDP formant ainsi la coiffe [14, 29, 33].

Comme mentionné précédemment, ce domaine est intimement lié à la polymérase. Ce domaine possède une boucle de 16 acides aminés qui se projette dans le site actif de la RdRp [19]. Étant donné que l'initiation de la synthèse des ARNs par le domaine RdRp se fait sans amorce, cette boucle pourrait servir de matrice d'initiation [19]. Il est également vraisemblable que cette boucle joue ensuite un rôle dans le contrôle de la fidélité de réplication puisque, pendant l'élongation, cette boucle reculerait dans le sillon de l'ARN naissant retardant l'élongation en cas d'erreur [19]. Ce retard permettrait alors la pyrophosphorolyse du nucléotide incorporé par erreur par l'ATP. Enfin, cette boucle étant adjacente au motif GXXT, il est également

L'activité PRNTase est propre aux mononégavirus et catalyse l'ajout de la coiffe de façon non conventionnelle (Figure 13) [14, 29]. Son site catalytique est formé de deux motifs très conservés, GXXT et HR, dans la CRV (Figure 14). Chez le VSV la réaction de synthèse de la coiffe à l'extrémité 5' de l'ARN naissant implique la formation d'un lien covalent entre l'histidine du motif HR et l'ARN monophosphate [30, 31]. Une molécule de GTP est hydrolysée en GDP de manière concomitante. Le domaine de la protéine L portant cette activité guanosine 5' triphosphatase (GTPase) n'est pas encore clairement identifié mais ferait partie des domaines C-terminaux, MTase et CTD (Figure 11) [32].

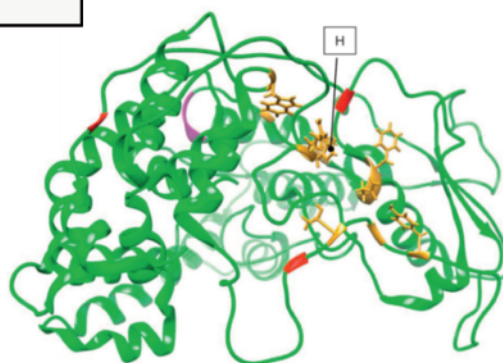


Figure 14 – Structure du domaine Cap du VSV

Le domaine Cap (en vert) (PDB: 5A22) catalyse la synthèse de la coiffe grâce à l'activité polyribonucléotidyltransférase (PRNTase). Cette activité possède un motif HR (en jaune) dans la région conservée CRV qui permet la fixation covalente de l'ARNm natif sur l'histidine. L'ARN est ensuite transféré sur une molécule de GDP contenue dans une poche grâce au motif GXXT (en jaune) pour former la coiffe. Tiré de Fearn R *et al.* (2017).

possible que l'ajout de la coiffe soit couplé à l'initiation de la polymérisation ou régule son passage en mode élongation. De fait, lorsque l'ajout de la coiffe est altéré, la transcription est prématurément interrompue [10]. En outre, chez VSV, cette réaction ne se fait *in vitro* que sur des ARNs d'au moins 31 nucléotides suggérant un remaniement structural de la protéine L permettant la poursuite de l'élongation et les modifications co-transcriptionnelles [34].

Cette hypothèse est supportée par exemple par le mode d'action des imidazoisoquinoline-diones qui inhibent l'ajout de la coiffe chez hRSV. Les mutations de résistance à ces composés localisées sur la protéine L corroborent l'hypothèse de remaniement et d'accessibilité au domaine Cap après le début de la polymérisation [35]. Un autre modèle alternatif de synthèse des structures coiffes a également été proposé chez le virus de la peste bovine (ou Rinderpest virus, RPV) où le GTP est hydrolysé en GMP qui est alors fixé de façon covalente sur la protéine L grâce à un domaine guanylyl transférase (GTase) [36]. Le GMP serait ensuite transféré sur l'ARN diphosphate généré à partir d'un ARN natif triphosphate hydrolysé par une ARN triphosphatase [36].

5. Activités méthyltransférases de la coiffe

Le domaine MTase possède un repliement Rossmann (alternance de brins β et d'hélices α) caractéristique des MTases catalysant le transfert de méthyl à partir de S-adenosyl méthionine (SAM) (**Figure 15A et 15B**) [19, 32, 37]. Les structures récemment décrites des MTases de VSV et de hMPV montrent que ces domaines sont formés d'une alternance de 7 brins β et de 4 hélices α . En plus de son repliement caractéristique, le motif GXGXG, très conservé au sein de l'ordre des *Mononegavirales* dans la CRVI, participe au recrutement du SAM dans la poche catalytique [38, 39]. Cette poche contient également la tétrade catalytique K-D-K-E caractéristique des 2'O MTases. Toutefois, les études biochimiques du domaine MTase des mononégavirus montrent que les deux activités N7- et 2'O MTases sont catalysées par ce même site actif [32, 40]. Ce domaine MTase est donc responsable de la méthylation des structures coiffes en position N7 de la guanosine et sur le 2'OH du ribose du premier nucléotide. Il semblerait que, chez VSV et hMPV, la séquence des réactions des méthylations soient non conventionnelle (**Figure 13**). Alors que chez les eucaryotes la 2'O méthylation suit la N7 méthylation, des expériences biochimiques indiquent que 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, facilitée par la 2'O méthylation de l'ARN, chez les mononégavirus. Ces méthylations, ainsi que l'ajout de la coiffe, nécessitent *in vitro* au moins 10 nucléotides dont la séquence est spécifique à chaque mononégavirus [32, 41].

Des expériences de génétique inverse ont démontré que, lorsque l'activité MTase est inhibée, les ARNs viraux sont hyperpolyadénylés ce qui suggère que l'activité MTase semble participer au contrôle de la polymérisation [10, 42, 43]. La structure actuellement disponible de la protéine L ne permet pas de comprendre comment l'ARN a accès au site actif de la MTase car celui-ci est masqué par l'encombrement des domaines CD et CTD [19]. La protéine L subirait donc vraisemblablement des changements de conformation importants entre l'initiation et l'élongation, permettant ainsi à l'ARN naissant d'être complètement mûré.

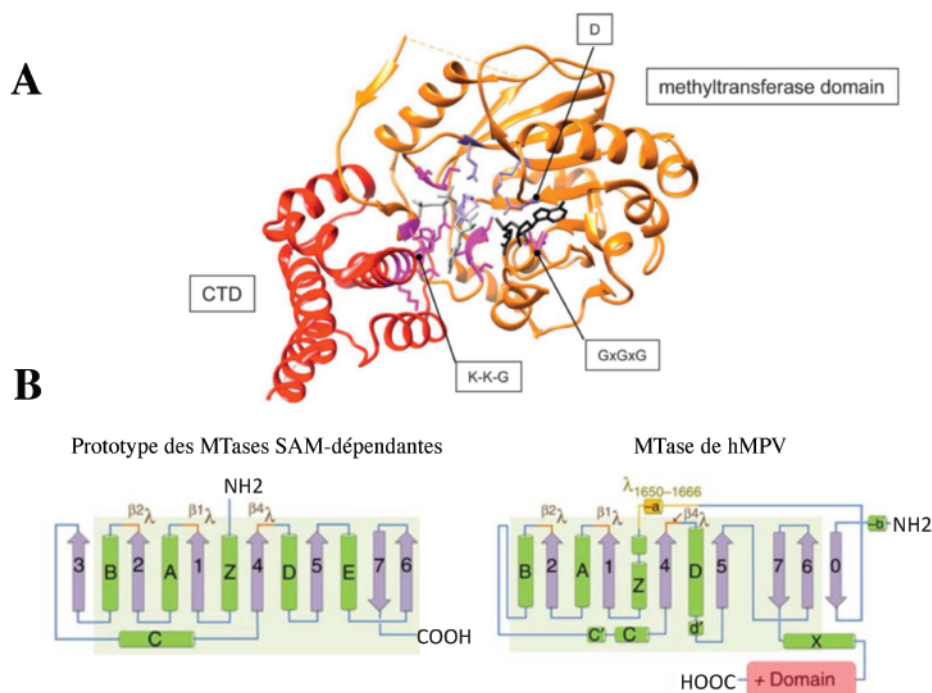


Figure 15 – Structure du domaine méthyltransférase de la protéine L de hMPV

La caractérisation structure/fonction du domaine méthyltransférase (MTase) de la protéine L des mononégavirus a été réalisée chez le hMPV (pneumovirus). (A) La structure du domaine MTase (en orange) et du domaine C-terminal (CTD) (en rouge) a été résolue à une résolution de 2,2 Å par diffraction aux rayons X (PDB: 4UCK). Ce domaine seul est capable de méthylater la coiffe des ARNm viraux sur la position 2'O du premier nucléotide de l'extrémité 5' de l'ARN puis sur la position N7 de la guanine de la coiffe grâce au seul site actif contenant le motif K-D-K-E (en gris) et un donneur de méthyle, la S-adenosyl méthionine (SAM), recrutée grâce au motif GXGXG (en noir). L'activité GTPase couplée à l'activité PRNTase du domaine Cap serait portée par le motif KKG (en violet). (B) Les MTases (SAM)-dépendantes ont un repliement dit de « Rossmann » qui correspond à l'alternance entre brins β (gris) et hélices α (en vert) entre lesquels certaines boucles (en orange) sont responsables du recrutement du SAM, $\beta^1\lambda$ et $\beta^2\lambda$, et de l'ARN substrat, $\beta^4\lambda$. Le domaine MTase porté par la protéine L du hMPV présente un repliement homologue suggérant donc un mécanisme conservé pour la méthylation des ARNm. Adapté de Fearn R *et al.* (2017) et Paesen GC *et al.* (2015).

6. Conjecture des remaniements structuraux

L'interconnectivité des différents domaines de la L dans la structure résolue chez VSV suggère que les activités portées par ceux-ci seraient couplées spécialement et temporellement grâce à des remaniements conformationnels importants (voir ci-dessus) [44]. Tous les mononégavirus codent également pour le cofacteur P de la polymérase mais certains virus comme les filovirus ou les pneumovirus (hRSV) ont des cofacteurs supplémentaires (VP30 et M2-1, respectivement). Ces partenaires sont nécessaires à la processivité de la polymérase. Collectivement, l'interconnexion des domaines et les interactions avec les cofacteurs pourraient opérer de larges remaniements structuraux affectant les différentes activités enzymatiques de la L. Etant donné la conservation et la divergence de certaines régions de la L à travers l'ordre des *Mononegavirales*, il est probable que ces remaniements structuraux soient communs chez les mononégavirus avec des mécanismes de régulation plus ou moins spécifiques.

En conclusion, la protéine L est indispensable pour le cycle viral des mononégavirus. Les activités PRNTase et MTases portées par la L catalysent chez les mononégavirus la formation d'ARNm coiffés et méthylés de façon non conventionnelle. Cette structure permet de protéger les ARNm des 5' exonucléases cellulaires et est également impliquée dans la traduction des ARNm et dans les mécanismes d'échappement des mononégavirus aux senseurs de l'immunité innée. Ce sont des rôles clefs dans le cycle viral, ce qui fait des activités PRNTase et MTases de la L des cibles originales et pertinentes pour le développement d'antiviraux. Cependant, ces activités restent encore peu documentées et n'ont pas encore été caractérisées chez les filovirus.

Bibliographie

- [1] : 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.
- [2] : Stec DS, Hill MG 3rd, Collins PL. Sequence analysis of the polymerase L gene of human respiratory syncytial virus and predicted phylogeny of nonsegmented negative-strand viruses. *Virology.* 1991 Jul;183(1):273-87.
- [3] : Mühlberger E, Sanchez A, Randolph A, Will C, Kiley MP, Klenk HD, Feldmann H. The nucleotide sequence of the L gene of Marburg virus, a filovirus: homologies with paramyxoviruses and rhabdoviruses. *Virology.* 1992 Apr;187(2):534-47.
- [4] : Emerson SU, Wagner RR. L protein requirement for *in vitro* RNA synthesis by vesicular stomatitis virus. *J Virol.* 1973 Dec;12(6):1325-35.
- [5] : Hunt DM, Smith EF, Buckley DW. Aberrant polyadenylation by a vesicular stomatitis virus mutant is due to an altered L protein. *J Virol.* 1984 Nov;52(2):515-21.
- [6] : Hercyk N, Horikami SM, Moyer SA. The vesicular stomatitis virus L protein possesses the mRNA methyltransferase activities. *Virology.* 1988 Mar;163(1):222-5.
- [7] : 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.* 1997 Apr;48(1):57-70.
- [8] : Barr JN, Whelan SP, Wertz GW. cis-Acting signals involved in termination of vesicular stomatitis virus mRNA synthesis include the conserved AUAC and the U7 signal for polyadenylation. *J Virol.* 1997 Nov;71(11):8718-25.
- [9] : Li J, Fontaine-Rodriguez EC, Whelan SP. Amino acid residues within conserved domain VI of the vesicular stomatitis virus large polymerase protein essential for mRNA cap methyltransferase activity. *J Virol.* 2005 Nov;79(21):13373-84.
- [10] : Li J, Rahmeh A, Brusic V, Whelan SP. Opposing effects of inhibiting cap addition and cap methylation on polyadenylation during vesicular stomatitis virus mRNA synthesis. *J Virol.* 2009 Feb;83(4):1930-40.
- [11] : Morin B, Kranzusch PJ, Rahmeh AA, Whelan SP. The polymerase of negative-stranded RNA viruses. *Curr Opin Virol.* 2013 Apr;3(2):103-10.
- [12] : Li J, Rahmeh A, Morelli M, Whelan SP. 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.* 2008 Jan;82(2):775-84.
- [13] : Sleat DE, Banerjee AK. Transcriptional activity and mutational analysis of recombinant vesicular stomatitis virus RNA polymerase. *J Virol.* 1993 Mar;67(3):1334-9.
- [14] : Ogino T, Banerjee AK. Unconventional mechanism of mRNA capping by the RNA-dependent RNA polymerase of vesicular stomatitis virus. *Mol Cell.* 2007 Jan 12;25(1):85-97.
- [15] : 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.* 2005 Feb 11;280(6):4429-35.
- [16] : Gopinath M, Shaila MS. Evidence for N⁷ guanine methyl transferase activity encoded within the modular domain of RNA-dependent RNA polymerase L of a Morbillivirus. *Virus Genes.* 2015 Dec;51(3):356-60.
- [17] : Dochow M, Krumm SA, Crowe JE Jr, Moore ML, Plemper RK. Independent structural domains in paramyxovirus polymerase protein. *J Biol Chem.* 2012 Feb 24;287(9):6878-91.
- [18] : Rahmeh AA, Schenk AD, Danek EI, Kranzusch PJ, Liang B, Walz T, Whelan SP. Molecular architecture of the vesicular stomatitis virus RNA polymerase. *Proc Natl Acad Sci USA.* 2010 Nov 16;107(46):20075-80.
- [19] : Liang B, Li Z, Jenni S, Rahmeh AA, Morin BM, Grant T, Grigorieff N, Harrison SC, Whelan SP. Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy. *Cell.* 2015 Jul 16;162(2):314-27.
- [20] : 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.* 1995 Nov 10;213(2):352-63.

- [21] : Smallwood S, Moyer SA. The L polymerase protein of parainfluenza virus 3 forms an oligomer and can interact with the heterologous Sendai virus L, P and C proteins. *Virology*. 2004 Jan 5;318(1):439-50.
- [22] : 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*. 2004 Oct 1;327(2):297-306.
- [23] : Cevik B, Smallwood S, Moyer SA. Two N-terminal regions of the Sendai virus L RNA polymerase protein participate in oligomerization. *Virology*. 2007 Jun 20;363(1):189-97.
- [24] : 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. *J Virol*. 2015 Oct 28;90(2):715-24.
- [25] : Smallwood S, Easson CD, Feller JA, Horikami SM, Moyer SA. Mutations in conserved domain II of the large (L) subunit of the Sendai virus RNA polymerase abolish RNA synthesis. *Virology*. 1999 Sep 30;262(2):375-83.
- [26] : Poch O, Sauvaget I, Delarue M, Tordo N. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J*. 1989 Dec 1;8(12):3867-74.
- [27] : Chattopadhyay A, Raha T, Shaila MS. Effect of single amino acid mutations in the conserved GDNQ motif of L protein of Rinderpest virus on RNA synthesis *in vitro* and *in vivo*. *Virus Res*. 2004 Feb;99(2):139-45.
- [28] : Iverson LE, Rose JK. Localized attenuation and discontinuous synthesis during vesicular stomatitis virus transcription. *Cell*. 1981 Feb;23(2):477-84.
- [29] : 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 USA*. 2010 Feb 23;107(8):3463-8.
- [30] : Koonin EV, Moss B. Viruses know more than one way to don a cap. *Proc Natl Acad Sci USA*. 2010 Feb 23;107(8):3283-4.
- [31] : 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*. 2016 Jan 8;44(1):330-41.
- [32] : Paesen GC, Collet A, Sallamand C, Debart F, Vasseur JJ, Canard B, Decroly E, Grimes JM. X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat Commun*. 2015 Nov 9;6:8749.
- [33] : Ogino T, Banerjee AK. Formation of guanosine(5')tetraphospho(5')adenosine cap structure by an unconventional mRNA capping enzyme of vesicular stomatitis virus. *J Virol*. 2008 Aug;82(15):7729-34.
- [34] : Tekes G, Rahmeh AA, Whelan SP. A freeze frame view of vesicular stomatitis virus transcription defines a minimal length of RNA for 5' processing. *PLoS Pathog*. 2011 Jun;7(6):e1002073.
- [35] : Liuzzi M, Mason SW, Cartier M, Lawetz C, McCollum RS, Dansereau N, Bolger G, Lapeyre N, Gaudette Y, Lagacé L, Massariol MJ, Dô F, Whitehead P, Lamarre L, Scouten E, Bordeleau J, Landry S, Rancourt J, Fazal G, Simoneau B. Inhibitors of respiratory syncytial virus replication target cotranscriptional mRNA guanylation by viral RNA-dependent RNA polymerase. *J Virol*. 2005 Oct;79(20):13105-15.
- [36] : Gopinath M, Shaila MS. RNA triphosphatase and guanylyl transferase activities are associated with the RNA polymerase protein L of rinderpest virus. *J Gen Virol*. 2009 Jul;90(Pt 7):1748-56.
- [37] : Galloway SE, Richardson PE, Wertz GW. Analysis of a structural homology model of the 2'-O-ribose methyltransferase domain within the vesicular stomatitis virus L protein. *Virology*. 2008 Dec 5;382(1):69-82.
- [38] : 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.
- [39] : 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.

[40] : Rahmeh AA, Li J, Kranzusch PJ, Whelan SP. 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.* 2009 Nov;83(21):11043-50.

[41] : Wang JT, McElvain LE, Whelan SP. Vesicular stomatitis virus mRNA capping machinery requires specific cis-acting signals in the RNA. *J Virol.* 2007 Oct;81(20):11499-506.

[42] : Rose JK, Lodish HF, Brock ML. Giant heterogeneous polyadenylic acid on vesicular stomatitis virus mRNA synthesized *in vitro* in the presence of S-adenosylhomocysteine. *J Virol.* 1977 Feb;21(2):683-93.

[43] : Galloway SE, Wertz GW. S-adenosyl homocysteine-induced hyperpolyadenylation of vesicular stomatitis virus mRNA requires the methyltransferase activity of L protein. *J Virol.* 2008 Dec;82(24):12280-90.

[44] : Ruedas JB, Perrault J. Putative domain-domain interactions in the vesicular stomatitis virus L polymerase protein appendage region. *J Virol.* 2014 Dec;88(24):14458-66.

CHAPITRE 7

Objectifs

Les filovirus suscitent une attention grandissante des agences de santé et un intérêt croissant pour la recherche, notamment du fait de leur importante pathogénicité. Bien que découverts depuis plus de 40 ans, la compréhension de certains mécanismes du cycle réplcatif de ces virus reste peu documentée notamment à cause du manque d'outils et de modèles d'études. Par exemple, il est acquis que la coiffe des ARNs messagers chez les mononégavirus est synthétisée par la protéine L. La synthèse de la coiffe a en effet été bien caractérisée chez le virus VSV. Cette synthèse fait intervenir différentes activités portées par la protéine L : l'activité polyribonucléotidyltransférase (PRNTase) qui permet l'ajout de la coiffe, portée par un domaine adjacent au domaine polymérase, et les activités N7 et 2'O méthyltransférases (N7 et 2'O MTases) catalysées par les domaines C-terminaux de la L. Ces données biochimiques ont été récemment renforcées avec la résolution de la structure de la L entière de VSV et l'étude structure/fonction des domaines MTase et CTD de hMPV, ce qui a ouvert la voie vers des études structure/fonction plus pertinentes [1,2]. En revanche, aucune donnée structurale ni fonctionnelle n'est actuellement disponible concernant les filovirus. Même si la comparaison de séquences de la protéine L des mononégavirus suggère la conservation de motifs et vraisemblablement de fonctions (PRNTase, méthyltransférase...), l'analyse de ces séquences n'est pas suffisante pour repérer de possibles différences mécanistiques. La récente épidémie associée au virus Ebola en Afrique de l'Ouest a souligné la nécessité de développer des outils pour la caractérisation structurale et fonctionnelle de la protéine L et la recherche d'inhibiteurs de fonctions essentielles au virus pour le développement de molécules antivirales.

La thèse que j'ai réalisée participe à cet effort par l'identification et la caractérisation de domaines de la protéine L impliqués dans la synthèse de la coiffe chez les filovirus. Afin d'atteindre ces objectifs, mon travail a, dans un premier temps, consisté à identifier le domaine porteur des activités MTases des filovirus et d'étudier ces activités par biochimie. Pour cela, plusieurs objectifs et tâches intermédiaires ont été définis. La protéine L étant une grosse polyprotéine difficile à produire en système d'expression hétérologue, j'ai du :

- identifier grâce à des outils bioinformatiques les modules conservés de la protéine L portant les motifs de la fonction MTase ;
- tester différentes constructions incluant ces motifs en système d'expression hétérologue pour obtenir de la protéine recombinante ;
- produire et purifier les protéines sélectionnées ;
- développer le test d'activité MTase ;
- définir les substrats des différentes activités MTases par quantification de l'activité sur différents ARNs;
- définir l'ordre réactionnel de la séquence de méthylation ;
- comparer le profil de méthylation des ARNm des filovirus à ceux d'autres mononégavirus (protéines recombinantes issues de collaborations avec GC Paesen et JM Grimes d'Oxford (UK) et de S Tchernyuk de Jouy-en-Josas (FR)).

Ensuite, il a été proposé de mettre en lien les résultats obtenus en biochimie avec des données structurales. Le second objectif a donc été d'obtenir des informations sur la structure du domaine responsable de l'activité MTase chez les filovirus. Cet objectif a été subdivisé en différentes tâches que j'ai réalisées :

- identifier des domaines protéiques stables grâce à différents outils bioinformatiques et biochimiques ;
- produire et purifier les protéines d'intérêt ;
- mettre en place une campagne de cristallogénèse ;
- résoudre la structure des domaines d'intérêt par cristallographie aux rayons X ;
- développer des outils permettant d'améliorer la cristallisation de la protéine.

Enfin, ce travail a été initié dans un contexte épidémique fort puisque la dernière épidémie à virus Ebola frappait encore en 2014 les pays d'Afrique de l'Ouest. La DGA a donc contribué aux financements de thèses pour le développement de stratégies antivirales encore manquantes contre ces virus. De ce fait, le troisième objectif proposé a été d'identifier des molécules inhibitrices de l'activité MTase des filovirus grâce à différentes tâches que j'ai mises en place :

- développer un test de criblage robuste
- identifier des molécules organiques inhibant l'activité MTase des filovirus ;
- déterminer l'efficacité des molécules intéressantes par différents tests biochimiques et biophysiques (inhibition et fixation à la protéine);
- valider le potentiel antiviral des molécules identifiées sur des systèmes cellulaires en collaboration (participation au projet DGA Astrid VMTaseIn (ANR-ST14-ASTR-0026) et collaboration avec M Yates et KL Seley-Radtke).

Ces travaux se sont donc attachés à démontrer l'importance du rôle des activités MTases de la protéine L des filovirus et plus largement chez les mononégavirus. La compréhension de leurs rôles dans la biologie de ces virus est cruciale pour sous-tendre le développement d'une nouvelle stratégie antivirale innovante.

Bibliographie

[1] : Liang B, Li Z, Jenni S, Rahmeh AA, Morin BM, Grant T, Grigorieff N, Harrison SC, Whelan SP. Structure of the L protein of Vesicular Stomatitis Virus from electron cryomicroscopy. *Cell*. 2015 Jul 16;162(2):314-27.

[2] : Paesen GC, Collet A, Sallamand C, Debart F, Vasseur JJ, Canard B, Decroly E, Grimes JM. X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat Commun*. 2015 Nov 9;6:8749.

CHAPITRE 8

Etude bioinformatique des domaines MTase et CTD de la protéine L des mononégavirus

L'objectif de ce chapitre est de comparer à l'aide d'outils bioinformatiques les différents domaines MTases et CTDs des protéines L des mononégavirus. Cette comparaison est particulièrement intéressante dans la mesure où la structure de ces domaines a été récemment résolue par cristallographie aux RXs (hMPV, [1]) et par cryo-microscopie (VSV, [2]). La comparaison des alignements permet donc de mieux comprendre les relations structure/fonction entre les domaines MTase+CTD des mononégavirus et d'identifier les résidus clefs pour l'activité MTase portée par ce domaine. De plus, sur base des alignements multiples, des modèles de la MTase du virus Soudan ont pu être construits. Cette analyse préliminaire a permis de définir les stratégies de clonage et d'expression des domaines de L mises en place dans ce travail.

1. Etude de la conservation des domaines MTase et CTD de la protéine L des mononégavirus

1.1. Alignements multiples et définition des domaines MTase et CTD

Les séquences en acides aminés des protéines L des filovirus et de virus d'intérêt médical dans les différentes familles de mononégavirus ont été obtenues sur le site du NCBI (USA) (liste et n° d'accès en **Annexe 1**). Dans un premier temps, un alignement multiple a été généré grâce au « méta-outil » M-Coffee [3] qui permet de combiner différents alignements conçus par d'autres outils d'alignements multiples. La séquence du VSV, pour laquelle la structure est connue, a été utilisée comme séquence de référence pour déterminer les bornes des différents domaines de la protéine L des autres virus. L'alignement obtenu (non présenté) a ainsi permis de définir une borne N-terminale du domaine MTase+CTD pour chaque virus étudié. Les séquences des domaines MTase+CTD ont ensuite été isolées et analysées par le logiciel Expresso [4] qui a permis la génération d'un nouvel alignement multiple prenant en considération les données structurales connues. Cet alignement est basé sur la séquence en acides aminés du domaine MTase+CTD du virus hMPV (plus proche phylogénétiquement des filovirus que le VSV) dont la structure et la caractérisation structure/fonction a défini les bornes entre le domaine MTase et le domaine CTD. De plus, les résidus catalytiques de l'activité MTase et ceux impliqués dans la reconnaissance du SAM et de l'ARN ont été identifiés.

L'alignement révèle différents éléments intéressants (**Annexe 1**). Tout d'abord, les séquences du domaine MTase+CTD s'alignent correctement avec la séquence de hMPV. En effet, les bornes N-terminales des domaines MTase et CTD suggérées par Whelan *et al.* chez le VSV correspondent aux bornes N-terminales de la structure des domaines MTase et CTD chez hMPV. Ces domaines sont donc assez bien définis ce qui a permis de mettre en place des constructions génétiques permettant l'expression de protéines recombinantes du virus Ebola (**Chapitre 9**). De plus, le motif conservé impliqué dans la fixation du SAM (GXGXG) est parfaitement aligné. Cependant, il est intéressant de noter que le motif GXGXG semble absent chez le virus Borna (BDV) qui présente une organisation assez lointaine par rapport aux autres mononégavirus. Cette observation peut s'expliquer par le fait que ces virus se répliquent dans le noyau. Il est possible qu'au cours de l'évolution les bornavirus aient gardé des vestiges non fonctionnels des différents domaines de la L. En outre, les résidus de la tétrade catalytique K-D-K-E des MTases de la coiffe sont strictement alignés illustrant également leur conservation et la qualité de l'alignement. Au contraire, la partie N-terminale du domaine MTase est variable en terme de longueur et de séquence, même au sein d'un genre viral donné (exemple des virus

du genre *Ebolavirus*). Le cœur du domaine est par contre bien conservé. Le CTD est, comme la partie N-terminale du domaine MTase, assez variable en longueur et en séquence. Toutefois, nous observons quelques régions homologues et un enrichissement significatif en résidus basiques.

1.2. Modélisation de la conservation des résidus

La structure des domaines MTase (orange) et CTD (rouge) chez hMPV a été utilisée afin de modéliser la conservation des résidus impliqués dans l'activité MTase (**Figure 16A**). Grâce à l'étude structure/fonction réalisée chez le virus hMPV, il est possible d'identifier sur le modèle les résidus importants pour l'activité MTase. Les résidus K1673, D1779, K1817 et E1848 forment la tétrade catalytiques (en bleu) où le D1779 a été montré comme clé pour les deux activités N7 et 2'O MTases chez hMPV. Les résidus S1668, S1669, R1662, H1659, S1842, K1843, K1991, K1992 et K1995 sont probablement impliqués dans le recrutement de l'ARN (en violet). Le motif G1696XG1698XG1700 et les résidus adjacents E1697 et T1670 constituent la poche de fixation du SAM, co-substrat de la réaction. L'alignement produit a ensuite permis de modéliser le domaine MTase+CTD sur base de la structure de hMPV permettant ainsi d'évaluer la conservation des résidus de façon spatiale. La **Figure 16B** souligne que les sites précédemment décrits sont remarquablement conservés (en rouge) ce qui

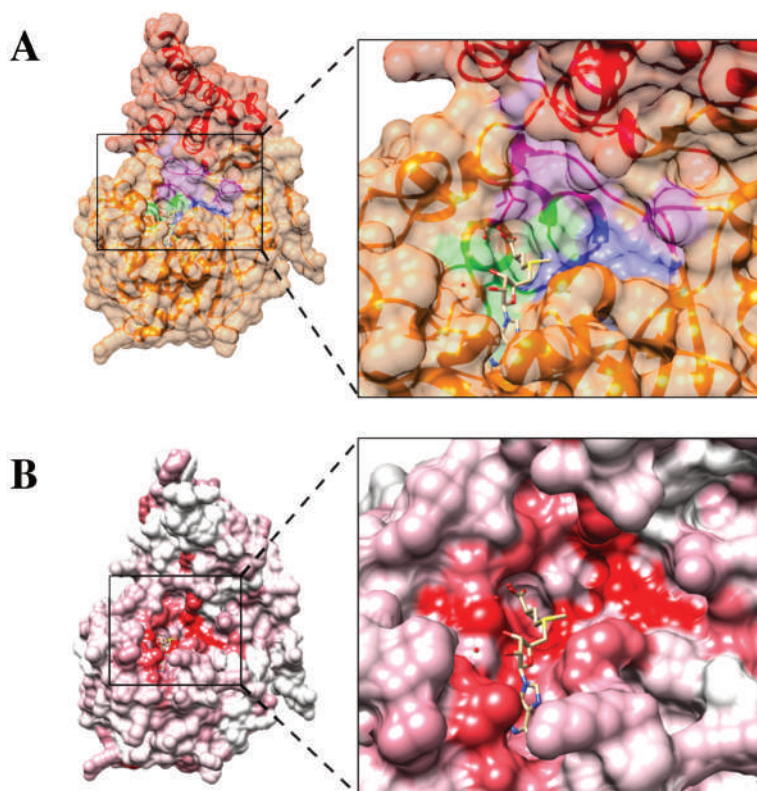


Figure 16 – Modélisation de la conservation structure/fonction du domaine méthyltransférase des mononégavirus

(A) La structure du domaine méthyltransférase (MTase, orange) et du domaine C-terminal (CTD, rouge) du virus hMPV par cristallographie aux rayons X à une résolution de 2,2 Å (PDB: 4UCK) [1]. L'étude biochimique associée a permis de modéliser les résidus impliqués dans le recrutement de l'ARN (en violet), la tétrade catalytique (en bleu) et le recrutement du co-substrat S-adenosyl méthionine (SAM, en vert) dans la structure. (B) Les séquences des domaines MTase et CTD des virus types pour chaque famille virale ont été étudiées par alignement multiple généré sur t-Coffee puis traité avec ESPript (**Annexe 1**). La conservation des résidus a été modélisée sur base de la structure du domaine MTase+CTD de hMPV, du moins conservé (blanc) au plus conservé (rouge). La région la plus conservée correspond aux résidus impliqués dans l'activité MTase.

suggère que les activités MTases et les mécanismes moléculaires de ces réactions sont également conservés.

1.3. Modélisation de la conservation de la structure

Les structures secondaires ont également été modélisées sur l'alignement grâce aux données structurales de hMPV (**Annexe 1**). La partie N-terminale du domaine MTase semble être peu structurée et sert vraisemblablement de connecteur flexible avec le CD et, par extension, avec les domaines RdRp+Cap. Les régions structurées qui constituent le cœur du domaine MTase sont quant à elles conservées en terme de séquence. Elles pourraient donc adopter un repliement de Rossmann typique des MTases canoniques et des MTases de VSV et hMPV. Par contre, le domaine CTD est beaucoup moins bien conservé bien que quelques régions homologues putativement impliquées dans la formation d'hélices α situées à l'interface entre le domaine MTase et le domaine CTD soient identifiées.

En outre, la superposition des structures de hMPV (domaine MTase+CTD seul) et de VSV (domaine MTase+CTD dans le contexte de la protéine L entière) (**Figure 17**) montre une forte homologie structurale. Le cœur du domaine MTase possède effectivement un repliement de Rossmann caractéristique dans les deux structures. Les deux domaines CTD sont aussi principalement structurés en hélice α . Cette homologie de structure indique que le domaine MTase+CTD est indépendant structurellement du reste de la protéine L. Les remaniements structuraux proposés au **Chapitre 6**, notamment lors de l'interaction avec la P, suggèrent cependant qu'une partie de la régulation, liée à la synchronisation des activités pendant la transcription, pourrait être perdue si le domaine MTase+CTD est isolé du reste de la protéine L.

La conservation des résidus impliqués dans l'activité MTase et cette homologie de structure suggèrent que l'activité MTase est portée par un domaine structurellement indépendant et conservé qui peut être isolé afin de caractériser les activités MTases chez d'autres mononégavirus. La variabilité de certaines régions de ce domaine MTase+CTD pourrait également se traduire par des variabilités fonctionnelles conférant quelque originalité fonctionnelle chez certains virus.

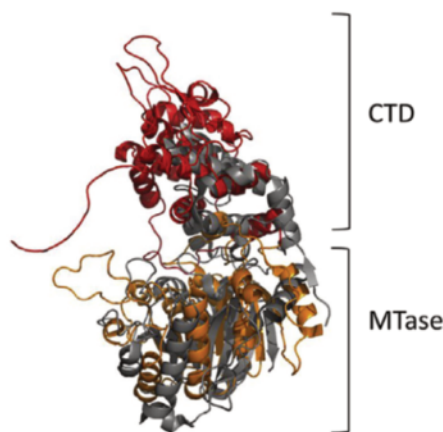


Figure 17 – Conservation structurale des domaines méthyltransférase et C-terminal de la protéine L des mononégavirus

Grâce aux études structurales chez le VSV (gris) et hMPV (couleur) [1,2], les structures des domaines méthyltransférase (MTase, en orange) et C-terminal (CTD, en rouge) ont été résolues à 3,8 Å de résolution par cryomicroscopie électronique (PDB: 5A22) et à 2,2 Å (PDB: 4UCK) par diffraction aux rayons X, respectivement. La superposition des structures de ces domaines chez VSV et hMPV montre une conservation structurale importante. Adapté de Fearn R *et al.* (2017).

2. Propriétés biochimiques des domaines MTase et CTD de la protéine L des mononégavirus

Suite à l'identification des bornes des domaines MTase et CTD grâce aux alignements multiples, leurs propriétés physico-chimiques ont été comparées de façon théorique grâce à l'outil ProtParam (Suisse) [5]. Certains paramètres ont révélé des propriétés physicochimiques intéressantes (Tableau 2).

Tableau 2 – Analyse des propriétés biochimiques des domaines méthyltransférases et C-terminaux des filovirus et autres virus de l'ordre des *Mononegavirales*

Ordre	Famille	Genre	Espèce type	Domaine	Nombre de résidus total	Nombre de résidus (+)	% résidus (+)	pI	ΔpI
Mononegavirales	Filoviridae	Ebolavirus	EBOV (AAG40171.1)	MTase	363	37	10	7,73	1,46
				CTD	171	24	14	9,19	
				MTase-CTD	534	61	11	8,81	
			SUDV (YP_138527.1)	MTase	346	35	10	6,49	3,22
				CTD	172	23	13	9,71	
				MTase-CTD	518	58	11	8,68	
			TAFV (ALT19766.1)	MTase	354	35	10	8,50	0,66
				CTD	169	20	12	9,16	
				MTase-CTD	523	55	11	8,93	
		BDBV (AKB09568.1)	MTase	354	39	11	7,86	1,12	
			CTD	169	20	12	8,78		
			MTase-CTD	523	59	11	8,53		
		RESTV (APA16576.1)	MTase	351	35	9	7,20	1,94	
			CTD	173	22	13	9,14		
			MTase-CTD	524	55	10	8,65		
	Marburgvirus	MARV (CAA82542.1)	MTase	445	53	12	8,48	0,88	
			CTD	176	24	14	9,36		
			MTase-CTD	621	77	12	8,95		
	Cuevavirus	LLOV (YP_004928143.1)	MTase	339	43	13	8,55	0,86	
			CTD	167	21	13	9,41		
			MTase-CTD	506	64	13	8,99		
	Paramyxoviridae	Respirovirus	SeV (AAB06283.1)	MTase	343	37	11	5,69	3,85
				CTD	205	31	15	9,54	
				MTase-CTD	548	68	12	7,26	
		Morbillivirus	MeV (BAB60955.1)	MTase	338	34	10	7,78	2,16
				CTD	165	27	16	9,94	
				MTase-CTD	503	61	12	9,32	
		Rubulavirus	NuV (BAA03432.1)	MTase	340	19	6	4,95	0,49
				CTD	227	26	11	5,44	
	MTase-CTD			567	45	8	5,13		
Henipavirus	HeV (O89344.3)	MTase	342	34	10	6,02	3,61		
		CTD	165	27	16	9,63			
		MTase-CTD	503	61	12	8,62			
Pneumoviridae	Metapneumovirus	hMPV (Q91L20)	MTase	285	34	12	8,57	0,92	
			CTD	124	19	15	9,49		
			MTase-CTD	409	53	13	8,97		
Orthopneumovirus	hRSV (AAJ23996.1)	MTase	323	36	11	8,94	0,48		
		CTD	127	16	13	9,42			
		MTase-CTD	450	52	12	9,14			
Rhabdoviridae	Vesiculovirus	VSV (5A22_A)	MTase	317	30	9	5,81	3,54	
			CTD	197	24	12	9,35		
			MTase-CTD	514	54	11	7,58		
Lyssavirus	RABV (ABZ81226.1)	MTase	283	27	9	6,98	1,36		
		CTD	217	24	11	8,34			
		MTase-CTD	502	51	10	8,02			
Bornaviridae	Bornavirus	BDV (P52639.2)	MTase	184	18	10	6,21	0,12	
			CTD	88	9	10	6,33		
			MTase-CTD	272	27	10	6,25		

Grâce à l'alignement de séquences généré par t-Coffee et les caractérisations structurales récentes [1,2], les séquences des domaines MTase, CTD et MTase+CTD de chaque filovirus et des virus types des autres familles de mononégavirus ont été délimités et analysés par ProtParam (n° d'accèsion entre parenthèses). Le nombre de résidus total, le nombre de résidus chargés positivement, le pourcentage de résidus chargés positivement dans la séquence et le point isoélectrique (pI). La différence entre le pI du domaine MTase et celui du domaine CTD associé a été calculée pour chaque virus. Chaque différence a été modélisée par une jauge bleue pour caractériser l'importance de cette différences relativement à celles des autres virus.

Afin de rappeler l'organisation de l'ordre des *Mononegavirales*, les différents virus ont été replacés dans leurs familles et genres respectifs. Le nombre total de résidus, le nombre et le pourcentage de résidus chargés positivement et le point isoélectrique (pI) des domaines MTase, CTD et MTase+CTD ont été reportés pour chaque virus. Enfin, la différence entre les pIs des domaines MTase et CTD a été calculée. Les pIs respectifs sont indiqués par une jauge bleue par rapport aux valeurs déterminées pour les autres virus.

Cette « méta-analyse », indique que le domaine CTD possède un pI toujours supérieur à celui du domaine MTase. Le CTD a donc une séquence primaire enrichie en acides aminés basiques, ce qui a été vérifié en calculant le pourcentage de résidus chargés positivement dans les séquences des domaines MTase et CTD. En se rapportant de nouveau à l'alignement multiple et à la structure des domaines MTase+CTD, il est intéressant de voir que la plupart des résidus basiques se trouvent à l'interface entre les domaines MTase et CTD, en particulier au sein de la dernière hélice α du CTD qui couvre en partie la poche catalytique du domaine MTase.

L'enrichissement des séquences du CTD en acides aminés chargés positivement suggère donc que ce domaine CTD est impliqué dans le recrutement de l'ARN. En effet, ces résidus étant à l'interface MTase/CTD qui intervient probablement dans le recrutement de l'ARN, leurs charges positives permettraient des interactions charge-charge avec l'ARN qui est chargé négativement. Cette hypothèse a été démontrée expérimentalement au laboratoire sur le domaine MTase+CTD de hMPV [1].

En conclusion, ce travail bioinformatique a permis de préparer le travail effectué chez les filovirus, en particulier SUDV. D'une part, l'alignement a permis de définir des bornes des domaines MTase et CTD afin de pouvoir construire une stratégie d'expression de ces domaines chez SUDV. D'autre part, la conservation de la relation structure/fonction des différents résidus cruciaux a montré qu'il était possible d'identifier des résidus clefs pour l'activité catalytique du domaine MTase+CTD chez le virus Soudan. Enfin, l'étude bioinformatique suggère que le domaine CTD joue un rôle clef dans le recrutement des ARNs substrats et leur positionnement dans le cœur catalytique de l'enzyme.

Bibliographie

- [1] : Paesen GC, Collet A, Sallamand C, Debart F, Vasseur JJ, Canard B, Decroly E, Grimes JM. X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat Commun.* 2015 Nov 9;6:8749.
- [2] : Liang B, Li Z, Jenni S, Rahmeh AA, Morin BM, Grant T, Grigorieff N, Harrison SC, Whelan SP. Structure of the L protein of Vesicular Stomatitis Virus from electron cryomicroscopy. *Cell.* 2015 Jul 16;162(2):314-27.
- [3] : Moretti S, Armougom F, Wallace IM, Higgins DG, Jongeneel CV, Notredame C. The M-Coffee web server: a meta-method for computing multiple sequence alignments by combining alternative alignment methods. *Nucleic Acids Res.* 2007 Jul;35(Web Server issue):W645-8.
- [4] : Armougom F, Moretti S, Poirot O, Audic S, Dumas P, Schaeli B, Keduas V, Notredame C. Espresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res.* 2006 Jul 1;34(Web Server issue):W604-8.
- [5] : Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, Hochstrasser DF. Protein identification and analysis tools in the ExPASy server. *Methods Mol Biol.* 1999;112:531-52.

CHAPITRE 9

Développement de la production, de la purification et du dosage enzymatique de l'activité MTase du virus Soudan

Avant de décrire la caractérisation des activités MTases de SUDV, il est important de comprendre le contexte technique dans lequel nous avons travaillé. Ce chapitre résume les techniques mises en place pour la production et la purification de protéines recombinantes et les outils biochimiques pour la caractérisation des activités MTases chez SUDV et d'autres mononégavirus. Ces techniques sont bien documentées dans les travaux antérieurs du laboratoire [1-4]. Seuls les principes et les particularités techniques sont donc détaillés ci-dessous. Dans un premier temps, l'ensemble du processus de purification et des techniques biochimiques utilisées est présenté afin de mettre en évidence les points critiques de ceux-ci. L'optimisation du tampon de lyse est ensuite détaillée avec l'utilisation des plans d'expérience comme outil de développement. Le principe du test d'activité MTase est également décrit avec les mises au point nécessaires pour l'obtention des conditions optimales d'activité pour le domaine MTase+CTD de SUDV. Enfin, les principes généraux des différentes techniques biochimiques et de la cristallographie sont introduits afin de te familiariser, cher lecteur, avec les résultats obtenus dans les chapitres suivants. Toutes les mises au point présentées ci-dessous sont basées sur l'étude du domaine MTase+CTD car il s'agit du domaine catalytiquement actif qui correspond malheureusement à la construction la plus difficile à exprimer.

1. Production et purification de protéines recombinantes

Le processus de production et de purification permettant d'obtenir les domaines MTase et MTase+CTD est schématisé dans la **Figure 18**. La séquence du domaine MTase+CTD de SUDV (résidus 1713 à 2211) a été synthétisée par la société GenScript. La séquence a été optimisée pour l'expression en système bactérien puis clonée dans le plasmide d'expression pET-14b avec une étiquette 6-His à l'extrémité N-terminale. Pour l'obtention du domaine MTase seul (résidus 1713 à 2046), deux codons STOP ont été insérés à la place des résidus 2047 et 2048 par mutagenèse dirigée (Pfu Turbo, Agilent).

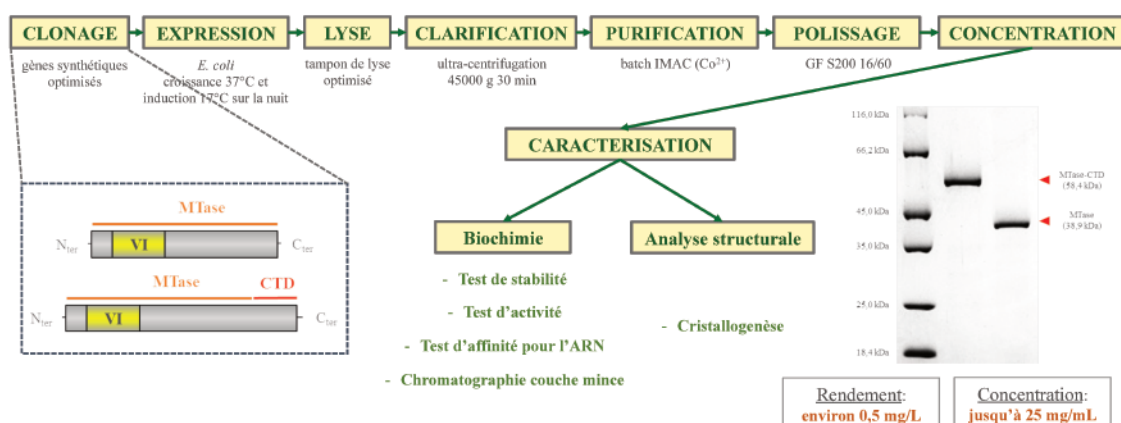


Figure 18 – Procédé global de production/purification du domaine méthyltransférase recombinant du virus Soudan

Les gènes synthétiques optimisés codant pour le domaine méthyltransférase (MTase) et le domaine MTase avec le petit domaine C-terminal (CTD) du virus Soudan (SUDV) ont été clonés dans un vecteur d'expression bactérien. Des bactéries de souche *E. coli* T7 ont été transformées avec les deux constructions, puis cultivées à 37°C jusqu'à une densité optique d'environ 0,6 UA. L'expression des protéines a alors été induite par ajout d'IPTG durant la nuit à 17°C. Les bactéries ont été récupérées puis lysées dans un tampon de lyse optimisé. La phase soluble a été clarifiée par centrifugation. Par la suite, les protéines d'intérêt ont pu être purifiées par chromatographie d'affinité (immobilized metal affinity chromatography ou IMAC) sur colonne de cobalt. Afin d'obtenir des protéines pures, les éluats obtenus ont été tamisés par chromatographie d'exclusion (ou gel filtration, GF). Les protéines ainsi purifiées ont été concentrées et contrôlées sur gel SDS-PAGE, puis caractérisées à la fois en biochimie et en analyse structurale grâce à différents outils disponibles au laboratoire.

Afin de produire ces domaines, des bactéries *E. coli* T7 sont transformées avec ces plasmides puis cultivées dans du milieu de culture Lysogeny Broth (LB) supplémenté en ampicilline comme agent de sélection. L'expression est induite par ajout d'isopropyl β -D-1-thiogalactopyranoside (IPTG). Après incubation durant la nuit, les bactéries sont récupérées par centrifugation puis congelées à -80°C . Les culots bactériens sont décongelés puis resuspendus dans un tampon de lyse optimisé. Les débris cellulaires et les protéines insolubles sont ensuite éliminés par centrifugation (clarification). Le surnageant est récupéré et incubé avec des billes d'agarose sur lesquelles est fixé du cobalt (Co-NTA). L'étiquette 6-His possède une forte affinité pour le cobalt ce qui permet de purifier la protéine recombinante surexprimée (méthode IMAC pour immobilized metal affinity chromatography). Les billes sont récupérées par centrifugation, puis lavées à forte concentration en NaCl (1 M) afin d'éliminer les acides nucléiques co-purifiés avec la protéine. La protéine est récupérée grâce à une élution à forte concentration en arginine (1 M) dans un tampon Tris pH 8 50 mM, NaCl 150 mM et glycérol 5% (v/v). L'éluat peut ensuite subir une seconde étape de purification par chromatographie d'exclusion sur une colonne de chromatographie d'exclusion (polissage) dont le principe est de séparer les protéines en fonction de leur taille. En effet, les colonnes de chromatographie d'exclusion sont constituées de billes poreuses où les petites protéines vont pouvoir passer mais pas les grosses protéines. De cette manière, les grosses protéines mettent moins de temps à traverser la colonne que les petites ce qui permet donc de les séparer. Cet outil de purification est monitoré grâce à un système Aktä (GE Healthcare) et la sortie des protéines est déterminée par mesure de l'absorbance à 280 nm au spectrophotomètre UV. La chromatographie d'exclusion permet d'éliminer les contaminants restants, de changer de tampon et de vérifier l'organisation quaternaire de la protéine en solution (monomère, dimère, etc.). Pour finir, la protéine pure (à plus de 95 %) est concentrée puis diluée dans 50% de glycérol (v/v) pour une conservation à -20°C . La protéine peut alors être utilisée pour les différentes études biochimiques. Pour l'étude structurale, les protéines sont préparées fraîchement en absence de glycérol.

Dans ce procédé, différents points critiques ont été optimisés afin d'obtenir les domaines MTase et MTase+CTD solubles en quantité suffisante pour une caractérisation biochimique. Différentes conditions d'expression ont été criblées : souche bactérienne, milieu et température de culture, température d'induction et concentration de l'inducteur. Le premier point critique a été de solubiliser le domaine MTase+CTD. Pour cela, la mise au point de la formulation du tampon de lyse a fait appel aux plans d'expériences afin de déterminer les éléments clés pour solubiliser la protéine (voir **2.1.**). Suite à l'optimisation du tampon de lyse, le second point critique a été d'identifier un tampon dans lequel la protéine reste soluble et stable. En effet, lors des premiers essais de purification par IMAC, la protéine présentait des signes d'agrégation ce qui entraînait une diminution importante du rendement de purification suite à sa précipitation. Usuellement, l'élution se fait avec de l'imidazole or, dans notre cas, l'élution est obtenue par addition d'arginine 1 M qui est un agent chaotropique permettant de garder la protéine stable en solution. La dernière étape critique a été de mettre au point le test MTase afin d'évaluer et de contrôler l'activité enzymatique de la purification (voir **2.2.**).

Grâce à ce procédé, les deux constructions des domaines MTase et MTase+CTD ont pu être exprimées, purifiées et caractérisées. Le rendement global a été optimisé et permet de purifier environ 0,5-1 mg de protéine recombinante par litre de culture. Les protéines peuvent également être concentrées sans agrégation ni précipitation jusqu'à une concentration de plus de 20 mg/mL, ce qui

illustre leur stabilité. Après purification, les protéines apparaissent pures sur gel SDS-PAGE et aucun produit de clivage n'est observé (Figure 11) (vérification par spectrométrie de masse, Annexe 2).

2. Optimisation des conditions expérimentales

2.1. Optimisation du tampon de lyse par plan d'expériences

La formulation du tampon de lyse pour récupérer une protéine d'intérêt est cruciale. En effet, c'est durant la lyse que la protéine est relarguée à l'extérieur de la bactérie. Elle a donc besoin de conditions physico-chimiques particulières favorisant sa solubilité et permettant la purification des protéines d'intérêt. Les conditions expérimentales doivent donc être optimisées afin de maximiser le rendement de purification. Or, ces facteurs sont nombreux et tester toutes les combinaisons possibles afin de déterminer la meilleure d'entre elles est un travail fastidieux. L'utilisation des outils de criblage en plan d'expériences nous a permis de formuler la composition du tampon de lyse pour le domaine MTase+CTD de façon plus efficace.

Pour commencer, le « domaine expérimental », c'est à dire les différents facteurs potentiellement influents sur la solubilisation de la protéine (tampon et pH, sels, détergents, etc.), a été défini (Tableau 3A). Chaque facteur a été décliné en plusieurs « niveaux » qui correspondent dans ce cas à la nature que peut prendre le facteur en question. Par exemple, le facteur « sels » peut être l'absence de sel, du NaCl, du KCl, du MgCl₂ ou de sulfate d'ammonium (AmSo). Une fois le domaine expérimental établi, la stratégie de criblage est choisie parmi les différents outils de plans d'expérience. Ces plans d'expériences reposent sur des matrices mathématiques, c'est à dire des séries d'équation, permettant de modéliser la réponse mesurée (la solubilisation de la protéine) en fonction des niveaux des différents facteurs testés. En utilisant cette stratégie de criblage, il est possible de modéliser la solubilisation du domaine MTase+CTD de SUDV en fonction de l'influence des facteurs, considérés comme indépendants les uns des autres, et ce, avec une organisation des expériences qui garantit la

Tableau 3 – Domaine et plan d'expérimentation pour l'optimisation par plan d'expériences du tampon d'extraction des protéines recombinantes du virus Soudan

A				B						
	Facteur	Nbre Niveaux	Niveaux	N°Exp	Réducteurs	Contre-ions	Tampons	Stabilisateurs	Sels	Détergents
U1	Réducteurs	3	none	1	none	none	Bis-Tris pH 6	none	none	none
			DTT	2	none	arginine	Tris pH 7	glycérol	NaCl	Triton X100
			B-mercaptoEtOH	3	none	glutathion	Bicine pH 8	SAM	KCl	NDSB195
U2	Contre-ions	3	none	4	none	arginine	CHES pH 9	SAH	MgCl ₂	OGP
			arginine	5	none	Bis-Tris pH 6	none	AmSo	CHAPSO	
			glutathion	6	DTT	none	Tris pH 7	SAM	MgCl ₂	CHAPSO
U3	Tampons	4	Bis-Tris pH 6	7	DTT	arginine	Bicine pH 8	SAH	AmSo	none
			Tris pH 7	8	DTT	glutathion	CHES pH 9	none	none	Triton X100
			Bicine pH 8	9	DTT	arginine	Bis-Tris pH 6	none	NaCl	NDSB195
			CHES pH 9	10	DTT	none	Bis-Tris pH 6	glycérol	KCl	OGP
U4	Stabilisateurs	4	none	11	B-mercaptoEtOH	none	Bicine pH 8	none	NaCl	OGP
			glycérol	12	B-mercaptoEtOH	arginine	CHES pH 9	none	KCl	CHAPSO
			SAM	13	B-mercaptoEtOH	glutathion	Bis-Tris pH 6	glycérol	MgCl ₂	none
			SAH	14	B-mercaptoEtOH	arginine	Bis-Tris pH 6	SAM	AmSo	Triton X100
U5	Sels	5	none	15	B-mercaptoEtOH	none	Tris pH 7	SAH	none	NDSB195
			NaCl	16	DTT	none	CHES pH 9	glycérol	AmSo	NDSB195
			KCl	17	DTT	arginine	Bis-Tris pH 6	SAM	none	OGP
			MgCl ₂	18	DTT	glutathion	Bis-Tris pH 6	SAH	NaCl	CHAPSO
			AmSo	19	DTT	arginine	Tris pH 7	none	KCl	none
U6	Détergents	5	none	20	DTT	none	Bicine pH 8	none	MgCl ₂	Triton X100
			Triton X100	21	none	none	Bis-Tris pH 6	none	KCl	Triton X100
			NDSB195	22	none	arginine	Bis-Tris pH 6	none	MgCl ₂	NDSB195
			OGP	23	none	glutathion	Tris pH 7	none	AmSo	OGP
			CHAPSO	24	none	arginine	Bicine pH 8	glycérol	none	CHAPSO
				25	none	none	CHES pH 9	SAM	NaCl	none

Afin d'optimiser le tampon de lyse pour l'extraction des protéines recombinantes du virus Soudan, l'utilisation d'un plan de criblage incomplet a été un outil puissant. (A) Dans un premier temps, les différents réactifs à tester (niveaux) ont été listés et regroupés par classe (facteur). Ainsi, chaque facteur peut prendre différents niveaux. (B) Les combinaisons de différents niveaux ont été déterminées grâce au logiciel Nemrodw. Les 25 combinaisons correspondent à une matrice de criblage permettant de déterminer les facteurs influents sur l'extraction des protéines recombinantes recherchées.

qualité de l'information obtenue. Grâce au logiciel Nemrodw (développé à Marseille), 25 combinaisons ont ainsi été générées et testées (Tableau 3B).

Dans ce travail, nous avons évalué l'effet des 25 tampons sur la solubilité de la protéine d'intérêt qui a été dosée par Western blot (Figure 19A). Pour cela, des culots bactériens issus de cultures pour l'expression du domaine MTase+CTD de SUDV ont été lysés dans les différents tampons, puis clarifiés par filtration sur filtres 0,4 μm et 0,22 μm . Ensuite, les protéines contenues dans les filtrats récupérés ont été hybridées sur membrane de nitrocellulose. Le dosage spécifique de la quantité de protéine d'intérêt récupérée a été réalisé grâce à un anticorps anti-5-His (Qiagen) qui reconnaît spécifiquement l'étiquette 6-His de la protéine recombinante. Cet anticorps est couplé à la horse radish

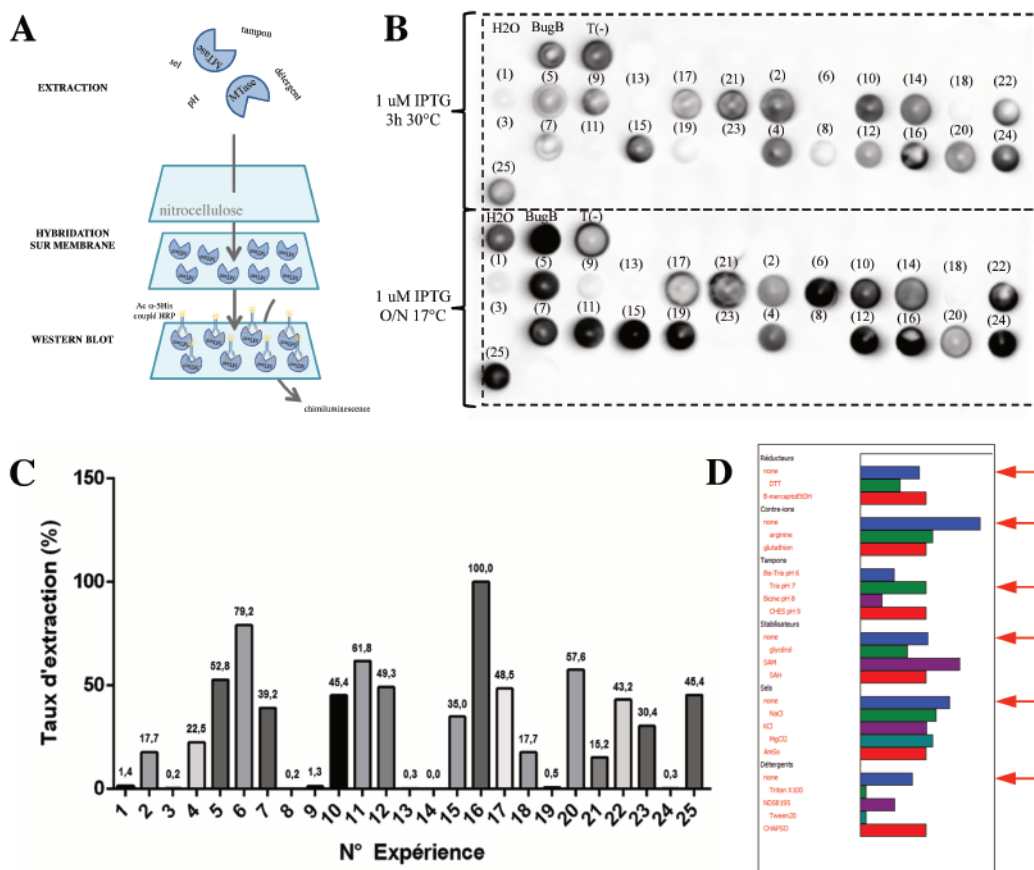


Figure 19 – Quantification de la réponse expérimentale pour l'optimisation par plan d'expériences du tampon d'extraction des protéines recombinantes du virus Soudan

Après préparation des tampons et lyse des bactéries dans les différentes formulations, les protéines recombinantes ont été dosées dans la partie soluble du lysat. (A) Après lyse et filtration des cultures de bactéries, le surnageant d'extraction a été récupéré puis filtré à travers une membrane de nitrocellulose pour hybrider les protéines contenues dans le surnageant sur cette membrane. Le dosage spécifique des protéines recombinantes a ensuite été réalisé par Western blot grâce à un anticorps anti-5His qui reconnaît spécifiquement l'étiquette 6-Histidine de ces protéines. Cet anticorps est couplé à l'enzyme Horse Radish Peroxydase (HRP) qui permet la détection et le dosage du complexe protéine/anticorps par chimioluminescence. (B) Western blot pour la détection du domaine MTase+CTD dans chaque condition testée pour deux lots de culture différents (induction à 1 μM IPTG 3h à 30°C (panel haut) ou induction à 1 μM IPTG à 17°C sur la nuit (panel bas)). Les témoins H₂O, BugB et T(-) correspondent à des tampons d'extraction contenant seulement de l'eau, seulement du BugBuster ou un tampon de base (Tris pH 8, NaCl 150 mM), respectivement. Le dosage des protéines de toutes les conditions a été réalisé sur le panel bas qui présente des quantités de protéines plus importantes que le panel haut. (C) Taux d'extraction mesuré (n=1) grâce au logiciel ImageJ, reflétant ainsi la quantité de protéine recombinante extraite. Un taux d'extraction a été calculé de manière relative par rapport à la condition au signal maximal (condition 16). (D) Analyse des réponses par Nemrodw pour le calcul des résidus, c'est-à-dire les écarts entre les réponses pour les différents niveaux de chaque facteur. Les écarts entre les résidus des différents niveaux pour un facteur donné sont représentés en valeur relative pour la détermination des conditions optimales (flèches rouges).

peroxidase (HRP) qui permet de visualiser et quantifier la protéine (MTase+CTD de SUDV) par chimiluminescence.

Les tampons de lyse ont été testés sur deux lots de culture (**Figure 19B**). Le premier lot (panel haut) correspond à une culture dont l'expression a été induite pendant 3h à 37°C avec 1 μ M IPTG alors que le second lot (panel du bas) correspond à une culture dont l'expression a été induite à 17°C pendant une nuit avec 1 μ M IPTG. La production est plus efficace sur des cultures induites sur la nuit. La quantification a donc été réalisée sur le panel bas dont le protocole a été retenu pour les productions en routine. Trois témoins ont été également inclus dans le test: un témoin H₂O qui correspond à une extraction avec de l'eau, un témoin « BugB » qui est un mélange commercial de détergents (BugBuster, EMD Millipore) qui permet de s'affranchir d'une étape de sonication pour la lyse des bactéries et qui est présent dans tous les tampons de lyse testés et un témoin négatif qui correspond à un tampon minimum (Tris pH 8 50 mM, NaCl 150 mM, 5% glycérol).

La **Figure 19C** révèle que certains tampons de lyse ne permettent pas d'extraire le domaine MTase+CTD alors que certains sont très efficaces. Le signal de chimiluminescence a été mesuré pour chaque condition grâce au logiciel ImageJ. Le taux d'extraction a ensuite été normalisé par rapport à la condition au signal le plus élevé (condition 16) (**Figure 19C**). Cependant, toutes les conditions n'ayant pas été testées, il est possible qu'il y ait une condition expérimentale plus optimale. L'interprétation du plan d'expériences permet d'identifier cette condition optimale sans pour autant l'avoir testée dans le plan. Pour cela, les « résidus », c'est à dire les différences entre les réponses (les taux d'extraction), ont été évalués afin de déterminer le « poids » de chaque facteur (**Figure 19D**). Ces différences ont été représentées de façon relative sous forme d'un graphique permettant d'identifier les niveaux des facteurs qui permettent de maximiser le rendement d'extraction. Dans le cas du domaine MTase+CTD, la composition choisie comprend seulement du Tris pH 7, du glycérol (5%) et du NaCl (150 mM). En effet, les agents réducteurs n'améliorent pas le rendement, les contre ions le diminuent, le SAM est cher et trop instable par rapport au bénéfice tiré (amélioration médiocre) et les détergents ne sont pas très efficaces. En fait, l'usage du BugBuster permet de remplacer leur usage. Supplémenter le tampon avec plus de détergent semble réduire le rendement. C'est pourquoi le BugBuster a été conservé dans la formulation finale. Afin d'attester que le tampon optimal ne dénature pas la protéine, l'activité MTase et la stabilité de la protéine ont été vérifiés en fin de purification.

2.2. Mise au point du test d'activité pour le domaine MTase+CTD de SUDV

Afin d'évaluer la qualité des lots de production et pour la caractérisation de l'enzyme, la protéine doit être testée en terme d'activité. Pour cela, un essai enzymatique préalablement établi au laboratoire a été optimisé pour l'étude de l'activité MTase de SUDV.

Le principe de ce test repose sur le transfert d'un groupement méthyl tritié (donc radioactif) porté par une molécule de SAM vers un ARN synthétique (**Figure 20A**) (liste des ARNs utilisés en **Annexe 3**). Pour cela, le domaine MTase+CTD de SUDV (4 μ M) est incubé avec le SAM radioactif et l'ARN substrat dans des conditions expérimentales optimisées (pH et d'additifs, voir ci-dessous). Après incubation, les produits de réaction sont filtrés à travers un filtre diéthylaminoéthyle (DEAE, chargé positivement) permettant la fixation des ARNs. Le filtre de DEAE est lavé afin d'éliminer la protéine, le SAM radioactif en excès et le co-produit de la réaction, le SAH. La radioactivité associée au filtre est ensuite mesurée par un compteur à scintillation et reflète le taux de méthylation de l'ARN testé.

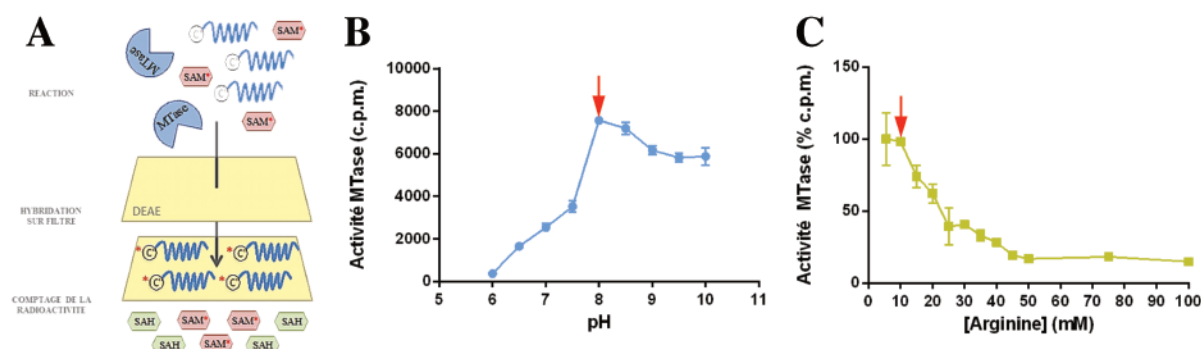


Figure 20 – Optimisation des conditions expérimentales pour le dosage enzymatique de l'activité méthyltransférase chez le virus Soudan

(A) Le test d'activité méthyltransférase (MTase) est basé sur la mesure du transfert de méthyl radiomarqué. La MTase d'intérêt est mélangé dans un tampon de réaction optimisé avec un substrat, un ARN synthétique d'intérêt (avec différents types de coiffe par exemple), et son co-substrat, la S-adenosyl méthionine (SAM), dont le groupement méthyl qui va être transféré porte un hydrogène tritié (*). Après incubation, les produits de réaction sont filtré à travers une membrane de diéthylaminoéthyle (DEAE) qui retient spécifiquement les ARNs. Après lavage de la membrane pour éliminer les co-produits de réaction, la S-adenosyl homocystéine (SAH), et l'excès de SAM*, la radioactivité associée au filtre est mesurée grâce à un compteur à scintillation. Cette radioactivité reflète le taux de méthylation des ARNs testés. Les conditions expérimentales pour le test d'activité méthyltransférase sont bien établies au laboratoire mais certains paramètres sont dépendants de la protéine testée. (B) Activité MTase en fonction du pH (n=3). L'activité est maximale à pH 8 (flèche rouge). Les données représentent les moyennes ± l'écart-type. (C) Activité MTase en fonction de la concentration en arginine (n=3). L'arginine inhibe l'activité MTase lorsque sa concentration est supérieure à 10 mM (flèche rouge). Les données représentent les moyennes ± l'écart-type.

Pour détecter une activité enzymatique optimale, les conditions de réaction doivent être optimisées. Dans le cas du domaine MTase+CTD de SUDV, deux paramètres ont été étudiés : le pH et la concentration en arginine. Le pH est un facteur important car il détermine notamment l'état des charges exposées par la protéine ce qui peut jouer un rôle clef dans la régulation de l'activité MTase. Une gamme de pH a donc été testée (Figure 20B). L'activité maximale est mesurée à pH 8 sur un substrat ARN coiffé non méthylé (GpppARN), pH auquel les tests de caractérisation ont été réalisés. Ce substrat permet les N7 et 2'O méthylations de la coiffe qui seront plus spécifiquement caractérisées par la suite (Chapitre 11). En outre, la stabilité de la protéine requiert une forte concentration en arginine (500 mM). Or, cet agent chaotropique peut altérer l'activité MTase. Il a donc fallu mesurer l'effet de l'arginine sur l'activité MTase afin de définir une concentration seuil à laquelle l'arginine présente un effet sur l'activité MTase. Cette valeur seuil a été évaluée à 10 mM (Figure 20C). En conséquence, lors de la purification, il est important de concentrer suffisamment la protéine afin d'avoir une dilution suffisante de l'arginine dans les essais enzymatiques (concentration inférieure à 10 mM) tout en ayant la concentration en protéine suffisante pour détecter l'activité (de l'ordre du μM).

3. Outils biochimiques pour la caractérisation des domaines MTase+CTD des mononégavirus

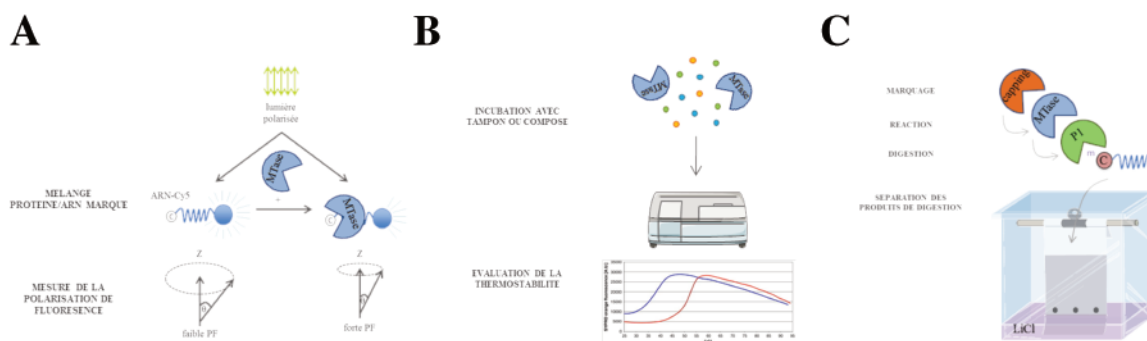


Figure 21 – Techniques de caractérisation biochimique utilisée pour l'étude du domaine méthyltransférase du virus Soudan

Le test d'activité méthyltransférase (MTase) renseigne les caractéristiques cinétiques et de spécificité de substrat. Cependant, d'autres paramètres biochimiques entrent en compte dans l'étude de cette réaction. (A) La polarisation de fluorescence (PF) permet de caractériser l'interaction entre un ARN marqué d'un fluorophore (cyanine 5, Cy5) et une protéine d'intérêt. Le suivi de la fluorescence pour une gamme de concentrations de protéine permet de déterminer un facteur de dissociation (K_d) reflétant l'affinité de l'ARN pour la protéine. (B) La thermostabilité de la protéine peut être évaluée par TSA (Thermo Stability Assay). Le principe est d'incuber la protéine d'intérêt avec le tampon ou les composés à tester puis de soumettre le mélange à un gradient de température qui va dénaturer la protéine plus ou moins vite en fonction de sa stabilité. La dénaturation peut être suivie par fluorescence ce qui permet d'évaluer un T_m reflétant la stabilité de la protéine. (C) Afin de déterminer la nature des méthylations, la structure coiffe a été analysée par chromatographie sur couche mince (CCM ou TLC). L'ARN synthétique d'intérêt est coiffé avec du GTP radioactif puis incubé avec la MTase étudiée. L'ARN est ensuite purifié et digéré par la nucléase PI qui libère la coiffe associée au premier nucléotide de l'extrémité 5'. Les différents types de coiffe sont enfin séparés par TLC en fonction de leurs méthylations.

3.1. Polarisation de fluorescence

La polarisation de fluorescence (PF) est une technique permettant par exemple de caractériser l'interaction entre une enzyme donnée et un ARN synthétique fluorescent en solution (Figure 21A) (liste des ARNs utilisés en Annexe 3). Pour cela, les ARNs d'intérêt sont couplés au fluorophore Cyanine5 (Cy5) à leur extrémité 3'. Ces ARNs marqués sont ensuite incubés en présence d'une gamme de concentrations en enzyme croissantes. L'interaction entre l'ARN et l'enzyme diminue la cinétique de rotation de l'ARN en solution ce qui se traduit par une augmentation du signal de polarisation de fluorescence. Au contraire, en absence d'interaction, la cinétique de rotation liée au mouvement brownien des ARNs induit une dépoliarisation de la fluorescence. En testant une gamme de concentration croissante en enzyme, une courbe d'interaction peut être tracée afin de calculer un coefficient de dissociation (K_d) apparent qui reflète l'affinité de l'ARN testé pour l'enzyme. La méthode de polarisation de fluorescence peut donc être utilisée pour caractériser la spécificité du substrat ARN. Il peut également permettre d'identifier les résidus responsables de l'interaction protéine/ARN en utilisant des protéines mutantes ou encore d'investiguer le mode d'action d'inhibiteurs qui pourraient entrer en compétition avec l'ARN.

3.2. Test de thermostabilité

Le test de thermostabilité (ou TSA) est un test basé sur la mesure du signal de fluorescence d'un fluorophore dont la fluorescence est induite dans un environnement hydrophobe induit par exemple

par la dénaturation des protéines (**Figure 21B**). Pour cela, l'enzyme est mélangée avec du tampon et le fluorophore, le SyproOrange. Le mélange est ensuite soumis à un gradient de température et la fluorescence de la solution est mesurée en temps réel dans un thermocycleur quantitatif. Plus la protéine est stable, plus elle mettra de temps à se « débobiner » ce qui se traduit par l'apparition tardive du signal de fluorescence à 602 nm. La courbe de thermostabilité permet ensuite de calculer un T_m (température pour laquelle 50% de la protéine est dépliée) qui est une mesure de la stabilité de la protéine.

Ce test peut être employé pour définir un tampon de conservation dans lequel la protéine est stabilisée. Il peut être également utilisé pour vérifier que des mutations ponctuelles introduites n'affectent pas la stabilité et donc le repliement de l'enzyme. Enfin, il s'agit d'un outil de choix pour le criblage d'inhibiteurs dans le but de déterminer si le composé testé stabilise ou déstabilise l'enzyme.

3.3. Chromatographie sur couche mince

Les essais enzymatiques décrits précédemment (par filtration sur membrane de DEAE) ne permettent pas d'identifier les positions de méthylation des structures coiffes (N7 ou 2'O). Afin de caractériser ces méthylations, les ARNs sont digérés par une nucléase afin de relarguer les structures coiffes (liste des ARNs utilisés en **Annexe 3**). Les coiffes sont ensuite séparées par chromatographie couche mince (CCM ou TLC) (**Figure 21C**). Cette technique est complexe car elle nécessite de nombreuses étapes critiques. Pour commencer, l'ARN à étudier doit être coiffé avec une guanosine radioactive marquée à l' α [^{32}P]. Ce marquage peut être réalisé en incubant des ARNs triphosphates en présence de α [^{32}P]-GTP et de l'enzyme de coiffage du virus de la vaccine (Epicentre). Après marquage, l'enzyme de coiffage du virus de la vaccine est éliminée par traitement à l'aide de billes de « Strataclean » (Agilent) et l'ARN radiomarqué est séparé de l' α [^{32}P]-GTP par filtration sur colonne de dessalage (GE Healthcare). Ces ARNs (Gp*ppARN) sont ensuite incubés avec la protéine dont l'activité MTase doit être caractérisée. L'ARN est purifié par précipitation à l'éthanol et incubé avec la nucléase P1 (Sigma) qui digère entièrement l'ARN sauf le lien 5'-5' triphosphate de la coiffe. Les produits de digestion sont déposés sur une plaque de silice puis séparés grâce à une phase mobile de chlorure de lithium ou de sulfate d'ammonium. Les produits de digestion sont alors détectés par fluorographie au Phosphor Imager. Les structures coiffes sont ensuite identifiées en fonction de la distance de migration (RF) qui dépend de l'hydrophobicité des produits de digestion.

4. Cristallographie aux rayons X des protéines

4.1. Principe général

La cristallographie aux rayons X (RXs) est une technique permettant de résoudre la structure d'une protéine. Afin de visualiser la disposition des atomes de la protéine étudiée, c'est à dire un objet extrêmement petit, il est nécessaire d'utiliser un rayonnement dont la longueur d'onde est de l'ordre de grandeur de ces atomes (soit 10^{-10} m ou Angström noté Å) ce qui correspond à la longueur d'onde des RXs. Cependant, l'interaction entre ces RXs et une molécule unique est beaucoup trop faible pour être mesurée. Afin d'amplifier ce signal, les protéines sont « empilées » de manière ordonnée dans un processus de cristallisation, ce qui permet aux ondes diffractées d'être cohérentes et de s'additionner pour avoir une intensité mesurable.

4.2. Cristallogénèse

La cristallogénèse des protéines est possible lorsque la protéine d'intérêt est dans des conditions physico-chimiques optimales lui permettant de s'assembler dans l'espace sous forme de cristal. La protéine à l'état soluble en solution doit passer dans un état de nucléation conduisant à une croissance cristalline. Pour obtenir cet état de nucléation, la concentration de la solution protéique augmente grâce à un procédé d'évaporation contrôlée jusqu'à atteindre un état de sur-saturation favorable à la formation de cristaux (**Figure 22A**). Or les protéines ont des structures biologiques fragiles et sensibles à leur environnement physico-chimique. Le passage progressif d'un état soluble à un état solide ordonné peut également conduire à la formation d'un précipité.

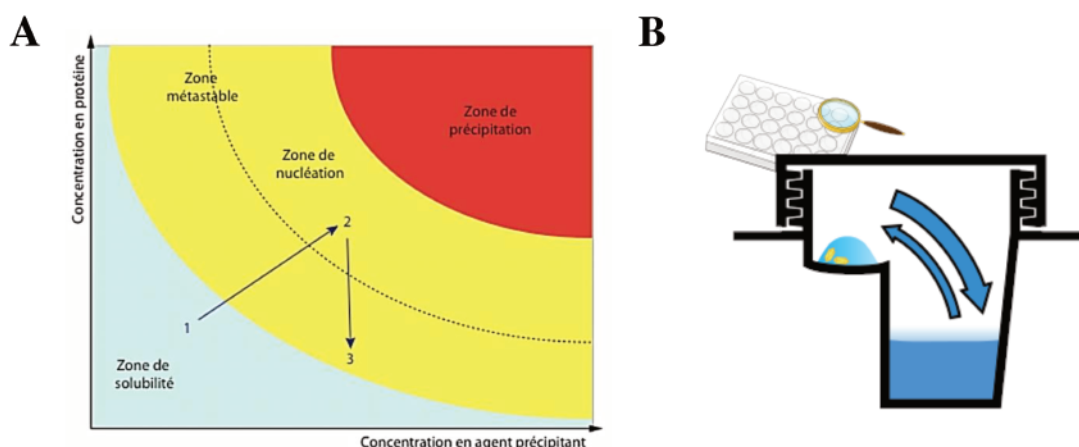


Figure 22 – Principe et mise en œuvre du processus de cristallogénèse

Afin de déterminer la structure d'une protéine, la cristallographie est une technique de choix et nécessite des cristaux de protéines. Ces cristaux sont générés grâce au processus de cristallogénèse qui repose sur le changement de phase de la protéine en fonction des conditions physico-chimiques dans lesquelles elle se trouve. (A) Diagramme de phase du processus de cristallogénèse. Les flèches indiquent l'évolution des concentrations lors d'une expérience de cristallisation d'une protéine par diffusion de vapeur. L'évaporation de la goutte va permettre à la protéine de passer de la position 1 à la position 2. Après nucléation, la croissance cristalline permet un retour à l'équilibre thermodynamique en position 3. (B) Principe de la cristallogénèse par diffusion de vapeur en goutte assise. Pour réguler finement la transition d'un état soluble vers un état de nucléation, le principe de la diffusion de vapeur est utilisé. Dans une plaque adaptée, du tampon de cristallisation est ajouté dans le puits pour constituer un « réservoir ». Une microgoutte contenant un mélange de tampon et de protéine à cristalliser est déposée dans une cupule au-dessus du réservoir. Le puits est ensuite fermé hermétiquement et la plaque est mise à incuber à la température choisie. Avec le temps, des échanges entre le réservoir et la microgoutte vont se faire par diffusion de vapeur (flèches bleues). Les variations subtiles de concentrations des différents éléments dans la microgoutte vont permettre petit à petit d'obtenir une phase de nucléation puis une croissance cristalline si les conditions physico-chimiques sont réunies.

Schémas tirés de Wikipédia.

Les conditions optimales pour obtenir une nucléation puis une croissance cristalline sont dépendantes de la protéine mais aussi d'un grand nombre de facteurs physico-chimiques. Ainsi, des possibilités infinies de conditions peuvent être testées. En effet, l'état de nucléation peut être dépendant du pH, de la nature du tampon, de la concentration en protéine, de la nature et de la concentration en sel, en agent précipitant et/ou en additif, de la température... Pour cela, des kits commerciaux regroupant une multitude de conditions de cristallisation sont testés en première intention. Ces kits permettent de dégager les tendances de comportement de la protéine en fonction des différents paramètres physico-chimiques cités ci-dessus voire d'identifier des conditions dans lesquelles la protéine peut cristalliser. Lorsqu'une ou des conditions permettent d'obtenir des

conditions favorables à la formation de micro-cristaux, celles-ci sont optimisées en faisant varier plus finement les paramètres identifiés dans la condition de criblage afin d'obtenir des cristaux de meilleure qualité.

La méthode par diffusion de vapeur en goutte assise a été utilisée afin d'obtenir un procédé d'évaporation contrôlée (**Figure 22B**). Dans un puits, 40 μL du tampon à tester sont déposés pour constituer le réservoir. Au-dessus de ce puits, 3 cupules permettent de déposer – d'où le terme goutte assise – trois nanogouttes de mélange protéine/tampon dont les concentrations varient pour maximiser les chances d'obtenir une nucléation (300/200/100 nL de protéine + 100 nL de tampon). Chaque cupule est fermée hermétiquement afin qu'un équilibre des concentrations s'établisse entre la goutte et le réservoir par diffusion des espèces volatiles entraînant une diminution progressive du volume de la goutte. La concentration en protéine et en agent précipitant augmentent jusqu'à atteindre l'état de sur-saturation ce qui favorise la formation de points de nucléation. Le retour à un état d'équilibre thermodynamique se fait par la croissance cristalline dans le meilleur des cas permettant ainsi de diminuer la concentration en protéine dans la goutte.

4.3. Diffraction des rayons X par un cristal

Suite à l'obtention de cristaux, ceux-ci doivent être congelés de manière à limiter les dommages causés par les rayonnements ionisants. Pour cela, les cristaux sont trempés dans une solution cryo-protectrice (tampon+glycérol, +TMAO, etc.) qui prévient la formation de cristaux de glace qui pourraient endommager le cristal de protéine.

Le cristal contient un assemblage ordonné de protéines qui forment un motif dans les trois dimensions de l'espace appelé maille élémentaire et qui est répété par translation pour former le cristal. Lorsque le cristal est soumis à un faisceau de RX, l'onde diffractée, qui est la somme cohérente des ondes diffusées par chaque maille élémentaire dont le contenu est identique d'une maille à l'autre, est observée sur un détecteur plan et mesurée par le logiciel de traitement de données associé à l'instrumentation. Il est nécessaire d'enregistrer les intensités diffractées à différents plans afin d'enregistrer les intensités diffractées dans un espace minimal dépendant de la symétrie cristalline et d'obtenir un jeu de données dit « complet ».

La densité électronique de la protéine est obtenue grâce à la transformée de Fourier des facteurs de structures résultant des intensités de diffraction mesurées. Or, les phases de ces facteurs de structure ne peuvent pas être déterminées ce qui rend le calcul de la densité électronique impossible. Afin de remédier à cela, des méthodes indirectes permettent de calculer ces phases comme le remplacement moléculaire par exemple.

Ce chapitre a donc permis de présenter les outils mis en place pour la caractérisation de l'activité MTase chez SUDV. Sur base des données bioinformatiques et des outils disponibles, nous avons pu adresser la question de la conservation de la relation structure/fonction de l'activité MTase identifiée chez VSV et hMPV chez les filovirus.

Bibliographie

- [1] : Paesen GC, Collet A, Sallamand C, Debart F, Vasseur JJ, Canard B, Decroly E, Grimes JM. X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat Commun.* 2015 Nov 9;6:8749.
- [2] : Coutard B, Barral K, Lichière J, Selisko B, Martin B, Aouadi W, Lombardia MO, Debart F, Vasseur JJ, Guillemot JC, Canard B, Decroly E. Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. *J Virol.* 2017 Feb 14;91(5).
- [3] : Aouadi W, Blanjoie A, Vasseur JJ, Debart F, Canard B, Decroly E. Binding of the methyl donor S-Adenosyl-1-Methionine to Middle East Respiratory Syndrome Coronavirus 2'-O-methyltransferase nsp16 promotes recruitment of the allosteric activator nsp10. *J Virol.* 2017 Feb 14;91(5).
- [4] : Bouvet M, Debarnot C, Imbert I, Selisko B, Snijder EJ, Canard B, Decroly E. *In vitro* reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS Pathog.* 2010 Apr 22;6(4):e1000863.

CHAPITRE 10

Caractérisation d'une activité MTase originale portée par le domaine MTase+CTD du virus Soudan

Suite à la mise en place du procédé de production et de purification des domaines MTase et MTase+CTD et des différents outils biochimiques, les activités biochimiques de SUDV ont été caractérisées afin de définir les propriétés et spécificités de ou des activité(s) MTase(s) de cette protéine chez le virus Ebola. Dans un premier temps, le rôle du domaine CTD dans l'activité MTase a été étudié. Nous avons ensuite identifié et caractérisé *in vitro* une activité MTase « non conventionnelle » chez SUDV qui pourrait remettre en question le dogme selon lequel la région C-terminale de la L chez EBOV serait confinée à un rôle de "cap-MTase", fonction déduite de la conservation de la séquence primaire d'autres MTase de la coiffe présentes notamment chez les virus hMPV et VSV (Chapitres 7 et 9).

1. Rôle du domaine CTD dans l'activité MTase chez SUDV

Comme mentionné au Chapitre 9, le domaine CTD est un petit domaine très variable en longueur et en séquence à travers les mononégavirus (de 120 résidus pour les pneumovirus jusqu'à environ 240 résidus chez les rhabdovirus). Ce domaine est toujours enrichi en acides aminés basiques. De plus, la structure du domaine MTase+CTD de hMPV révèle que ce domaine forme une gorge à proximité du site catalytique de la MTase, suggérant un rôle dans le recrutement de l'ARN. Afin de mieux comprendre le rôle du CTD dans la régulation de l'activité MTase chez les filovirus, deux constructions protéiques ont été produites et purifiées (Figure 23A): le domaine MTase+CTD (résidus

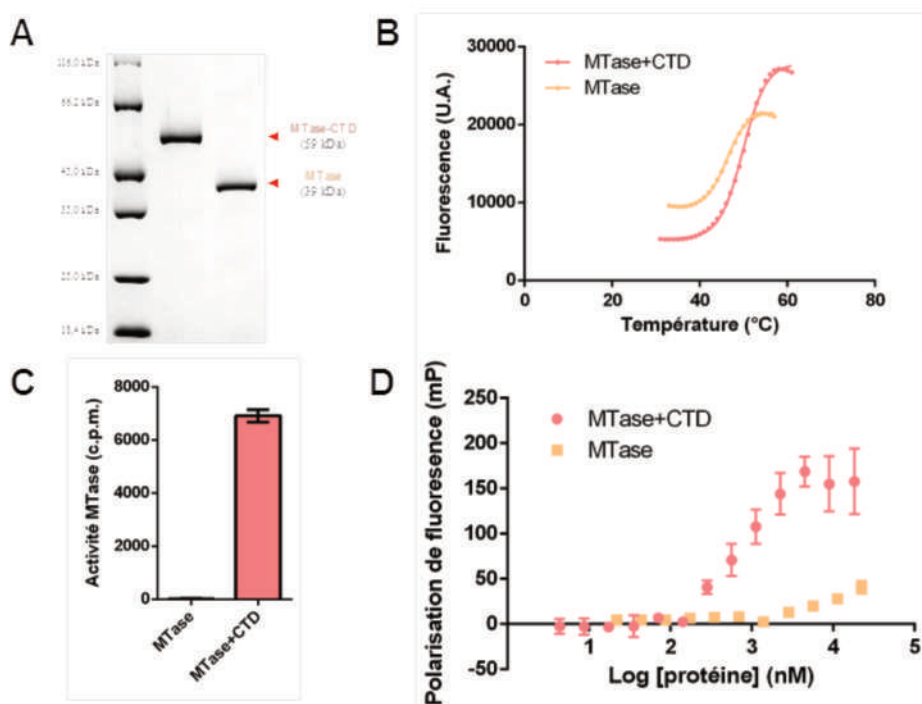


Figure 23 – Le CTD participe au recrutement de l'ARN pour assurer l'activité MTase chez SUDV.

Le domaine C-terminal (CTD) est un domaine variable dont le rôle dans l'activité méthyltransférase (MTase) n'a pas été élucidé chez les virus de l'ordre des *Mononegavirales*. (A) Analyse SDS-PAGE des domaines MTase+CTD (59 kDa) et MTase (39 kDa) purifiés. (B) Test de thermostabilité des domaines MTase+CTD et MTase (n=3). Les points représentent les moyennes. (C) Test d'activité MTase des domaines MTase+CTD et MTase sur un ARN coiffé non méthylé de séquence spécifique à l'extrémité 5' des ARNm de SUDV (GpppG-SUDV₁₂) (n=3). Les données représentent les moyennes ± les écart-types. (D) Mesure du recrutement de l'ARN par les domaines MTase+CTD et MTase (n=3). Les points représentent les moyennes brutes ± les écart-types.

1713 à 2211; 58,4 kDa) et le domaine MTase (résidus 1713 à 2046; 38,9 kDa).

La stabilité des deux constructions protéiques a été évaluée par TSA (**Figure 23B**). Chaque domaine présente un profil de transition "foldée"/dénaturée suggérant que ces protéines sont correctement repliées. Le domaine MTase+CTD est toutefois plus stable que le domaine MTase seul avec des T_m estimés à 57°C et 53°C respectivement, suggérant que i) les deux domaines ne se comportent pas de manière indépendante (une seule phase de transition) et ii) le domaine CTD participe à la stabilisation du domaine MTase (augmentation du T_m). Les T_m des deux protéines sont relativement élevés (> 50°C), attestant de la stabilité des deux protéines dans le tampon utilisé. L'activité MTase de ces deux domaines a été ensuite mesurée en incubant les protéines purifiées avec des ARNs synthétiques coiffés non méthylés de séquence mimant l'extrémité 5' des ARNm du virus Soudan (GpppG-SUDV₁₂). La **Figure 23C** indique que le domaine MTase+CTD possède une activité MTase significative sur cet ARN, alors que le domaine MTase seul ne présente aucune activité, ce qui montre que le domaine CTD est impliqué dans l'activité MTase. L'interaction entre l'ARN a ensuite été mesurée par polarisation de fluorescence (PF). La **Figure 23D** indique que le domaine MTase seul n'interagit pas avec l'ARN aux concentrations testées (< 10 μM) alors que le domaine MTase+CTD interagit fortement avec ce substrat (K_d estimé à 610 nM).

Ces résultats suggèrent que les domaines MTase et CTD fonctionnent de manière concertée et que le domaine CTD participe au recrutement de l'ARN chez SUDV afin de positionner les ARNs à proximité de la tétrade catalytique de la MTase. Nos données biochimiques et les alignements de séquences multiples présentés dans le **Chapitre 8** confirment l'implication du domaine CTD dans le recrutement de l'ARN substrat. Cette fonction est probablement conservée chez l'ensemble des mononégavirus dans la mesure où les différents domaines CTD sont fortement enrichis en acides aminés basiques.

2. Identification d'une activité MTase non conventionnelle

La construction contenant le domaine MTase+CTD porte l'activité MTase. Cette protéine a donc été utilisée par la suite pour caractériser l'activité MTase chez les filovirus en utilisant des ARNs synthétiques portant différents types de coiffe (liste des ARNs utilisés en **Annexe 3**) produits par nos collaborateurs (Debart F, IBMM). Ce travail a permis de mettre en évidence que le domaine MTase+CTD de SUDV possède une activité de méthylation interne des ARNs jusqu'à présent non décrite chez les virus de l'ordre des *Mononegavirales*.

2.1. Validation des ARNs synthétiques

Afin de caractériser l'activité MTase de la protéine MTase+CTD, la protéine a été incubée avec des ARNs synthétiques coiffés de séquence correspondant à la séquence consensus de 13 nucléotides retrouvée à l'extrémité 5' des ARNm de SUDV [1]. Ces ARN sont coiffés et des groupements méthyls ont été ajoutés sur différentes positions (N7 et/ou 2'O) afin d'identifier les positions de méthylation sur la coiffe. Dans un premier temps, les ARNs ont été incubés avec des MTases préalablement caractérisées au laboratoire (**Figure 24A**): la N7 MTase humaine (hN7), qui ne réalise que la méthylation en position N7 de la coiffe, et la 2'O MTase du virus de la vaccine (VV2'O), qui ne catalyse que la méthylation en position 2'O de la coiffe. La **Figure 24A** indique que les ARNs simplement coiffés peuvent être méthylés par la hN7, la VV2'O ou le mélange des deux enzymes

puisque les positions N7 et 2'O de la coiffe sont libres. A l'inverse, lorsque la position N7 est occupée (coiffe-0), la hN7 n'est pas active. Il en est de même pour la VV2'O qui ne présente plus d'activité lorsque l'ARN synthétique est préalablement méthylé en position 2'O. Enfin, lorsque les deux positions sont préalablement méthylées (coiffe-1), aucune activité n'est détectée, confirmant la spécificité des enzymes pour la structure coiffe. Ce profil de méthylation nous permet donc de valider l'utilisation de cette méthode expérimentale pour la caractérisation des méthylations de la coiffe par

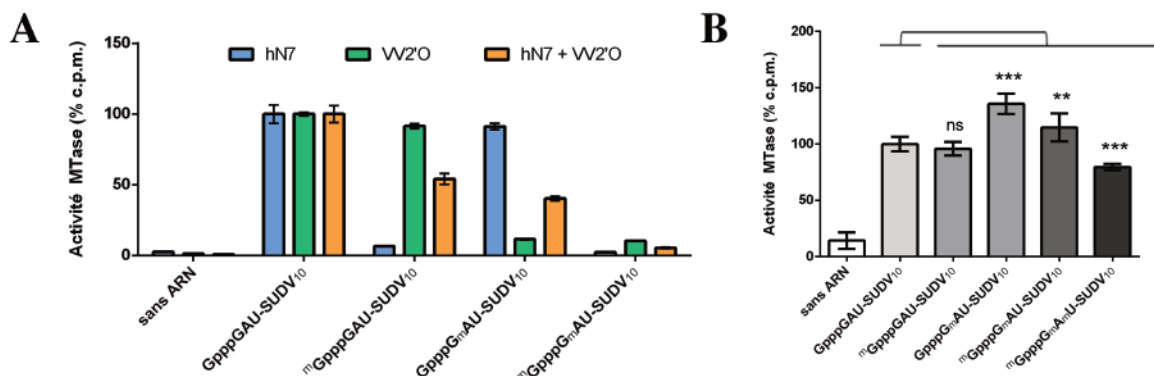


Figure 24 – Le domaine MTase+CTD de SUDV n'a pas une activité MTase conventionnelle.

Afin de caractériser les activités MTases portées par le domaine MTase+CTD, différents types de coiffe portées par des ARNs synthétiques ont été testés (^mX: coiffe N7-méthylée, X_m: résidu 2'O-méthylé). (A) Validation des différents types de coiffe par test d'activité MTase avec des MTases canoniques capables de catalyser soit la N7 méthylation (hN7, N7 MTase humaine, bleu), soit la 2'O méthylation (VV2'O, 2'O MTase du virus de la vaccine, vert) soit les deux types de méthylation (hN7+VV2'O, orange) (n=3). Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé ± les écart-types. (B) Test d'activité MTase du domaine MTase+CTD de SUDV sur les différents types de coiffe (du moins méthylé au plus méthylé, niveaux de gris) (n=6). Les groupes ont été statistiquement testés contre le groupe contrôle (coiffé non méthylé) avec un test de Student-Newman-Keuls. ns, non significatif; **, p<0,01 et ***, p<0,001. Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé ± les écart-types.

l'enzyme MTase+CTD de SUDV.

2.2. Identification d'une activité MTase non conventionnelle

L'activité MTase sur ces ARNs synthétiques a ensuite été analysée avec le domaine MTase+CTD de SUDV (Figure 24B). Le témoin sans ARN montre une faible activité résiduelle (environ 10% de l'activité mesurée sur un ARN coiffé), suggérant la présence de molécules d'ARNs bactériens copurifiés avec la protéine MTase+CTD. La valeur de ce bruit de fond a été retranchée dans les essais d'activité MTase présentés dans les figures suivantes. La Figure 24B présente le profil de méthylation du domaine MTase+CTD sur les différents ARNs substrats. L'ARN coiffé (GpppG-SUDV₁₂) utilisé précédemment est utilisé comme référence (100%). Un taux de méthylation similaire est mesuré avec les ARNs portant une structure coiffe-0 (^mGpppG-SUDV₁₂). Par contre, l'activité MTase est significativement augmentée quand la position 2'O de la coiffe est déjà méthylée (135%) et sur des ARNs portant une structure coiffe-1 (120% sur ^mGpppG_m-SUDV₁₂). Ces observations suggèrent un profil de méthylation « non conventionnel » puisque, contrairement aux autres MTases virales (flavivirus, alphavirus, mononégavirus), l'activité principale n'est pas observée sur les positions N7 et/ou 2'O de la coiffe. Afin de mieux caractériser cette activité, nous avons incubé l'enzyme avec un ARN portant une coiffe-1 et méthylé sur la position 2'O du second nucléotide (coiffe-2). La Figure

24B indique que l'activité est réduite de 20% sur cet ARN suggérant que le second nucléotide peut être méthylé par le domaine MTase+CTD de SUDV et que d'autres résidus de l'ARN sont méthylés.

Le recrutement de l'ARN est un mécanisme régulateur dans les activités MTases. En effet, chez différents virus à ARN positif étudiés au laboratoire, les méthylations de la coiffe induisent des modifications de l'affinité de l'ARN pour la MTase permettant le relargage des produits de réaction et favorisant le turnover de la réaction [2,3]. Afin d'évaluer si une régulation similaire est également observée chez SUDV, l'affinité d'ARNs synthétiques de séquence spécifique à SUDV et portant différents types de coiffe a été analysée par PF (**Figure 25**). De façon surprenante, tous les ARNs testés sont capables de se fixer avec une affinité similaire sur le domaine MTase+CTD de SUDV ce qui suggère que la présence de structures coiffées (coiffe-0, -1 ou -2) n'est pas un facteur discriminant pour le recrutement du substrat. Il semble donc que la reconnaissance ARN/protéine se fasse grâce à l'ARN et non grâce à la coiffe comme chez les flavivirus ou les coronavirus par exemple [3-5]. Ces résultats sous-tendent donc, comme les tests d'activité, un mécanisme de réaction non conventionnel. Nous garderons cependant à l'esprit que les ARN testés ne font que 13 nucléotides de long et qu'ils constituent un modèle minimal qui ne prend pas en compte certains paramètres tels que les structures secondaires présentes dans les régions 5'UTR des ARNm de SUDV (voir **Chapitre 11**).

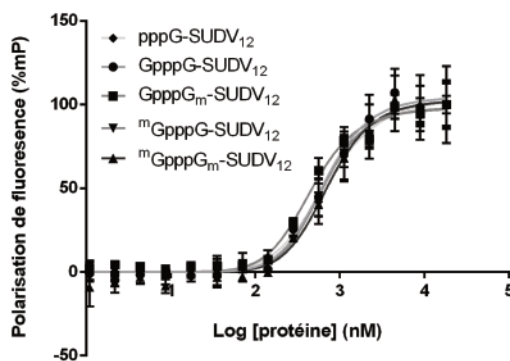


Figure 25 – Le recrutement de l'ARN ne dépend pas du type de méthylation sur la coiffe.

Mesure du recrutement de différents ARNs coiffés ou non et plus ou moins méthylés (niveaux de gris) par le domaine MTase+CTD de SUDV (^mX: coiffe N7-méthylée, X_m: résidu 2'O-méthylé) (n=3). Les points représentent les moyennes normalisées par rapport au maximum de polarisation de fluorescence pour chaque condition ± les écarts-types.

2.3. Une nouvelle activité MTase chez les mononégavirus

L'ensemble de ces observations suggère deux hypothèses. La première hypothèse propose que l'ARN coiffé soit méthylé sur une autre position de la coiffe. En effet, des triméthylguanosines ou d'autres modifications plus « exotiques » ont déjà été décrites chez d'autres organismes [6-8]. Une seconde possibilité est que cette activité est portée sur des résidus en interne dans la séquence ARN. Cette seconde hypothèse pourrait expliquer la diminution d'activité observée en présence d'une structure coiffe-2 (**Figure 24B**). Afin de déterminer si ces méthylations ciblent la structure coiffe ou des positions internes de l'ARN, nous avons incubé des ARNs non coiffés (méthylés ou non en position N1) avec le domaine MTase+CTD de SUDV. La **Figure 26** indique que l'activité détectée est réduite d'environ 35% lorsque l'extrémité 5' est monophosphate ou triphosphate par rapport à un ARN coiffé. La réduction du taux de méthylation est moindre (25%) lorsque la position 2'O du premier nucléotide de l'ARN est préalablement méthylée soulignant de nouveau le rôle stimulateur de la 2'O méthylation sur cette activité. La méthylation d'ARNs non coiffés indique une activité « non conventionnelle » qui pourrait résulter d'un processus de méthylation interne de la séquence ARN. De plus, cette observation est corroborée par les expériences de PF montrant que la coiffe n'est pas nécessaire au recrutement de l'ARN.

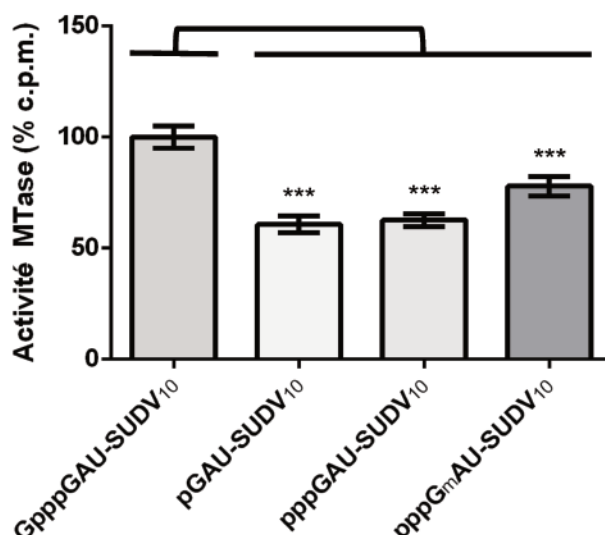


Figure 26 – Le domaine MTase+CTD méthyle des ARNs non coiffés.

L'activité MTase supplémentaire peut être portée sur une autre position de la coiffe ou sur des positions internes. Des ARNs synthétiques non coiffés (p: monophosphate, ppp: triphosphate, X_m: résidu 2'O-méthylé) ont été testés pour vérifier la première hypothèse. Test d'activité MTase sur des ARNs non coiffés plus ou moins méthylés (niveaux de gris) (n=6). Les groupes ont été statistiquement testés contre le groupe contrôle (coiffé non méthylé) avec un test de Student-Newman-Keuls. ***, p<0,001. Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé ± les écart-types.

3. Caractérisation des méthylations internes chez SUDV

L'observation de méthylations internes des ARNs est une nouveauté chez les virus de l'ordre des *Mononegavirales*. Afin de mieux caractériser ces méthylations, des ARNs de séquences variables ont été synthétisés et testés afin d'identifier les résidus méthylés.

3.1. Activité interne et séquence ARN

Afin de mieux définir si l'activité MTase interne de SUDV dépend de la séquence des ARNs substrats, l'activité MTase a été évaluée sur des ARNs synthétiques coiffés de séquences variables disponibles au laboratoire. La **Figure 27A** indique que les ARNs ne mimant pas l'extrémité 5' des ARNm de SUDV sont moins bien méthylés que ceux correspondant à l'extrémité 5' de SUDV. Certains de ces ARNs ne sont pas méthylables comme l'ARN GpppAC₄ [9] et l'ARN mimant l'extrémité 5' des ARNm du SARS-CoV (coronavirus). Les ARNs correspondant à l'extrémité 5' de flavivirus (WNV et DENV) sont faiblement méthylés. Au contraire, les ARNs respectifs du MERS-CoV (coronavirus), de RABV (rhabdovirus) et de hMPV (pneumovirus) sont significativement méthylés (entre 35 et 60% d'activité par rapport à l'activité observée sur l'ARN de SUDV). Cette observation suggère d'une part que la présence de la séquence consensus de l'extrémité 5' des ARNs de SUDV n'est pas indispensable pour observer l'activité MTase. Toutefois, il est évident que certaines séquences sont préférentiellement méthylées. En comparant les occurrences des différentes bases à chaque position de ces ARNs calculées par l'outil Weblogo (**Figure 27B**) [10], il est intéressant de noter qu'il existe une corrélation entre l'occurrence de résidus adénosines et le niveau de méthylation observé. Cette observation est confirmée si l'on porte l'activité MTase mesurée pour chaque ARN viral en fonction du nombre de A contenu dans sa séquence (**Figure 27C**). En effet,

l'analyse de cette courbe et la régression linéaire démontrent qu'il existe une relation entre l'activité MTase du domaine MTase+CTD de SUDV et le nombre de A présents au sein de la séquence ARN. Cette observation suggère que les méthylations observées concernent principalement les résidus adénosine des ARNs.

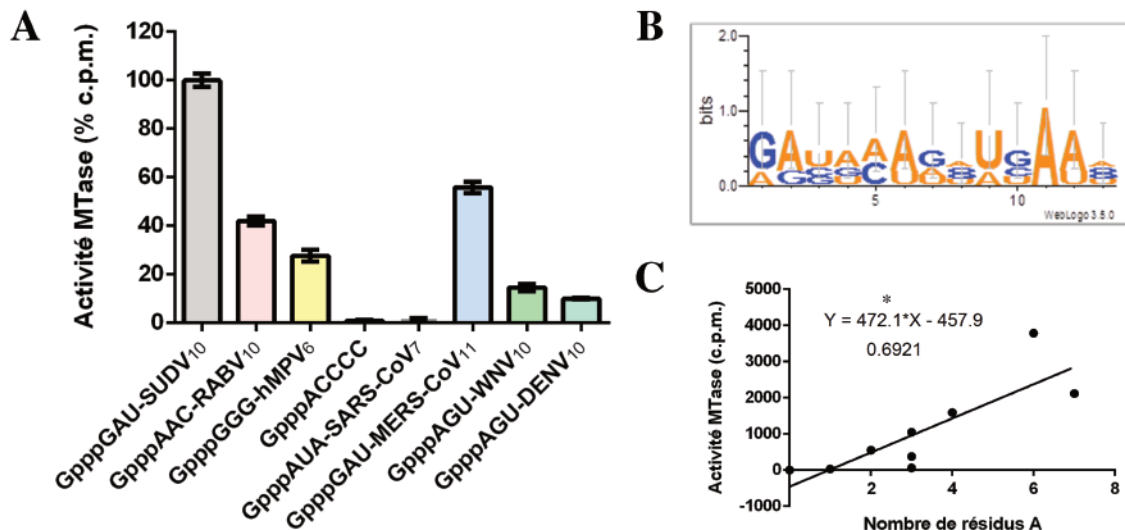


Figure 27 – L'activité MTase interne ne dépend pas de la séquence consensus des filovirus.

L'activité supplémentaire est vraisemblablement portée sur des positions internes. (A) Test d'activité MTase sur des ARNs synthétiques coiffés de séquences spécifiques de différents virus (virus Soudan, SUDV, gris; virus de la rage, RABV, rouge; métapneumovirus humain, hMPV, jaune; ARN synthétique produit par transcription *in vitro*, GpppACCCC, noir; coronavirus du syndrome respiratoire aigüe sévère, SARS-CoV, bleu foncé; coronavirus du syndrome respiratoire du Moyen-Orient, MERS-CoV, bleu clair; virus du Nil Occidental, WNV, vert pomme; virus de la dengue, DENV, vert pastel) (n=3). Les données représentent les moyennes normalisées par rapport à l'activité sur l'ARN coiffé de SUDV \pm les écart-types. (B) Recherche de consensus dans les séquences des extrémités 5' des ARNs viraux testés par l'outil Weblogo [7]. (C) Corrélation entre l'activité MTase et le nombre de résidus A dans les séquences des différents ARNs viraux testés. Les points représentent les moyennes brutes. La modélisation a été réalisée par régression linéaire de $R^2 = 0,6921$. Le coefficient a de l'équation de régression ($a=472,1$) est significatif (*, $p < 0,05$).

3.2. Identification d'une activité adénosine-2'O MTase interne

Afin de vérifier l'hypothèse selon laquelle l'activité interne serait portée préférentiellement sur les résidus adénosines internes, l'activité MTase du domaine MTase+CTD a été mesurée sur des ARNs homopolymériques commerciaux de 27 nucléotides (Biomers). La **Figure 28A** démontre que l'activité MTase est détectée sur un ARN polyadénosines et pas sur les ARNs polyguanosines, polycytidines et polyuridines.. Afin d'identifier le type de méthylation catalysé par le domaine MTase+CTD, l'activité MTase a été ensuite évaluée sur des ARNs polyadénosines de 27 mers méthylés sur différentes positions 2'OH du ribose (Biomers) (**Figure 28B**). Alors que les ARNs méthylés sur la première ou les deux premières positions ne présentent pas de diminution significative de l'activité MTase par rapport au polymère non méthylé, l'ARN méthylé en position 2'O sur l'ensemble les résidus n'est pas méthylé par le domaine MTase+CTD de SUDV. Cette observation suggère que l'activité MTase interne cible spécifiquement les positions 2'OH des riboses des résidus adénosines (ou adénosine-2'O MTase interne). Afin de confirmer ces résultats, cette activité A-2'O MTase interne a ensuite été mesurée en présence d'ARNs hétéropolymériques de séquence correspondant à l'extrémité 5' des ARNs de SUDV. Dans ces expériences, les ARNs portent une structure coiffe-1 et tous les G, tous les U ou tous les A de la séquence sont méthylés en position 2'OH (**Figure 28C**). En procédant de cette

manière, seule l'activité MTase interne sur les différents résidus peut être détectée. La **Figure 28C** montre que l'activité MTase est abolie si les résidus A internes sont préalablement méthylés sur les positions 2'OH des riboses. Il est donc clair que cette activité MTase originale portée par le domaine MTase+CTD chez SUDV est une MTase interne spécifique des résidus adénosines qui cible probablement le 2'OH du ribose. Une caractérisation par HPLC devrait permettre de confirmer cette identification.

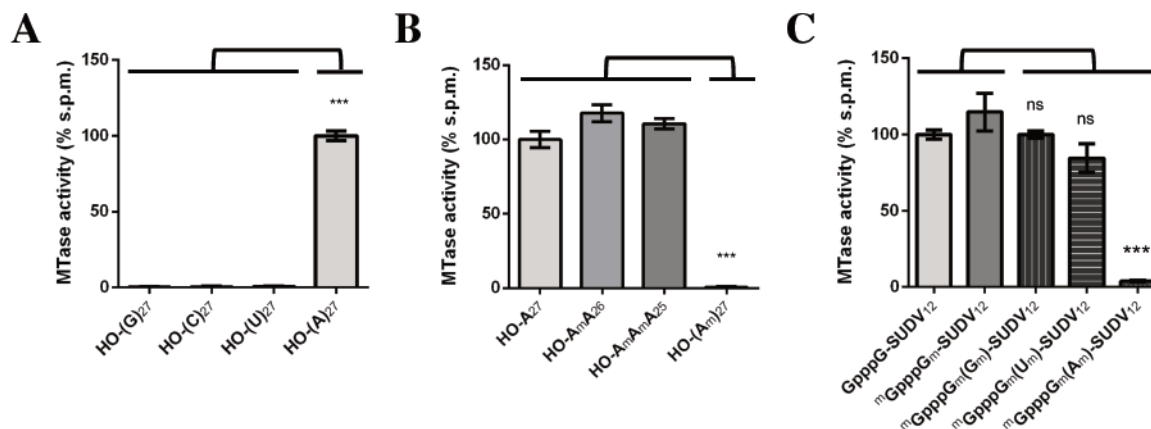


Figure 28 – Le domaine MTase+CTD est actif sur les résidus A internes.

L'activité du domaine MTase+CTD de SUDV sur les positions internes semble porter sur les résidus A. (A) Test d'activité MTase sur des ARNs homopolymériques (n=6). Les groupes ont été statistiquement testés contre le groupe présentant la plus forte activité avec un test de Student-Newman-Keuls. ***, p<0,001. Les données représentent les moyennes normalisées par rapport au groupe présentant la plus forte activité ± les écart-types. (B) Test d'activité MTase sur des ARNs polyadénosines plus ou moins méthylés (niveaux de gris) (n=6). Les groupes ont été statistiquement testés contre le contrôle non méthylé avec un test de Student-Newman-Keuls. ***, p<0,001. Les données représentent les moyennes normalisées par rapport au contrôle non méthylé ± les écart-types. (C) Test d'activité MTase sur des ARNs hétéropolymériques de séquence spécifique aux extrémités 5' des ARNs du virus SUDV, plus ou moins méthylés sur la coiffe (^mX: coiffe N7-méthylée, X_m: résidu 2'O-méthylé; niveaux de gris) et sur les résidus internes ((X_m): ARN méthylé sur tous les résidus X en aval de la coiffe; les G, rayures verticales; les U, rayures horizontales; les A, rayures transversales) (n=6). Les groupes ont été statistiquement testés contre le contrôle coiffé non méthylé et coiffé et méthylé en positions N7 et 2'O de la coiffe avec un test de Student-Newman-Keuls. ns, non significatif; ***, p<0,001. Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé ± les écart-types.

3.3. Identification des résidus catalytiques

Il a été montré que le site catalytique des 2'O MTases porte un motif K-D-K-E caractéristique [12,-14]. Les mononégavirus ne font pas exception car ce motif est à la fois décrit dans des études bioinformatiques [15] et biochimiques [16,17]. Sur base de l'alignement présenté en **Annexe 1**, les résidus K1813, D1924, K1959 et E1996 du domaine MTase+CTD de SUDV ont été proposés comme potentielle tétrade catalytique. Une étude par mutagenèse dirigée a été menée afin de vérifier le rôle de ces résidus dans l'activité MTase décrite. Pour cela, les résidus catalytiques ont chacun été remplacés par une alanine. Après production et purification des simples mutants K1813A, D1924A, K1959A et E1996A, leur stabilité a été évaluée par TSA. La **Figure 29A** montre que les différents mutants présentent une courbe de dénaturation similaire à celle de la protéine non mutée indiquant que ces mutations n'affectent pas la stabilité et donc probablement pas la conformation du domaine MTase+CTD de SUDV. L'activité MTase a ensuite été évaluée en incubant les différentes protéines mutantes avec les ARNs, GpppG-SUDV₁₂ pour l'activité globale et ^mGpppG_m-SUDV₁₂ et HO-(A)₂₇ pour l'activité interne (**Figure 29B**). Les activités ont été comparées à l'activité du domaine non muté (wild type, WT). Aucune activité n'est détectée sur les ARNs testés. Il semble que chaque mutation

altère l'activité MTase globale et interne. Ces résultats suggèrent donc que les quatre résidus K1813, D1924, K1959 et E1996 constituent la tétrade catalytique nécessaire pour les deux types de méthylations. Etant donné que cette tétrade est un motif caractéristique des 2'O MTases, cette observation étaye l'idée que le domaine MTase+CTD de SUDV porte une activité adénosine-2'O MTase ciblant les nucléotides interne de la séquence ARN.

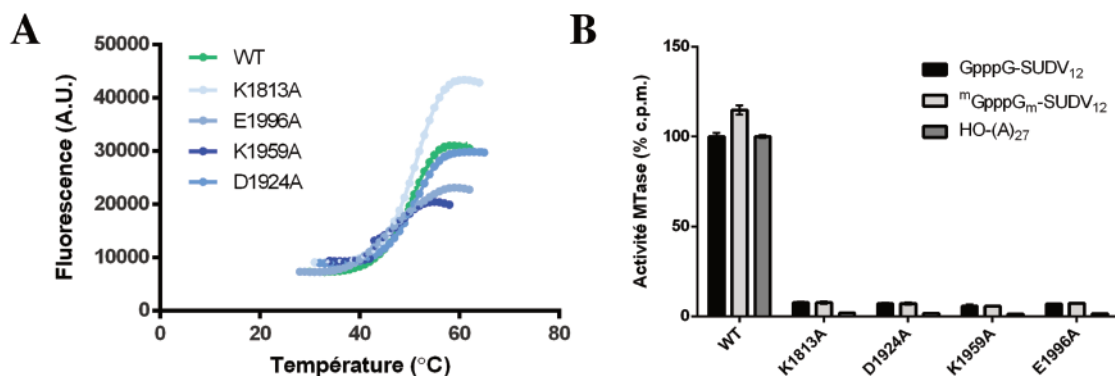


Figure 29 – La tétrade K-D-K-E catalyse l'activité adénosine-2'O MTase interne.

Le motif K-D-K-E est conservé et suggère que le domaine MTase+CTD de SUDV possède une activité 2'O MTase sur les résidus A. (A) Test de thermostabilité des mutants catalytiques K1813A, D1924A, K1959A et E1996A (n=3). Les points représentent les moyennes. (B) Test d'activité MTase du domaine MTase+CTD de SUDV (wild type, WT) et des mutants catalytiques K1813A, D1924A, K1959A et E1996A sur des ARNs homo- ou hétéropolymériques (n=3). Les données représentent les moyennes normalisées par rapport au groupe contrôle WT pour chaque ARN testé \pm les écart-types.

3.4. Méthylations internes et recrutement de l'ARN

Les méthylations sur les positions 2'O des résidus adénosines internes pourraient entraîner une modification de l'affinité de l'ARN pour la protéine. Le recrutement de différents ARNs polyadénosines plus ou moins méthylés par le domaine MTase+CTD de SUDV a donc été analysé par PF (Figure 30). Quel que soit le taux de méthylation de ces ARNs, l'affinité de ces ARNs pour le domaine MTase+CTD de SUDV n'est pas modifiée par rapport au recrutement de l'ARN coiffé hétéropolymérique de séquence spécifique à SUDV. Ces observations suggèrent que le groupement 2'OH du ribose des résidus ARNs n'intervient pas dans l'interaction ARN/protéine. Des données similaires ont été obtenues sur d'autres ARNs hétéropolymériques de différentes séquences et avec différents taux de méthylations (non présenté). Nous pouvons donc conclure que ce domaine MTase+CTD de SUDV ne porte pas de spécificité de séquence ce qui explique que des ARNs qui ne sont pas spécifiques de SUDV puissent être méthylés, en particulier sur des A internes. Il s'agit d'une

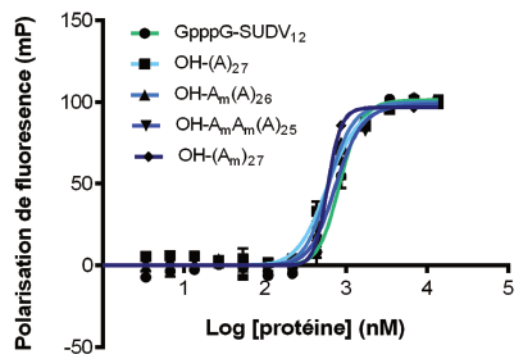


Figure 30 – Les méthylations internes n'ont pas d'effet sur le recrutement de l'ARN.

Mesure du recrutement de différents ARNs polyadénosines plus ou moins méthylés en interne par le domaine MTase+CTD de SUDV (X_m: résidu 2'O-méthylé, (X_m): ARN méthylé sur tous les résidus X en aval de la coiffe) par rapport à un ARN coiffé de SUDV (n=3). Les points représentent les moyennes normalisées par rapport au maximum de polarisation de fluorescence pour chaque condition \pm les écart-types.

différence majeure entre SUDV et d'autres virus de l'ordre des *Mononegavirales* tels que le VSV et hMPV, ou il semble que les méthylations de la coiffe nécessitent la séquence spécifique de chacun des virus et aucune activité MTase interne n'a été décrite à ce jour [16,17]. Etant donné que l'ARN peut être recruté quel que soit son état de méthylation, il est également possible que des régulations allostériques de l'activité MTase existent dans le contexte de la protéine L entière ou que ce domaine MTase puisse réaliser un « scanning » de l'ARN natif permettant d'ajouter des méthylations internes au fur et à mesure de sa synthèse.

3.5. Activité sur de l'ARN db

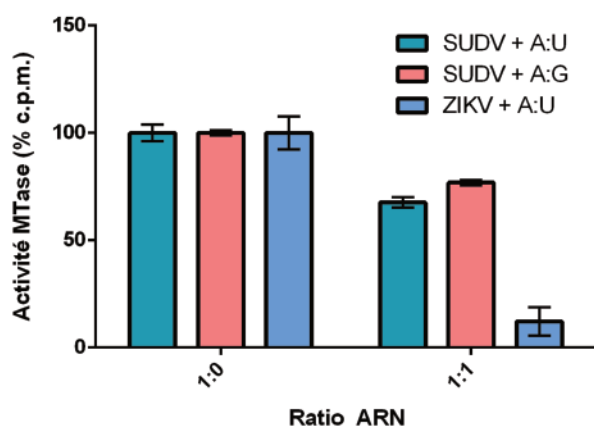


Figure 31 – Le domaine MTase+CTD est capable de méthyler l'ARN double brin.

L'ARN simple brin peut être méthyler mais les ARNs du virus Soudan forment des structures secondaires double brin. Test d'activité MTase sur des ARNs polyadénosines (A_{27}) plus ou moins hybridés en fonction du ratio d'ARN du brin complémentaire (U_{27}) ou d'un ARN non complémentaire (G_{27}) ($n=3$). Les données représentent les moyennes normalisées par rapport à l'activité sur l'ARN polyadénosine seul (ratio 1:0) \pm les écart-types.

Ces résultats ont été obtenus sur des ARNs courts et sb. Or les ARNs du virus Soudan forment des structures secondaires en tige/boucle impliquant l'existence de portions d'ARN db (discuté au **Chapitre 11**). Afin de déterminer si le domaine MTase+CTD catalyse également cette activité adénosine-2'O MTase sur de l'ARN double-brin, l'activité MTase a été évaluée sur des mélanges de différents ARNs homopolymériques (**Figure 31**). Deux mélanges ont été testés : un mélange d'ARNs polyadénosines et polyuridines qui, selon le rapport A:U, forme des ARNs db et un mélange d'ARNs polyadénosines et polyguanosines qui ne forment pas d'ARN db quel que soit le ratio. La mesure de l'activité MTase sur le mélange A:U de rapport 1:1 (c'est à dire à concentration équimolaire et donc une hybridation en db théoriquement totale) montre que 70% de l'activité MTase est conservée par rapport à un mélange A:U au

ratio 1:0 (uniquement de l'ARN polyadénosines). La mesure de l'activité MTase sur les mélanges A:G présente des résultats similaires aux résultats obtenus avec les complexes A:U suggérant que la faible diminution de l'activité observée résulte d'un phénomène de compétition. Il est également intéressant de noter que contrairement à l'activité MTase de SUDV, l'activité de méthylation interne de la MTase du virus Zika cible essentiellement les ARNs sb et cette enzyme n'est pas capable de méthyler le complexe A:U (voir article en **Annexe 5**). Ces expériences indiquent donc que la MTase du SUDV ne discrimine pas les ARN sb des ARN formant des structures db.

3.6. Activité de méthylation interne chez les MTases d'autres mononégavirus

Parmi les mononégavirus, l'activité A-2'O MTase interne n'a été décrite à ce jour que chez les filovirus. Dans la mesure où l'activité de méthylation interne chez SUDV ne concerne qu'un domaine isolé de la protéine L (MTase+CTD), il est possible qu'une partie de la régulation de l'activité MTase

soit manquante. Toutefois, d'autres études également réalisées avec un domaine MTase+CTD chez hMPV n'ont pas mis en évidence une telle activité [17]. Nous avons donc recherché si cette activité de méthylation interne était bien une signature spécifique des filovirus. Différents domaines MTase+CTD de virus des différentes familles constituant l'ordre des *Mononegavirales* ont été, produits et purifiés. Les domaines MTase+CTD de RABV (rhabdovirus) et hMPV (pneumovirus) ont été produits par nos collaborateurs à Oxford (GC Paesen) et celui de hRSV (pneumovirus) a été produit à Jouy-en-Josas par S Tchernyuk. Les activités MTases ont ensuite été évaluées sur différents substrats ARNs. Dans un premier temps, des ARNs courts (coiffés, coiffe-1 ou triphosphate) dont la séquence ARN est spécifique de la famille virale ont permis de caractériser le profil de méthylation de ces protéines sur la coiffe (**Figure 32A**). Chez SUDV, rappelons que la méthylation majoritaire observée concerne les A en interne de la séquence ARN expliquant la forte activité portée sur les trois ARNs testés. Le domaine MTase+CTD de hMPV est capable de méthyler un ARN coiffé et triphosphate mais l'activité est largement diminuée lorsque l'ARN porte une coiffe-1. Pour hRSV et RABV, seul l'ARN coiffé est méthylyé. Ces résultats montrent donc que l'activité principale des domaines MTase+CTD de RABV, hMPV et hRSV est portée sur la structure coiffe à la différence de SUDV qui porte principalement une activité de méthylation interne. D'autre part, l'activité de méthylation interne des différents domaines MTases a également été évaluée sur l'ARN polyadénosines HO-(A)₂₇ sur (**Figure 32B**). Seul SUDV est capable de méthyler significativement cet ARN. Il est également remarquable que hMPV exerce également une faible activité sur ce substrat, qui correspondrait à l'activité résiduelle observée sur l'ARN avec une coiffe-1.

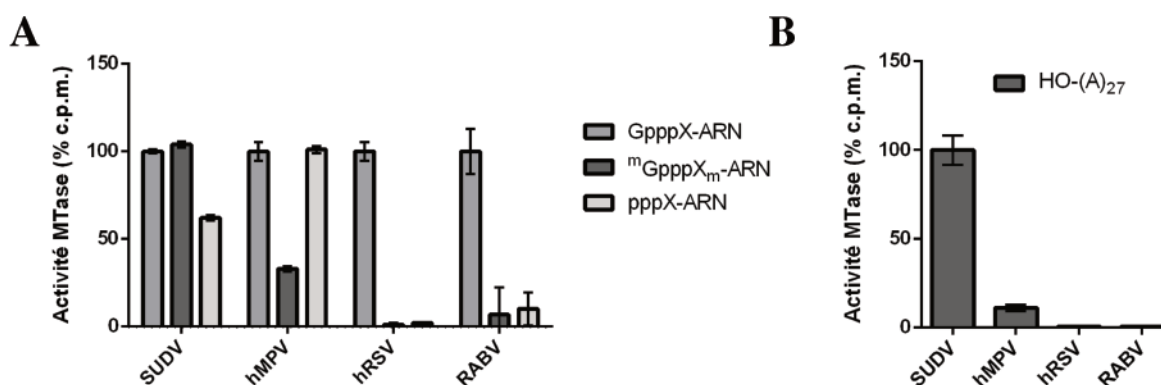


Figure 32 – Les domaines MTase+CTD des mononégavirus présentent une diversité fonctionnelle.

Le domaine MTase+CTD des mononégavirus renferme des régions très conservées comme le site actif mais d'autres régions sont beaucoup plus variables suggérant des originalités structurales et fonctionnelles. (A) Test d'activité MTase des domaines MTase+CTD du virus Soudan (SUDV, filovirus), du métapneumovirus humain (hMPV, pneumovirus), du virus respiratoire syncytial humain (hRSV, pneumovirus) et du virus de la rage (RABV, rhabdovirus) sur de courts ARNs synthétiques coiffés ou non, plus ou moins méthylés (niveaux de gris) et spécifiques à chaque virus (SUDV 13 mers pour SUDV, hMPV 9 mers pour hMPV et hRSV et RABV 13 mers pour RABV) (^mX: coiffe N7-méthylée, X_m: résidu 2'O-méthylé) (n=3). Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé pour chaque virus ± les écart-types. (B) Test d'activité MTase des domaines MTase+CTD de ces mêmes virus sur un ARN homopolymérique polyadénosine (n=3). Les données représentent les moyennes normalisées par rapport à SUDV ± les écart-types.

Il est donc clair que les domaines MTase+CTD isolés de différents mononégavirus n'ont pas la même capacité de méthylation interne. Etant donné que les constructions ont des bornes similaires, il

semble que l'activité A-2'O MTase interne portée par le domaine MTase+CTD de SUDV soit une spécificité des filovirus. En outre, comme présenté dans le **Chapitre 8**, il semble que le fait d'isoler le domaine MTase+CTD de l'ensemble de la protéine L n'affecte pas sa structure ce qui renforce l'idée que cette activité, bien qu'originale, soit effectivement présente chez les filovirus.

3.7. Perspective d'identification des A-2'O méthylations internes chez SUDV

Cette première étude a donc permis d'identifier une activité MTase interne ciblant spécifiquement les positions 2'O des résidus adénosines (ou adénosine-2'O MTase interne notée A-2'O MTase interne). Cette activité implique la tétrade catalytique K-D-K-E et est capable de méthyler des ARNs sb et db. Cette activité interne surprenante semble spécifique aux filovirus et soulève de nouvelles hypothèses biologiques intéressantes (voir 4.). Cependant, la méthylation de la position 2'OH des adénosines doit encore être démontrée par analyse HPLC des ARNs substrats après hydrolyse. Le but de ce type d'expérience sera d'identifier formellement *in vitro* l'activité A-2'O MTase interne chez SUDV. L'enjeu ultérieur sera de vérifier si ces méthylations internes existent sur les ARNs synthétisés par le virus. Pour cela, deux méthodes de détection et d'identification de résidus internes ont été envisagées et sont en cours de développement au laboratoire.

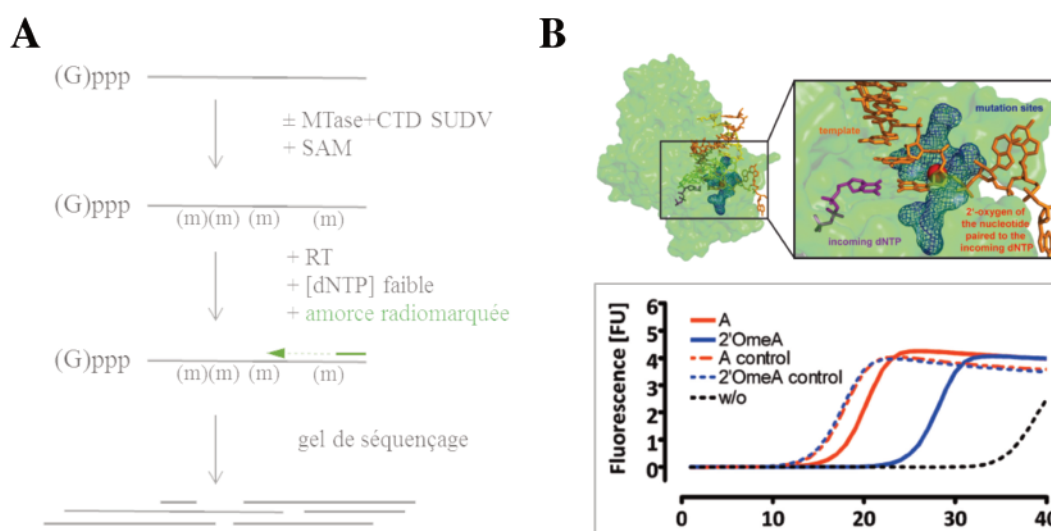


Figure 33 – Les 2'O méthylations internes peuvent être détectées chez les virus par différentes techniques biochimiques.

L'activité adénosine-2'O MTase portée par le domaine MTase+CTD de SUDV a été caractérisée *in vitro*. L'enjeu est maintenant de mettre en évidence ces 2'O méthylations internes au sein des ARN viraux. Pour ce faire, deux méthodes sont à l'étude au laboratoire. **(A)** Les 2'O méthylations peuvent être détectées par extension d'amorce grâce à la transcriptase inverse du virus de la myéloblastose aviaire (AMV) à partir d'une amorce radiomarquée. A faible concentration en dNTPs, cette transcriptase inverse est sensible aux 2'O méthylations et donne donc des produits abortifs permettant de déterminer les positions méthylées sur gel de séquençage. En choisissant des amorces pertinentes, il sera possible d'étudier les ARNs viraux purifiés. **(B)** Une deuxième technique repose sur l'utilisation en qPCR d'une polymérase thermorésistante modifiée par génie génétique. Cette polymérase est capable de polymériser de l'ADN grâce à une matrice ADN ou ARN. Une mutation supplémentaire sur cette protéine permet de la sensibiliser aux 2'O méthylations. De fait, par analyse différentielle en qPCR, il est possible de quantifier le taux de 2'O méthylation d'un ARN d'intérêt dans un extrait de culture.

La première méthode repose sur le principe d'extension d'amorce par transcription inverse (RT) à différentes concentrations en dNTPs (**Figure 33A**) [18]. Le principe de cette méthode repose sur le fait que la transcriptase inverse du virus de la myéloblastose aviaire (RT AMV, Promega) est sensible aux

2'O méthylations internes lorsque la concentration en dNTPs est faible. Grâce à une réaction de RT dans ces conditions avec une amorce radiomarquée ou fluorescente, il est possible de détecter les 2'O méthylations d'ARNs purifiés sur gel de séquençage et ainsi déterminer les positions sur lesquelles elles sont portées.

La seconde méthode utilise le principe de la RT-qPCR avec une polymérase thermorésistante issue de la bactérie *Thermophilus aquaticus* (Taq) modifiée par génie génétique [19,20] (**Figure 33B**). Les mutations des résidus L459M, S515R, I638F et M747K permettent à cette protéine de synthétiser de l'ADN à partir d'une matrice ADN ou ARN, alors que l'enzyme sauvage utilise uniquement l'ADN comme matrice. Une mutation supplémentaire V669L permet de sensibiliser cette polymérase aux 2'O méthylations des ARNs. Grâce au différentiel entre les résultats de qPCR avec ces deux formes de la polymérase sur des extraits de culture de virus, il doit être possible de quantifier le taux de méthylation d'ARNs cibles.

Ces techniques, en cours de développement au laboratoire, permettront donc d'identifier les régions méthylées des ARNs viraux et ainsi de mieux comprendre leur implication dans le cycle viral des filovirus. Pour l'heure, plusieurs rôles ont été proposés sur base des informations reportées dans la littérature. Ces questions pourront être adressées par la suite grâce à des études d'interaction hôte/pathogène dans différents modèles cellulaires.

4. Implication des méthylations internes dans la biologie des filovirus

Des 2'O méthylations internes ont déjà été mises en évidence chez d'autres virus à ARN, notamment les flavivirus [21,22]. Cependant, si leur identification a été rapportée, leur rôle biologique dans le cycle de replication des virus n'a pas encore été décrit. Plusieurs rôles de ces modifications épitranscriptomiques sont concevables. En effet, il est possible que ces modifications jouent un rôle dans l'échappement des virus à leur détection par les senseurs de l'immunité innée de l'hôte ou régulent des mécanismes de transcription, de réplication et d'encapsulation.

4.1. Une nouvelle stratégie d'échappement à l'immunité innée ?

Une de nos hypothèses est que les A-2'O méthylations internes des ARNs viraux seraient un des moyens d'échapper à la détection par le senseur de l'immunité innée de l'hôte. Elles pourraient, dans ce cas, être impliquées dans l'interférence des filovirus avec la réponse à l'interféron et ce, à différents niveaux de la voie de signalisation.

Ces 2'O méthylations internes pourraient interférer avec les senseurs cytoplasmiques de l'immunité innée de l'hôte que sont RIG-I et MDA5. Alors que RIG-I a été largement décrit dans la littérature comme senseur du non-soi via la détection des ARNs non méthylés en 2'O de la coiffe [23,24], les propriétés biologiques du senseur MDA5 restent moins bien comprises. Ce senseur peut détecter les ARNs viraux db [25]. Sa structure en complexe avec de l'ARN db a en outre été récemment résolue par cristallographie aux RXs à 3,5 Å [25] (**Figure 34A**). Cette structure révèle que MDA5 se lie à l'ARN db comme une pince grâce à une interaction avec le squelette ribose+phosphate de l'ARN. Il semble que ces interactions n'impliquent pas la position 2'OH du ribose. Toutefois, la structure révèle que si le 2'OH des ARNs db est méthylée, cela entraînerait un encombrement stérique qui ne permettrait pas à l'ARN d'être reconnu par MDA5. Etant donné nos résultats, une hypothèse serait que ces méthylations internes permettent aux filovirus de masquer leurs ARNs vis-à-vis des senseurs de

l'immunité innée, en particulier MDA5. Différentes études vont également dans ce sens. En effet, les A-2'O méthylations ont été identifiées chez d'autres virus comme les flavivirus [21] (**Annexe 5**). L'inhibition des activités MTases virales *in vivo* se traduisent par une restauration de l'immunité innée médiée majoritairement par MDA5 [26]. Ces quelques observations indirectes tendent à indiquer que l'activité A-2'O MTases internes pourrait participer à l'échappement des virus à la détection des ARNm par le senseur MDA5.

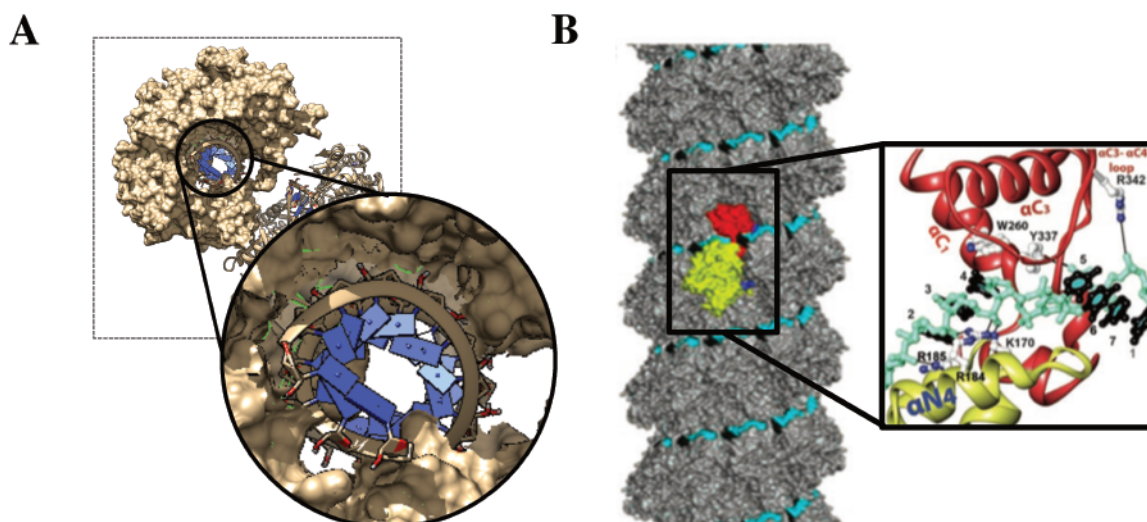


Figure 34 – Les méthylations internes pourraient être impliquées dans différents mécanismes de régulation chez les filovirus.

L'ARN simple brin de SUDV peut être méthylé sur les positions 2'O des résidus adénosines internes. Ces méthylations peuvent avoir un rôle dans différents mécanismes de régulation du cycle viral de ce virus. (A) Le senseur cytosolique de l'immunité MDA5 agit comme une pince qui enserre l'ARN cible. Son rôle est encore peu documenté mais lorsqu'il détecte un ARN double brin, il déclenche une cascade de signalisation entraînant la synthèse d'interféron. La structure de cette protéine a été résolue par cristallographie aux RXs à une résolution de 3,56 Å en complexe avec un ARN double brin (beige et bleu, PDB: 4GL2). Cette structure révèle que les positions 2'OH du ribose des nucléotides de l'ARN se projettent en direction de la protéine. Un encombrement stérique est possible dans le cas où ces positions seraient méthylées ce qui constituerait une nouvelles stratégie d'échappement des filovirus aux senseurs de l'immunité innée. (B) La nucléoprotéine (NP) des filovirus forme une structure hélicoïdale qui lierait l'ARN génomique viral du côté extérieur comme les pneumovirus. La structure de la nucléoprotéine N du virus respiratoire syncytial humain (hRSV) a été résolue par cristallographie aux RXs à 3,59 Å de résolution (PDB: 2WJ8). De manière semblable aux observations décrites pour MDA5, l'ARN est intimement lié à la structure de la N. Il est donc possible que des 2'O méthylations internes altèrent cette interaction chez les filovirus.

Adapté de Wu *et al.* (2013) et de Tawar *et al.* (2009).

De plus, ces 2'O méthylations internes pourraient avoir des interférences avec des effecteurs activés suite à une stimulation par l'interféron. Pour rappel (**Chapitre 5**), suite à la détection du virus par l'immunité innée, les senseurs activés entraînent une cascade de signalisation conduisant à la production d'INFs. Les INFs produits sont sécrétés permettant d'activer via leur récepteur IFNAR les cellules avoisinantes de façon paracrine et autocrine. En réponse à cette stimulation, différents effecteurs sont synthétisés afin de limiter la réplication du virus. L'un de ces effecteurs est le facteur «interferon induced protein 1 » (IFIT1) dont le rôle est de séquestrer les ARNs viraux afin de bloquer la réplication virale [27]. Cette séquestration concerne également les ARNm qui sont détectés s'ils portent une coiffe-0 [28-30]. Or, il a été montré qu'une coiffe-1 n'est pas suffisante non plus pour échapper à la reconnaissance des ARNs par IFIT1 [30] ce qui pourrait impliquer des méthylations sur des positions plus internes dans la séquence ARN. En effet, le résidu N2 (coiffe-2) a été montré

comme déterminant pour la reconnaissance ARN/IFIT1. Etant donné que le second nucléotide de l'extrémité 5' des ARNs chez le virus Ebola est un A et que les ARNs d'EBOV ne sont pas détectés par IFIT1 *in vivo* [31], il est donc envisageable que les A-2'O méthylations internes aient un rôle à jouer en échappant à ce facteur de restriction.

Enfin, il est également possible que ces méthylations internes ciblent d'autres effecteurs protéiques stimulés en réponse à l'interféron tels que l'isoforme p150 d'ADAR1 interagissent avec l'ARN viral [32]. Les ADARs portent une activité déaminase qui conduit à la transformation d'une adénosine en inosine. Lorsque cette édition se fait sur le génome ou l'antigénome, cette édition conduit à des mutations respectives A>C ou A>G sur le génome. Ce mécanisme d'édition détecté chez MARV et EBOV [33,34] a un rôle important dans l'évolution des filovirus et pourrait avoir des effets proviral et/ou antiviral. Les effets de ce mécanisme d'édition sur la réplication, l'apoptose et la réponse à l'interféron ont été évalués chez différents virus à ARN (revus dans [35]). La 2'O méthylation a été montrée comme étant un mécanisme de protection vis-à-vis de l'activité déaminase d'ADAR [36].

4.2. Implication dans les mécanismes de régulation de la réplication virale

Les modifications épitranscriptomiques des ARNs viraux peuvent également avoir un impact sur les différents mécanismes impliquant le recrutement de ces ARNs comme le cycle de réplication/transcription ou la régulation de l'encapsidation.

Ces A-2'O méthylations étant portées sur des ARNs destinés à la réplication et/ou au cycle de transcription/traduction, il est possible que ces modifications épitranscriptomiques aient un impact sur la régulation de ces mécanismes. Cette hypothèse a été évaluée chez les flavivirus grâce à des systèmes réplicons [21]. Ces expériences ont montré d'une part que les 2'O méthylations diminuent l'efficacité de la traduction des ARNs viraux et, d'autre part, que ces méthylations atténuent la réplication de l'ARN durant le premier cycle de réplication virale. Toutefois, il reste à évaluer plus précisément si les effets observés résultent de l'absence de la 2'O de la coiffe plutôt qu'aux 2'O méthylations internes. L'intérêt d'une telle régulation pour le cycle de réplication du virus reste finalement encore mal compris.

Une autre hypothèse serait que les A-2'O méthylations seraient impliquées dans la régulation de l'encapsidation. Sur base des données bioinformatiques et structurales disponibles sur la nucléoprotéine des mononégavirus, il a été proposé que la NP des filovirus recrute l'ARN à sa face extérieure comme les pneumovirus [37,38]. Chez hRSV par exemple [37], la structure N/ARN révèle que l'ARN est intimement lié à la N, notamment grâce à des interactions avec le squelette ribose+phosphate des nucléotides (**Figure 34B**). Il est alors possible d'imaginer, comme pour MDA5, que les 2'O méthylations internes altèrent l'interaction NP/ARN chez les filovirus. Ce mécanisme pourrait même sous-tendre l'hypothèse que les 2'O méthylations internes permettraient à la NP de discriminer les ARNm, libres dans le cytoplasme pour leur traduction, et les ARNg qui, eux, sont encapsidés au fur et à mesure de leur synthèse. Il est possible que l'activité MTase interne soit donc modulée par les remaniements structuraux de la protéine L selon ses interactions avec ses différents cofacteurs afin de réguler en aval l'encapsidation des ARNs selon la phase de réplication ou de transcription.

4.3. Conclusion générale sur l'activité MTase interne de SUDV

L'activité A-2'O MTase interne chez SUDV a été identifiée et caractérisée *in vitro* sur des ARNs synthétiques. L'enjeu actuel est de démontrer que ces méthylations sont également détectées sur des ARNs viraux, et sur lesquels (génome, antigénome, et/ou ARNm). De telles méthylations pourraient avoir des implications dans des mécanismes importants pour la réplication virale. Cette activité n'est pas commune chez les mononégavirus alors qu'il s'agit de l'activité principale chez SUDV. De plus, quel que soit le taux de méthylation des ARNs synthétiques, l'ARN est toujours capable d'interagir avec le domaine MTase+CTD. Etant donné nos résultats biochimiques et la structure de la protéine L, il est possible qu'au cours de la synthèse de l'ARN, celui-ci soit continuellement recruté par le domaine MTase qui méthylerait les résidus adénosines internes. Suivant les interactions engagées avec les différents cofacteurs de la L, il est également possible que cette méthylation puisse être régulée entre la transcription et la réplication afin de ne méthyler que les ARNm qui ne seront pas encapsidés. Dans le cytoplasme, ces ARNm sont libres mais ne seraient pas pour autant détectés par les senseurs de l'immunité grâce à ces méthylations. En conclusion, il semble donc que l'activité 2'O de la coiffe déjà décrite chez d'autres mononégavirus ait évolué chez les filovirus pour donner une activité 2'O interne qui pourrait jouer un rôle majeur dans le cycle viral de ces virus. Afin de définir si ce glissement évolutif dérive effectivement des activités MTases de la coiffe, nous avons évalué et caractérisé ces activités pour le domaine MTase+CTD de SUDV.

Bibliographie

- [1] : Sanchez A, Rollin PE. Complete genome sequence of an Ebola virus (Sudan species) responsible for a 2000 outbreak of human disease in Uganda. *Virus Res.* 2005 Oct;113(1):16-25.
- [2] : Li C, Guillén J, Rabah N, Blanjoie A, Debart F, Vasseur JJ, Canard B, Decroly E, Coutard B. mRNA Capping by Venezuelan Equine Encephalitis Virus nsP1: Functional Characterization and Implications for Antiviral Research. *J Virol.* 2015 Aug;89(16):8292-303.
- [3] : Aouadi W, Blanjoie A, Vasseur JJ, Debart F, Canard B, Decroly E. Binding of the Methyl Donor S-Adenosyl-l-Methionine to Middle East Respiratory Syndrome Coronavirus 2'-O-Methyltransferase nsp16 Promotes Recruitment of the Allosteric Activator nsp10. *J Virol.* 2017 Feb 14;91(5).
- [4] : Ma Y, Wu L, Shaw N, Gao Y, Wang J, Sun Y, Lou Z, Yan L, Zhang R, Rao Z. Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. *Proc Natl Acad Sci U S A.* 2015 Jul 28;112(30):9436-41.
- [5] : Zhao Y, Soh TS, Lim SP, Chung KY, Swaminathan K, Vasudevan SG, Shi PY, Lescar J, Luo D. Molecular basis for specific viral RNA recognition and 2'-O-ribose methylation by the dengue virus nonstructural protein 5 (NS5). *Proc Natl Acad Sci U S A.* 2015 Dec 1;112(48):14834-9.
- [6] : Motorin Y, Helm M. RNA nucleotide methylation. *Wiley Interdiscip Rev RNA.* 2011 Sep-Oct;2(5):611-31.
- [7] : Jäschke A, Höfer K, Nübel G, Frindert J. Cap-like structures in bacterial RNA and epitranscriptomic modification. *Curr Opin Microbiol.* 2016 Apr;30:44-9.
- [8] : Liu J, Jia G. Methylation modifications in eukaryotic messenger RNA. *J Genet Genomics.* 2014 Jan 20;41(1):21-33.
- [9] : Peyrane F, Selisko B, Decroly E, Vasseur JJ, Benarroch D, Canard B, Alvarez K. High-yield production of short GpppA- and 7MeGpppA-capped RNAs and HPLC-monitoring of methyltransfer reactions at the guanine-N7 and adenosine-2'O positions. *Nucleic Acids Res.* 2007;35(4):e26. Epub 2007 Jan 26.
- [10] : Schneider TD, Stephens RM. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 1990 Oct 25;18(20):6097-100.
- [11] : Kennedy EM, Courtney DG, Tsai K, Cullen BR. Viral Epitranscriptomics. *J Virol.* 2017 Apr 13;91(9).
- [12] : Hodel AE, Gershon PD, Shi X, Quijcho FA. The 1.85 Å structure of vaccinia protein VP39: a bifunctional enzyme that participates in the modification of both mRNA ends. *Cell.* 1996 Apr 19;85(2):247-56.
- [13] : Bujnicki JM, Rychlewski L. Prediction of a novel RNA 2'-O-ribose methyltransferase subfamily encoded by the Escherichia coli YgdE open reading frame and its orthologs. *Acta Microbiol Pol.* 2000;49(3-4):253-60.
- [14] : Bujnicki JM, Rychlewski L. Reassignment of specificities of two cap methyltransferase domains in the reovirus lambda 2 protein. *Genome Biol.* 2001;2(9):RESEARCH0038.
- [15] : 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.
- [16] : Rahmeh AA, Li J, Kranzusch PJ, Whelan SP. 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.* 2009 Nov;83(21):11043-50.
- [17] : Paesen GC, Collet A, Sallamand C, Debart F, Vasseur JJ, Canard B, Decroly E, Grimes JM. X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat Commun.* 2015 Nov 9;6:8749.
- [18] : Motorin Y, Muller S, Behm-Ansmant I, Branlant C. Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol.* 2007;425:21-53.
- [19] : Blatter N, Bergen K, Nolte O, Welte W, Diederichs K, Mayer J, Wieland M, Marx A. Structure and function of an RNA-reading thermostable DNA polymerase. *Angew Chem Int Ed Engl.* 2013 Nov 4;52(45):11935-9.

- [20] : Aschenbrenner J, Marx A. Direct and site-specific quantification of RNA 2'-O-methylation by PCR with an engineered DNA polymerase. *Nucleic Acids Res.* 2016 May 5;44(8):3495-502.
- [21] : Dong H, Chang DC, Hua MH, Lim SP, Chionh YH, Hia F, Lee YH, Kukkaro P, Lok SM, Dedon PC, Shi PY. 2'-O methylation of internal adenosine by flavivirus NS5 methyltransferase. *PLoS Pathog.* 2012;8(4):e1002642.
- [22] : Coutard B, Barral K, Lichière J, Selisko B, Martin B, Aouadi W, Lombardia MO, Debart F, Vasseur JJ, Guillemot JC, Canard B, Decroly E. Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives. *J Virol.* 2017 Feb 14;91(5). pii: e02202-16.
- [23] : Schubert-Wagner C, Ludwig J, Bruder AK, Herzner AM, Zillinger T, Goldeck M, Schmidt T, Schmid-Burgk JL, Kerber R, Wolter S, Stümpel JP, Roth A, Bartok E, Drosten C, Coch C, Hornung V, Barchet W, Kümmerer BM, Hartmann G, Schlee M. A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2'-O-Methylated Self RNA. *Immunity.* 2015 Jul 21;43(1):41-51.
- [24] : Devarkar SC, Wang C, Miller MT, Ramanathan A, Jiang F, Khan AG, Patel SS, Marcotrigiano J. Structural basis for m7G recognition and 2'-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proc Natl Acad Sci U S A.* 2016 Jan 19;113(3):596-601.
- [25] : Wu B, Peisley A, Richards C, Yao H, Zeng X, Lin C, Chu F, Walz T, Hur S. Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5. *Cell.* 2013 Jan 17;152(1-2):276-89.
- [26] : Züst R, Cervantes-Barragan L, Habjan M, Maier R, Neuman BW, Ziebuhr J, Szretter KJ, Baker SC, Barchet W, Diamond MS, Siddell SG, Ludewig B, Thiel V. 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 Feb;12(2):137-43.
- [27] : Pichlmair A, Lassnig C, Eberle CA, Górna MW, Baumann CL, Burkard TR, Bürckstümmer T, Stefanovic A, Krieger S, Bennett KL, Rülcke T, Weber F, Colinge J, Müller M, Superti-Furga G. IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. *Nat Immunol.* 2011 Jun 5;12(7):624-30.
- [28] : Kimura T, Katoh H, Kayama H, Saiga H, Okuyama M, Okamoto T, Umemoto E, Matsuura Y, Yamamoto M, Takeda K. Ifit1 inhibits Japanese encephalitis virus replication through binding to 5' capped 2'-O unmethylated RNA. *J Virol.* 2013 Sep;87(18):9997-10003.
- [29] : Habjan M, Hubel P, Lacerda L, Benda C, Holze C, Eberl CH, Mann A, Kindler E, Gil-Cruz C, Ziebuhr J, Thiel V, Pichlmair A. Sequestration by IFIT1 impairs translation of 2'-O-unmethylated capped RNA. *PLoS Pathog.* 2013;9(10):e1003663.
- [30] : Kumar P, Sweeney TR, Skabkin MA, Skabkina OV, Hellen CU, Pestova TV. Inhibition of translation by IFIT family members is determined by their ability to interact selectively with the 5'-terminal regions of cap0-, cap1- and 5'ppp- mRNAs. *Nucleic Acids Res.* 2014 Mar;42(5):3228-45.
- [31] : Pinto AK, Williams GD, Szretter KJ, White JP, Proença-Módena JL, Liu G, Olejnik J, Brien JD, Ebihara H, Mühlberger E, Amarasinghe G, Diamond MS, Boon AC. Human and Murine IFIT1 Proteins Do Not Restrict Infection of Negative-Sense RNA Viruses of the Orthomyxoviridae, Bunyaviridae, and Filoviridae Families. *J Virol.* 2015 Sep;89(18):9465-76.
- [32] : Pfaller CK, Li Z, George CX, Samuel CE. Protein kinase PKR and RNA adenosine deaminase ADAR1: new roles for old players as modulators of the interferon response. *Curr Opin Immunol.* 2011 Oct;23(5):573-82.
- [33] : Shabman RS, Jabado OJ, Mire CE, Stockwell TB, Edwards M, Mahajan M, Geisbert TW, Basler CF. Deep sequencing identifies noncanonical editing of Ebola and Marburg virus RNAs in infected cells. *MBio.* 2014 Nov 4;5(6):e02011.
- [34] : Brody T, Yavatkar AS, Park DS, Kuzin A, Ross J, Odenwald WF. Flavivirus and Filovirus EvoPrinters: New alignment tools for the comparative analysis of viral evolution. *PLoS Negl Trop Dis.* 2017 Jun 16;11(6):e0005673.
- [35] : Samuel CE. ADARs: viruses and innate immunity. *Curr Top Microbiol Immunol.* 2012;353:163-95.

[36] : Yi-Brunozzi HY, Easterwood LM, Kamilar GM, Beal PA. Synthetic substrate analogs for the RNA-editing adenosine deaminase ADAR-2. *Nucleic Acids Res.* 1999 Jul 15;27(14):2912-7.

[37] : Tawar RG, Duquerroy S, Vornrhein C, Varela PF, Damier-Piolle L, Castagné N, MacLellan K, Bedouelle H, Bricogne G, Bhella D, Eléouët JF, Rey FA. Crystal structure of a nucleocapsid-like nucleoprotein-RNA complex of respiratory syncytial virus. *Science.* 2009 Nov 27;326(5957):1279-83.

[38] : Peng R, Zhu T, Oladejo BO, Musyoki AM, Cui Y, Shi Y, Wang P, Gao GF. In vitro assembly of Ebola virus nucleocapsid-like complex expressed in *E. coli*. *Protein Cell.* 2016 Dec;7(12):888-898.

CHAPITRE 11

Etude des méthylations de la coiffe par le domaine MTase+CTD du virus Soudan

Les MTases de virus à ARN caractérisées jusqu'à présent sont impliquées dans le processus de méthylation de la coiffe. Il est donc naturel que, suite à la caractérisation *in vitro* de l'activité A-2'O MTase interne, le rôle de la MTase de SUDV dans les méthylation de la coiffe soit étudié. Afin de pouvoir observer uniquement la ou les activité(s) MTase(s) cibant la coiffe, nos collaborateurs (Debart F, IBMM) ont synthétisé des ARNs dont les résidus A ont tous été préalablement méthylés (notés $^{(m)}(G)pppX_{(m)}(A_m)$ -SUDV₁₂). Les résultats présentés dans cette partie sont des données préliminaires. Elles nous ont cependant permis de construire des hypothèses de travail intéressantes voire intrigantes qui sont discutées à la fin de ce chapitre.

1. Méthylation de la coiffe et pH

Les activités N7 et 2'O MTases sont des activités sensibles au pH. En effet, chez le virus de la dengue (DENV) par exemple, les activités N7 et 2'O MTases, qui sont portées par la protéine NS5, ont des pHs optimaux respectifs de 6,0 et de 8,0-9,0 [1]. Chez les mononégavirus, cette même tendance a été observée pour VSV [2]. Afin de déterminer si le pH a également un rôle dans la régulation des activités MTases de la coiffe chez SUDV qui, rappelons-le, sont aussi catalysées par un seul site actif, l'activité MTase a été évaluée sur différents ARNs synthétiques plus ou moins méthylés pour une gamme de pH de 6,0 à 10,0 (**Figure 35A**).

1.1. Profils de méthylation de différents ARNs en fonction du pH

L'activité du domaine MTase+CTD a dans un premier temps été mesurée sur un ARN coiffé non méthylyé (GpppG-SUDV₁₂, en noir). Les pHs bas, entre 6 et 7,5, ne semble pas favorables à l'activité MTase au contraire des pHs plus élevés. Comme décrit précédemment lors du développement du test d'activité (**Chapitre 9**), le pH optimal est de 8,0. La gamme de pH a été également testée sur un ARN avec une coiffe-1 (m GpppG_m-SUDV₁₂, violet clair) afin de vérifier le profil de l'activité adénosine-2'O MTase interne en fonction du pH. Le profil est similaire à celui obtenu sur un ARN coiffé avec un pH optimal de 8,5, ce qui confirme que l'activité majoritaire observée en présence de l'ARN GpppG-SUDV₁₂ est portée sur les positions internes. Il est intéressant de noter que, pour un pH donné, l'activité est toujours plus élevée sur l'ARN avec une coiffe-1 suggérant que la présence d'une structure coiffe-1 pourrait stimuler la méthylation interne. Afin de suivre uniquement la N7 ou la 2'O méthylation de la coiffe, l'activité du domaine MTase+CTD de SUDV a été déterminée sur des ARNs coiffés méthylés en position N7 ou 2'O de la coiffe et sur tous les résidus A internes ($^{(m)}GpppX_{(m)}(A_m)$ -SUDV₁₂). Les extrémités 5' des ARNm de SUDV commençant par un G ou par un A [3], les deux formes d'ARNs ont été synthétisées (voir les séquences dans l'**Annexe 3**). De cette manière, les ARNs GpppG_m(A_m)-SUDV₁₂ et GpppA_m(A_m)-SUDV₁₂ ont permis de suivre spécifiquement l'activité N7 MTase (en vert, **Figure 35A**). Aucune activité n'a été détectée sur l'ARN GpppA_m(A_m)-SUDV₁₂ quel que soit le pH suggérant que cet ARN n'est pas un bon substrat pour l'activité N7 MTase de SUDV. Au contraire, une activité est détectée sur l'ARN GpppG_m(A_m)-SUDV₁₂ à un pH entre 6,5 et 7,5. Cette activité n'est pas observée à pH plus élevé. Par analogie avec les résultats décrits sur la MTase de DENV, cette activité optimale entre pH 6,5 et 7 reflète probablement la N7 méthylation du substrat indiquant que le domaine MTase+CTD de SUDV porte une activité N7 MTase.

L'activité 2'O MTase de la coiffe a quant à elle été évaluée en incubant des ARNs m GpppG(A_m)-SUDV₁₂ et m GpppA(A_m)-SUDV₁₂ (en violet foncé, **Figure 35A**) en présence du domaine MTase+CTD

de SUDV. Aucune activité 2'O MTase significative n'est détectée sur l'ARN $mGpppG(A_m)$ -SUDV₁₂, alors qu'une activité MTase est détectée sur les $mGpppA(A_m)$ -SUDV₁₂. Cette activité MTase a un profil en fonction du pH similaire aux profils observés avec les ARNs GpppG-SUDV₁₂ et $mGpppG_m$ -SUDV₁₂. Cette observation indiquerait que le domaine MTase+CTD de SUDV porterait une activité 2'O MTase de la coiffe, détectable uniquement sur des ARN de 13 nucléotides commençant par un A. Il est intéressant de noter que le profil d'activité en fonction du pH est similaire pour les activités de méthylation interne et cette activité 2'O MTase de la coiffe. Cette observation étaye le fait que l'activité 2'O MTase et l'activité interne concernent la position 2'O des résidus adénosines.

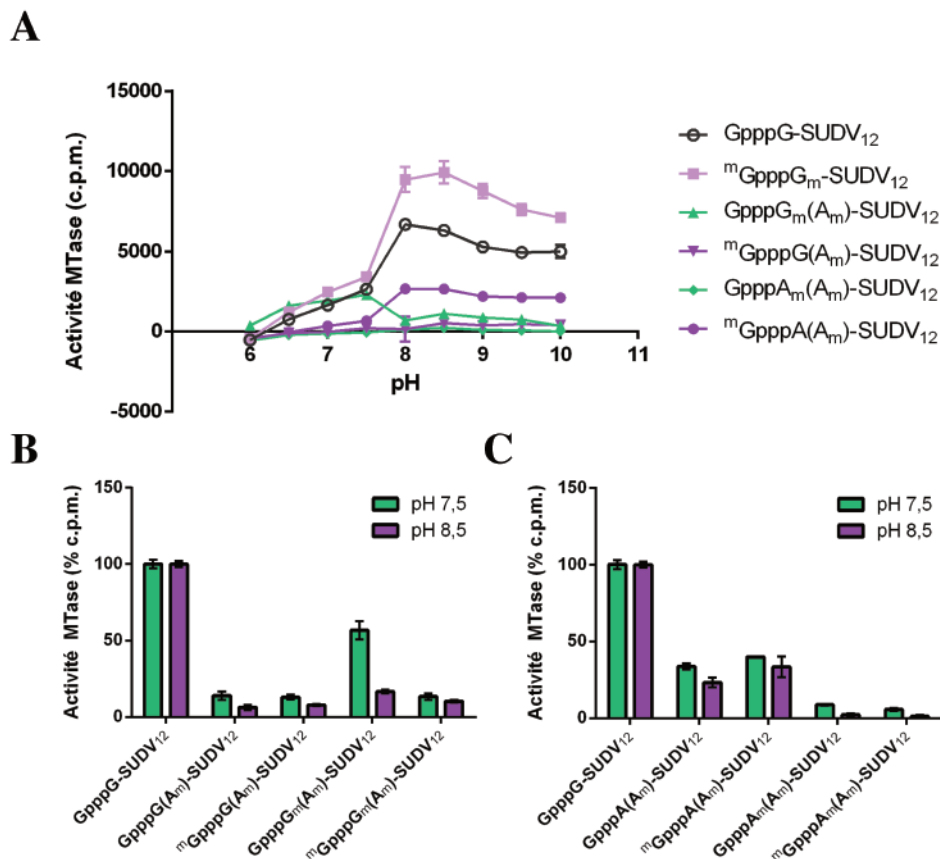


Figure 35 – Les activités MTases de coiffe de SUDV sont dépendantes du pH.

Bien que l'activité MTase en interne semble l'activité principale du domaine MTase+CTD de SUDV, ce domaine présente aussi les activités N7 et 2'O MTases de la coiffe. (A) Test d'activité MTase sur différents ARNs spécifiquement synthétisés pour détecter la N7 méthylation de la coiffe (vert), la 2'O méthylation de la coiffe (violet foncé), ou l'adénosine-2'O méthylation interne (violet clair) à différents pHs (mX : coiffe N7-méthylée, X_m : résidu 2'O-méthylé, (X_m) : ARN méthylé sur tous les résidus X en aval de la coiffe) (n=3). L'ARN coiffé non méthylé commençant par un G a été utilisé comme contrôle. Les points représentent les moyennes brutes \pm les écart-types. (B) Test d'activité MTase à pH 7,5 et 8,5 sur des ARNs synthétiques coiffés plus ou moins méthylés commençant par un G (mX : coiffe N7-méthylée, X_m : résidu 2'O-méthylé, (X_m) : ARN méthylé sur tous les résidus X en aval de la coiffe) (n=3). L'ARN coiffé non méthylé commençant par un G a été utilisé comme contrôle. Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé \pm les écart-types. (C) Test d'activité MTase à pH 7,5 et 8,5 sur des ARNs synthétiques coiffés plus ou moins méthylés commençant par un A (mX : coiffe N7-méthylée, X_m : résidu 2'O-méthylé, (X_m) : ARN méthylé sur tous les résidus X en aval de la coiffe) (n=3). L'ARN coiffé non méthylé commençant par un G a été utilisé comme contrôle. Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé \pm les écart-types.

1.2. Activités MTases sur la coiffe

Le processus de formation de la coiffe a été très peu étudié chez les filovirus. Il a cependant été démontré que les ARNm des filovirus possèdent une coiffe [4] et que les structures secondaires localisées en 5' des ARNm n'ont pas la capacité à elles seules d'initier la traduction comme dans le cas d'IRES [5]. La présence de la coiffe sur les ARNm a été détectée à l'aide d'un anticorps anti ^mGTP, ce qui ne permet pas de conclure si les ARNs des filovirus portent une structure de type coiffe-0 ou coiffe-1.

Les résultats préliminaires obtenus sur la MTase de SUDV en fonction du pH suggèrent que le domaine MTase+CTD de SUDV catalysent deux activités vraisemblablement portées sur les positions N7 et 2'O dont les pHs optimaux sont respectivement de 7,5 et de 8,5. Cependant, les deux activités sont observées sur deux ARNs de séquences différentes (un commençant par un G et l'autre par un A) suggérant que soit les conditions optimales ne sont pas réunies pour observer ces activités sur un même ARN ou que les coiffes des ARNm de SUDV puissent être différemment méthylées selon leur séquence. Afin de caractériser plus en détail les profils de méthylation sur les ARNs commençant par G ou A, deux séries d'ARNs ont été testées aux pHs 7,5 et pH 8,5 (**Figure 35B et 35C**). Chaque série est composée d'ARNs coiffés, méthylés sur tous les A en interne et plus ou moins méthylés sur la coiffe.

Pour les ARNs commençant par un G (**Figure 35B**), aucune activité n'est détectée sur les ARNs coiffés, avec une coiffe-0 ou une coiffe-1 quel que soit le pH. Cependant, une activité significative est détectée sur l'ARN coiffé préalablement 2'O-méthylé pour un pH à 7,5. Ces résultats confirment les observations précédentes qui suggèrent donc qu'une N7 méthylation serait catalysée spécifiquement à pH neutre. En outre, il est intéressant de noter que l'activité N7 MTase n'est observée que sur les ARNs préalablement méthylés en 2'OH du premier nucléotide. Cette observation suggère que la 2'O méthylation des structures coiffes doit précéder leur N7 méthylation. Cet ordre réactionnel obligatoire a déjà été proposé chez d'autres mononegavirus (présenté au **Chapitre 6**). Toutefois, il est étonnant que la 2'O méthylation du premier G de l'extrémité 5' de l'ARN n'est pas réalisable dans une large gamme de conditions testées. Plusieurs raisons peuvent expliquer cette observation. Il est en effet possible que les ARNs utilisés ne soient pas adéquats en terme de séquence et/ou de longueur pour permettre la 2'O méthylation de la coiffe sur les ARNs commençant par un G. Il est également possible que chez SUDV, la 2'O méthylation puisse nécessiter l'intervention d'une 2'O MTase cellulaire.

Pour les ARNs commençant par un A (**Figure 35C**), une activité est détectée sur les ARNs coiffés et avec une coiffe-0 quel que soit le pH. Il est d'ailleurs intéressant de noter que le pourcentage de méthylation par rapport au contrôle est équivalent d'un pH à un autre. Aucune activité n'est mesurée sur l'ARN coiffé et 2'O-méthylé, ni sur l'ARN avec une coiffe-1, indiquant qu'il s'agirait d'une méthylation en position 2'O du ribose du premier A. De plus, il semble que l'activité 2'O MTase soit bien la première réaction de méthylation de la coiffe dans la mesure où cette activité ne discrimine pas les ARNs coiffés des ARNs avec une coiffe-0. En outre, bien que cette activité soit augmentée à pH 8 (**Figure 35A**), l'activité portée sur ce premier A suit l'activité globale portée sur les A internes (**Figure 35C**) suggérant que cette activité fasse appel au même mécanisme catalytique que l'activité 2'O interne. Il est cependant surprenant de ne pas observer d'activité N7 MTase sur l'ARN GpppA_m(A_m)-SUDV₁₂. Ce substrat est de même longueur mais pas la même séquence que son

homologue commençant par un G. Ces observations suggèrent que la N7 méthylation nécessiterait des conditions particulières comme des ARNs de taille optimale ou formant des structures secondaires particulières (voir 4.). Afin de confirmer cette hypothèse il serait intéressant d'étudier l'activité N7 MTase sur des ARNs de longueur croissante capables de former des structures tige/boucle conservées à l'extrémité des ARNs viraux. Enfin, il est également possible que cette activité soit régulée au sein de la protéine L entière par d'autres domaines ou couplée à l'activité polymérase.

1.3. Mécanismes réactionnels des N7 et 2'O MTases de la coiffe

Il a été possible de détecter l'activité N7 MTase sur un ARN commençant par un G et l'activité 2'O MTase sur un ARN commençant par un A. L'activité N7 MTase nécessite un pH plutôt neutre et un ARN préalablement méthylé en 2'O suggérant que la 2'O méthylation précède la N7 méthylation. L'activité 2'O est optimale à un pH plutôt élevé (> 7,5). Des données similaires ont été décrites chez d'autres mononégavirus comme le VSV (rhabdovirus) et le virus Sendai (SeV, paramyxovirus). Chez VSV, l'activité N7 MTase est sensible aux pHs alcalins au contraire de l'activité 2'O MTase qui elle supporte une large gamme de pH (de 7 à 10) [2]. Chez SeV, seule la N7 MTase a été détectée et son activité optimale a été mesurée pour un pH autour de 6,0-6,5 [6]. Ces résultats sont donc très similaires à ceux observés pour les activités MTases chez SUDV et les flavivirus mentionnés précédemment [1].

Cette différence entre les pHs optimaux des activités N7 et 2'O MTases suggère donc que ces activités impliquent des mécanismes réactionnels différents. Cette hypothèse a également déjà été proposée chez les flavivirus pour lesquels des observations similaires ont été rapportées [1,7]. Des alignements structuraux ont révélé que les domaines MTases des flavivirus ont des similarités avec la N7 MTase d'*Encephalitozoon cuniculi* (Ecm1) et la 2'O MTase VP39 du virus de la vaccine [8-10]. Or, il a été montré que ces deux enzymes impliquent deux mécanismes réactionnels différents. En effet, la N7 méthylation de Ecm1 est catalysée grâce à une proximité directe et la géométrie des deux substrats [9] tandis que la 2'O MTase implique une déprotonation du groupement 2'OH qui peut ensuite réagir avec le méthyl du SAM par attaque nucléophile de type S_N2 [11, 12]. De plus, il a été montré chez hMPV que la N7 méthylation dépend majoritairement du résidu D de la tétrade catalytique contrairement à la 2'O méthylation qui nécessite l'ensemble de la tétrade [13]. Des résultats similaires ont été obtenus chez les flavivirus [7]. Les différents pHs optimaux obtenus pour les activités N7 et 2'O MTases de SUDV témoigneraient donc de ces mécanismes réactionnels qui seraient également conservés chez les mononégavirus.

2. Identification des activités N7 et 2'O MTases

2.1. Caractérisation par TLC des produits de méthylation de la coiffe

Afin de confirmer que les activités MTase observées sur les ARNs GpppG_m(A_m)-SUDV12 et GpppA(A_m)-SUDV12 sont bien des activités N7 MTase et 2'O MTase respectivement, les coiffes ont été caractérisées par chromatographie sur couche mince (ou TLC) (**Figure 36**). Ces expériences ont été conduites sur des ARNs de séquences similaires à ceux décrits précédemment avec l'ajout d'une structure coiffe radiomarquée au ³²P (**Chapitre 9**). Ces ARNs sont incubés avec le domaine MTase+CTD de SUDV en présence de SAM et, après méthylation, ils sont digérés par la nucléase P1 puis les coiffes ont été caractérisées par TLC.

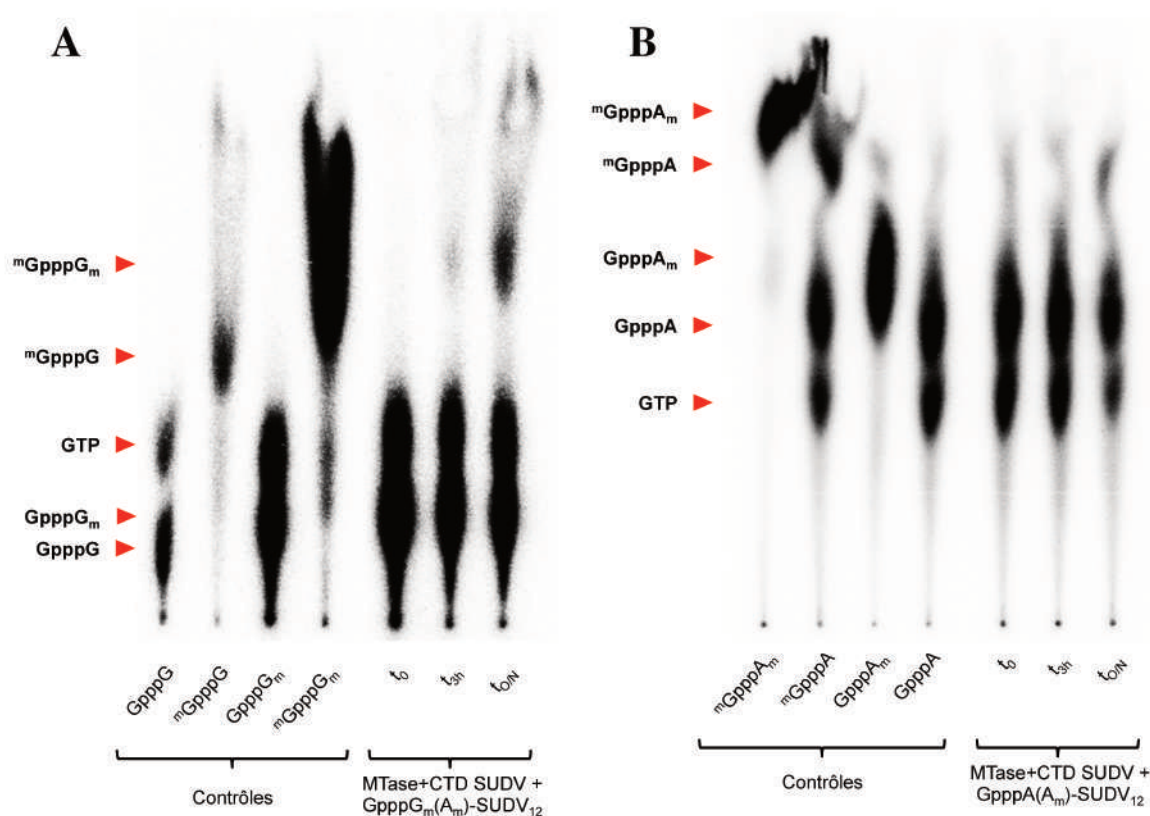


Figure 36 – Les activités MTases de coiffe sont portées sur les positions N7 de la coiffe et 2'O du premier nucléotide de l'extrémité 5' de l'ARN.

L'activité N7 MTase semble ressortir sur un ARN commençant par un G alors que l'activité 2'O MTase semble ressortir sur un ARN commençant par un A. (A) Chromatographie sur couche mince des différents types de coiffe d'un ARN commençant par un G et des coiffes générées en fonction du temps (0h, 3h et sur la nuit) par le domaine MTase+CTD de SUDV sur un ARN coiffé commençant par un G déjà 2'O-méthylé. (B) Chromatographie sur couche mince des différents types de coiffe d'un ARN commençant par un A et des coiffes générées en fonction du temps (0h, 3h et sur la nuit) par le domaine MTase+CTD de SUDV sur un ARN coiffé commençant par un A.

2.2. N7 méthylation

Dans un premier temps, les coiffes ($mGpppG_m$) sont séparées grâce à une phase mobile de chlorure de lithium (Figure 36A). Afin de préparer les contrôles, un ARN $pppG(A_m)$ -SUDV₁₂ a été coiffé avec du $\alpha[^{32}P]$ -GTP pour d'obtenir une coiffe simple $GpppG$ ou coiffé puis incubé avec la hN7 pour obtenir une coiffe-0 $mGpppG$. Les contrôles $GpppG_m$ et coiffe-1 ont été obtenus avec le même procédé sur un ARN $pppG_m(A_m)$ -SUDV₁₂ qui a également été utilisé pour réaliser une cinétique de l'activité MTase du domaine MTase+CTD de SUDV à pH 7,5. Après migration et fluorographie, il a été possible d'identifier les différents types de coiffe. Alors que la coiffe $GpppG$ et la coiffe 2'O méthylée migrent peu, les coiffes -0 et -1 migrent davantage et sont détectées dans le haut de la TLC. Il semble que plus la coiffe est hydrophobe, plus celle-ci migre par rapport à la ligne de dépôt. Cependant, pour un même taux de méthylation ($mGpppG$ et $GpppG_m$), la position de celle-ci fait également varier la distance de migration ce qui montre que cette technique permet d'identifier avec certitude la nature de la méthylation. Lorsque les ARNs portant une structure coiffe $GpppG_m$ sont incubées avec le domaine MTase+CTD, une tâche co-migrant avec la structure coiffe-1 contrôle apparaît. Ce produit de réaction est majoritairement détecté après une nuit d'incubation dans les

expériences cinétiques. De fait, l'activité sur l'ARN GpppG_m(A_m)-SUDV₁₂ est bien une activité N7 MTase de la coiffe, comme suggéré précédemment.

2.3. 2'O méthylation

Afin d'identifier la méthylation observée sur l'ARN commençant par un A, le même type d'expérience a été réalisé sur les ARNs pppA(A_m)-SUDV₁₂ et pppA_m(A_m)-SUDV₁₂ pour générer les coiffes contrôles (**Figure 36B**). L'ARN pppA(A_m)-SUDV₁₂ a été incubé avec le domaine MTase+CTD de SUDV à pH 8,5 afin de suivre la cinétique réactionnelle. Après libération des structures coiffes, leur séparation par TLC a été réalisée en présence de chlorure de lithium puis de sulfate d'ammonium afin de mieux discriminer les coiffes simples des coiffes 2'O-méthylées. Les différents types de coiffe migrent plus loin de la ligne de dépôt que les coiffes impliquant un G mais le même ordre de migration est observé. Cependant, les structures GpppA et GpppA_m ne sont pas complètement résolues. Malgré cette difficulté technique, la **Figure 36B** montre qu'un produit minoritaire se détachant légèrement de la tâche correspondant au GpppA est observable après plus de 3h d'incubation. Etant donné que l'activité mesurée en test MTase est faible, il n'est pas étonnant que la coiffe GpppA_m générée par le domaine MTase+CTD de SUDV soit difficilement détectable.

2.4. Conclusions

En résumé, l'activité observée par test d'activité sur l'ARN GpppG_m(A_m)-SUDV₁₂ à pH 7,5 et donnant une coiffe qui migre comme une coiffe-1_mGpppG_m en TLC montre que le domaine MTase+CTD de SUDV porte une activité N7 MTase. De même, l'activité mesurée sur un ARN GpppA(A_m)-SUDV₁₂ dont le pH optimal est de 8-8,5 et qui donne une structure qui est vraisemblablement GpppA_m. Ainsi, le domaine MTase+CTD de SUDV porte bien les activités N7 et 2'O MTases déjà identifiées chez VSV et hMPV [2,13]. Cependant, ces réactions restent faibles et la 2'O méthylation de la coiffe n'a pas encore été clairement élucidée. Il est possible que le substrat nécessaire à cette activité soit un ARN non coiffé et que cette 2'O MTase précède l'ajout de la coiffe comme il l'a été suggéré chez hMPV [13]. Cette hypothèse a été testée par la suite.

3. Implication de la coiffe pour la 2'O méthylation

Nos résultats préliminaires suggèrent que la 2'O méthylation précéderait la N7 méthylation chez les filovirus. L'activité 2'O MTase de la coiffe n'étant pas clairement détectée, il est possible que celle-ci précède également l'ajout de la coiffe.

Afin de vérifier cette hypothèse, l'activité MTase a été évaluée sur des ARNs triphosphates commençant par un G ou un A aux pHs 7,5 et 8,5 (**Figure 37**). Aucune activité significative n'a été détectée sur ces ARNs suggérant que l'activité 2'O MTase de la coiffe ne précède pas l'ajout de la coiffe. Cette expérience met aussi en évidence qu'un ARN pppA(A_m)-SUDV₁₂ ne peut pas être méthylé sur le premier A (**Figure 37B**). L'absence d'activité sur un ARN non coiffé et méthylé en interne indique que la méthylation du premier nucléotide, notamment s'il s'agit d'un résidu adénosine, dépend probablement de la présence de la coiffe. Il semble donc que la 2'O méthylation observée précédemment sur le GpppA(A_m)-SUDV₁₂ soit à la fois une 2'O méthylation interne et une méthylation de la coiffe et que cette 2'O méthylation succède l'ajout de la coiffe.

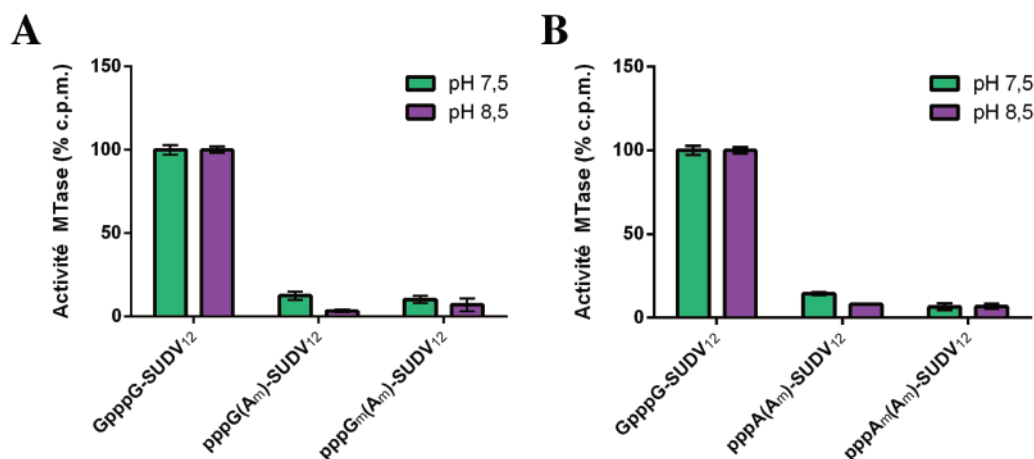


Figure 37 – La 2'O méthylation ne précède pas l'ajout de la coiffe chez SUDV.

La 2'O méthylation de la coiffe précède la N7 méthylation de la coiffe et pourrait même précéder l'ajout de la coiffe. (A) Test d'activité MTase à pH 7,5 et 8,5 sur des ARNs synthétiques non coiffés plus ou moins méthylés commençant par un G (X_m : résidu 2'O-méthylé, (X_m): ARN méthylé sur tous les résidus X en aval de la coiffe) (n=3). L'ARN coiffé non méthylé a été utilisé comme contrôle. Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé \pm les écart-types. (C) Test d'activité MTase à pH 7,5 et 8,5 sur des ARNs synthétiques non coiffés plus ou moins méthylés commençant par un A (X_m : résidu 2'O-méthylé, (X_m): ARN méthylé sur tous les résidus X en aval de la coiffe) (n=3). L'ARN coiffé non méthylé a été utilisé comme contrôle. Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé \pm les écart-types.

Grâce à cette analyse biochimique des méthylations d'ARNs portant différentes structures coiffes, la synthèse de la coiffe pourrait suivre l'ordre démontré chez VSV [2] : ajout de la coiffe, puis 2'O méthylation et N7 méthylation. Cependant, dans ce travail, nous n'avons pas réussi à identifier les conditions expérimentales permettant la détection de l'activité 2'O MTase de la coiffe sur les ARNs commençant par un G alors que cette activité est détectable sur des ARNs commençant par un A. Plusieurs hypothèses sont envisageables pour la poursuite de cette caractérisation. En effet, les ARNm des différentes souches du virus Ebola forment des structures tige/boucle qui pourraient favoriser les méthylations de la coiffe.

4. Structuration des ARNs de SUDV et méthylations

4.1. Extrémités 5' des ARNm des mononégavirus

L'étude du génome et des transcrits de SUDV [3] a révélé que les extrémités 5' des ARNm possèdent une structure tige/boucle (**Figure 38**). Ces structures sont variables en longueur (de 24 à 72 nucléotides) et en séquence mais possèdent toutes une séquence consensus de 12 nucléotides (GA(U/G)GAAGAUUAA) vraisemblablement impliquée dans la transcription comme chez les autres mononégavirus [14-16]. Il semble également que cette séquence participe de manière significative à la formation de ces structures tige/boucle puisque 6 à 11 résidus de cette séquence s'apparient avec une séquence complémentaire en aval. Enfin, il est remarquable que ces boucles ont à l'extrémité 5' soit une séquence commençant par un G qui est directement impliqué dans la formation de la structure tige soit une séquence commençant par un A qui lui n'est pas apparié. Etant donné les différences observées en test d'activité MTase sur des ARNs commençant par un A ou un G, il est possible que ces structures jouent un rôle dans la régulation des activités MTases de la coiffe.

Chez VSV, il a été montré *in vitro* que les N7 et 2'O méthylations de la coiffe avec la protéine L entière sont détectées sur des ARNs coiffés d'au moins 10 nucléotides [2]. Cependant, lorsque la synthèse complète de la coiffe (ajout de la coiffe et méthylations) est reproduite *in vitro* avec la protéine L entière, un ARN d'au moins 31 nucléotides de long est nécessaire [17]. Or, lorsque ces 31 nucléotides de l'extrémité 5' des différents ARNs du VSV sont étudiés pour leur structuration grâce à l'outil mFold [18], il est remarquable d'observer que ces extrémités forment également une structure tige/boucle (non illustré). Ces structures pourraient donc être impliquées dans la réaction PRNTase qui conditionne les réactions MTases. Des structures similaires peuvent également être décrites chez d'autres mononégavirus tels que hMPV, hRSV ou RABV, confortant l'hypothèse d'un rôle probable de ces structures tige/boucle dans la synthèse de la coiffe des ARNm des mononégavirus.

4.2. N7 méthylation et structure tige/boucle

De telles structures ont été décrites chez d'autres virus tels que les flavivirus qui sont des virus à ARN positif. Chez les flavivirus, l'ARN génomique est également l'ARNm qui permet la synthèse des différentes protéines virales. Cet ARN génomique est coiffé et méthylé sur les positions N7 et 2'O de la coiffe. Les activités N7 et 2'O MTase requièrent des conditions biochimiques différentes. Alors que la 2'O méthylation est observée sur des ARNs courts (≥ 4 nucléotides) [19] et des ARNs plus longs (74 nucléotides) [20], la N7 méthylation n'est détectée que sur des ARNs d'au moins 74 nucléotides [20,21]. Ces 74 nucléotides du génome viral forment une structure tige/boucle nommée SLA [20]. Chez les flavivirus, la N7 précède la 2'O méthylation [7, 22] ce qui suggère que la boucle SLA est nécessaire pour les méthylations de la coiffe des flavivirus.

La présence de structures tige/boucle similaires chez SUDV pourrait aussi être nécessaire pour réaliser la N7 méthylation et/ou la 2'O méthylation du premier nucléotide. Cette hypothèse sera évaluée prochainement au laboratoire en étudiant par TLC les méthylations sur des ARNs de taille croissante, capable ou non de former ces structures tige/boucle. L'enjeu technique dans cette approche sera de disposer d'ARN méthylés sur les A en positions internes afin de permettre la ou les méthylations efficaces de la coiffe.

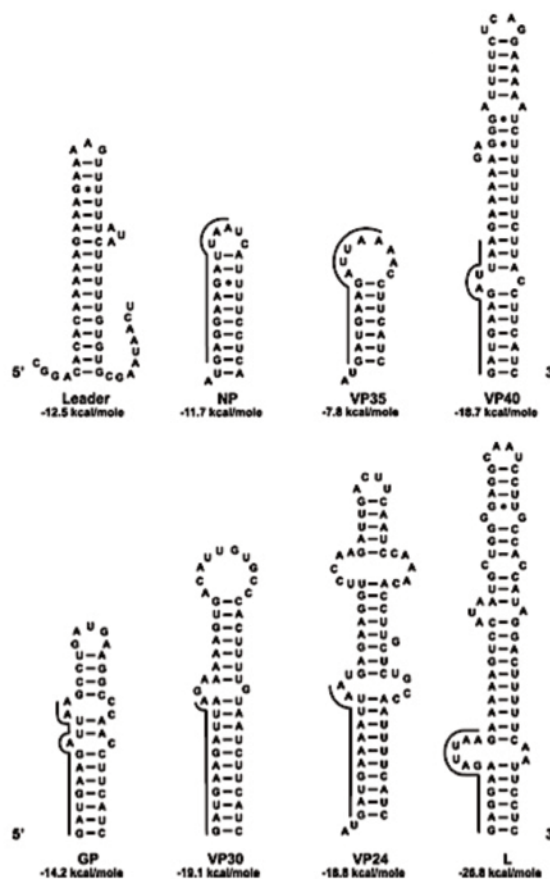


Figure 38 – Les extrémités des ARNm de SUDV forment des structures secondaires en tige/boucle.

L'analyse bioinformatique des extrémités 5' des ARNm de SUDV a suggéré qu'ils formeraient tous une structure tige/boucle plus ou moins longue. La séquence consensus de 12 nucléotides a été soulignée. Cette séquence très conservée (GA(U/G)GAAGAUUAA) forme l'extrémité 5' de 4 ARNm (VP40, GP, VP30 et L) qui est totalement impliquée dans la formation de la boucle au contraire des 3 autres (NP, VP35 et VP24) qui eux ont un motif AU sortant.

Adapté de Sanchez *et al.* (2005).

4.3. Présence ou absence de 2'O méthylation de la coiffe ?

Les alphavirus, virus à ARN positif comme les flavivirus, possèdent également une structure tige/boucle appelée CSE (conserved sequence element) ou SL (stem loop) à l'extrémité 5' de l'ARN viral qui est coiffée mais méthylée seulement en position N7 [23]. Lorsque cette structure est altérée, les senseurs de l'immunité innée sont capables de détecter les ARNs viraux et de déclencher les voies de la réponse immunitaire innée ce qui souligne l'implication de cette structure CSE ou SL dans l'échappement des alphavirus à l'immunité de l'hôte [24].

Chez SUDV, les résultats concernant la 2'O méthylation suggèrent que seuls les ARNs commençant par un A sont méthylés en position 2'O de la coiffe. Or, les structures tige/boucle de SUDV montrent que les ARNm commençant par un A (NP, VP35 et VP24) ont une structure tige/boucle dont l'extrémité 5' a deux résidus sortants (AU). Au contraire, les ARNm commençant par un G possèdent une structure tige/boucle sans résidu sortant. Il est donc possible que les ARNm commençant par un A ou un G soient discriminés afin de ne masquer que les extrémités sortantes grâce à la 2'O méthylation de la coiffe, les extrémités sans résidu sortant étant protégées des senseurs de l'immunité innée grâce à la structure secondaire de l'ARN. Afin de vérifier une telle hypothèse, une analyse des méthylations de la coiffe sur des ARNs plus longs permettra de confirmer les observations décrites sur des ARNs courts et une étude en génétique inverse permettra d'évaluer l'implication des structures secondaires des ARNs de SUDV dans l'échappement du filovirus à l'immunité innée.

Il est à noter par ailleurs que, chez les filovirus, il pourrait exister une diversité de séquences à l'extrémité 5' des ARNm. En effet, chez le virus Marburg, tous les ARNm commencent par la séquence consensus GA(U/G)GAAGAUUAA [25], contrairement aux ARNm de EBOV ou SUDV [26]. Les implications de ces possibles différences devront faire l'objet d'une attention particulière notamment pour évaluer l'existence d'une activité 2'O MTase de la coiffe spécifique et le rôle de cette activité 2'O MTase sur la méthylation en position N7.

5. Vers des mécanismes de méthylation plus complexes

Les résultats préliminaires obtenus ont montré que le domaine MTase+CTD de SUDV est capable de catalyser les N7 et 2'O méthylations. Contrairement à hMPV et VSV, les substrats d'ARNs courts ne permettent pas de reconstituer efficacement la séquence de méthylations de la coiffe. Des structures tige/boucle à l'extrémité 5' des ARNm de certains virus sont impliquées dans la régulation des méthylations de la coiffe chez d'autres familles de virus. Des structures semblables ont été mises en évidence chez SUDV et pourraient donc être nécessaires pour la synthèse de la coiffe. L'étude de ces structures secondaires chez les filovirus pourrait révéler des fonctions complémentaires dans la régulation des modifications co-transcriptionnelles des ARNm et de leurs rôles fonctionnels associés. Cette étude préliminaire met en relief la diversité des mécanismes réactionnels des activités MTases chez les mononégavirus qui pourtant sont portées par la protéine la plus conservée au sein de cet ordre viral. La compréhension de cette diversité des relations structure/fonction au sein des différents domaines de la protéine L souligne le manque de données structurales qui permettront de faire la lumière sur les bases moléculaires de cette variabilité.

Bibliographie

- [1] : Dong H, Chang DC, Xie X, Toh YX, Chung KY, Zou G, Lescar J, Lim SP, Shi P. Biochemical and genetic characterization of dengue virus methyltransferase. *Virology*. 2010 Sep 30;405(2):568-78.
- [2] : Rahmeh AA, Li J, Kranzusch PJ, Whelan SP. 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*. 2009 Nov;83(21):11043-50.
- [3] : Sanchez A, Rollin PE. Complete genome sequence of an Ebola virus (Sudan species) responsible for a 2000 outbreak of human disease in Uganda. *Virus Res*. 2005 Oct;113(1):16-25.
- [4] : Weik M, Modrof J, Klenk HD, Becker S, Mühlberger E. Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *J Virol*. 2002 Sep;76(17):8532-9.
- [5] : Shabman RS, Hoenen T, Groseth A, Jabado O, Binning JM, Amarasinghe GK, Feldmann H, Basler CF. An upstream open reading frame modulates ebola virus polymerase translation and virus replication. *PLoS Pathog*. 2013 Jan;9(1):e1003147.
- [6] : 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*. 2005 Feb 11;280(6):4429-35.
- [7] : Ray D, Shah A, Tilgner M, Guo Y, Zhao Y, Dong H, Deas TS, Zhou Y, Li H, Shi PY. West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O methylations by nonstructural protein 5. *J Virol*. 2006 Sep;80(17):8362-70.
- [8] : Egloff MP, Benarroch D, Selisko B, Romette JL, Canard B. An RNA cap (nucleoside-2'-O)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. *EMBO J*. 2002 Jun 3;21(11):2757-68.
- [9] : Fabrega C, Hausmann S, Shen V, Shuman S, Lima CD. Structure and mechanism of mRNA cap (guanine-N7) methyltransferase. *Mol Cell*. 2004 Jan 16;13(1):77-89.
- [10] : Shiryayev SA, Ratnikov BI, Chekanov AV, Sikora S, Rozanov DV, Godzik A, Wang J, Smith JW, Huang Z, Lindberg I, Samuel MA, Diamond MS, Strongin AY. Cleavage targets and the D-arginine-based inhibitors of the West Nile virus NS3 processing proteinase. *Biochem J*. 2006 Jan 15;393(Pt 2):503-11.
- [11] : 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.
- [12] : Hager J, Staker BL, Bugl H, Jakob U. Active site in RrmJ, a heat shock-induced methyltransferase. *J Biol Chem*. 2002 Nov 1;277(44):41978-86.
- [13] : Paesen GC, Collet A, Sallamand C, Debart F, Vasseur JJ, Canard B, Decroly E, Grimes JM. X-ray structure and activities of an essential Mononegavirales L-protein domain. *Nat Commun*. 2015 Nov 9;6:8749.
- [14] : Stillman EA, Whitt MA. Mutational analyses of the intergenic dinucleotide and the transcriptional start sequence of vesicular stomatitis virus (VSV) define sequences required for efficient termination and initiation of VSV transcripts. *J Virol*. 1997 Mar;71(3):2127-37.
- [15] : Whelan SP, Wertz GW. Regulation of RNA synthesis by the genomic termini of vesicular stomatitis virus: identification of distinct sequences essential for transcription but not replication. *J Virol*. 1999 Jan;73(1):297-306.
- [16] : Stillman EA, Whitt MA. Transcript initiation and 5'-end modifications are separable events during vesicular stomatitis virus transcription. *J Virol*. 1999 Sep;73(9):7199-209.
- [17] : Tekes G, Rahmeh AA, Whelan SP. A freeze frame view of vesicular stomatitis virus transcription defines a minimal length of RNA for 5' processing. *PLoS Pathog*. 2011 Jun;7(6):e1002073.
- [18] : Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. 2003 Jul 1;31(13):3406-15.

- [19] : Selisko B, Peyrane FF, Canard B, Alvarez K, Decroly E. Biochemical characterization of the (nucleoside-2'O)-methyltransferase activity of dengue virus protein NS5 using purified capped RNA oligonucleotides (7Me)GpppAC(n) and GpppAC(n). *J Gen Virol.* 2010 Jan;91(Pt 1):112-21.
- [20] : Dong H, Ray D, Ren S, Zhang B, Puig-Basagoiti F, Takagi Y, Ho CK, Li H, Shi PY. Distinct RNA elements confer specificity to flavivirus RNA cap methylation events. *J Virol.* 2007 May;81(9):4412-21.
- [21] : Chung KY, Dong H, Chao AT, Shi PY, Lescar J, Lim SP. Higher catalytic efficiency of N-7-methylation is responsible for processive N-7 and 2'-O methyltransferase activity in dengue virus. *Virology.* 2010 Jun 20;402(1):52-60.
- [22] : Zhou Y, Ray D, Zhao Y, Dong H, Ren S, Li Z, Guo Y, Bernard KA, Shi PY, Li H. Structure and function of flavivirus NS5 methyltransferase. *J Virol.* 2007 Apr;81(8):3891-903.
- [23] : Li C, Guillén J, Rabah N, Blanjoie A, Debart F, Vasseur JJ, Canard B, Decroly E, Coutard B. mRNA Capping by Venezuelan Equine Encephalitis Virus nsP1: Functional Characterization and Implications for Antiviral Research. *J Virol.* 2015 Aug;89(16):8292-303.
- [24] : Hyde JL, Gardner CL, Kimura T, White JP, Liu G, Trobaugh DW, Huang C, Tonelli M, Paessler S, Takeda K, Klimstra WB, Amarasinghe GK, Diamond MS. A viral RNA structural element alters host recognition of nonself RNA. *Science.* 2014 Feb 14;343(6172):783-7.
- [25] : Mühlberger E, Trommer S, Funke C, Volchkov V, Klenk HD, Becker S. Termini of all mRNA species of Marburg virus: sequence and secondary structure. *Virology.* 1996 Sep 15;223(2):376-80.
- [26] : Sanchez A, Kiley MP, Holloway BP, McCormick JB, Auperin DD. The nucleoprotein gene of Ebola virus: cloning, sequencing, and in vitro expression. *Virology.* 1989 May;170(1):81-91.

CHAPITRE 12

Vers la caractérisation structurale du domaine MTase du virus Soudan

Etant donné les originalités biochimiques que présente le domaine MTase+CTD de SUDV, une campagne de cristallogénèse a été initiée afin de tenter de résoudre la structure de ce domaine et de comprendre la relation structure/fonction de l'activité MTase chez SUDV. Une étude par protéolyse ménagée combinée à une analyse bioinformatique a été conduite dans un premier temps. Cette étude a permis d'identifier les domaines de la protéine les plus stables. A ce jour, seule une construction correspondant au domaine MTase a été cristallisée.

1. Identification d'un domaine stable par protéolyse ménagée et analyse bioinformatique

Les domaines MTase et MTase+CTD de SUDV caractérisés biochimiquement *in vitro* ne sont pas des constructions adaptées pour la cristallogénèse. En effet, le rendement de production/purification nécessaire pour une campagne de cristallogénèse requiert quelques dizaines de milligrammes de protéines. Or, notre procédé actuel de production et de purification ne permet pas d'avoir des quantités de protéines suffisantes ou demanderait un investissement trop lourd. Le processus de cristallisation peut être long et s'effectue à 20°C en première intention. Un premier criblage a été tenté sur le domaine MTase+CTD mais ces essais non pas été concluants. Dans ces conditions, nos constructions ne sont pas suffisamment stables pour permettre la mise en place d'un état cristallin en particulier à cause du tampon de conservation contenant de l'arginine, peu propice à la cristallisation. Par conséquent, de nouvelles constructions plus stables et plus faciles à exprimer en système bactérien ont été recherchées.

1.1. Protéolyse ménagée

Une étude des produits de digestion du domaine MTase+CTD par protéolyse ménagée a été réalisée (**Figure 39**). Cette technique permet, grâce à une gamme de dilutions de protéase, de digérer de façon plus ou moins « agressive » la protéine d'intérêt. Si les conditions de digestion sont douces, la protéase n'hydrolyse que les fragments désordonnés et superficiels de la protéine. Dans des conditions plus drastiques, seuls les fragments les plus stables résistent à la protéolyse. Après digestion, les produits sont analysés sur gel SDS-PAGE (**Figure 39A**).

1.1.1. Analyse sur gel SDS-PAGE

Dans notre étude, deux protéases ont été utilisées : la trypsine, une endoprotéase qui coupe après une lysine ou une arginine, et la chymotrypsine, une endoprotéase qui coupe après des résidus hydrophobes (Y, F, M, L et W). A partir d'un kit commercial (Hampton Research), une gamme de dilution de 100 µg/mL à 100 ng/mL pour chaque protéase a été préparée puis incubée pendant 20 min à 37°C en présence du domaine MTase+CTD de SUDV (2 mg/mL). Les deux protéases ont généré quatre produits de digestion majoritaires très similaires en fonction de la concentration en trypsine ou chymotrypsine. Une forme de 59 kDa, qui correspond au domaine MTase+CTD entier, est digérée pour donner la forme de 50 kDa à de faibles concentrations en protéase (1 µg/mL et 100 ng/mL). La protéine est digérée aux concentrations intermédiaires (10 µg/mL et 1 µg/mL) pour donner le fragment de 47 kDa. Le domaine MTase+CTD subit une digestion supplémentaire pour devenir la forme de 41 kDa qui résiste aux concentrations en protéases les plus élevées (100 µg/mL et 10 µg/mL). Ces observations indiquent que le domaine MTase+CTD entier utilisé pour l'étude biochimique est rapidement digéré en une forme plus courte à de faibles concentrations en protéase. Ce domaine n'est

donc effectivement pas un domaine suffisamment stable pour une étude structurale. En revanche, un domaine de 41 kDa résistant à des concentrations élevées de protéase a été identifié. Ce fragment est donc une forme très stable, intéressante pour des essais en cristallogénèse.

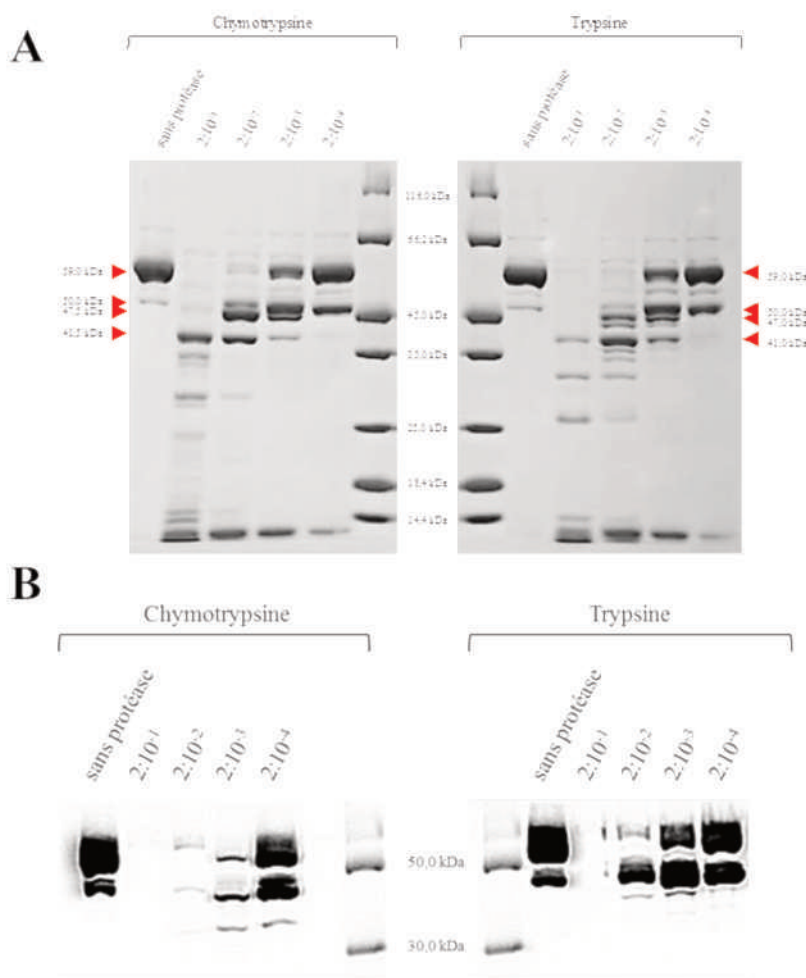


Figure 39 – La protéolyse ménagée a permis d’identifier un domaine stable pour la cristallogénèse.

Les domaines MTase+CTD et MTase de SUDV utilisés en biochimie ne sont pas assez stables et nécessitent des concentrations en arginine qui font que ces constructions ne sont pas adaptées pour mener une campagne de cristallogénèse. (A) Analyse SDS-PAGE des produits de digestion par protéolyse ménagée du domaine MTase+CTD de SUDV à 2 mg/mL par une gamme de dilution de 100 μ g/mL à 100 ng/mL (ratios protéine:protéase) des protéases chymotrypsine (à gauche) et trypsine (à droite) issues d’un kit commercial (Hampton Research). (B) Western blot des produits de digestion par protéolyse ménagée du domaine MTase+CTD de SUDV à 2 mg/mL par une gamme de dilution de 100 μ g/mL à 100 ng/mL (ratios protéine:protéase) des protéases chymotrypsine (à gauche) et trypsine (à droite) issues d’un kit commercial (Hampton Research). La détection a été réalisée par un anticorps anti-5-His (Qiagen), qui reconnaît spécifiquement l’étiquette 6-His du domaine MTase+CTD à son extrémité N-terminale.

1.1.2. Analyse par Western blot

Ces produits de digestion ont également été caractérisés par Western blot avec un anticorps anti-5-His (Qiagen) permettant de détecter spécifiquement le domaine MTase+CTD par son étiquette 6-His N-terminale (Figure 39B). Ainsi, il est possible de détecter si les produits de digestions contiennent encore l’étiquette 6-His et donc de déterminer si les digestions sont portées sur l’extrémité C-terminale ou N-terminale. De fait, si la protéine est digérée par l’extrémité C-terminale, les produits de digestion

portent toujours l'étiquette 6-His et sont donc détectés par Western blot. Au contraire, des produits digérés en N-terminal ont perdus l'étiquette et ne sont donc plus détectés par l'anticorps. Les fragments détectés par Western blot correspondent aux fragments de 59, 50 et 47 kDa observés sur gel SDS-PAGE. Conjointement avec l'analyse précédente par SDS-PAGE, il a été possible de montrer qu'un fragment total de 12 kDa est digéré du côté C-terminal pour générer le fragment de 47 kDa. Le fragment de 41 kDa n'est pas détecté par Western blot indiquant que l'hydrolyse de la forme de 47 kDa vers la forme de 41 kDa est portée sur l'extrémité N-terminale et potentiellement sur l'extrémité C-terminale aussi. Ce produit de 41 kDa est donc le produit d'une digestion aux deux extrémités, en particulier du côté C-terminal, qui donne un fragment très stable.

1.2. Analyse bioinformatique

Afin d'identifier les bornes N- et C-terminales de ce fragment de 41 kDa pour clonage et expression en système bactérien, une étude par spectrométrie de masse a été menée (non présentée). La séquence montre qu'il s'agit du domaine MTase délété de quelques dizaines de résidus à l'extrémité N-terminale et avec un fragment du domaine CTD en C-terminal. Cette séquence a été analysée à l'aide de MeDor (non illustré), un « méta-serveur » qui combine les analyses de différents outils bioinformatiques pour la prédiction des domaines structurés et désordonnés d'une protéine [1]. Cette analyse a confirmé la grande structurabilité de ce sous-domaine qui contient le cœur du domaine MTase. Or, pour éviter les artefacts structuraux et donc plus de cohérence, le fragment conservé du domaine CTD a été enlevé. La séquence générée pour l'étude structurale correspond donc à une forme du domaine MTase de 36 kDa où l'extrémité N-terminale désordonnée a été retirée (1744-2046) (contrôlée par spectrométrie de masse, **Annexe 2** (voir construction CRVI)).

2. Cristallogénèse du domaine MTase de SUDV seul

La séquence de ce domaine de 36kDa a été clonée dans le plasmide d'expression pDEST14 grâce au système Gateway (Invitrogen) avec une étiquette 6-His à l'extrémité N-terminale (6HN) ou C-terminale (6HC). Ces constructions ont été produites et purifiées avec un rendement supérieur aux constructions précédemment décrites (2 à 3 mg de protéine par litre de culture) et dans des conditions compatibles avec la cristallogénèse (conservation dans un tampon minimum Tris pH 8 50 mM, NaCl 150 mM, 5% glycérol). L'analyse par SDS-PAGE coloré au bleu de Coomassie indique que la protéine issue des deux étapes de purification est pure (à plus de 95%) (**Figure 40A**). En outre, l'analyse du profil de purification par chromatographie d'exclusion (**Figure 40B**) montre également que la protéine est déjà hautement purifiée dans l'éluat de la purification par affinité (IMAC). Le volume d'élution sur colonne Superdex S200 16/60 (89 mL) indique que ce domaine forme des dimères en solution (masse moléculaire apparente d'environ 75 kDa).

Afin de réaliser notre campagne de cristallogénèse sur le domaine MTase de SUDV, les formes 6HN et 6HC ont été concentrées respectivement à 7,1 et 5,0 mg/mL. Ces concentrats ont permis de cribler les conditions de cristallisation de trois kits commerciaux : Stura (48 conditions, Molecular Dimensions), Wizard (96 conditions, Molecular Dimensions) et JCSG core 1 (96 conditions, Qiagen). Les plaques de criblage ont ensuite été placées à 20°C. Après 2 jours d'incubation, des petits cristaux en forme de baguette sont apparus dans les conditions imidazole malate pH 7 0,2 M, PEG 4000 10-15-20% du kit Stura. Après 6 jours d'incubation, des baguettes plus épaisses se sont formées dans la

condition Bicine pH 9 PEG 6000 20% du kit JCSG core 1 (**Figure 40C**). Ces baguettes ont été photographiées en lumière blanche et en UV. Les cristaux sont détectés en UV indiquant que ces cristaux ne sont pas de nature saline donc potentiellement de nature protéique. Ces cristaux sont trop petits pour leur diffraction aux RXs mais ils ont permis de trouver des conditions de cristallisation favorables pour des optimisations. Malheureusement, à ce jour, aucun cristal de la forme 6HC n'a été obtenu.

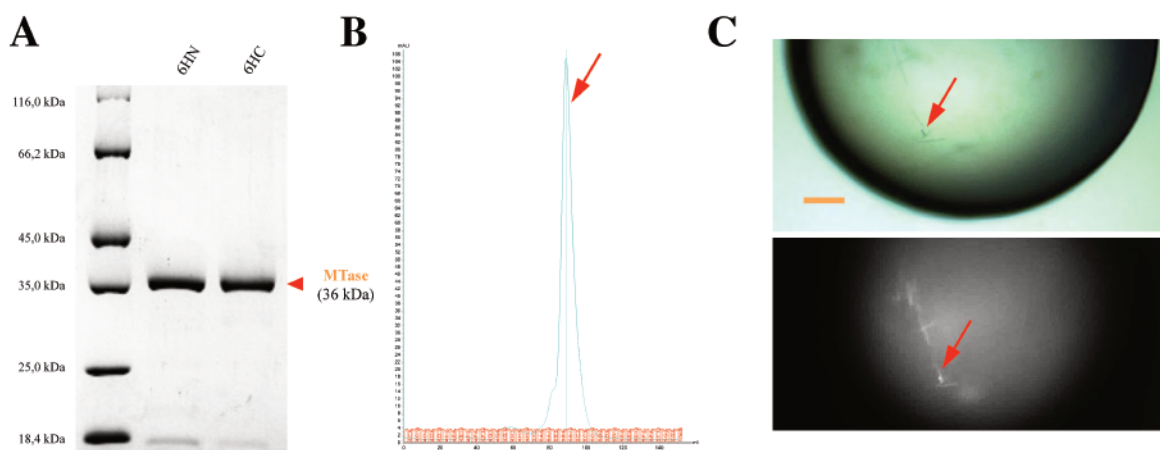


Figure 40 – Le domaine MTase de SUDV est cristallisable.

Un domaine MTase a été identifié comme étant plus stable que les formes générées pour l'étude biochimique. (A) Gel SDS-PAGE du domaine MTase stable avec une étiquette 6-His à l'extrémité N-terminale (6HN) ou C-terminale (6HC). (B) Profil de purification du domaine MTase 6HN de SUDV par chromatographie d'exclusion sur colonne Superdex200 16/60 dans un tampon Tris pH 8 50 mM, NaCl 150 mM, glycérol 5% (v/v). Le même profil est observé pour la forme 6HC. L'élution des protéines est suivie par la mesure de l'absorbance à 280 nm pour détecter les protéines (en bleu). Un seul pic (flèche rouge) est mesuré à un volume d'élution de 89 mL qui correspond à une forme dimérique du domaine MTase de SUDV. (C) Cristaux du domaine MTase 6HN de SUDV (flèches rouges) en lumière blanche (en haut) et en UV (en bas). Ces baguettes ont été obtenues dans la condition Bicine pH 9 PEG 6000 20% du kit JCSG core 1 pour la forme 6HN à 7,1 mg/mL après 6 jours d'incubation à 20°C. Echelle 100 μ m (orange).

3. Développement de VHHs anti-domaine MTase de SUDV

Pour faciliter la cristallisation de certaines protéines, une des stratégies envisageables est de développer des VHHs (ou nanobodies) contre la protéine d'intérêt. En effet, ces domaines hypervariables des anticorps de camélidés portent le paratope qui définit leur spécificité. Au contraire des anticorps humains dont le paratope est porté par deux chaînes (environ 25 kDa), le nanobody n'est constitué que d'une seule chaîne de moins de 15 kDa ce qui lui confère une très grande stabilité. Il est donc facile à purifier, permet de former des complexes protéiques avec sa cible qui sont plus stables et potentiellement plus faciles à cristalliser et aide à déterminer les phases nécessaires pour la résolution de la densité électronique [2].

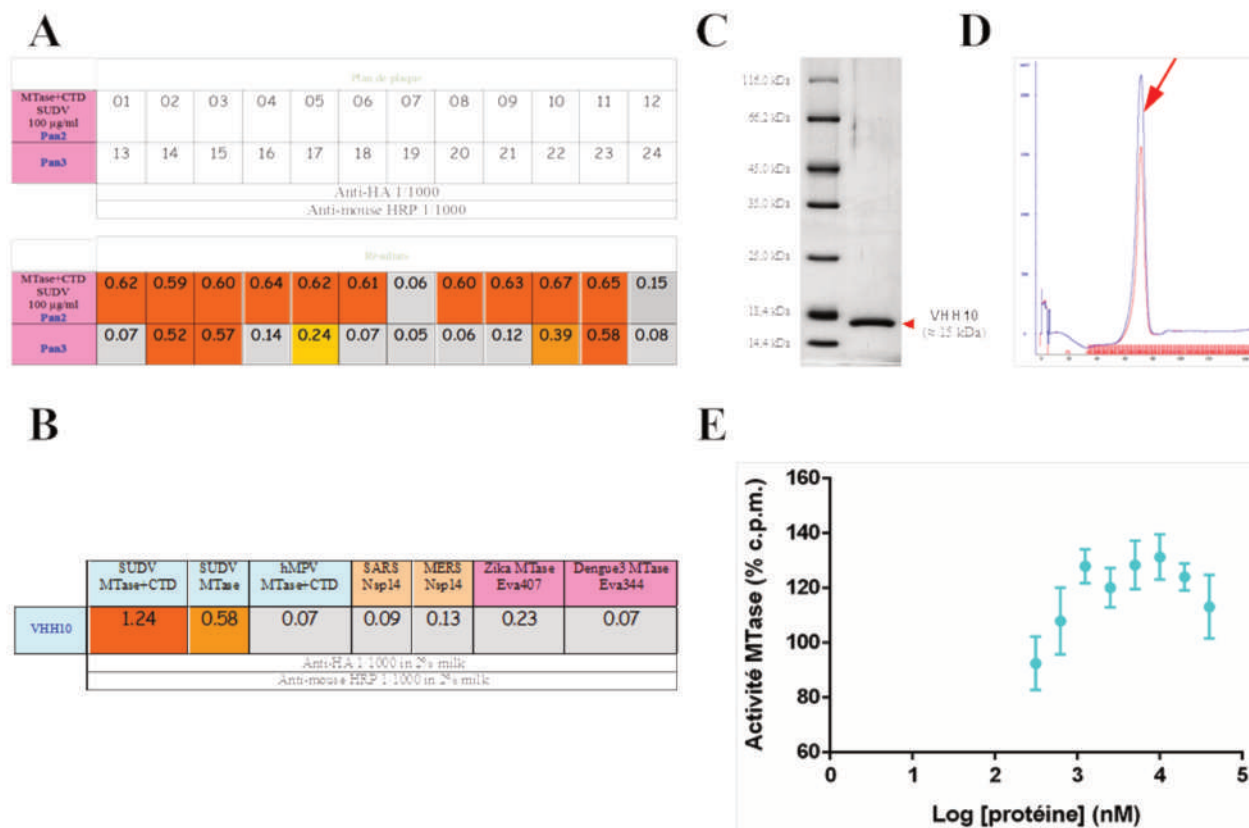


Figure 41 – Un VHH contre le domaine MTase a été identifié puis produit comme outil de cristallisation.

La cristallisation des protéines est facilitée lorsqu’elles sont complexées avec des nanobodies (ou VHH). (A) Identification de VHHs reconnaissant le domaine MTase+CTD de SUDV par test ELISA (n=1) après leur sélection dans une banque développée à partir d’un lama immunisé avec la forme 6HN du domaine MTase. L’intensité du signal a été modélisée par une gamme de couleurs froide (peu de signal) à chaude (signal fort). (B) Caractérisation de la spécificité du VHH10 par test ELISA sur différentes MTases virales avec une étiquette 6-His: les domaines MTase et MTase+CTD de SUDV et le domaine MTase+CTD de hMPV (mononégavirus, en bleu ciel), les protéines nsp14 (N7 MTase) de SARS-CoV et MERS-CoV (coronavirus, en orange) et les domaines MTases des virus de la Dengue et Zika (flavivirus, en rose). L’intensité du signal a été modélisée par une gamme de couleurs froide (peu de signal) à chaude (signal fort). (C) Analyse SDS-PAGE du VHH10 purifié. (D) Profil de purification du VHH10 par chromatographie d’exclusion sur colonne Superdex75 16/60 dans du PBS. L’élution des protéines est suivie par la mesure de l’absorbance à 280 nm pour détecter les protéines (en bleu). Un seul pic (flèche rouge) est mesuré à un volume d’élution de 71 mL qui correspond à une forme monomérique du VHH10. (E) Etude de l’effet du VHH10 sur l’activité MTase globale du domaine MTase+CTD de SUDV sur un ARN coiffé (GpppG-SUDV₁₂) (n=3). Les points représentent les moyennes normalisées par rapport au contrôle sans VHH ± les écart-types.

Un lama a été immunisé avec le domaine MTase de 36 kDa. Une banque de VHHs a ensuite été clonée à partir d’un extrait de sang de l’animal grâce à la plateforme Nanobodies du laboratoire (Desmyter A) comme décrit dans [3]. Après sélection et enrichissement, vingt-quatre clones (12 suite au second tour de sélection Pan2 et 12 suite à l’enrichissement Pan3) ont été choisis puis les VHHs ont été exprimés dans le périplasma bactérien (Figure 41A). La capacité de reconnaissance de la MTase de SUDV par les VHHs enrichis dans les extraits périplasmiques a ensuite été analysée par test ELISA avec un anticorps anti-HA. Treize des vingt-quatre extraits présentent un signal de chimiluminescence significatif indiquant l’interaction entre les VHHs et le domaine MTase+CTD préalablement fixé sur la plaque ELISA. Le séquençage des différents clones positifs a révélé qu’il s’agit d’un unique VHH.

La spécificité du clone 10 (VHH10) a été évaluée sur d'autres MTases virales étiquetées avec un 6-His (**Figure 41B**). Pour cela, l'extrait périplasmique brut du VHH10 a été testé par ELISA (comme précédemment décrit) sur les domaines MTase et MTase+CTD de SUDV et le domaine MTase+CTD de hMPV (mononégavirus, en bleu ciel) produite par nos collaborateurs (Grimes JM et Paesen GC, Oxford, UK), les protéines nsp14 (N7 MTase) de SARS-CoV et MERS-CoV (coronavirus, en orange) produites au laboratoire et les domaines MTases des virus de la Dengue et Zika (flavivirus, en rose) produits au laboratoire également. Seules les protéines de SUDV sont reconnues par le VHH10. Ce VHH reconnaît donc spécifiquement le domaine MTase de SUDV et aucune réaction croisée avec l'étiquette 6-His n'est détectable. De plus, le VHH10 reconnaît les différentes formes du domaine MTase utilisées dans ce travail. Ce VHH a donc été cloné dans un plasmide d'expression pHEN6 adapté pour la production périplasmique du VHH avec une étiquette 6-His à son extrémité C-terminale. Après expression et purification comme décrites dans [4], le VHH10 a été obtenu avec un degré de pureté estimé à 99% après analyse par SDS-PAGE coloré au bleu de Coomassie (**Figure 41C**). Enfin, le profil de chromatographie d'exclusion montre que le volume d'élution est de 71 mL sur une colonne Superdex75 16/60 (GE Healthcare) indiquant que le VHH est purifié sous forme d'un monomère d'environ 15 kDa en solution (**Figure 41D**).

Son interaction avec les domaines MTase et MTase+CTD permet de suggérer que le VHH10 reconnaît un épitope structurellement conservé au sein du domaine MTase. Une gamme de concentrations croissantes du VHH10 a donc été testée afin d'évaluer l'effet inhibiteur éventuelle du VHH sur l'activité MTase de SUDV (**Figure 41E**). Pour cela, la gamme de concentration a été préparée dans du tampon de réaction (Tris pH 8 50 mM, NaCl 150 mM, glycérol 5%) puis testée sur l'activité du domaine MTase+CTD sur un ARN coiffé spécifique de SUDV (GpppG-SUDV₁₂). L'activité MTase globale du domaine MTase+CTD n'est pas altérée même aux plus fortes concentrations en VHH utilisées. Au contraire, l'activité MTase est faiblement stimulée (environ 30%) pour des concentrations en VHH entre 1 et 10 μ M. Il semble donc que le VHH stabilise l'enzyme sans interagir directement avec un domaine de la protéine directement impliqué dans l'activité MTase (poche catalytique, site de fixation du SAM, gorge de fixation des ARNs).

4. Cristallogénèse du domaine MTase de SUDV en complexe avec le VHH10

Avec le développement d'un VHH contre le domaine de SUDV, il a été possible de former les complexes 6HN/VHH10 et 6HC/VHH10. Pour cela, les deux formes de la MTase ont été mélangées avec le VHH10 dans un rapport de concentration 1:2. Les mélanges ont ensuite été séparés par chromatographie d'exclusion (**Figure 42A**). Le profil observé sur colonne S200 16/60 est le même pour les deux complexes et présente 3 pics qui correspondent respectivement au complexe MTase/VHH (75 mL), au complexe contaminé par d'autres protéines très minoritaires (84 mL) et au VHH10 en excès (97 mL). Il semble donc que la complexation du VHH soit stoechiométrique et efficace à 100% puisque le domaine MTase n'est pas retrouvé seul. Le complexe élué à un volume de 75 mL indique que sa taille apparente correspond au dimère du domaine MTase associé à deux VHHs (un par domaine vraisemblablement). Les complexes 6HN/VVHH10 et 6HC/VHH10 contenus dans le premier pic de chaque chromatographie d'exclusion ont été concentrés à 7,6 et 6,7 mg/mL respectivement. Leur taux de pureté a été évalué à plus de 95% après concentration des fractions du pic 1 et analyse par SDS-PAGE coloré au bleu de Coomassie (**Figure 42B**).

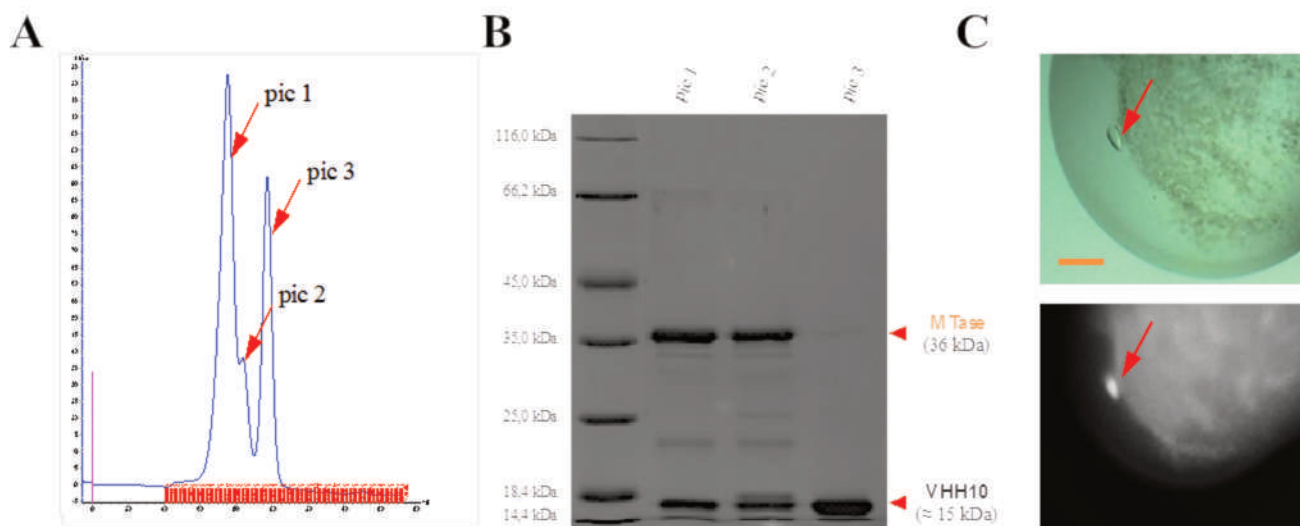


Figure 42 – Le complexe MTase/VHH est cristallisable.

Le complexe MTase/VHH10 est prometteur pour l'obtention de la première structure d'un domaine de la protéine L chez les filovirus. (A) Profil de purification du complexe entre le domaine MTase 6HC de SUDV et le VHH10 (6HC/VHH10) par chromatographie d'exclusion sur colonne Superdex200 16/60 dans un tampon Tris pH 8 50 mM, NaCl 150 mM, glycérol 5% (v/v). Le même profil est observé pour le complexe 6HN/VHH10. L'élué des protéines est suivie par la mesure de l'absorbance à 280 nm pour détecter les protéines (en bleu). Trois pics (flèches rouges) sont mesurés à des volumes d'élué de 75, 84 et 97 mL. Le pic 1 correspond au complexe 6HC/VHH10 le plus pur avec un volume d'élué indiquant un complexe de 4 protéines (dimère de 6HC portant 2 VHHs). (B) Analyse SDS-PAGE des différents pics d'élué en gel filtration. Le pic 1 correspond au complexe 6HC/VHH10 le plus pur. Le pic 2 contient également du complexe mais avec un taux de pureté moins important. Le pic 3 correspond au VHH10 excédentaire. (C) Cristaux du complexe 6HC/VHH10 (flèches rouges) en lumière blanche (en haut) et en UV (en bas). Ce cristal lenticulaire a été obtenu dans la condition potassium sodium tartrate tétrahydrate 0,8 M, sodium HEPES pH 7,5 0,1 M du kit MDL pour le complexe 6HC/VHH10 à 6,7 mg/mL après 6 jours d'incubation à 20°C. Echelle 100 μ m (orange).

La cristallisation du complexe 6HN/VHH10 a été testée à l'aide de cinq kits commerciaux : Stura, MDL et Wizard (Molecular Dimensions), JCSG core 1 et JCSG + (Qiagen). La quantité de complexe 6HC/VHH10 n'a permis de tester que trois kits commerciaux : Stura, MDL, et Wizard (Molecular Dimensions). Les plaques de criblage ainsi préparées ont été incubées à 20°C. Après 6 jours d'incubation, des cristaux de forme lenticulaire sont apparus dans de nombreuses conditions des kits Stura, MDL et JCSG +, en particulier pour le complexe 6HN/VHH10. Ces conditions ont été reportées dans le **Tableau 4**. La majorité des cristaux obtenus sont de trop petite taille pour tester leur diffraction des RXs. Cependant, un cristal assez gros (environ 30 μ m) a été photographié en lumière blanche et en UV confirmant la présence de protéines (**Figure 42C**). Ce cristal, obtenu avec le complexe 6HC/VHH10 dans la condition potassium sodium tartrate tétrahydrate 0,8 M, sodium HEPES pH 7,5 0,1 M du kit MDL a été étudié au Synchrotron SOLEIL (Gif-sur-Yvette) pour une première tentative de diffraction. Ce cristal a été congelé dans du TMAO pur puis soumis à un rayonnement RX sur la ligne microfocus Proxima2. Des clichés de diffraction ont été mesurés avec une résolution de plus de 20 Å. Malheureusement, cette résolution n'est pas suffisante pour calculer des densités électroniques exploitables pour la résolution de la structure. Cependant, ces tests préliminaires ont permis d'identifier des conditions de cristallisation favorables pour obtenir des cristaux qui diffractent. Avec l'analyse des différentes conditions qui conduisent à une cristallisation (**Tableau 4**), il a été possible de dégager une gamme de paramètres favorables à la formation de cristaux (tampon

HEPES, gamme de pH entre 7,0 et 8,0, chlorure de magnésium ou tartrate de sodium potassium). Leur analyse future dans le processus d'optimisation des cristaux devrait permettre l'amélioration de la espérant ainsi pouvoir bientôt résoudre la structure de ce domaine MTase de SUDV en complexe avec le VHH10.

Tableau 4 – Criblage des conditions de cristallisation des complexes 6HN/VHH10 et 6HC/VHH10.

Complexe	Kit	Tampon			Sel		Agent précipitant		Cristallisation
		nature	[c]	pH	nature	[c]	nature	[c]	
6HC/VHH10 (6,7 mg/mL)	MDL (Molecular Dimensions)	HEPES	0,1 M	7,5	potassium sodium tartrate tétrahydraté	0,8 M	-	-	cristaux lenticulaires
		Tris	0,1 M	8,5	-	-	PEG 8000	8% (m/v)	précipité cristallin
6HN/VHH10 (7,6 mg/mL)	MDL (Molecular Dimensions)	HEPES	0,1 M	7,5	chlorure de magnésium hexahydraté	0,2 M	PEG 400	30% (v/v)	cristaux grains de riz
		HEPES	0,1 M	7,5	sulfate de lithium	1,5 M	-	-	cristaux lenticulaires
		-	-	-	tartrate de potassium sodium tétrahydraté	0,4 M	-	-	cristaux grains de riz
		-	-	-	formate de magnésium dihydraté	0,2 M	-	-	cristaux grains de riz
		Tris	0,1 M	8,5	chlorure de magnésium hexahydraté	0,2 M	1,6-hexanediol	3,4 M	précipité cristallin
		HEPES	0,1 M	7,5	-	-	PEG 6000 MPD	10% (m/v) 5% (v/v)	précipité cristallin
		HEPES	0,1 M	7,5	-	-	PEG 8000 éthylène glycol	10% (m/v) 8% (v/v)	cristaux grains de riz
		MES	0,1 M	6,5	sulfate de magnésium heptahydraté	1,6 M	-	-	précipité cristallin
		citrate de sodium	0,1 M	5,6	chlorure de sodium	0,5 M	polyéthylèneimine	4% (v/v)	cristaux grains de riz
	Wizard (Molecular Dimensions)	imidazole	0,1 M	8,0	chlorure de sodium	0,2 M	potassium sodium tartrate tétrahydraté	1 M	cristaux grains de riz
		CHES	0,1 M	9,5	chlorure de sodium	0,2 M	PEG 8000	10% (m/v)	cristaux grains de riz
	JCSG core 1 (Qiagen)	HEPES	0,1 M	7,5	-	-	PEG 4000 isopropanol	20 (m/v) 10% (v/v)	précipité cristallin
		-	-	-	sulfate de sodium	0,2 M	PEG 3350	20% (m/v)	précipité cristallin
	JCSG + (Qiagen)	-	-	-	citrate de potassium tribasique monohydraté	0,2 M	PEG 3350	20% (m/v)	précipité cristallin
		-	-	7,0	acide succinique	0,8 M	-	-	précipité cristallin
-		-	7,0	acide malique DL	2,1 M	-	-	précipité cristallin	
HEPES		0,1 M	7,0	malonate de sodium dibasique monohydraté	1,1 M	Jeffamine® ED-2003	0,5% (v/v)	précipité cristallin	
HEPES		0,1 M	7,0	acide succinique	1,0 M	PEG 2000 MME	1% (m/v)	précipité cristallin	

Ce tableau récapitule les résultats obtenus lors de la campagne de criblage des conditions de cristallogénèse pour les complexes entre les formes 6HN et 6HC du domaine MTase de SUDV et le VHH10, nanobody spécifique du domaine MTase de SUDV. De nombreuses conditions ont permis d'obtenir de petits cristaux ou un précipité cristallin, en particulier pour le complexe 6HN/VHH10. En considérant l'ensemble de ces conditions, il est possible de voir que certaines conditions se retrouvent: tampon HEPES, pH entre 7 et 8, tartrate de potassium sodium tétrahydraté, chlorure de magnésium. Ces observations permettront de mettre en place une stratégie d'optimisation afin d'obtenir des cristaux de meilleure qualité.

5. Perspectives

En conclusion, cette première étude de criblage des conditions de cristallogénèse pour la résolution de la structure du domaine MTase de SUDV a permis de montrer que le domaine MTase est cristallisable avec ou sans le VHH10. Ces conditions devront être optimisées afin d'obtenir des cristaux diffractant à une résolution suffisante pour déterminer la structure de ce domaine. Avec plus de 30% d'homologie avec les séquences des structures déjà connues pour VSV et hMPV, il devrait être possible de déterminer les phases de diffraction grâce au remplacement moléculaire.

Le travail de cristallogénèse a été initié avec le domaine MTase+CTD malgré sa faible stabilité. Différentes dialyses ont été tentées afin de diminuer la concentration en arginine du tampon de conservation. Ce domaine ne reste pas soluble quel que soit le type de tampon et de pH testé (PBS,

HEPES, etc.) s'il ne contient pas au moins 300 mM d'arginine. C'est pourquoi le développement de VHH a été conduit sur les domaines MTase (immunisation) et MTase+CTD (sélection). L'obtention d'un VHH spécifique du domaine MTase sous les différentes formes décrites dans ce travail permet d'envisager une caractérisation de domaines plus larges et moins stables. Les résultats du test MTase en présence de différentes concentrations en VHH suggèrent que le VHH stabilise le domaine MTase+CTD de SUDV. Le test de spécificité indique également que le VHH10 reconnaît mieux le domaine MTase+CTD que le domaine MTase seul. Il est donc possible de former des complexes MTase+CTD/VHH10 et d'évaluer si la stabilité de ceux-ci pourrait aider au processus de cristallo-genèse. Des tests sont en cours mais les résultats sont peu concluants à l'heure actuelle. Il serait également possible d'améliorer la stabilité du domaine MTase+CTD en figeant les domaines flexibles par mutagenèse dirigée. La banque de mutants développée pour la caractérisation biochimique de certains résidus pourrait également être utilisée dans cette optique.

Pour finir, le développement du VHH10 a été une stratégie efficace. En effet, le criblage de conditions de cristallisation des complexes a donné des résultats plus larges par rapport aux études menées sur les deux formes 6HN et 6HC seules. Le VHH10 semble donc augmenter la capacité de cristallisation du domaine MTase de SUDV. Avec une structure très bien caractérisée dans la littérature, c'est également un outil intéressant permettant de déterminer les phases de diffraction indispensables à la résolution de la structure.

Bibliographie

[1] : Lieutaud P, Canard B, Longhi S. MeDor: a metaserver for predicting protein disorder. *BMC Genomics*. 2008 Sep 16;9 Suppl 2:S25.

[2] : Tereshko V, Uysal S, Koide A, Margalef K, Koide S, Kossiakoff AA. Toward chaperone-assisted crystallography: protein engineering enhancement of crystal packing and X-ray phasing capabilities of a camelid single-domain antibody (VHH) scaffold. *Protein Sci*. 2008 Jul;17(7):1175-87.

[3] : Desmyter A, Farenc C, Mahony J, Spinelli S, Bebeacua C, Blangy S, Veessler D, van Sinderen D, Cambillau C. Viral infection modulation and neutralization by camelid nanobodies. *Proc Natl Acad Sci USA*. 2013 Apr 9;110(15):E1371-9.

[4] : Conrath K, Pereira AS, Martins CE, Timóteo CG, Tavares P, Spinelli S, Kinne J, Flaudrops C, Cambillau C, Muyldermans S, Moura I, Moura JJ, Tegoni M, Desmyter A. Camelid nanobodies raised against an integral membrane enzyme, nitric oxide reductase. *Protein Sci*. 2009 Mar;18(3):619-28.

CHAPITRE 13

A la recherche d'inhibiteurs contre les MTases des mononégavirus

Le développement d'inhibiteurs spécifiques des activités MTases des filovirus pourrait être une stratégie à la fois innovante et efficace (**Chapitre 5**). En effet, leur inhibition entrainerait d'une part l'inhibition de la traduction des ARNs viraux et donc l'inhibition de la réplication virale et, d'autre part, la détection de ces ARNm viraux par les senseurs de l'immunité innée, restaurant ainsi la réponse immunitaire précoce. Dans le but d'identifier des molécules inhibitrices des activités MTases chez SUDV, nous avons évalué notre test d'activité radioactif pour le criblage de molécules organiques. Nous avons ensuite participé à deux études. Dans le cadre de l'ANR-ST14-ASTR-0026 pour le projet DGA Astrid VMTaseIn, une campagne de criblage sur le domaine MTase+CTD de SUDV a été menée sur des composés sélectionnés à l'issue d'un criblage virtuel basé sur la structure de la protéine NS5 qui porte les activités MTases du virus de la dengue. D'autre part, nous avons été contactés par une équipe de chimie médicinale de Baltimore (Seley Radtke KL, UMBC, USA) pour caractériser des composés originaux sur les activités MTases de différents mononégavirus.

1. Validation du système expérimental

Le test enzymatique radioactif développé pour la caractérisation des activités MTases a été utilisé afin de tester le potentiel inhibiteur des composés organiques. Pour cela le test a été réévalué afin d'obtenir des conditions de criblage optimales.

1.1. Calcul des paramètres d'évaluation du criblage

Différentes conditions expérimentales ont été testées afin de mettre en place un système expérimental robuste (**Figure 43**). L'activité MTase a été mesurée en utilisant un ARN coiffé GpppG-SUDV₁₂ et l'effet du DMSO sur l'essai enzymatique a été évalué dans un premier temps (**Figure 43A**). La condition sans ARN correspond à l'activité enzymatique résiduelle de l'enzyme en absence d'ARN substrat. En normalisant les valeurs de la condition sans ARN par rapport au contrôle positif, il est possible de calculer le bruit de fond (**Figure 43B**) exprimée en pourcentage d'activité par rapport au contrôle. Ce bruit de fond a été évalué entre 6 et 10% dans les différents criblages effectués. Cette valeur montre que le bruit de fond n'influe pas de façon significative sur l'activité ($\leq 10\%$). Les composés utilisés étant solubilisés dans du DMSO, leur dilution lors du test entraîne une concentration finale en DMSO de 5% (v/v) dans le test d'activité (**Figure 43A**). La condition DMSO 5% (v/v) correspond à la condition de criblage sans composé. Les résultats indiquent que le DMSO à une concentration de 5% (v/v) n'a pas d'effet significatif sur les activités MTases de SUDV. Ces contrôles ont également permis de déterminer le facteur Z' (**Figure 43B**) [1]. Ce facteur Z' reflète la qualité du test de criblage. Il correspond en effet à la fenêtre de variation entre le témoin positif et le témoin négatif du test. Il est calculé à partir des moyennes et écart-types des contrôles positif et négatif [1]. Dans notre cas, le contrôle négatif a été considéré comme étant le test sans ARN puisqu'une inhibition de 100% du signal correspondrait à la valeur du bruit de fond. La condition en DMSO a été considérée comme contrôle positif car elle reflète les conditions expérimentales du criblage. Ce facteur Z' est généralement compris entre 0 et 1. Lorsque sa valeur est supérieure à 0,60, la qualité du test est considérée comme excellente. Pour le test d'activité développé pour le domaine MTase+CTD de SUDV, les valeurs calculées du facteur Z' lors des différents essais sont comprises entre 0,75 et 0,80. Les conditions expérimentales de cet essai permettent donc la caractérisation d'une inhibition de l'activité MTase de SUDV. Cependant, pour des considérations techniques et «écologiques»

(utilisation de $^3\text{H-SAM}$), cet essai n'est adapté que pour des criblages à petite ou moyenne échelle (de quelques composés à quelques dizaines de composés). Il serait intéressant de développer un test non radioactif basé sur la technologie HTRF (voir Annexe 6).

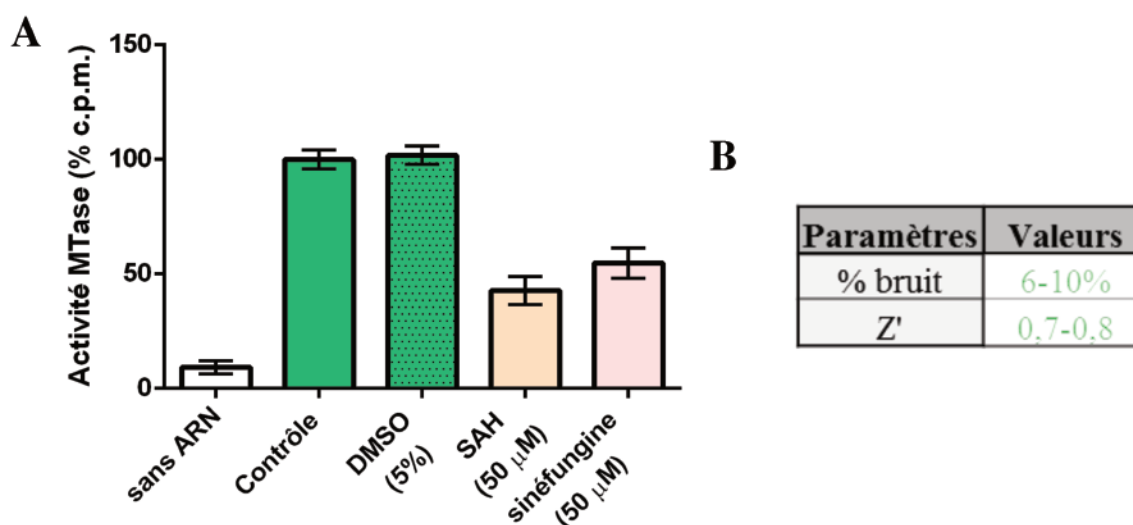


Figure 43 – Le test d'activité radioactif est adapté au criblage d'inhibiteurs.

Le test radioactif pour l'évaluation de l'activité MTase a été proposé pour le criblage d'inhibiteurs. (A) Test d'activité MTase du domaine MTase+CTD de SUDV sur un ARN coiffé spécifique de SUDV (GpppG-SUDV₁₂) pour la validation des conditions de criblage (n=6). DMSO: diméthyl-sulfoxyde. SAH: S-adenosylhomocystéine. Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé \pm les écart-types. (B) Paramètres calculés pour la validation du test MTase radioactif comme modèle de criblage d'inhibiteurs (n=3). % bruit: pourcentage du bruit de fond par rapport à l'activité totale. Z': évaluation de la qualité du test de criblage.

1.2. Evaluation d'inhibiteurs de MTases connus

Dans un premier temps, des composés connus pour inhiber l'activité de différentes MTases virales ou cellulaires ont été testés pour valider le système expérimental (Figure 43A). Le SAH, co-produit de la réaction MTase, est un inhibiteur de la réaction car il entre en compétition avec le co-substrat SAM lorsque sa concentration augmente (rétrocontrôle négatif). La sinéfungine est un analogue du SAM qui inhibe donc la plupart des MTases SAM-dépendantes car le soufre sur lequel est porté le méthyle transférable est remplacé par un carbone portant une fonction amine primaire qui ne permet donc pas le transfert de méthyle (Figure 44A). Cette molécule a été isolée du champignon *Streptomyces griseolus* qui la produit naturellement et s'en sert comme antifongique [2]. La sinéfungine a par la suite suscité un intérêt particulier pour son pouvoir antiviral [3] puis plus largement sur les MTases [4].

Le potentiel d'inhibition de ces composés a d'abord été évalué pour une concentration en inhibiteur de 50 μM (Figure 43A). À cette concentration, le SAH et la sinéfungine inhibent respectivement 60% et 45% de l'activité MTase de SUDV. Ces résultats confirment que notre système expérimental permet l'évaluation de l'inhibition de composés organiques et suggèrent que les activités MTases de SUDV ne sont pas très sensibles ni au SAH ni à la sinéfungine. Pour illustrer ce propos, l'IC₅₀ de la sinéfungine, c'est-à-dire la concentration en sinéfungine pour laquelle 50% de l'activité MTase est inhibée, a été déterminée pour différentes MTases (Figure 44B). Pour ce faire, une gamme de concentrations croissante de sinéfungine a été testée sur la hN7 (N7 MTase humaine, 100 nM,

considérée ici comme un contrôle de spécificité dans les études suivantes) et les domaines MTase+CTD des virus SUDV (4 μ M, en jaune), hMPV (4 μ M, en bleu) et hRSV (4 μ M, en rose). A ces concentrations, le taux de conversion du substrat a été évalué à environ 80% ce qui permet d'être dans des conditions idéales pour observer une inhibition de l'activité enzymatique. Les protéines de hMPV et hRSV ont été produites par nos collaborateurs d'Oxford (UK) et de Jouy-en-Josas (FR), respectivement. Les IC₅₀ ont été évaluées à 21 nM pour la hN7, 66 μ M pour la MTase de SUDV, 10 μ M pour la MTase de hMPV et 11 μ M pour la MTase de hRSV. Alors que l'IC₅₀ de la sinéfungine est très faible pour la hN7, il est intéressant de remarquer que les MTases virales sont moins sensibles à la sinéfungine. Les IC₅₀ de la sinéfungine pour les domaines MTase+CTD de hMPV et hRSV sont très proches et significativement plus faibles que l'IC₅₀ pour la MTase de SUDV, ce qui confirme que le domaine MTase+CTD de SUDV n'est pas très sensible à ce composé. Les virus hMPV et hRSV appartenant tous deux à la famille des pneumovirus, il est donc probable que leurs domaines MTase+CTD soient très proches du point de vue structural et fonctionnel expliquant les résultats très similaires obtenus. De même, ces observations semblent indiquer, comme les données de caractérisation (**Chapitre 10**), que des différences structure/fonction existent entre le domaine MTase+CTD de SUDV et les domaines MTase+CTD des pneumovirus. Il est possible que ces différentes MTases n'engagent pas les mêmes interactions moléculaires lors du recrutement du SAM. En effet, la sinéfungine étant un analogue du SAM qui ne diffère que sur un des groupements fonctionnels, il est possible que le phénomène de compétition soit différent du fait de la reconnaissance ou non du soufre et des atomes adjacents (**Annexe 5**). Il est intéressant de noter que la hN7 a une haute affinité pour la sinéfungine comparativement aux MTases virales testées. Il est possible que cette affinité importante résulte de la présence de résidus hydrophobes au sein du site de

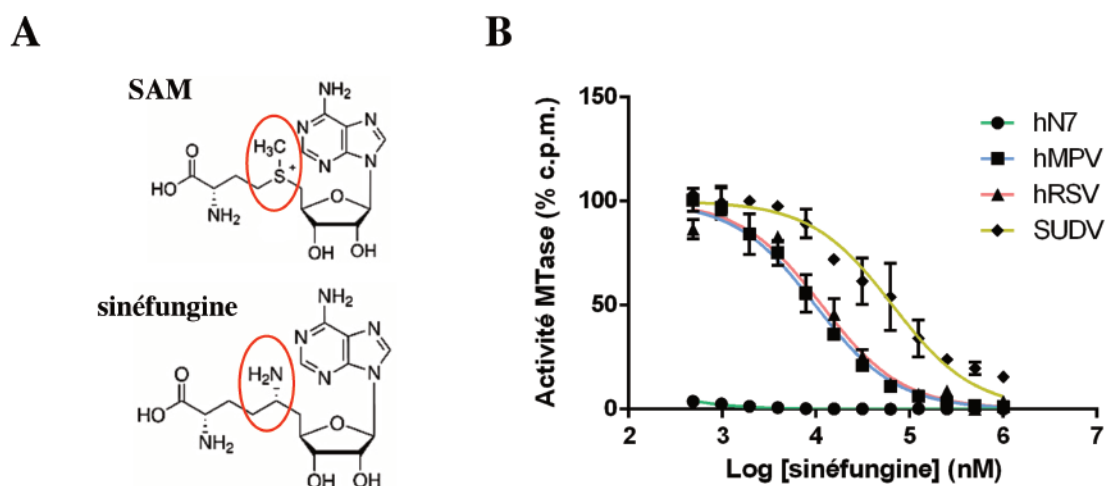


Figure 44 – La sinéfungine inhibe les activités MTases humaine et virales.

(A) Structure chimique de la S-adénosyl méthionine (SAM) et de la sinéfungine. La sinéfungine diffère du SAM par le remplacement du soufre portant le méthyl transféré par un carbone portant une fonction amine primaire. (B) Test d'activité MTase du domaine MTase+CTD de SUDV sur un ARN coiffé spécifique de SUDV (GpppG-SUDV₁₂) avec une gamme de concentration croissante de sinéfungine pour la validation du test d'activité comme modèle de criblage (n=3) et la détermination de l'IC₅₀ (concentration inhibitrice 50%) de la sinéfungine pour les différentes protéines utilisées. Quatre enzymes ont été testées: le N7 MTase humaine (hN7, vert) et les domaines MTase+CTD du métapneumovirus humain (hMPV, bleu), du virus respiratoire syncytial humain (hRSV, rose) et du virus Soudan (SUDV, jaune). Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé sans sinéfungine \pm les écart-types. Les IC₅₀ ont été déterminées grâce à une modélisation d'inhibition dose-réponse sur le logiciel Prism : 21 nM pour la hN7, 66 μ M pour la MTase de SUDV, 10 μ M pour la MTase de hMPV et 11 μ M pour la MTase de hRSV.

fixation du SAM de la hN7, absents chez les MTases virales testées. Il est donc envisageable que d'autres modifications autour du soufre du SAM puissent modifier les propriétés biophysiques de la sinéfungine afin d'augmenter sa spécificité pour les MTases virales.

En conclusion, ces premières analyses ont permis de valider le test d'activité MTase radioactif comme modèle de criblage pertinent. Ce test permet seulement d'évaluer le potentiel inhibiteur de composés organiques. En outre, les essais biophysiques (TSA, PF et mutagenèse) développés précédemment permettront de mieux comprendre le mode d'action des composés ayant des propriétés inhibitrices.

2. Criblage d'une petite banque de composés développée par la société Prestwick Chemical

Ce criblage s'est inscrit dans le projet DGA Astrid VMTaseIn (ANR-ST14-ASTR-0026) dont le but a été d'identifier un sous-ensemble de molécules capables d'inhiber des MTases virales afin de construire une chimiothèque dédiée à la recherche d'inhibiteurs en cas d'émergence virale. Ce projet a été réalisé en collaboration avec Prestwick Chemical (Strasbourg, FR), un partenaire industriel spécialisé en chimie médicinale et dans l'assemblage de chimiothèques dédiées. Cette étude entrant dans le cadre de ce partenariat, la structure des molécules citées et les détails du criblage virtuel réalisé par Prestwick Chemical sont confidentiels.

2.1. Criblage virtuel (réalisé par Prestwick Chemical)

Un criblage virtuel de molécules capable d'accomoder le site de fixation du SAM du domaine MTase de la protéine NS5 des flavivirus (virus de la dengue et virus du Nil Occidental) a été mené dans une première étape. La première étape a été de construire un site pharmacophorique du SAM à partir de 18 structures de ce domaine chez différents sérotypes du virus de la dengue et du virus du Nil Occidental. Les composés issus de différentes bases de données ont ensuite été criblés *in silico* à l'aide du site pharmacophorique. Différents hits, c'est-à-dire différentes molécules pouvant théoriquement se lier à la place du SAM, ont ainsi été obtenus. En étudiant le mode d'interaction de la molécule de SAM (ou de SAH selon les structures utilisées), il a été possible de réaliser un tri de ces différents hits selon différents filtres afin de ne garder que les composés les plus intéressants structuralement. A l'issue de ce processus, 78 composés ont été « sourcés » et rassemblés afin de former une chimiothèque dédiée contre les MTases des flavivirus.

2.2. Test d'inhibition *in vitro*

Les composés identifiés par criblage virtuel ont été évalués *in vitro* sur les domaines MTases de différents virus. Le criblage primaire de ces composés a été réalisé à 50 μM avec le domaine MTase du virus de la dengue et sur la protéine nsp14 (N7 MTase) du virus du SARS mais sans résultat concluant. Malgré cela, cette petite chimiothèque a été testée pour rechercher des d'inhibiteurs éventuels des activités MTases de SUDV. La **Figure 45A** résume l'ensemble des résultats obtenus pour un criblage à 50 μM d'inhibiteur, 4 μM d'enzyme pendant 3h à 30°C sur un ARN GpppG-SUDV₁₂. Une majorité de composés (82% soit 64 composés) n'ont aucun effet inhibiteur sur l'activité MTase de SUDV dans ces conditions expérimentales. Environ 15% soit 12 composés présentent une inhibition entre 0 et 25%. Pour une concentration en composé de 50 μM , ces composés ne peuvent être

considérés comme intéressants car leur IC_{50} seraient bien supérieures à $50 \mu\text{M}$. Cependant, 2 composés présentent un taux d'inhibition entre 26 et 50% et 1 composé un taux d'inhibition entre 76 et 100%. Un des deux premiers composés correspond au contrôle d'inhibition par la sinéfungine qui, à $50 \mu\text{M}$, a un taux d'inhibition d'environ 45%. Ce taux d'inhibition étant équivalent aux tests précédents, les données de criblage sont donc pertinentes. Le taux d'inhibition des deux autres composés a été reporté dans la **Figure 45B** par rapport au contrôle positif DMSO 5% (v/v). Le composé STL066314 a un taux d'inhibition d'environ 45% et le composé 7860798 d'environ 80%. Ces composés constituent donc des hits intéressants qui devront être optimisés pour le développement ultérieur d'inhibiteurs.

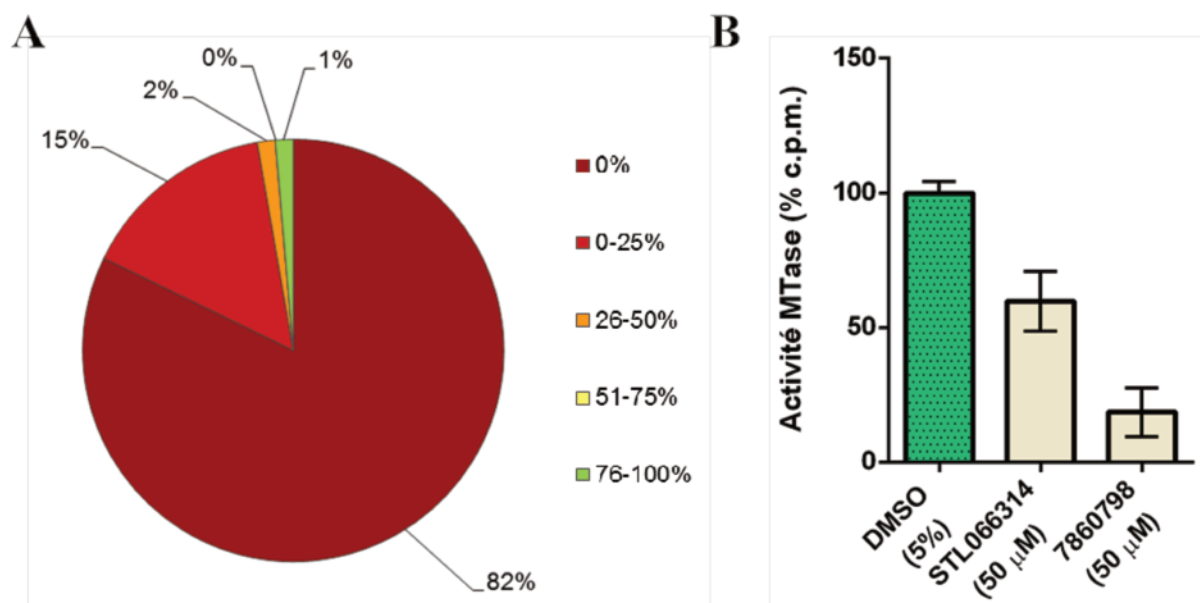


Figure 45 – Deux hits ont été identifiés grâce au criblage virtuel de composés sur la MTase de la dengue.

Le criblage virtuel basé sur les interactions du SAM dans la structure de la MTase du virus de la dengue a permis d'identifier 78 composés. **(A)** Distribution des 78 composés et de la sinéfungine en fonction de leur taux d'inhibition à $50 \mu\text{M}$ sur l'activité MTase du domaine MTase+CTD de SUDV avec un ARN coiffé spécifique de SUDV (GpppG-SUDV₁₂) ($n=3$). Les données représentent la proportion du nombre de composés pour chaque gamme d'inhibition (0% d'inhibition en rouge foncé, 0 à 25% d'inhibition en rouge, 26 à 50% d'inhibition en orange, 51 à 75% d'inhibition en jaune et 76 à 100% d'inhibition en vert). **(B)** Test d'inhibition de l'activité MTase des composés intéressants identifiés (STL066314 et 7860798) à $50 \mu\text{M}$ sur l'ARN GpppG-SUDV₁₂ ($n=3$). Les données représentent les moyennes normalisées par rapport au contrôle DMSO \pm les écarts-types.

2.3. Conclusions

Cette étude a permis de mettre en évidence deux composés ayant un potentiel inhibiteur contre l'activité MTase de SUDV. Etant donné les résultats de l'étude biochimique présentée précédemment, il est possible de conclure que cette inhibition est majoritairement portée sur l'activité de méthylation interne. Comme ces composés sont des analogues pharmacophoriques du SAM, il est probable que ces inhibiteurs affectent l'ensemble des activités MTases du domaine MTase+CTD de SUDV. Par la suite, au cours du processus d'optimisation des hits, il sera nécessaire de confirmer l'effet des inhibiteurs sur les activités N7 et 2'O MTases de la coiffe. Malheureusement, les quantités de composés disponibles au laboratoire étant limitées, des criblages sur d'autres matrices ARN et la détermination des IC_{50} n'ont pas encore pu être effectués. Il est cependant intéressant de remarquer que ces hits n'ont pas été

identifiés sur d'autres enzymes virales ce qui indique que les activités MTases de mononégavirus ont probablement des repliements structuraux particuliers qui ouvrent la voie au développement d'inhibiteurs spécifiques des virus de l'ordre des *Mononegavirales*. Il est également intéressant de noter qu'aucune molécule ne se soit révélée efficace contre les activités MTases des flavivirus, bien que les composés aient été sélectionnés sur base des structures disponibles des domaines MTases de ces virus. Il est donc possible que les filtres utilisés n'aient pas été suffisamment stringents pour permettre d'identifier des molécules mimant le SAM. Les composés identifiés contre les activités MTases de SUDV seront exploités bientôt pour une caractérisation plus complète lorsque de nouveaux lots de molécules seront disponibles.

3. Participation au développement de nouvelles molécules innovantes, les « fleximers »

3.1. Présentation des fleximers

Les fleximers sont des composés développés par le laboratoire du Pr. Katherine L Seley Radtke (UMBC, Baltimore, USA). Ces composés sont des dérivés nucléosidiques innovants. En effet, les analogues nucléosidiques actuels ne sont basés que sur la variation des substitutions portées par les différents résidus comme la base qui a une géométrie plane donc rigide. Les fleximers sont des nucléosides modifiés afin d'introduire de la flexibilité dans la structure de la base (**Figure 46**). Par exemple, la structure de l'adénine a été modifiée en séparant le noyau purine en un noyau imidazole et un noyau pyrimidine flexibles l'un par rapport à l'autre [5]. Ces noyaux peuvent porter différents substituants afin de modifier l'affinité des composés développés pour la cible étudiée.

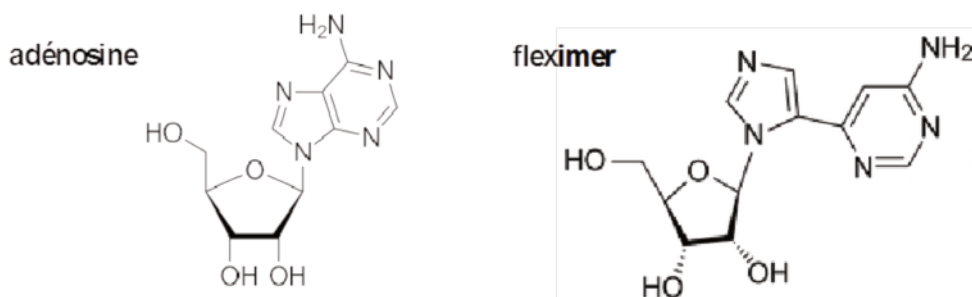


Figure 46 – Les « fleximers » sont des analogues nucléosidiques flexibles.

Les fleximers sont des analogues nucléosidiques des purines. Par exemple, l'adénine peut être séparé en un noyau imidazole et un noyau pyrimidine qui sont alors flexibles l'un par rapport à l'autre. Les deux noyaux peuvent également porter différents substituants pour modifier l'affinité de ceux-ci à la cible étudiée.

3.2. Description des composés

Quatre composés ont été synthétisés par le laboratoire de KL Seley Radtke sur base de la guanosine (**Figure 47A**). La guanine a été modifiée comme précédemment présenté avec le remplacement de la fonction cétone par un groupement méthoxyle sur le carbone 6 du noyau pyrimidine. Différentes modifications ont également été apportées sur les autres parties de la molécule. Le ribose a été remplacé par un éther-oxyde reliant la base modifiée à différents groupements fonctionnels : un alcool (HP083), un acide carboxylique (HP-105), un groupement triphosphate (2MR04) ou un groupement

phosphate portant lui-même différentes fonctions (MR064). La synthèse des composés HP083, HP-105 et MR064 a été reportée précédemment dans la littérature [6].

Des tests d'inhibition en modèles cellulaires ont permis d'évaluer le potentiel de ces molécules et leur cytotoxicité [6]. L' EC_{50} , c'est-à-dire la concentration efficace pour diminuer de moitié le taux de réplication virale, a été déterminée pour les composés HP083 et MR064 sur des réplicons de différents virus responsables de fièvres hémorragiques (virus Ebola, virus de Lassa et virus de la vallée du Rift). Seul EBOV est sensible à ces composés avec des EC_{50} de 2,2 μ M pour le composé HP083 et 27,2 μ M pour le composé MR064. Ces résultats suggèrent donc une spécificité de ces composés contre les filovirus. Pour compléter ces observations, les EC_{50} ont été déterminées sur un modèle de cellules infectées pour EBOV, SUDV et MARV. Le composé HP083 ne présente pas d'inhibition. Cependant, les composés HP-105 et MR064 présentent une activité inhibitrice intéressante (44 μ M sur EBOV, 20 μ M sur SUDV et 70 μ M sur MARV pour HP-105 et 29 μ M sur EBOV, 7 μ M sur SUDV et 62 μ M sur MARV pour MR064). En outre, ces composés présentent une faible cytotoxicité car leur CC_{50} , c'est-à-dire la concentration en composé pour laquelle 50% des cellules meurent, est supérieure 100 μ M. Ce résultat fait de ces molécules des candidats intéressants pour le développement d'antiviraux. Fort de ces résultats, KL Seley Radtke a fait appel à notre équipe pour identifier la cible potentielle de ces molécules.

3.3. Tests *in vitro* et IC_{50}

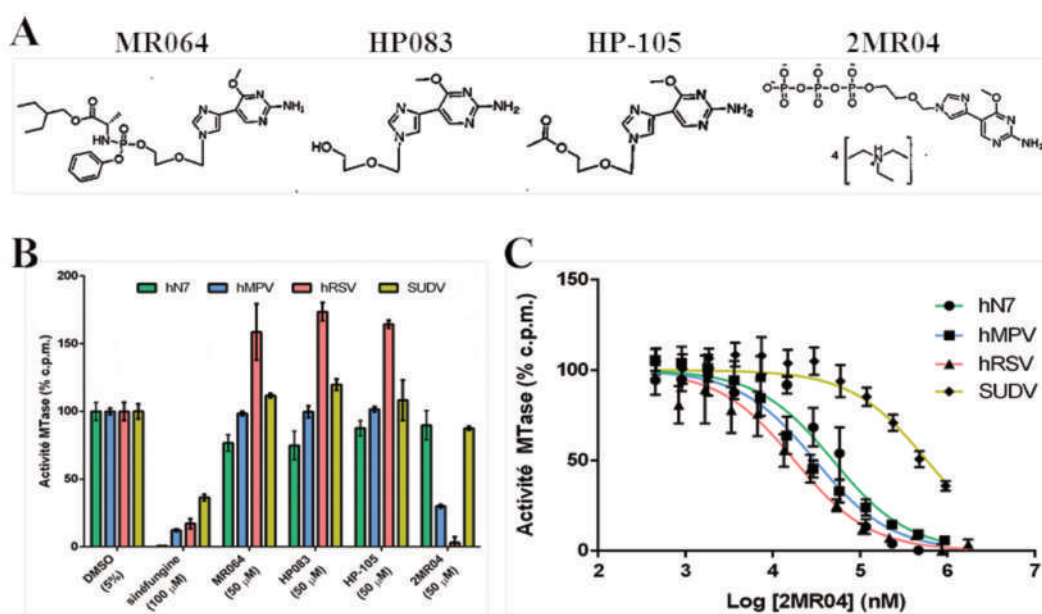


Figure 47 – Les « fleximers » sont des composés intéressants pour l'inhibition des MTases virales.

Les fleximers ont été synthétisés par nos collaborateurs (Seley-Radtke K, Baltimore). Certains de ces nucléotides modifiés sont actifs contre différents filovirus en modèle cellulaire mais leur cible n'est pas clairement identifiée. (A) Structures chimiques des composés MR064, HP083, HP-105 et 2MR04 basés sur l'adénosine. (B) Test d'inhibition de l'activité MTase des composés 2MR04, MR064, HP083 et HP-105 à 50 μ M pour différentes protéines avec leur substrat coiffé spécifique (n=3). Quatre enzymes ont été testées: la N7 MTase humaine (hN7, vert) sur GpppAC₄ et les domaines MTase+CTD du métapneumovirus humain (hMPV, bleu) sur GpppG-hMPV₈, du virus respiratoire syncytial humain (hRSV, rose) sur GpppG-hMPV₈ et du virus Soudan (SUDV, jaune) sur GpppG-SUDV₁₂. Les données représentent les moyennes normalisées par rapport au contrôle DMSO \pm les écart-types. (C) Test d'activité MTase des protéines hN7 et MTase+CTD des virus hMPV, hRSV et SUDV sur leurs ARNs coiffés spécifiques avec une gamme de concentration croissante de composé 2MR04 pour la détermination de son IC_{50} (concentration inhibitrice 50%) sur chaque protéine (n=3). Les données représentent les moyennes normalisées par rapport au contrôle coiffé non méthylé sans composé \pm les écart-types. Les IC_{50} ont été déterminées grâce à une modélisation d'inhibition dose-réponse sur le logiciel Prism.

Ces composés étant des analogues de nucléosides, ils pourraient inhiber les activités RdRp et de synthèse de la coiffe de la protéine L qui impliquent des nucléosides. Il est donc possible que ces composés puissent avoir un impact sur l'activité MTase des filovirus. Nous avons donc utilisé notre test d'activité radioactif afin d'évaluer l'activité inhibitrice des fleximers sur l'activité MTase de différents mononégavirus.

Dans un premier temps, un criblage primaire des composés MR064, HP083, HP-105 et 2MR04 – ce dernier correspond à la forme triphosphate du fleximer décrit ci-dessus – à une concentration de 50 μM a été réalisé sur les activités MTases de la hN7 (100 nM) et des domaines MTase+CTD des virus SUDV, hMPV et hRSV (4 μM) dans les conditions précédemment décrites pour la détermination de l' IC_{50} de la sinéfungine (voir 1.2.) (Figure 47B). L'activité maximale (100% d'activité) a été évaluée grâce à un test d'activité en absence de composé et la sinéfungine a été utilisée comme témoin d'inhibition. A 50 μM , la sinéfungine inhibe efficacement les différentes activités MTases. L'étude des composés à 50 μM a permis de montrer que les composés MR064, HP083 et HP-105 n'ont pas d'effet inhibiteur sur les différentes activités MTases testées. Cependant, il semble que ces composés augmentent l'activité MTase de hRSV ce qui pourrait indiquer qu'ils stimulent ou stabilisent le domaine MTase+CTD de hRSV. En outre, il semble que le composé 2MR04 ait une activité inhibitrice significative sur les pneumovirus, en particulier hRSV.

Afin de caractériser plus précisément l'effet du composé 2MR04, les IC_{50} pour les différentes enzymes ont été déterminées (Figure 47C). Pour cela, une gamme de concentration du composé a été préparée puis testée sur les MTases. Les IC_{50} ont été estimées à 49 μM pour la hN7, plus de 100 μM pour SUDV, 30 μM pour hMPV et 18 μM pour hRSV. Ces résultats confirment que ce composé est potentiellement une molécule intéressante contre les pneumovirus. Au contraire, il semble que cette molécule soit faiblement efficace contre l'activité MTase de SUDV, suggérant que leur cible est plutôt la polymérase ou la PRNTase. Une caractérisation plus complète permettra de confirmer ces résultats et de comprendre leur mécanisme d'action sur les activités MTases ou polymérase lorsque les essais seront disponibles.

3.4. Conclusions

Les fleximers apparaissent donc comme des inhibiteurs potentiels des mononégavirus. D'une part, les filovirus sont inhibés dans différents systèmes cellulaires avec des EC_{50} relativement faibles notamment pour le composé MR064. D'autre part, les tests *in vitro* ont révélé que ces composés ne sont pas actifs sur l'activité MTase des filovirus mais contre les activités MTases des pneumovirus. Il semble également que ces composés puissent être actifs contre la hN7.

Ces résultats suggèrent plusieurs conclusions concernant le potentiel antiviral des fleximers. Concernant les filovirus, les fleximers inhibent leur répllication mais pas leur activité MTase ce qui laisse supposer que la cible chez les filovirus n'est pas l'activité MTase mais probablement l'activité polymérase ou PRNTase. Cependant, les différences entre les résultats obtenus sur les différents modèles concernant les composés MR064 et HP-105 pourraient s'expliquer par le fait que les fleximers puissent être métabolisés par les cellules pour donner un composé actif. De fait, en fonction de la forme du fleximer (alcool, acide carboxylique ou autre), le fleximer serait plus ou moins métabolisable. Le métabolite actif serait alors plus ou moins disponible ce qui se traduirait par des EC_{50} différentes. Une approche métabolomique permettrait d'adresser cette question et d'identifier le

métabolite actif. Cette hypothèse est soutenue par les résultats obtenus sur les pneumovirus. Seuls des résultats préliminaires sur l'activité MTase *in vitro* sont disponibles mais ceux-ci indiquent que seule la forme triphosphate du fleximer est active sur l'activité MTase. Ces observations suggèrent que les composés sont métabolisés en une forme triphosphate fonctionnellement active. Ce processus est connu dans le développement d'analogues nucléosidiques. En effet, la ribavirine, la molécule antivirale la plus utilisée, doit être métabolisée de sa forme monophosphate en sa forme triphosphate pour acquérir son potentiel inhibiteur [7].

Pour finir, les différents tests sur cellules ont montré que les composés HP083, HP-105 et MR064 ne présentent pas de cytotoxicité. Ces résultats indiquent que ces composés n'interfèrent pas de façon significative avec le métabolisme cellulaire. Cependant, il semble que le composé 2MR04 inhibe l'activité de la hN7 *in vitro*. De plus, des composés similaires ont montré une activité inhibitrice significative sur les SAH hydrolases cellulaires [8]. L'inhibition des SAH hydrolases permet d'augmenter la concentration cellulaire en SAH ce qui inhibe les MTases par rétro-contrôle négatif. Cette stratégie a fait ses preuves contre le virus Ebola *in vivo* mais peu de données sont disponibles concernant la cytotoxicité des inhibiteurs de SAH hydrolases [9,10]. Si la caractérisation des fleximers confirme leur effet sur la hN7 et montre également une activité contre les SAH hydrolases cellulaires, des contrôles de cytotoxicité plus poussés seront nécessaires pour assurer l'innocuité des fleximers car les méthylations sont des régulateurs impliqués dans des mécanismes cellulaires critiques dont la régulation de l'expression génétique [11].

4. Perspectives

En conclusion, différentes molécules inhibitrices de l'activité MTase de différentes MTases de mononégavirus ont été identifiées. Les composés STL066314 et 7860798 sont actifs contre l'activité MTase de SUDV et les fleximers dérivés de l'adénosine semblent être de bons inhibiteurs contre les mononégavirus. Une caractérisation biochimique plus approfondie de ses composés permettra de définir un mode d'action plus précis de ces molécules. En effet, une étude par TSA permettra de définir si les composés stabilisent ou déstabilisent les domaines MTases. Des études de compétition avec le SAM ou l'ARN permettront de déterminer le type d'inhibition. En outre, il sera également envisageable de réaliser des expériences de trempage avec des cristaux du domaine MTase+CTD de hMPV. Ces expériences permettront de réaliser une optimisation des composés par SAR (Structure/Activity Relationship) qui permet d'augmenter l'affinité des composés en modifiant leurs interactions sur base de la structure de la molécule en complexe avec la protéine. Les différentes sensibilités relevées sur les différents domaines MTase+CTD des mononégavirus étudiés sont un exemple supplémentaire de la diversité fonctionnelle de ces domaines chez les virus de l'ordre des *Mononegavirales*. La caractérisation transversale d'inhibiteurs au sein des mononégavirus souligne ici encore le manque de connaissances structurales et fonctionnelles chez les différentes familles qui constituent l'ordre des *Mononegavirales*.

Bibliographie

- [1] : Zhang JH, Chung TD, Oldenburg KR. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J Biomol Screen*. 1999;4(2):67-73.
- [2] : Hamil RL, Hoehn MM. A9145, a new adenine-containing antifungal antibiotic. I. Discovery and isolation. *J Antibiot (Tokyo)*. 1973 Aug;26(8):463-5.
- [3] : Pugh CS, Borchardt RT, Stone HO. Sinefungin, a potent inhibitor of virion mRNA(guanine-7-)-methyltransferase, mRNA(nucleoside-2'-)-methyltransferase, and viral multiplication. *J Biol Chem*. 1978 Jun 25;253(12):4075-7.
- [4] : Borchardt RT, Eiden LE, Wu B, Rutledge CO. Sinefungin, a potent inhibitor of S-adenosylmethionine: protein O-methyltransferase. *Biochem Biophys Res Commun*. 1979 Aug 13;89(3):919-24.
- [5] : Seley KL, Zhang L, Hagos A, Quirk S. "Fleximers". Design and synthesis of a new class of novel shape-modified nucleosides(1). *J Org Chem*. 2002 May 17;67(10):3365-73.
- [6] : Yates MK, Raje MR, Chatterjee P, Spiropoulou CF, Bavari S, Flint M, Soloveva V, Seley-Radtke KL. Flex-nucleoside analogues - Novel therapeutics against filoviruses. *Bioorg Med Chem Lett*. 2017 Jun 15;27(12):2800-2802.
- [7] : Streeter DG, Miller JP, Robins RK, Simon LN. The enzymic conversion of 1,2,4-triazole-3-carboxamide to ribavirin-5'-phosphate and its relationship to the proposed mechanism of action. *Ann N Y Acad Sci*. 1977 Mar 4;284:201-10.
- [8] : Seley KL, Quirk S, Salim S, Zhang L, Hagos A. Unexpected inhibition of S-adenosyl-L-homocysteine hydrolase by a guanosine nucleoside. *Bioorg Med Chem Lett*. 2003 Jun 16;13(12):1985-8.
- [9] : Bray M, Driscoll J, Huggins JW. Treatment of lethal Ebola virus infection in mice with a single dose of an S-adenosyl-L-homocysteine hydrolase inhibitor. *Antiviral Res*. 2000 Feb;45(2):135-47.
- [10] : Bray M, Raymond JL, Geisbert T, Baker RO. 3-deazaneplanocin A induces massively increased interferon-alpha production in Ebola virus-infected mice. *Antiviral Res*. 2002 Jul;55(1):151-9.
- [11] : Prokhorchuk AV, Ruzov AS. Genome methylation and its role in functioning of the eukaryotic organism. *Genetika*. 2000 Nov;36(11):1475-86.

CHAPITRE 14

Conclusions et perspectives

Durant cette thèse, nous nous sommes intéressés à l'activité MTase portée par la protéine L des filovirus. Les domaines MTase et MTase+CTD du virus SUDV ont été clonés et exprimés en bactérie afin de caractériser l'activité MTase chez ces virus. Notre étude est la première mise en évidence d'un domaine actif catalysant des activités MTases chez SUDV, un virus de la famille des *Filoviridae*. L'analyse biochimique de l'activité MTase a permis de démontrer que ce domaine MTase est original dans la mesure où il catalyse des méthylations internes sur des résidus adénosine en plus de son activité de méthylation de la structure coiffe. En comparant l'activité de différentes MTases d'autres mononégavirus, il apparaît que cette activité de méthylation interne est dominante chez SUDV comparativement à d'autres mononégavirus. Cette observation remet en question le dogme de la conservation stricte des structures et des fonctions de la protéine L chez les virus de l'ordre des *Mononegvirales*.

Cette étude a, dans un premier temps, permis d'identifier le rôle fonctionnel du domaine CTD de la polymérase L chez SUDV. En comparant les activités des protéines MTase et MTase+CTD, nous avons démontré que le domaine CTD est un régulateur de l'activité MTase. En effet, nos analyses biochimiques ont permis de montrer que ce domaine joue un rôle clé dans le recrutement des ARNs méthylés par le domaine MTase. En outre, l'activité MTase du domaine MTase+CTD de SUDV a été observée sur des ARNs coiffés et non coiffés, indiquant que cette protéine porte une activité non conventionnelle. En utilisant des ARNs homopolymériques et hétéropolymériques de séquences variables, nous avons démontré que cette MTase méthyle spécifiquement des résidus adénosines internes de la séquence ARN plutôt que la structure coiffe. Cette méthylation interne semble être catalysée sur les positions 2'OH des groupements riboses des résidus adénosines mais cette position de la méthylation reste à démontrer formellement par des méthodes analytiques de type HPLC ou spectrométrie de masse. Une activité similaire a déjà été rapportée chez les MTases des virus à ARN positif de la famille des *Flaviviridae* mais leurs rôles biologiques restent à démontrer. Fort de ces observations biochimiques, il est maintenant essentiel de mettre en évidence ces méthylations internes sur des ARNs viraux de filovirus. Cette tâche est ardue chez des virus uniquement manipulables en laboratoire P4. Nous avons donc commencé à développer des outils biochimiques sensibles utilisables en laboratoire P4 afin d'adresser cette question avec nos collaborateurs. Dans l'hypothèse où ces méthylations internes seraient confirmées, il sera intéressant de comprendre leur rôle fonctionnel et leur implication éventuelle dans la pathogenèse virale. Ces modifications épitranscriptomiques pourraient en effet jouer différents rôles afin d'assurer une réplication optimale du virus. Notre hypothèse principale est que ces modifications des ARNs pourraient permettre l'échappement aux senseurs de l'immunité innée tels que MDA5 ou limiter la restriction induite par les ISGs de la famille IFIT. Par ailleurs il serait également intéressant de déterminer dans quelle mesure ces modifications épitranscriptomiques ont un rôle régulateur dans le cycle réplication/transcription ou encore dans le processus d'encapsidation des ARNs viraux. Ces questions pourront être adressées dans différents modèles biochimiques et cellulaires.

Ce travail a également permis de mieux caractériser les méthylations de la structure coiffe des ARNm des filovirus *in vitro*. Il a été précédemment montré que ces ARNs portent une coiffe méthylée en position N7 [1]. La présence de 2'O méthylation n'a pas été démontrée expérimentalement à ce jour et les enzymes impliquées dans la méthylation de ces structures n'ont pas été caractérisées biochimiquement. En utilisant des ARNs synthétiques, nous avons montré que le domaine

MTase+CTD de SUDV porte l'activité N7 MTase de la coiffe. Cette activité et la méthylation interne ont des pHs optimaux différents, suggérant que ces deux types de méthylation impliquent deux mécanismes réactionnels distincts portés par un site catalytique unique. En ce qui concerne la 2'O méthylation des structures coiffées, elle a été observée uniquement sur des ARNs coiffés commençant par un A. Cette méthylation n'a pas été observée sur des ARNs courts commençant par un G, suggérant que des structures présentes sur les substrats naturels pourraient jouer un rôle dans ce processus. Cette hypothèse est actuellement à l'étude au laboratoire sur de nouveaux ARNs synthétiques capables de former ces structures tige/boucle.

La caractérisation biochimique des activités MTases portées par le domaine MTase+CTD de SUDV a révélé que le profil de méthylation des filovirus est tout à fait original au sein de l'ordre des mononégavirus malgré la conservation apparente de la protéine L. Ces résultats remettent en question le postulat selon lequel le domaine MTase+CTD de la protéine L des mononégavirus est une enzyme spécifique de la coiffe. Afin de comprendre les bases moléculaires et d'établir un modèle des relations structure/fonction des activités MTases chez les virus de l'ordre des *Mononegavirales*, des études structurales plus poussées sont nécessaires. Suite à l'identification d'un domaine MTase stable de la protéine L de SUDV, une campagne de cristallogénèse a été menée au laboratoire. Un cristal de ce domaine en complexe avec un VHH développé au laboratoire a été obtenu et diffracté à une faible résolution (20 Å). Nous espérons que l'optimisation de ces cristaux permettra prochainement d'obtenir les premières données structurales sur la protéine L des filovirus qui pourront peut être apporter de nouveaux arguments au modèle proposé dans ce travail.

Enfin, la caractérisation *in vitro* du domaine MTase+CTD de SUDV a permis d'ouvrir de nouvelles perspectives de criblage de molécules antivirales. Le rôle pléiotropique jouée par cette MTase dans les méthylations de la coiffe et les méthylations internes des adénosines suggère que l'inhibition des activités MTases chez les filovirus pourrait agir à différents niveaux. En effet, l'inhibition de l'activité N7 MTase de la coiffe devrait limiter la réplication virale en bloquant la synthèse des protéines virales alors que l'inhibition de l'activité 2'O MTase pourrait restaurer la détection précoce des ARNs viraux induisant une immunité antivirale protectrice. Différentes molécules ont été identifiées dans un travail préliminaire. Il est maintenant nécessaire de mieux les caractériser et d'évaluer leur potentiel d'inhibition en cellules infectées.

Pour conclure, ces travaux ont mis en relief l'importance des activités MTases chez les filovirus. Ces activités pourraient en effet avoir un rôle central dans leur cycle réplication et leur pathogénicité. Elles illustrent avec justesse l'aphorisme de Robert Sabatier qui rapporte que « se montrer nu, c'est cacher l'essentiel de soi-même ».

Bibliographie

[1] : Weik M, Modrof J, Klenk HD, Becker S, Mühlberger E. Ebola virus VP30-mediated transcription is regulated by RNA secondary structure formation. *J Virol.* 2002 Sep;76(17):8532-9.

ANNEXES

LISTE DES ANNEXES

ANNEXE 1 - Alignement de séquences multiples pour l'étude de la conservation de la séquence protéique du domaine MTase+CTD de la protéine L chez les mononégavirus

ANNEXE 2 - Contrôle des protéines produites par spectrométrie de masse

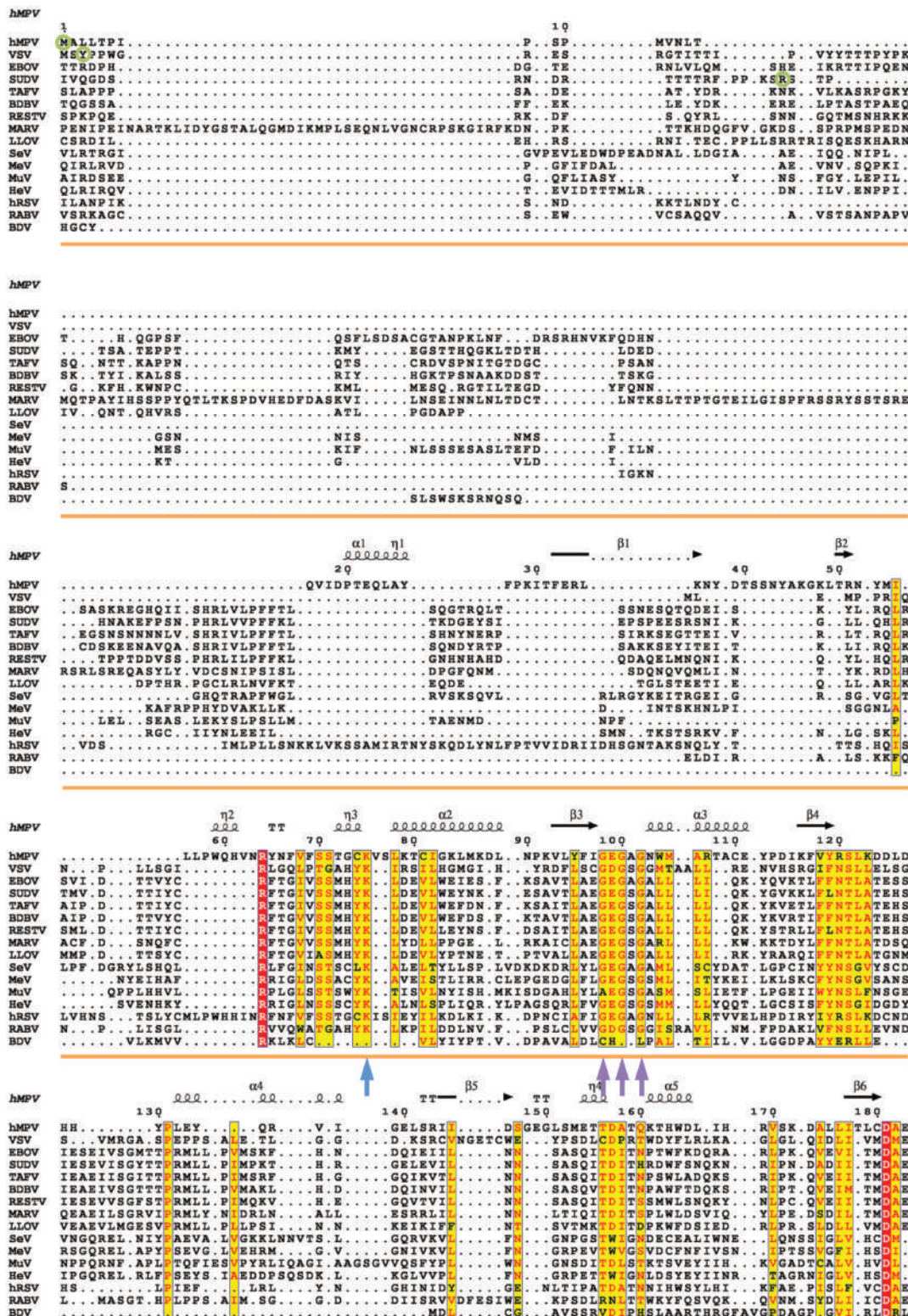
ANNEXE 3 – Liste des ARNs synthétiques

ANNEXE 4 – Comparaison des propriétés électrostatiques du site de fixation du SAM chez les MTases humaines et virales

ANNEXE 5 – Article 1

ANNEXE 6 – Article 2

ANNEXE 1 – Alignement de séquences multiples pour l'étude de la conservation de la séquence protéique du domaine MTase+CTD de la protéine L chez les mononégavirus



hMPV α6 α7 β7 β8 β9

190 200 210 220 230 240 250

hMPV .FKDRDDFFK.VILWRKH.LS.....CRCTTYGTD...LYLF.AKYHAKDCNVKLP.FVRSVATFIMQGGKLSGSECYI
VSV .VRDSSTSLK.ETNVRNY.HRIL.DE.QGV.LYTYGTICSE.KN.VT.I...GP.M.K.TVD.VQTEFSSQTSSEVYM
EBOV .TTENINRSK.YEAVYKI.LHHI.DPSVLKA.VLVFLSDT.EGM.LW.ND.N...AP.F.A.TGY.IKPITSAKSSSWYI
SUDV .TTENIDRSR.YEAVYTI.CNHI.NPKTLKV.LLVFLSDI.DGM.CW.NN.Y...AP.M.G.SGY.IKPITSAKSSSWYI
TAFV .TTENINRSK.YEAVQQL.VSHI.DPNALKV.LLVFLSDI.DGI.LW.ND.N...TP.L.G.LGY.IKPITSAKSSSWYI
BDBV .TTENINRSK.YEAIQQL.VSHI.DPRVTKI.LLVFLSDI.EGL.LW.ND.H...AP.L.G.SGY.IKPITSAKSSSWYI
REBTV .TTENINRSK.YRAVYNL.LDHI.DPQYLVK.LLVFLSDI.EGI.LW.ND.Y...AP.L.G.AGY.IKPITSAKSSSWYI
MARV .TTKDRTRQL.YKTIVNIW.TRT.SPNIPIK.SIKVFLLDY.EGT.LF.MR.NA...IQ.Y.G.QVQ.LKPYSAKNSSEWYI
LLOV .TIGNSIRIP.LYNAVIKC.DRDK.MQDKPKC.VVLFLLNDY.LGT.CW.LN.HI...TL.R.N.KVE.IKPLSNARSSSEWYI
SeV .EGGDHKDDQV.LHEHYSV.RIAYL.VGDRD.VV.LSIA.P.R.LGT.DWTRQLS.LY...LR.Y.D.EVN.IVLKTSNPASTEMYL
MeV .ETLPDKDTIEK.LEELAAL.SMALL.LGKIGSI.LVIXLMPFSG.DFV.QG.FIS.YV...GS.H.R.EVN.VYPRYSNFISTEYI
MuV .EGVPGSMNSM.ERAQVHALITVT.VLKPGLL.LLXASWEFF.NRF.SF.LT.I...WQ.F.SS.TIR.LRSSYSDPNHXYI
HeV .ESGIDKQVEE.MIEHSHL.SIAIN.VMIEDGV.LVXIAFAP.FPI.SR.LN.MY...RS.Y.G.LVLVCFPVYSPSTEVYI
hRSV .LPVTYVNSK.IIEWSKH.RCKKYCSSVNKCT.LVYHAQD.D.D.FKLD.NI...TI.L...KTYVCLGSKLGSSEVYI
RABV .VVDIASINR.TLLMSDF.LS.I.DG.PLY.LVFTYGTMLVNFPE.YR.IQ.HI...SR.A.P.SVTGFITQMSFSEELI
BDV .ES.....VCYAHPCLEE.EFNAY.....LDELVDI.SDM.CC.....PLA.P.CKA.FRPIYRSIQFRLA

hMPV α8

260 270

hMPV .LLTLGE.HN.....NLPCHEIQN...SKK...IA
VSV .VCKGLK.KLIDE...PNP.....D.....WSSINESWKN...YAFQSSSEQEFARAKKVSTYFTLTGIPSQFIP
EBOV .CLTNFL.STTRKM...PHQ.....N.....HLSCKQVILT...A...Q...L
SUDV .CLSNLL.STLRIT...QHQ.....T.....QANCLHVVC...A...Q...Q
TAFV .CLSNLL.SSRRI...PHQ.....S.....HTCMBVLOT...A...Q...L
BDBV .CLSNFL.SASRR...PHQ.....G.....HATCMQVIQT...R...L...A
REBTV .CLSNLI.STNRRS...AHQ.....T.....HKACLGVIIRD...A...Q...A
MARV .CCGKRR.IQRLKI...DFS.....T.....DQVGFILICK...A...S...R
LLOV .KAEQPVLSSELNINVNWE...G.....H.TTKLIKK...A...H...C
SeV .LSRHPK.SDIIED...SKTVLASLLPLSKEDSIKI...EK...W...L...I
MeV .MRTDLK.ANRLMN...PEKIKQIIESSVRTSPGLGHLSIKQLSCI...QA...G...D
MuV .IATLAV.DPTTSS...FTALNRRARTLNEQ...GFS.....P...P...P
HeV .ICLQKT.IKTIP...PQRVLDHSYLSDEINDQGITSVIFKIKNIQSK...QPHEDL...KH...Y
hRSV .VLTIGP.AN.....I...IFFVNVVQN...K...I...L
RABV .RFSKRG.KFRDA...EYL.....T.....SSTLRMSLV...FNCSSPKSEMQR
BDV .M.....DNYSFVMDLIM.....

hMPV α9

280 290 300

hMPVVCND.FY.AK...AAL.....K.DNKSEANCKSLLSGLRI...FI...
VSV .DFFVNIETMLQIFGV.PT.GVSHAAAL.....KS...SDRPADL...T...MAI
EBOVQIQRS.PY.WSHHTQY.....AD...CEHLSTR...L...GF...EKKVL
SUDVQIQRS.SY.WSHHTQY.....TT...SRHMSYAF...L...GF...EKKVL
TAFVQIQRS.SY.WSHHTVQY.....AN...HMHLDYNI...L...GF...ERVL
BDBVQIQRS.SY.WSHHTVQY.....AD...INHLSTN...L...GF...EKKVL
REBTVQVQRG.VY.WSHHTAQY.....AT...XNHCEYGL...L...GF...EKKVL
MARVQRQAI.PY.WKHEKN.....YP...ASHKFF...L...GF...ESSF
LLOVQLSRQ.SY.WTHMSY.....AN...LDHREYRM...L...GF...TSVL
SeVEKAKA.HE.WTRELRE.....GSSSSGMLRP.YHQALQTF...GF...NLYKLSRDF
MeVAVSRG.DINPTKKTPI...EQ...V...INCGWAI...NG...PKL...CKEL
MuVE.LVS.EY.WRRHEGQIQQDCIDKVISVCVRDQYLDN...N...LIL...QAG...GT...RKLW
HeVQVEQP.FF.VPSHTCD...EK...L...MQAGKM...NG...PEI...LKNE
hRSVSRTKNF.IPK...KADKES.DANKSLIPFCY...FIT...KKG...
RABVIRG.LD.IRP...EF...RS...NQD...V...R...GF...E...
BDVIRG.LD.IRP...EF...DE...L...VVQ...H...GQ...SVL...VEVV

hMPV α10

310

hMPVISYYN.NHIRVGP.P.PNPPSDGIAQNVGIAITGISFWLSLMEKDIPLYQQCLAVIQQSFPIRWE...VVS...NKK...ELNRQR
VSVYHRYN.VDSRGGP.V.S...TQHLAHLRA...E...RE...KGGY...KQKWS
SUDVYHRYN.VDSRGGP.V.S...TQHLAHLRA...E...RE...NDY...NQQRQ
TAFVYHRYN.VDSQKGP.T.S...IVQHLAHLQ...E...RE...FDY...NQQRQ
BDBVYHRYN.VDSRGGP.V.S...ILYHLTHLQA...E...RE...LDY...NQQRQ
REBTVYHRYN.VDTGLGP.S.S...VIRHTNLQQA...E...RD...LDY...NMRRE
MARVCHRYT.PFSEKAL.FHK...VQSYVRQGGQ...H...HS...LDY...ENNSPL
LLOVHYRYN.VATESSP.L.T...IRRHNLNLLQ...E...RE...RDY...FLSRC
SeVLSTMN.ADTHNCM.A.F...NRVL...KDT...FEW...RITESDKRL.K
MeVI.H.HD.VASQDGL.N.S...ILIL...YRE...L...RFK...DNQR.
MuV .DLPDYLSPFNELQSEMAR.I.T...IH...KEV...EILK...GQS.
HeV .VG.YD.GSDINTLRS.T...IIL...NEA...NYF...DDE.
hRSVINTALSLSV...SGD...LSY...SIAGR.
RABVYVVG...VRRRPV.L.A...RH...TTLSDVSEF...HK...DDL...ELQRG.
BDVYVVG...VRRRPV.L.A...RH...TTLSDVSEF...HK...DDL...ELQRG.

hMPV α11

320 330 340 350

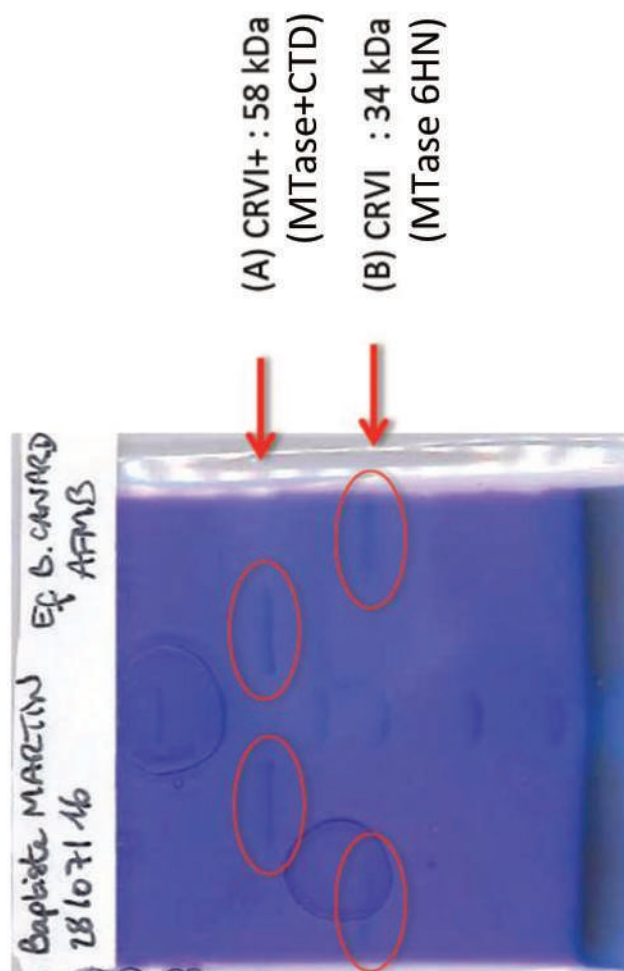
hMPV .RLTLQSNHSSVATVGGSKVIESKWLTKANTIIDWLEHI...LN
VSV .RG...DGLPKDTRI...SDSL
EBOV .RT...QT...YHF...R
SUDV .RT...QT...YHF...K
TAFV .RT...QT...YHF...K
BDBV .RT...QT...YHF...K
REBTV .RT...QT...YHF...K
MARV .LD...LR...NHF...IC
LLOV .RT...QA...FHF...K
SeV .LT...IT...GKYDLYPVRDVGK...K
MeV .QT...Q...GMFHAYPVLVSSR...R
MuV .DH...DT...LLFTSYNVGVLGK...N
HeV .EV...FS...HLEFPVLEKTR...K
hRSV .L...SMPYEM...NKI...N
RABV .L...SKM...SII...AIMIVFSNRVFNVSXPLTDPLFYPPSDPKILRHFNICCS
BDV .L...SKM...SII...AIMIVFSNRVFNVSXPLTDPLFYPPSDPKILRHFNICCS

ANNEXE 2 – Contrôle des protéines produites par spectrométrie de masse

Spectrométrie de masse pour la vérification de la séquence des protéines recombinantes (réalisé par Patrick Fourquet)

Les bandes ont été découpées, digérées à la trypsine et analysées par Maldi-TOF (ULTRAFLEX III, Bruker).

Les pics obtenus ont été comparés à la liste des pics de la digestion tryptique théorique de la séquence attendue. Les peptides du spectre retrouvés dans la séquence sont représentés en rouge dans la diapositive suivante. Les rectangles en nuances de gris en dessous délimitent les peptides et leurs intensités relatives sur le spectre (gris clair pour les moins intenses jusqu'à noir pour les peptides pris comme 100%). Le pourcentage global de couverture est encadré.



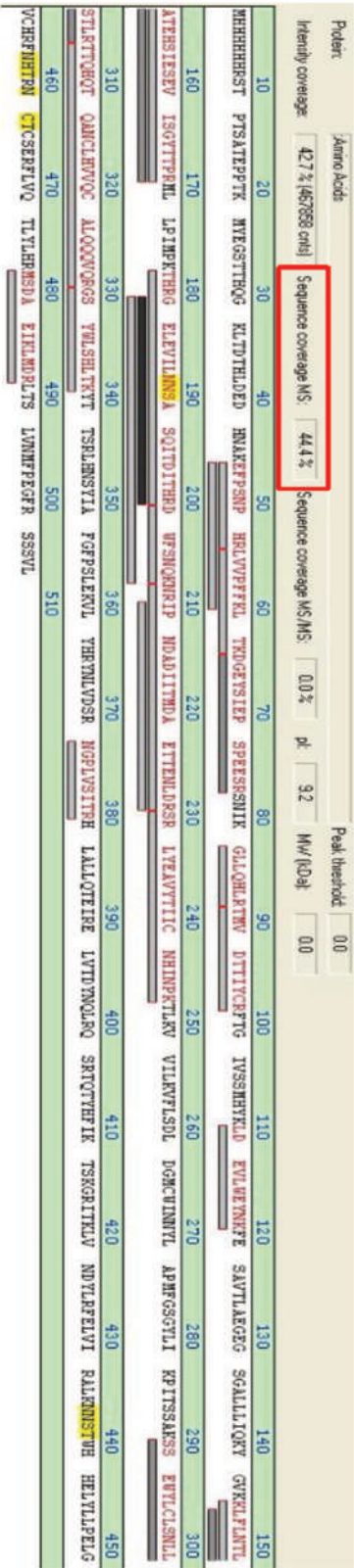
Séquences :

➤ (A) CRVI+
 MHHHHHRSTPTSATEPPTKMYEGSTTHQGKLTDTHTLDEHNAKEFPSNPHRLVVPFFKLTKDGEYSIEPSPE
 ESRNIKGLLQHLRMTVDTTIYCRFTGIVSSMHYKLLDVLWEYNKFESAVTLAEGEGSGALLIQKYGVKKLFLN
 TLATEHSIESEVIGYTTTPRMLLPMPKTHRGELEVLNNSASQITDTHRDFWFSNQKNRIPNDADIITMDAETTE
 NLDRSRLYEAVYTIICNHINPKTLKVVILKVFSLDGMWCWINNYLAPMFGSGYLKIPITSSAKSSEWYLCLSNLLS
 TLRTTQHQTOANCLHVVQCALQQQVQRGSYWLSHLTKYTTSLRHLNSYIAFGFSPLEKVLVHRYNLVDSRNGP
 LV/SITRHLALLOQTEIRVELVDYNQLRQSRQTQYHFIKTSKGRITKLVNDYLRFELVIRALKNNSTWHHELYLLPELG
 VCHRFRNHTRNCTCSEFLVQTLYLHRMSD AEIKLMDRLTSLVNMFPPEGFRSSSVL

➤ (B) CRVI
 MHHHHHHEDHNAKEFPSNPHRLVVPFFKLTKDGEYSIEPSPEESRSNIKGLLQHLRMTVDTTIYCRFTGIVSSM
 HYKLDENVLWEYNKFESAVTLAEGEGSGALLIQKYGVKKLFLNLTATEHSIESEVIGYTTTPRMLLPMPKTHRGE
 LEVLNNSASQITDTHRDFWFSNQKNRIPNDADIITMDAETTENDRSRLYEAVYTIICNHINPKTLKVVILKVFLLS
 DLDGMWCWINNYLAPMFGSGYLKIPITSSAKSSEWYLCLSNLLSLTRTTQHQTOANCLHVVQCALQQQVQRGS
 Y WLSHLTKYTT

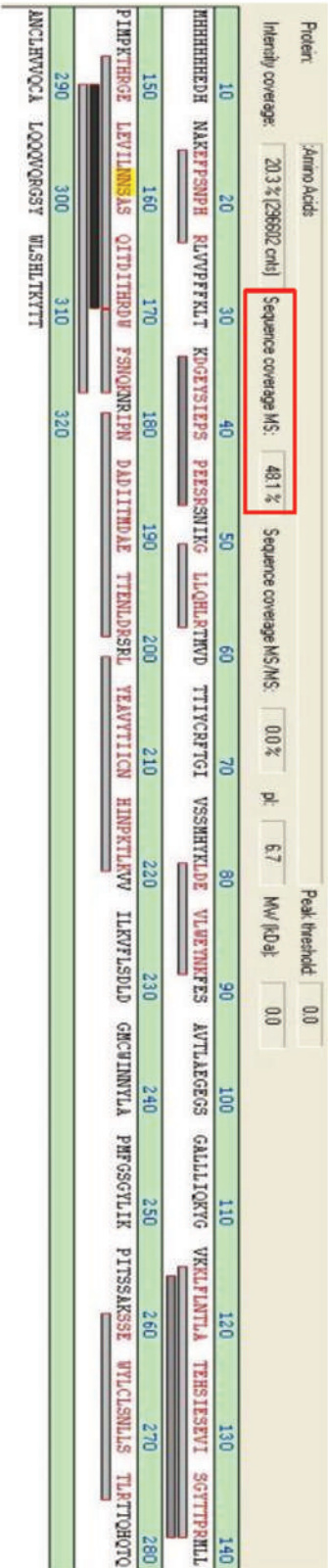
Bande A = CRVI +

(MTase+CTD)



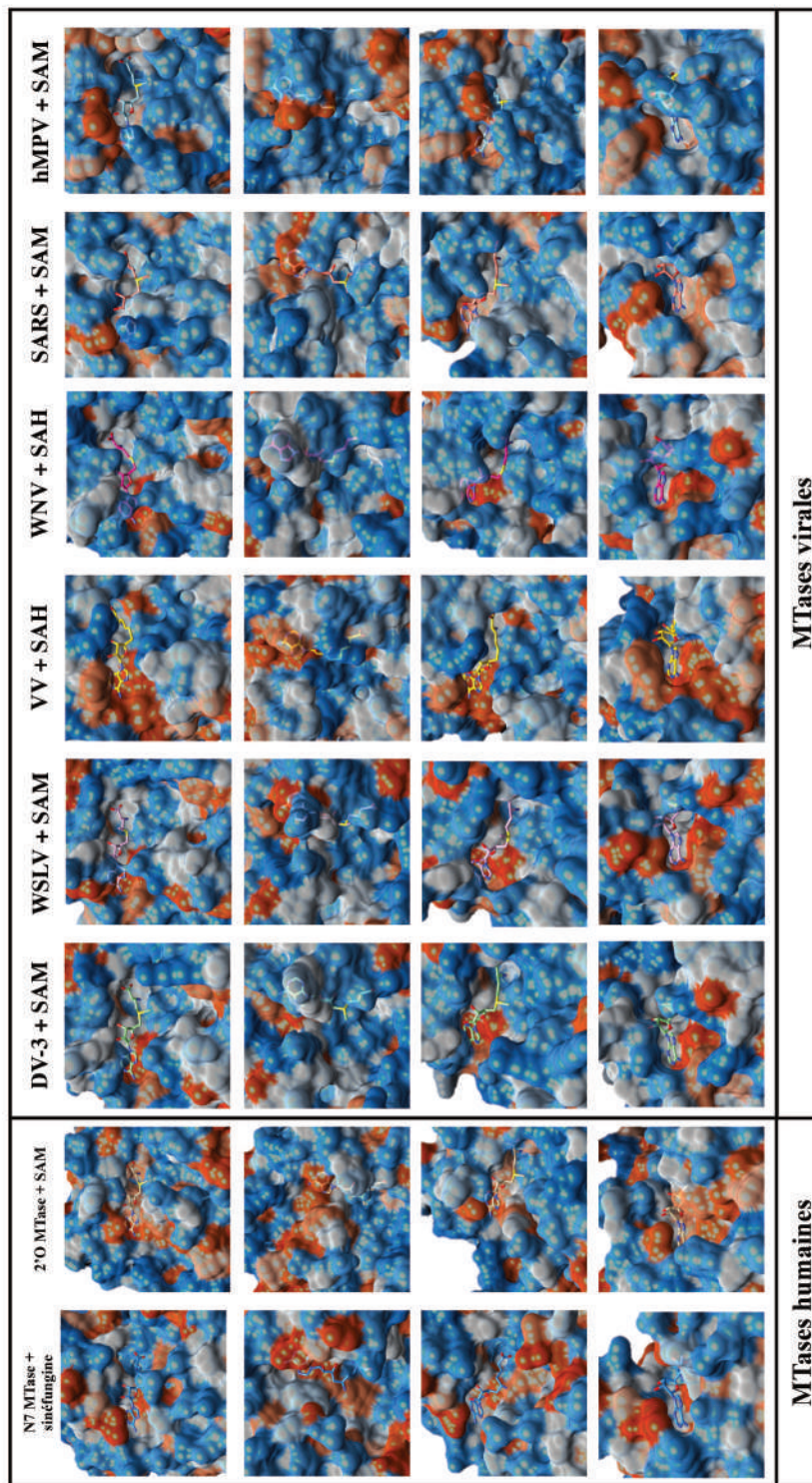
Bande B = CRVI

(MTase 6HN)



Les 2 bandes analysées correspondent aux protéines attendues

**ANNEXE 4 – Comparaison des propriétés électrostatiques du site de fixation du SAM
chez les MTases humaines et virales**



ANNEXE 5 – Article 1

Titre : Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives.

Auteurs : Coutard B, Barral K, Lichière J, Selisko B, Martin B, Aouadi W, Lombardia MO, Debart F, Vasseur JJ, Guillemot JC, Canard B, Decroly E.

Référence : J Virol. 2017 Feb 14;91(5). pii: e02202-16. doi: 10.1128/JVI.02202-16. Print 2017 Mar 1.

Le flavivirus Zika (ZIKV) est l'agent pathogène responsable de troubles neurologiques tels que des microcéphalies chez le nouveau-né et de syndromes de Guillain-Barré. Sa protéine NS5 porte un domaine méthyltransférase (MTase) impliqué dans la formation de la coiffe des ARNm viraux. Nous avons caractérisé les propriétés structurales et fonctionnelles de la MTase du virus Zika. Nous avons montré que la MTase de ZIKV peut méthyler en position N7 de la coiffe et en position 2'O du ribose du premier nucléotide de l'extrémité 5' des ARNs formant une structure coiffe-1. De plus, la MTase de ZIKV est capable de méthyler les positions 2'O du ribose des résidus adénosines internes. La structure cristallographique de la MTase de ZIKV décrite à 2,01 Å de résolution a révélé une organisation cristallographique en homodimère. Une chaîne est liée au donneur de méthyl (S-adénosyl méthionine ou SAM) et présente une forte similarité structurale avec le domaine MTase du virus de la dengue (DENV). La seconde chaîne ne contient pas de SAM et présente des différences conformationnelles notamment au niveau de l'hélice αX contribuant au recrutement du SAM et de l'ARN. Ces modifications structurales révèlent un mécanisme probable du turnover enzymatique impliquant un motif Ser/Arg conservé. Dans la seconde chaîne, la poche de fixation du SAM accomode un sulfate proche d'une molécule de glycérol qui pourrait servir de base pour le développement d'antiviraux. En outre, des composés connus pour leur inhibition de l'activité MTase de DENV montrent une inhibition similaire sur le domaine MTase de ZIKV. L'ensemble de ces résultats contribuent donc à une meilleure compréhension de l'activité MTase de ZIKV, un acteur central dans la réplication virale et l'échappement à la réponse immunitaire innée de l'hôte, et de servir de base pour le développement de futures molécules antivirales.

Dans ce travail, j'ai participé à la caractérisation biochimique du domaine MTase de ZIKV. En effet, j'ai évalué l'activité de cette protéine sur différents ARNs substrats synthétiques et j'ai mis en évidence une activité méthyltransférase sur les résidus adénosines internes d'ARN simple brin. J'ai également participé au processus de révision.



Zika Virus Methyltransferase: Structure and Functions for Drug Design Perspectives

Bruno Coutard,^a Karine Barral,^b Julie Lichière,^a Barbara Selisko,^a Baptiste Martin,^a Wahiba Aouadi,^a Miguel Ortiz Lombardia,^c Françoise Debart,^d Jean-Jacques Vasseur,^d Jean Claude Guillemot,^a Bruno Canard,^a Etienne Decroly^a

Aix Marseille Université, CNRS, AFMB UMR 7257, Marseille, France^a; Aix Marseille Université, CNRS, INSERM, Institut Paoli-Calmettes, CRCM, Marseille, France^b; Structural and Genomic Information Laboratory, UMR 7256 (IMM FR 3479) CNRS Aix-Marseille Université, Luminy Campus, Marseille, France^c; IBMM, UMR 5247 CNRS, Université de Montpellier, ENSCM, Department of Nucleic Acids, Montpellier, France^d

ABSTRACT The *Flavivirus* Zika virus (ZIKV) is the causal agent of neurological disorders like microcephaly in newborns or Guillain-Barre syndrome. Its NS5 protein embeds a methyltransferase (MTase) domain involved in the formation of the viral mRNA cap. We investigated the structural and functional properties of the ZIKV MTase. We show that the ZIKV MTase can methylate RNA cap structures at the N-7 position of the cap, and at the 2'-O position on the ribose of the first nucleotide, yielding a cap-1 structure. In addition, the ZIKV MTase methylates the ribose 2'-O position of internal adenosines of RNA substrates. The crystal structure of the ZIKV MTase determined at a 2.01-Å resolution reveals a crystallographic homodimer. One chain is bound to the methyl donor (*S*-adenosyl-L-methionine [SAM]) and shows a high structural similarity to the dengue virus (DENV) MTase. The second chain lacks SAM and displays conformational changes in the α X α -helix contributing to the SAM and RNA binding. These conformational modifications reveal a possible molecular mechanism of the enzymatic turnover involving a conserved Ser/Arg motif. In the second chain, the SAM binding site accommodates a sulfate close to a glycerol that could serve as a basis for structure-based drug design. In addition, compounds known to inhibit the DENV MTase show similar inhibition potency on the ZIKV MTase. Altogether these results contribute to a better understanding of the ZIKV MTase, a central player in viral replication and host innate immune response, and lay the basis for the development of potential antiviral drugs.

IMPORTANCE The Zika virus (ZIKV) is associated with microcephaly in newborns, and other neurological disorders such as Guillain-Barre syndrome. It is urgent to develop antiviral strategies inhibiting the viral replication. The ZIKV NS5 embeds a methyltransferase involved in the viral mRNA capping process, which is essential for viral replication and control of virus detection by innate immune mechanisms. We demonstrate that the ZIKV methyltransferase methylates the mRNA cap and adenosines located in RNA sequences. The structure of ZIKV methyltransferase shows high structural similarities to the dengue virus methyltransferase, but conformational specificities highlight the role of a conserved Ser/Arg motif, which participates in RNA and SAM recognition during the reaction turnover. In addition, the SAM binding site accommodates a sulfate and a glycerol, offering structural information to initiate structure-based drug design. Altogether, these results contribute to a better understanding of the *Flavivirus* methyltransferases, which are central players in the virus replication.

KEYWORDS methyltransferase, RNA processing, Zika virus, flavivirus, protein structure-function

Received 7 November 2016 Accepted 16 December 2016

Accepted manuscript posted online 28 December 2016

Citation Coutard B, Barral K, Lichière J, Selisko B, Martin B, Aouadi W, Lombardia MO, Debart F, Vasseur J-J, Guillemot JC, Canard B, Decroly E. 2017. Zika virus methyltransferase: structure and functions for drug design perspectives. *J Virol* 91:e02202-16. <https://doi.org/10.1128/JVI.02202-16>.

Editor Michael S. Diamond, Washington University School of Medicine

Copyright © 2017 American Society for Microbiology. All Rights Reserved.

Address correspondence to Etienne Decroly, Etienne.Decroly@afmb.univ-mrs.fr.

First identified in 1947 in Uganda, Zika virus (ZIKV) reemerged in Brazil in early 2015, and the virus has continued to disseminate to South and North America in 2016. More than 2 million suspected cases have been reported (1), and ZIKV is now considered a major public health threat by the World Health Organization (WHO). Whereas the bite of an infected *Aedes aegypti* mosquito is the main cause of ZIKV transmission, direct human-to-human transmissions by sexual contacts have also been reported (2). ZIKV infection was initially considered a minor infection, often asymptomatic, causing mainly Zika fever with rash, arthralgia, and conjunctivitis. However, this outbreak revealed that ZIKV infecting pregnant women can be transmitted to the fetus, causing fetal loss, microcephaly, and other serious brain defects in newborns (3). In addition, Guillain-Barre syndrome, a rare autoimmune disease affecting the peripheral nervous system, and other severe neurological disorders have also been associated with ZIKV infection in adults (4). In the absence of a vaccine and antiviral drugs, the only way to reduce virus transmission is to limit the propagation of *Aedes* mosquito vectors. In view of the current expanding outbreak, it is urgent to develop a vaccine and antiviral strategies blocking essential enzymes involved in ZIKV replication.

ZIKV belongs to the *Flavivirus* genus, which contains other important human pathogens such as dengue virus (DENV), West Nile virus (WNV), or yellow fever virus (YFV). Flaviviruses are enveloped viruses containing a single-stranded positive-sense RNA genome of around 11,000 nucleotides, which is decorated by a cap structure. The cap structure consists of a guanosine residue linked through a 5'-5' triphosphate bond to the 5' end of mRNA (GpppN) (5). In flaviviruses, the cap structure is methylated at the nitrogen in position 7 of the guanosine (cap-0 structure, or ^mGpppN) and at the 2'-oxygen atom (2'-O) of the N₁ ribose (cap-1 structure, or ^mGpppN_m) (6, 7). The N-7 methylation of the cap structure is essential for RNA stability and stimulates their translation into viral protein by recognizing the translation initiation factor eIF4E (8). The 2'-O-methylation of the cap structure was demonstrated to protect viral RNA from being recognized by host cell sensors such as RIG-I and MDA5 that stimulate the production of interferons (9–11). In turn, interferon-stimulated genes, such as IFIT-1, can detect miscapped RNA and restrict their translation into proteins (12, 13). DENV mutant viruses lacking N-7-methyltransferase (MTase) activity show strong replication defects in infected cells, whereas those altering the 2'-O-MTase activity show only attenuated phenotypes (7, 14), as most cell lines used in the laboratory for virus replication are deprived of the RIG-I/MDA5 antiviral pathway. In contrast, it was demonstrated that a 2'-O-MTase knockout virus barely replicates in infected mice and elicits a strong humoral and cellular antiviral response (15, 16). Thus, viral RNA capping represents an attractive antiviral strategy (17, 18) since it should inhibit viral replication and/or accelerate virus clearance upon stimulation of the innate antiviral response.

In *Flavivirus*, the genome replication is driven by NS5, the RNA-dependent RNA polymerase (RdRp), which is associated with other viral nonstructural proteins in a membrane-bound replication-transcription complex. In addition to the RdRp domain, NS5 also harbors an N-terminal MTase domain, which methylates the cap structure of nascent RNA at both the N-7 and 2'-O positions (6, 7). Initially, biochemical studies performed with recombinant NS5 proteins or MTase domains (NS5-MTase) of DENV, WNV and other flaviviruses have confirmed the 2'-O-MTase activity using short synthetic RNA substrates in the presence of the methyl donor S-adenosyl-L-methionine (SAM, or AdoMet) (6, 19). The N-7-MTase activity was next demonstrated using longer RNA, with an optimal activity detected when a RNA containing the highly conserved stem loop A (SLA) hairpin structure was used (7, 20). The DENV MTase domain was later reported to methylate the 2'-O-ribose position of internal adenosines of the RNA (21). Nevertheless, the role of these methylations is still poorly understood. The structures of several *Flavivirus* NS5-MTases have been solved (6, 7), and the structure of ZIKV MTase has recently been determined by Coloma et al. (22). They adopt a canonical MTase Rossmann fold with a seven-stranded β -sheet surrounded by four α -helices. The MTase core structure closely resembles the catalytic domain of other viral and cellular SAM-dependent MTases with a conserved SAM binding site in close vicinity of the conserved

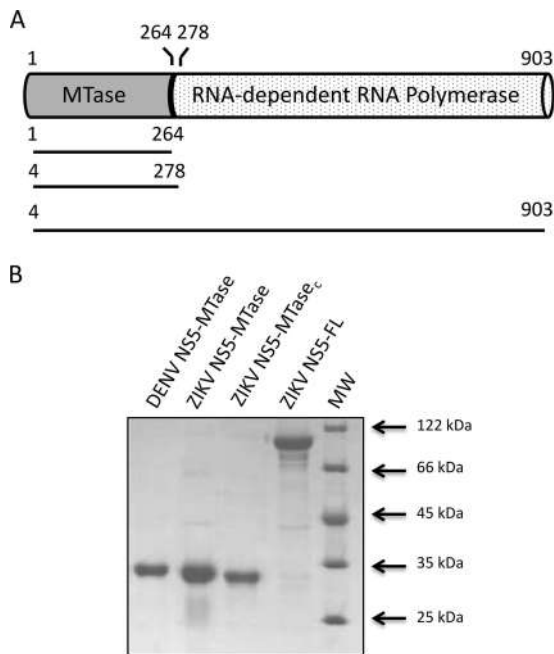


FIG 1 Purification of ZIKV NS5 2'-O-MTase. (A) The NS5 protein of ZIKV is a bifunctional protein with an MTase domain (N terminus) and an RdRp domain (C terminus). Two constructs, aa 4–903 and aa 4–278, were used for the functional studies, whereas construct aa 1–264 was used for crystallogenes and structure determination. (B) The recombinant proteins were separated by SDS-PAGE and visualized by Coomassie blue staining prior their functional characterization. Arrows indicate the molecular mass markers.

K-D-K-E catalytic tetrad, the latter being a signature of 2'-O-MTases. In addition, structures determined in the presence of cap analogues, GTP or RNA reveal a cap binding pocket where the cap guanosine is held during 2'-O-methylation by stacking interactions with a phenylalanine (F24 residue in the DENV NS5 sequence) and an RNA binding zone where the RNA chain downstream of the cap structure is accommodated by a basic groove (4, 6, 7, 22–25).

Enzyme-based screenings or structure-based drug design have identified DENV, WNV, and YFV MTase inhibitors targeting either the SAM/SAH (*S*-adenosyl-*L*-homocysteine, the coproduct of the methylation) binding pocket (26–28), the cap binding pocket (29), or allosteric sites (30, 31). As ZIKV is closely related, we initiated a comparative study of ZIKV MTase with other *Flavivirus* MTases with the aim of subsequently developing original inhibitors or repurposing inhibitors targeting DENV MTase for antiviral research on ZIKV. Here we describe the structural and functional analysis of the ZIKV MTase domain as well as the evaluation of inhibitors developed for DENV MTase on ZIKV MTase.

RESULTS

Production and purification of ZIKV MTase constructs. The ZIKV genome encodes the bifunctional NS5 protein, with an ~264-amino-acid (aa) MTase domain at the N terminus followed by an RNA-dependent RNA polymerase (RdRp) domain (Fig. 1A). The MTase domain contains the canonical 2'-O-MTase (K-D-K-E) catalytic tetrad and possesses ~65% and 70% similarity to those of the four DENV serotypes and WNV, respectively. In this study, different constructs were designed. For functional assays, a construct, aa 4–278, encompassing the putative MTase domain from aa 4 to aa 278 of NS5, and full-length NS5 (aa 4 to 903) were produced and purified. From NS5 preparations, both the full-length protein and the MTase domain resulting from cleavage by bacterial proteases (here called NS5-MTase_c) were recovered after size exclusion chromatography (SEC). The purified proteins were detected as single bands at their expected molecular masses (~30 to 35 kDa and ~100 kDa, respectively) (Fig. 1B) upon

TABLE 1 Data collection and refinement statistics

Parameter	Value(s) for SAM complex ^a
Data collection and phasing statistics	
Space group	P1
Unit cell parameters	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	37.6, 64.1, 72.0
α , β , γ (°)	113.1, 97.8, 92.0
Resolution range (Å)	65.41–2.01 (2.02–2.01)
No. of unique reflections	40,055 (398)
Completeness (%)	99.3 (93.9)
Multiplicity	3.3 (2.5)
R_{meas}^b	0.157 (0.454)
$I/\sigma(I)$	5.8 (2.0)
Refinement statistics	
Resolution range (Å)	65.41–2.01
No. of reflections:	
Used for refinement	40,054
Used for R_{free} calculation	2,105
R_{work} (%) ^c	16.0
R_{free} (%) ^d	19.1
No. of atoms	
Protein chain A	2,077
Protein chain B	1,963
Ligand/ion (SAM/glycerol/others)	27/6/21
Water	375
B factors (Å ²)	
Protein chain A	28.0
Protein chain B	28.2
Ligand/ion (SAM/glycerol/others)	25.3/49.5/80.1
Water	37.9
RMSD from ideality	
Bond lengths (Å)	0.010
Bond angles (°)	0.96
Ramachandran plot (%)	
Residues in most favored regions	97.40
Residues in disallowed regions	0

^aValues in parentheses are for the outermost resolution shell of data.

^b $R_{\text{meas}} = \sum |I_{\text{obs}} - I_{\text{av}}| / \sum I_{\text{av}}$, over all symmetry-related observations.

^c $R_{\text{work}} = \sum |F_{\text{obs}} - \langle |F_{\text{calc}}| \rangle| / \sum |F_{\text{obs}}|$, over all reflections included in the refinement.

^d R_{free} is calculated with 5% of reflections excluded from the refinement. In this formula, $\langle |F_{\text{calc}}| \rangle$ denotes the expectation value of $|F_{\text{calc}}|$ under the probability distribution used to define the likelihood function that is maximized in the refinement.

SDS-PAGE analysis. The apparent molecular mass observed for NS5-MTase_C is similar to those of the ZIKV and DENV MTase domains (Fig. 1B).

ZIKV MTase crystallizes as a homodimer in which the two protomers show conformational differences. As the original MTase construct (aa 4–278) did not crystallize, we generated another that includes the complete N terminus but is devoid of the linker between the MTase and RdRp domains (aa 1–264; PDB code 5M5B). This construct allowed determination of the crystal structure of the ZIKV MTase domain. The structure was solved by molecular replacement using the DENV3 NS5 MTase/SAM (PDB code 3P97) as the template model and refined to a resolution of 2.01 Å, with R_{work} and R_{free} values of 16.0% and 19.1%, respectively. The crystal belongs to space group P1 with the following unit cell parameters: $a = 37.6$ Å, $b = 64.1$ Å, and $c = 72.0$ Å and $\alpha = 113.1^\circ$, $\beta = 97.8^\circ$, and $\gamma = 92.0^\circ$. Data collection and refinement statistics are reported in Table 1. The ZIKV MTase domain crystallized with two molecules in the asymmetric unit (Fig. 2A), whereas size exclusion chromatography indicates that the protein behaves as a monomer in solution. The low-interaction surface area (about 500 Å²) suggests that dimerization occurs during crystallization as previously observed for Modoc virus (MODV) MTase (32). The two protomers (A and B chains) are similarly folded, with significant differences arising from disorder in residues 43 to 58 of chain B (Fig. 2A). A superimposition of 244 C α atoms in the A and B chains of the protein

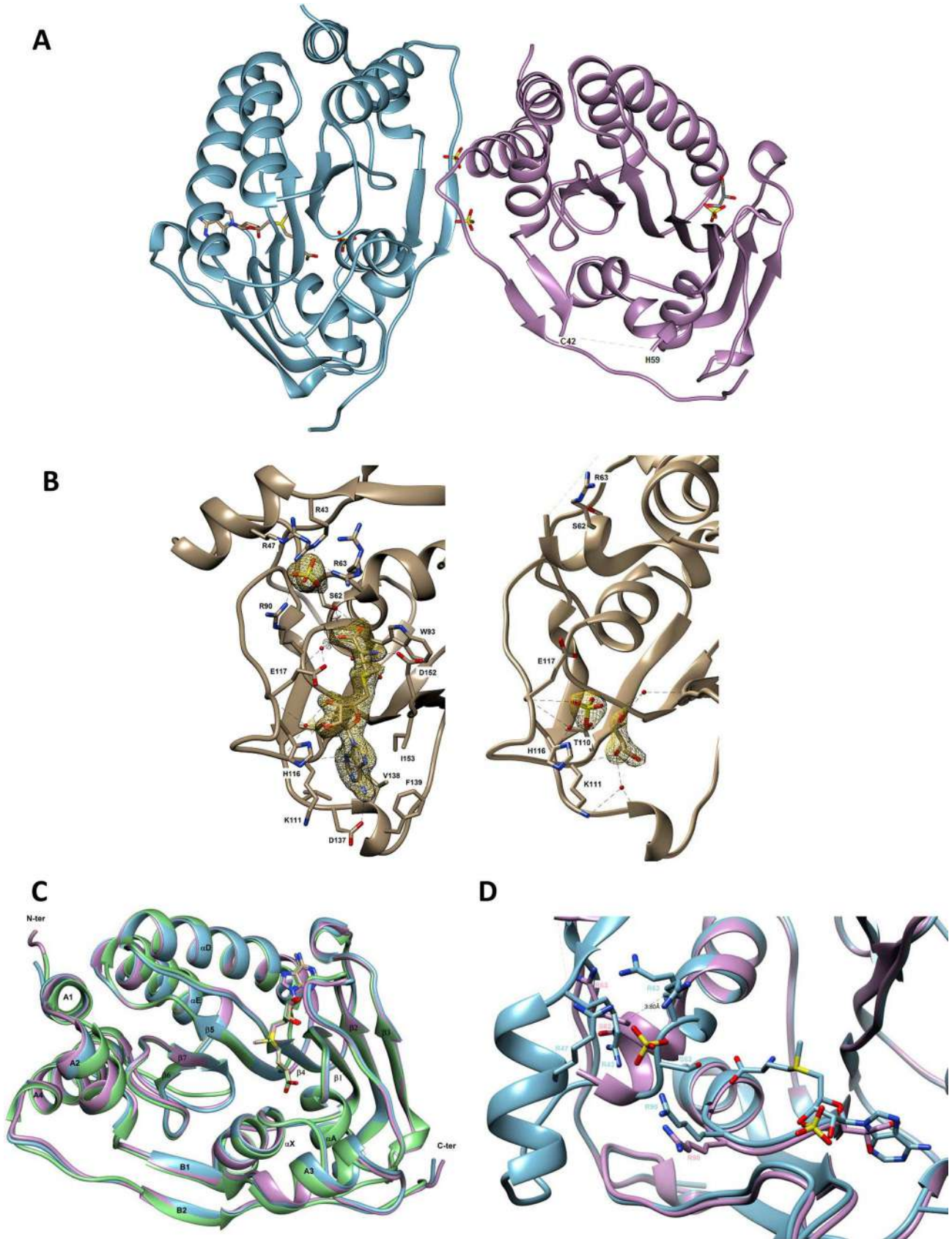


FIG 2 Crystal structure of ZIKV NS5-MTase. (A) Crystal structure of the ZIKV MTase two monomers (A and B chains). ZIKV MTase is drawn as a ribbon: chain A is colored blue, and chain B is colored pink. (B) To the left is shown the structure of the ZIKV MTase chain A in complex with the cofactor SAM and a (Continued on next page)

resulted in an RMSD of 0.52 Å. Chain A consists of a canonical MTase core (residues 61 to 229, construct numbering PDB code [5M5B](#)) folded into a seven-stranded β -sheet ($\beta 1$ to $\beta 7$) surrounded by four α -helices (αX , αA , αD , and αE), as shown in Fig. 2A and C. Appended to the core are an N-terminal extension (residues 8 to 60) and a C-terminal extension (residues 230 to 272). The N-terminal subdomain comprises a helix-turn-helix motif followed by a β -strand and an α -helix (A1, A2, B1, and A3). The C-terminal subdomain consists of an α -helix followed by a β -strand (A4 and B2). Inspection of the electron density maps reveals strong additional electron density in the SAM binding pocket, allowing the unambiguous modeling of SAM. Because no SAM is added during purification or crystallization, its presence in the structure must have originated from *E. coli*. The SAM molecule is bound in the central cleft formed of β -strands $\beta 1$, $\beta 2$, and $\beta 4$ and α -helices αX and αA (Fig. 2A and B). The SAM adenine base is accommodated within a hydrophobic pocket defined by the side chains of Val138, Phe139, and Ile153 and stabilized by hydrogen bonds with a carboxylic oxygen from Asp137 and the nitrogen main-chain atoms of Lys111 and Val138 (Fig. 2B, left panel). The ribose moiety is hydrogen bonded to the Glu117 carboxyl group. The sugar is also bridged via a water molecule to the side chain of Glu117. The methionine tail of the SAM molecule is positioned by interactions with the side chains of Ser62, Trp93, and Asp152, as well as with the main chain of Gly92. The methionine carboxylate is also bridged via a single water molecule to the side chain of Glu117 and to the backbone of Arg90. Three molecules of sulfate interact with chain A. One of them is bound via a cluster formed by arginines Arg43, Arg47, Arg63, and Arg90 (Fig. 2D). The structure of chain A is highly similar to those of DENV-3 MTase (RMSD of 0.79 Å with PDB code [3P97](#) chain A) and of ZIKV MTase recently determined by Coloma et al. (RMSD of 0.29 Å with PDB code [5KQR](#) chain A) (22). There is no difference in SAM binding with the latter structure (PDB code [5KQR](#)), and the sulfate bound via a cluster formed by arginines Arg43, Arg47, Arg63, and Arg90 overlaps perfectly with the phosphate described by Coloma et al. (22). Structural differences from DENV-3 MTase are limited to small changes in the conformation of three solvent-exposed loops (residues 55 to 58, 178 to 183, and 251 to 254), as shown by superimposition of several structures (Fig. 2C).

We observe significant conformational changes between chain A and chain B of ZIKV MTase (RMSD of 0.52 Å) likely favored by their different crystal environments. Indeed, the chain B environment does not allow this loop to adopt the conformation observed in chain A. However, it is unlikely that crystal packing is the only factor involved since similar disorder has been observed, under different crystal packing, for the MODV MTase (PDB code [2WA2](#)) (32). We conclude that the intrinsic flexibility of this loop, at least in the absence of the RNA substrate, has also contributed to building of this crystal form. As illustrated in Fig. 2A and D, three main differences are noteworthy in chain B, as follows. (i) There is no supporting electron density for residues 43 to 58 (part of the B1 β -strand, the A3 α -helix, and the following loop (Fig. 2A)). A similar absence of electron density has already been described for one of the two chains of the MODV MTase (32), suggesting a certain degree of flexibility in this region. (ii) Downstream of the aa 43–58 flexible region, the αX α -helix is kinked (Fig. 2D), with its N-terminal part (residues 61 to 67) taken away from this helix axis. (iii) There is no additional electron density characterizing the methyl donor SAM in the SAM binding site. Interestingly, in chain B of the MODV MTase, a SAM molecule is present, even if the

FIG 2 Legend (Continued)

sulfate ion. To the right is shown the structure of ZIKV MTase chain B in complex with glycerol- and sulfate-bound molecules. ZIKV MTase is drawn as a ribbon and colored gray. Small red spheres represent water molecules around the binding sites. Hydrogen bond interactions are shown as black dotted lines. The amino acid residues shown are those establishing Van der Waals interactions with bound molecules. In the insets, electron density maps are shown as a yellow mesh contoured at 1.5 σ around bound molecules. (C) Superposition of the ZIKV MTase chain A (PDB code [5M5B](#) [in blue]), of the ZIKV MTase described by Coloma et al. (22) (PDB code [5KQR](#) [in pink]), and of the DENV MTase (PDB code [3P97](#) [in green]) structures bound with SAM. (D) Superposition of ZIKV MTase chain A (drawn as a blue ribbon) and of ZIKV MTase chain B (drawn as a pink ribbon). Shown is a view of the cluster formed by arginines Arg43, Arg47, Arg63, and Arg90 and of the SAM binding site. The distance between Arg63 from chain A and chain B is shown as black dotted lines. For panels A through D, MTases are drawn as ribbons. The small molecules, SAM, glycerol, and sulfate ions are shown in cylinder representation with gray indicating the carbon atom. Nitrogen, oxygen, and sulfur are colored blue, red, and yellow, respectively.

homologous region aa 35–56 appears to also be disordered. However, the kink observed in the α X α -helix of ZIKV MTase chain B is absent from MODV MTase chain B. The kink has two main structural consequences. First, the side chain of Ser62 cannot interact anymore with the methionine tail of the SAM, as observed in chain A (Fig. 2B, left panel). Second, a cluster formed by arginines of the RNA binding groove (Arg43, Arg47, Arg63, and Arg90) in chain A is distorted in chain B and does not accommodate any sulfate molecule (Fig. 2D). In the absence of SAM, two polar molecules, a sulfate and a glycerol close by (~ 3.2 Å), are held in the SAM binding pocket of chain B (Fig. 2A and B). The sulfate molecule is positioned by hydrogen bonds with the side chain of His116, the main chain of Glu117, and with the bound glycerol nearby. It is also bridged via a single water molecule to the main chains of Gly112 and Glu117, as well as to the side-chain hydroxyl of Thr110. Moreover, the glycerol molecule participates in van der Waals interactions with the main chains of Gly87 and Lys111 and is stabilized by hydrogen bonds with the sulfate and the main-chain amide of Lys111. The glycerol is also bridged via two water molecules to the main chains of Asp137 and Val138 for the first one and to the main chains of Gly87 and Asp152 for the second one. Such a structure with compounds bound in the SAM binding site could serve as a basis for structure-based drug design.

Full-length ZIKV NS5 displays a 2'-O-MTase activity that is not regulated by the polymerase domain. The 2'-O-MTase activity was measured comparatively using the three ZIKV proteins (Fig. 1B). Proteins were incubated with a short RNA substrate bearing an unmethylated cap structure, GpppAC₄, in the presence of a radiolabeled methyl donor ([³H]SAM). The amount of [³H]CH₃ transferred onto RNA was quantified with a DEAE filter binding assay. ZIKV NS5, NS5-MTase_C, and the ZIKV NS5-MTase domain exhibit similar enzymatic activities (Fig. 3A). The level of detected enzymatic activities is about 30% lower than the control activity of DENV NS5-MTase. ZIKV NS5-MTase activity is thus not stimulated by the presence of the RdRp domain in *cis*. In addition, we compared the enzymatic activity of the ZIKV MTase domain using the unmethylated substrate GpppAC₄ and a substrate methylated at the N-7 position, mGpppAC₄. The methyl transfer efficiency is similar (Fig. 3B), suggesting that both substrates can be methylated at the 2'-O position of the adenosine ribose, as already reported for DENV and WNV MTases (19, 33). Additionally, the ZIKV MTase domain shows an optimal activity between pH 8 and 9 (Fig. 3C). Time course experiments indicate that the 2'-O-MTase functions without additional magnesium (not shown). We next addressed sequence specificity using short RNAs consisting of the 5' ends of various viral RNAs starting either with A or G (Fig. 3D). Although RNA methylation efficiency by ZIKV MTase varies, we could not observe any obvious preference for *Flavivirus* GpppA-RNAs or GpppG-RNAs.

The ZIKV MTase requires long structured RNA to methylate the N-7 position of the cap. An SLA hairpin structure is present at the 5' end of *Flavivirus* genomes. This structure is required for the detection of the N-7-MTase activity of WNV and DENV (7, 20). First, we compared the 5' ends of ZIKV and DENV. The seven first residues of the ZIKV RNA are strictly identical to those of DENV, and ZIKV RNA forms a hairpin structure similar to that of DENV. N-7-methylation assays driven by ZIKV NS5-MTase were thus conducted on SLA-containing DENV RNAs with various 5'-capped ends (GpppA-RNA₇₅, mGpppA-RNA₇₅, GpppA_m-RNA₇₅, and mGpppA_m-RNA₇₅). mGpppA-DENV₇₅ and GpppA_m-DENV₇₅ are both significantly better methylated than mGpppA_m-DENV₇₅ ($P < 0.001$), indicating that N-7- and 2'-O-methylations occur (Fig. 3E). In addition, the high methylation level observed with mGpppA-DENV₇₅ suggests that the 2'-O-methylation is more efficient than the N-7-methylation under the applied experimental conditions. Interestingly, MTase activity using RNA already methylated at the N-7- and 2'-O positions (mGpppA_m-DENV₇₅) was also detected. This observation suggests that additional methylations downstream of the cap structure might occur (described below). Altogether, these results show that, as observed for other flaviviruses, the ZIKV MTase can perform cap 2'-O-methylation on short RNAs without sequence specificity, whereas

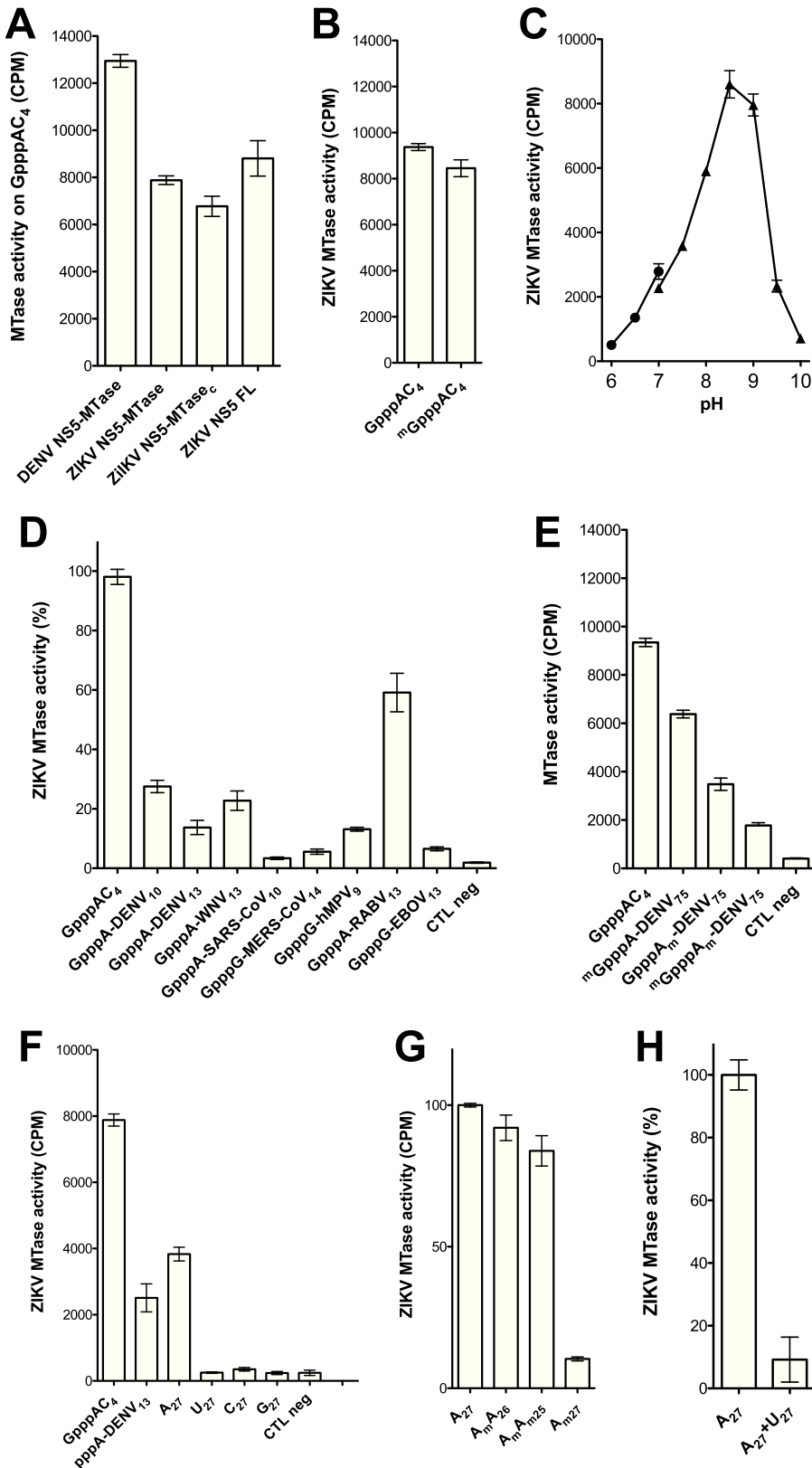


FIG 3 Characterization of the three MTase activities carried by ZIKV MTase. The ZIKV MTase activity was determined by monitoring the transfer of tritiated methyl groups (in counts per minute) from SAM to various synthetic RNAs. The bar graph presents the mean value and the standard deviation from 3 independent experiments. (A) Comparison of the MTase activities carried by the recombinant proteins using GpppAC₄ RNA as methyl acceptor. Shown is the enzymatic activity of ZIKV MTase domain on (B) small (Continued on next page)

Downloaded from http://jvi.asm.org/ on May 17, 2017 by Bibliotheque - Univ. de la Mediterranee

cap N-7-methylation does indeed require longer RNAs containing the *Flavivirus* conserved SLA structure.

The ZIKV MTase methylates single-strand RNA at the 2'-OH group of internal adenosines. Since the ZIKV MTase is active on double-methylated cap-1 RNAs (Fig. 3E), the methyl transfer on internal residues of RNAs was further investigated using uncapped RNAs in the presence of a 2-times-higher ZIKV MTase concentration than in the previous experiments. Figure 3F indicates that a 13-mer RNA bearing a 5'-triphosphate corresponding to the nascent 5' end of DENV RNA can be methylated, confirming that ZIKV MTase methylates uncapped RNA putatively on internal positions. In order to further characterize this activity, homopolymeric RNAs bearing a 5'-OH group were tested (Fig. 3F). ZIKV NS5-MTase can methylate a 5'-OH-A₂₇ RNA, but not 5'-OH-U₂₇, -G₂₇, and -C₂₇ RNAs, indicating that ZIKV MTase methylates adenosine residues. Since ZIKV harbors a 2'-O-MTase catalytic tetrad, we inferred that the MTase activity targets the 2'-O of the ribose of the adenosine residues. This activity was confirmed by demonstrating that an RNA previously methylated at all 2'-O positions (A_{m27}), is not methylated by the ZIKV MTase (Fig. 3G). In addition, the internal adenosine 2'-O-methylation does not specifically occur in the vicinity of the RNA 5' end since RNAs already methylated on the first (A_mA₂₆) or the two first (A_mA_mA₂₅) adenosines were methylated (Fig. 3G).

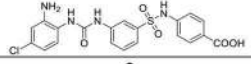
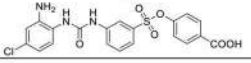
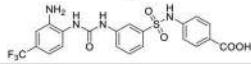
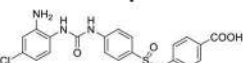
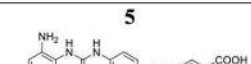
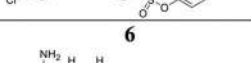
The 5' ends of *Flavivirus* mRNAs are organized in an SLA structure, and many adenosines are engaged in double-strand stretches. We thus investigated whether the internal methylation could be selective for single-strand RNA (ssRNA) or double-strand RNA (dsRNA). We compared the MTase activities on A₂₇ and on A₂₇ previously annealed with a complementary U₂₇ RNA. No MTase activity was detected using dsRNA (A₂₇ + U₂₇) (Fig. 3H) confirming that ZIKV MTase methylates preferentially ssRNA. In conclusion, the ZIKV MTase preferentially methylates the 2'-O position of internal adenosine residues that are not involved in dsRNA structures.

Selected DENV MTase inhibitors show increased potency on ZIKV MTase. Our structural and biochemical analyses indicate that the ZIKV MTase shares common structural and functional characteristics with DENV MTase (RMSD of 0.79 Å for PDB code 3P97 chain A). We thus tested whether some compounds already documented to inhibit DENV MTase might also be active against the ZIKV MTase. SAM analogues and cap analogues were tested as well as allosteric inhibitors designed from a fragment-based screening campaign performed against DENV MTase (31). For this purpose, both internal methylation and cap 2'-O-methylation were monitored on GpppAC₄ and A₂₇ RNA substrates, respectively, with increasing concentrations of the investigated inhibitors. We found that SAM analogues such as SAH and sinefungin inhibit cap-dependent ZIKV 2'-O-MTase activity with 50% inhibitory concentrations (IC₅₀) in the low micromolar range (Fig. 4A; Table 2). These compounds also abrogate internal methylation, but SAH shows a lower efficacy against internal methylation than sinefungin, which inhibits both methylations in a similar manner. Cap analogues (GpppA, ^mGpppA, GpppG, and ^mGpppG) bind to the DENV cap binding site by stacking interactions (34) and might therefore block RNA cap recognition, thus limiting cap methylation. This might also interfere with recognition and binding of uncapped RNA for internal methylation since the cap binding site might extend the RNA binding zone. GpppG and ^mGpppG show more potent inhibitions than GpppA and ^mGpppA on cap 2'-O-methylation, with IC₅₀ of 72 μM and 184 μM, respectively, versus 491 μM and 293 μM (Fig. 4B and C; Table 2). In addition, cap analogues inhibit internal methylations, to a lesser extent though, suggesting that the MTase cap binding site forms a part of the RNA binding zone but is not essential for internal methylation of an A₂₇ RNA.

FIG 3 Legend (Continued)

GpppAC₄ and ^mGpppAC₄ RNA, (C) 75 nucleotides of capped RNA with various 5'-end modifications (GpppA-RNA, ^mGpppA-RNA, GpppA_mA-RNA and ^mGpppA_mA-RNA), (D) different uncapped RNAs, (E) uncapped poly(A) oligomers with various 2'-O-methylated positions (A₂₇, no methylation; A_mA₂₆, first position 2'-O-methylated; A_mA_mA₂₅, first and second positions 2'-O-methylated; A_{m27}, all positions 2'-O-methylated), and (F) ssRNA A₂₇ and dsRNA A₂₇-U₂₇. CTL neg, enzymatic assay performed in the absence of RNA.

TABLE 2 Inhibition of the ZIKV 2'-O-MTase activity by various DENV MTase inhibitors^a

	ZIKV 2'-O-MTase activity IC ₅₀ (μM)		DENV 2'-O-MTase activity IC ₅₀ (μM)
	GpppAC ₄	A ₂₇	GpppAC ₄
SAH	0.43 ± 0.013	7.30 ± 0.44	0.34 ± 0.02
Sinefungin	1.18 ± 0.05	1.62 ± 0.08	0.63 ± 0.04
GpppA	491 ± 42	1082 ± 123	nd
mGpppA	293 ± 22	526 ± 47	nd
GpppG	72 ± 3.0	405 ± 22	nd
mGpppG	184 ± 8.0	351 ± 40	nd
1 	221 ± 19	409 ± 47	452 ± 38
2 	33.0 ± 2.7	138 ± 16	369 ± 14
3 	24.0 ± 1.2	79.0 ± 16	91.0 ± 8.0
4 	87.0 ± 9.7	168 ± 25	435 ± 26
5 	31.0 ± 6.7	127 ± 22	368 ± 11
6 	67.0 ± 5.1	230 ± 27	110 ± 11

^aShown is a summary of the IC₅₀ and the standard deviation (SD) deduced from Fig. 4 using the log (inhibitor) versus response variable slope equation and compared with the IC₅₀ previously published for DENV MTase. The IC₅₀ are an average from two to three independent measurements. nd, not determined.

Finally, we tested whether six compounds (1 to 6 in Table 2) targeting an allosteric site of DENV MTase (30, 31) could inhibit ZIKV MTase. All of the compounds selected on DENV MTase inhibit ZIKV cap 2'-O-MTase activity (Fig. 4D and Table 2). Compounds 1 to 6 exhibit lower IC₅₀ for ZIKV 2'-O-MTase activity than those for DENV MTase, suggesting that they are more potent on ZIKV MTase. In addition, all compounds inhibiting the cap-2'-O-MTase activity also inhibit the internal methylation, but with higher IC₅₀ (Table 2), indicating that internal and cap 2'-O-MTase activity use the same catalytic mechanism. These data also indicate that the structural and functional similarities between ZIKV and DENV MTase pave the way for the development of pan-viral inhibitors targeting the members of the *Flavivirus* genus.

DISCUSSION

The development of therapeutics against emerging arboviruses, such as ZIKV, is a major challenge. Acquisition of basic scientific knowledge on the emerging viral agent is often slow, whereas outbreaks can be sudden and unpredictable. In this work, we focus on the structure/function analysis of ZIKV NS5-MTase, closely related to that of DENV, and involved in the cap structure methylation. The latter is essential for viral RNA translation into proteins and the hiding of viral genomes from detection by cellular antiviral pathways (9, 15). The ZIKV MTase presented here crystallizes as a homodimer. Both protomers show a canonical MTase fold similar to those of DENV3 or WNV MTases, as well as those of other flaviviruses with SAM and cap binding sites close to the K-D-K-E catalytic tetrad. The first protomer of the ZIKV MTase crystallographic dimer

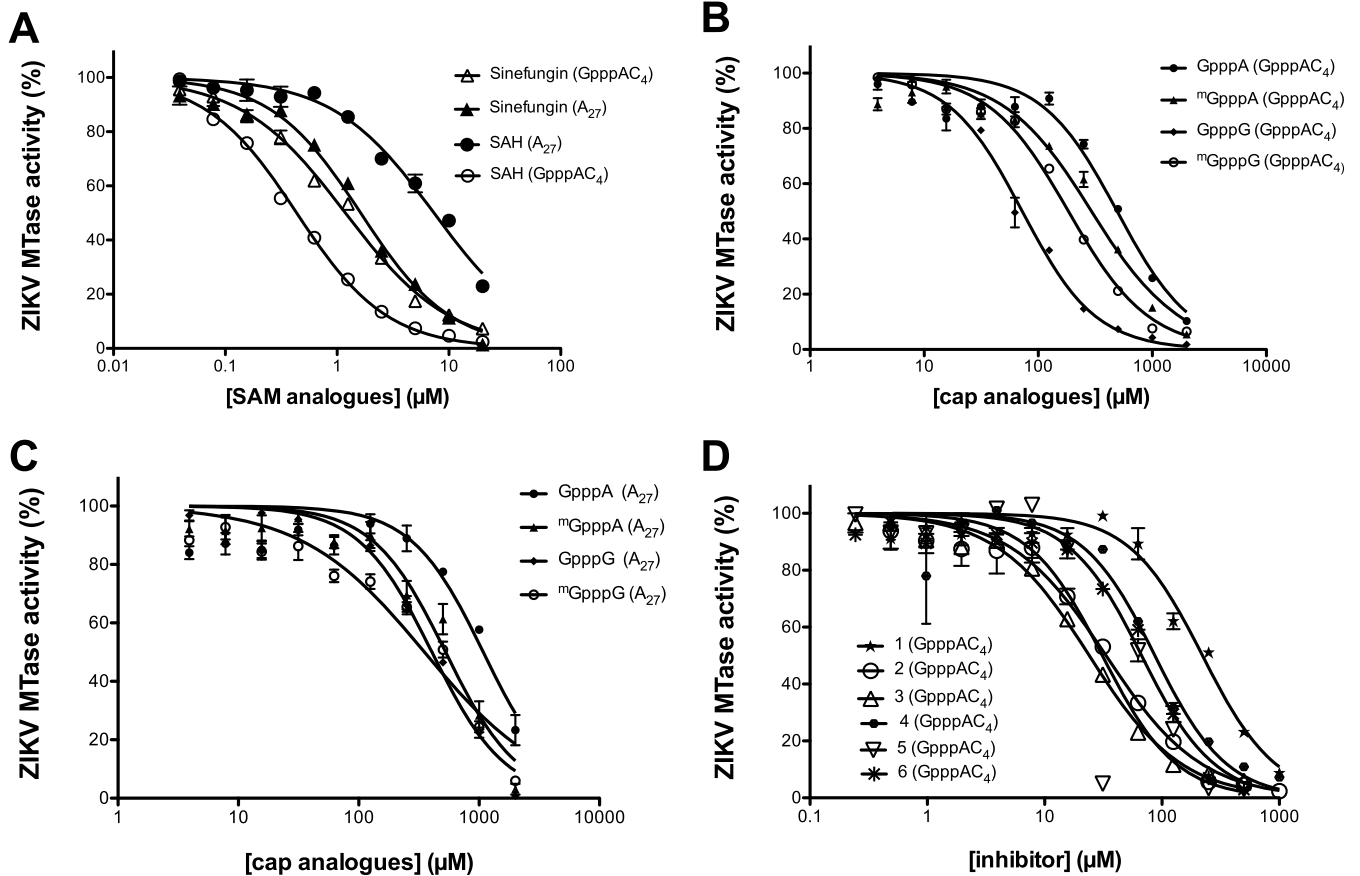


FIG 4 ZIKV MTase inhibition. The MTase assay was performed as indicated in Materials and Methods by incubating ZIKV NS5-MTase with GpppAC₄ (2'-O-cap methylation) or A₂₇₇ RNA (internal 2'-O-methylation). The methyl transfer onto RNA is measured by filter binding assay in the presence of increasing concentrations of inhibitors. The graph presents the mean value and standard deviation from 3 independent experiments. Shown is inhibition of cap 2'-O-MTase and internal methylation activity by (A) SAH and sinefungin, (B and C) cap analogues, and (D) compounds having allosteric inhibition (31).

contains the methyl donor SAM. It shows a high structural similarity to the ZIKV MTase structure recently determined by Coloma et al. (22). In this protomer, the SAM adenine base is held within a conserved hydrophobic pocket formed of β -strands β_1 , β_2 , and β_4 and α -helices α_X and α_A . The leaving methyl group of the SAM is closely positioned to the catalytic site. The structure reveals that an arginine cluster, localized at the RNA binding site, binds a sulfate molecule (Fig. 2D). The homologous arginine residues were already reported to play a key role in the recruitment of RNA (25). Especially, residues corresponding to ZIKV MTase Arg63 and Arg90 (Arg57 and Arg84 in DENV, respectively) bind to the phosphate groups of the RNA during the 2'-O-methylation (25). The sulfate molecule bound via a cluster formed by arginines Arg43, Arg47, Arg63, and Arg90 (Fig. 2D) is closely located in between two phosphodiester groups of bound RNA in the structure described by Zhao et al. (25) (namely P3 U and P4 U from chain B, with distances for S-P3 of 3 Å and for S-P4 of 4 Å). In this cluster formed by arginines Arg43, Arg47, Arg63, and Arg90, the main chain of ZIKV MTase (chain A) and DENV MTase (chain B [23]) structures overlap perfectly. Only slight conformational changes are noted for some side chains (mostly Arg63 and Arg90) due to the accommodation of bound RNA. This slight difference in positioning between sulfate and phosphate is probably due to the fact that sulfate can accommodate more freely than RNA. We thus speculate that the sulfate identified in the ZIKV protomer A mimics a phosphate group of viral RNA and holds the arginines in an RNA-bound conformation. In chain B, these arginines are not organized in a cluster and the sulfate ion is absent, possibly mimicking the RNA-free enzyme (Fig. 2D). In this conformation, the region encompassing residues 43 to 58 is disordered and the α_X helix is kinked, moving Arg63 away from the cluster. As

a consequence, Ser62 can no longer stabilize the SAM (or *S*-adenosyl-L-homocysteine [SAH]) molecule bound to the MTase. The structural differences between two protomers could thus provide possible snapshots of the ZIKV MTase after the recruitment of a capped RNA together with a SAM molecule (chain A) or after the release of the methylated capped RNA and the SAH molecule (chain B). This hypothesis would propose a key role for Arg63 in the enzymatic reaction turnover.

In addition, the structure highlights a hydrophobic cavity adjacent to the SAM binding site, which is conserved among *Flavivirus* MTases. This site can serve for structure-based drug design approaches. Indeed, this cavity can accommodate SAH derivatives substituted at the N6 position of the adenosine base, leading to inhibition of DENV MTase (27). The second protomer (chain B) presents significant differences. Instead a glycerol molecule and a sulfate ion were closely accommodated into the SAM binding site, providing a unique opportunity to serve as a basis for structure-based drug design and more especially for a “fragment-linking” strategy (35).

The functional characterization demonstrates that the ZIKV MTase harbors N-7 and 2'-*O*-methyltransferase activities contributing to the formation of a cap-1 structure on the viral mRNA. The interface between the MTase and the RdRp domains has been reported to be essential for the replication regulation of dengue virus (36–38) or Japanese encephalitis virus (39). It is also known that the MTase domain contributes to an efficient initiation and elongation of the RNA polymerization by the dengue virus RdRp (40). Conversely, in this study, we show that the 2'-*O*-MTase activity of ZIKV MTase on short capped RNAs is not activated by the RdRp domains. However, it is noteworthy that our biochemical assay might not be optimal to look for a possible interplay between the two domains as the RNA used by the MTase is not neo-synthesized by the RdRp. As observed for its close relatives DENV and WNV, the 2'-*O*-MTase activity can be detected using different short RNA substrates with no obvious sequence specificity (19, 33). In contrast the N-7-MTase activity requires an RNA substrate forming a hairpin structure mimicking the SLA structure conserved among flaviviruses (7, 20). We infer that both cap N-7- and 2'-*O*-MTase activities play an essential role in the virus life cycle. The N-7-methylation of the viral RNA cap by ZIKV NS5-MTase allows recognition by the translation initiation factor eIF4E and stimulation of translation into viral proteins (8). On the other hand, 2'-*O*-methylation of the cap is expected to mask the presence of exogenous viral RNAs from host cell sensors such as RIG-I and MDA5 (9–11), which induce the production of interferons. In addition, 2'-*O*-methylation of viral RNA is thought to prevent translation restriction by interferon-stimulated genes (ISGs) such as the IFIT-1 gene (12, 13). Thus, molecules inhibiting ZIKV MTase activities are expected to translate into a strong antiviral effect (41). Apart from the RNA-cap MTase activities, we observed that ZIKV NS5-MTase also methylates the 2'-*O* position of adenosine residues located at internal positions of the RNA substrate. This activity is mainly detected on adenosine residues using homopolymeric ssRNAs but not dsRNAs, which barely interact with the ZIKV NS5-MTase. Similar internal methylation on adenosine residues was already reported for DENV MTase (21), but the role of such methylations is still elusive.

The close structural and functional similarity of ZIKV MTase to other *Flavivirus* MTases suggests that work previously made on DENV MTase may help in the development of antiviral compounds targeting ZIKV MTase. To further assess this, we evaluated three classes of DENV MTase inhibitors, cap analogues, SAM analogues, and allosteric inhibitors on ZIKV MTase. Our results show that these molecules can serve as a starting point to identify MTase inhibitors. As expected, we observe that cap and SAM analogues, already described to inhibit viral MTases, show a similar inhibitory effect on ZIKV MTase. However, when targeting the SAM binding site, the main challenge remains compound selectivity to avoid inhibition of cellular MTases. In this respect, the structure of ZIKV MTase reveals that a specificity increase might be achievable using the conserved hydrophobic cavity adjacent to the SAM binding site, as in the case of DENV MTase (27). Another option to develop selective compounds is to target allosteric—i.e., noncatalytic—sites. We thus tested compounds previously discovered using DENV MTase and targeting an allosteric site structurally conserved between ZIKV and DENV

enzymes (31). Similarly, higher inhibition effects were observed on ZIKV MTase cap 2'-O-MTase activity.

In summary, we present here the first structure/function study of the ZIKV MTase. This study reveals the high similarities between DENV and ZIKV MTase structures. In addition we also describe an original structure of the ZIKV MTase in the absence of SAM showing conformational changes in the RNA binding groove together with the SAM binding pocket. The comparison of both structures might reveal the molecular basis driving the enzymatic turnover of flavivirus MTases. Finally, we provide a strong basis for the development of small molecules targeting MTase activities carried by ZIKV NS5. This study highlights the need to improve the basic science knowledge and antiviral research on virus families with high emerging potential.

MATERIALS AND METHODS

Plasmid constructs. For the functional study, the coding sequence of ZIKV NS5 (aa 4–903) and the ZIKV NS5 MTase (aa 4–278) domain were synthesized (Genscript) based on the sequence of the ZIKV strain H/PPF/2013 (GenBank accession no. [KJ776791.2](https://www.ncbi.nlm.nih.gov/nuccore/KJ776791.2)) and then cloned into a pQE30 (Qiagen) plasmid with an N-terminal His₆ tag. For the structural study, the coding sequence of ZIKV NS5 MTase (aa 1–264) was amplified by PCR starting from the synthetic gene and subcloned in pDEST14 (Thermo Fisher Scientific) with an N-terminal His₆ tag.

Expression and purification of NS5 and NS5-MTase domain proteins. The MTase domains (aa 4–278 and aa 1–264) were produced in *Escherichia coli* T7 Express Iq (New England BioLabs). Cells were grown in Terrific Broth at 37°C until the optical density at 600 nm (OD₆₀₀) reached 0.6. Protein expression was then induced by 0.5 mM IPTG (isopropyl-β-D-thiogalactopyranoside) at 17°C overnight. Bacteria were harvested by centrifugation. The bacterial pellets from a 2-liter bacterial culture were resuspended in 100 ml lysis buffer (50 mM Tris-HCl [pH 8], 300 mM NaCl, 5% glycerol, 0.1% Triton, 10 μg/ml DNase I, 2 tablets of EDTA-free antiprotease cocktail [Roche], 0.25 mg/ml lysozyme). After 30 min of incubation at 4°C, the cells were sonicated and clarified by centrifugation prior to immobilized metal affinity chromatography (IMAC) purification on a 5-ml His prep column (GE Healthcare), with elution in 50 mM Tris-HCl, 300 mM NaCl, and 250 mM imidazole (pH 8.0). The eluted protein was then loaded on a 16/60 Superdex 200 (GE Healthcare) equilibrated in a mixture of 10 mM HEPES, 500 mM NaCl, glycerol 5%, and 1 mM dithiothreitol (DTT [pH 7.5]).

The protocol for the production and purification of the ZIKV NS5 was adapted from the protocol developed for DENV NS5 (40). After the IMAC purification, a size exclusion chromatography (SEC) step was applied to separate the cleaved MTase and the full-length NS5 using a Superdex S75 HR 16/20 column (GE Healthcare) pre-equilibrated in a mixture of 50 mM HEPES (pH 7.5), 750 mM NaCl, 10% glycerol, and 10 mM DTT. Proteins were concentrated and stored at –20°C after adding glycerol to a final concentration of 40%.

Radioactive methyltransferase assay. The enzymatic assays were carried out in 40 mM Tris-HCl (pH 8.0), 1 mM DTT, 2 μM SAM, and 0.33 μM [³H]SAM (PerkinElmer) in the presence of 0.7 μM synthetic RNAs with various 5'-end modifications (triphosphorylated, pppRNA; unmethylated cap, GpppARN; cap-0, mGpppARN, or 2'-O-methylated cap, GpppG_mRNA). Purified ZIKV NS5 and NS5-MTase were added to final concentrations of 0.75 μM, except for the internal methylation assay, which was performed in the presence of 1.5 μM MTase.

Reaction mixtures were incubated at 30°C and stopped after 30 min except for the time course experiments by a 10-fold dilution of the reaction mixture in 20 μM ice-cold SAH. Samples were then transferred to DEAE cellulose filters (PerkinElmer) by using a Filtermat Harvester apparatus (Packard Instruments). The unincorporated [³H]SAM was removed from the filter by several washing steps with 10 mM ammonium formate (pH 8.0), with H₂O, and with absolute ethanol, before drying the DEAE filters. The filters transferred into plastic bags, BetaplateScint (Wallac) scintillation fluid was added before quantification of [³H]methyl groups transferred onto RNA substrates in counts per minute using a Wallac 1450 MicroBetaTriLux liquid scintillation counter.

For the MTase inhibition assays, 0.75 μM ZIKV NS5-MTase was mixed with the inhibitor candidate before the addition of SAM and RNA substrate to start the reaction. The final concentration of dimethyl sulfoxide (DMSO) in the reaction mixtures was below 5%, and control reactions were performed in the presence of similar DMSO concentrations. Reaction mixtures were incubated at 30°C for 30 min and analyzed by filter binding assay as described above.

The IC₅₀s of SAH, sinefungin, cap analogues, and compounds 1 to 6 were determined with GraphPad Prism using the log (inhibitor) versus response variable slope equation.

Synthesis of RNA substrates. Various short RNA sequences were chemically synthesized on a solid support using an ABI 394 synthesizer. After RNA elongation with 2'-O-pivaloyloxymethyl phosphoramidite monomers (42) (Chemgenes, USA), the 5'-hydroxyl group of RNA still anchored to the solid support was phosphorylated, and the resulting *H*-phosphonate derivative was oxidized and activated into a phosphoroimidazolide derivative to react with either pyrophosphate to give a 5'-triphosphate RNA (pppRNA) (43) or GDP to give GpppRNA (44). After deprotection and release from the solid support, the GpppRNA sequences were purified by ion-exchange–high-performance liquid chromatography (IEX-HPLC) and validated to be >95% pure by IEX-HPLC analysis, and they were unambiguously characterized by matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometry. Subsequent N-7-methylation of the purified GpppRNAs was performed enzymatically using N-7-hMTase (44).

Crystallization conditions. The ZIKV MTase (aa 1–264) was concentrated to 5 mg/ml for crystallization screenings in sitting drops. After 2 weeks, crystals were obtained and optimized in 150 mM sodium citrate (pH 5.6) and 1.5 M ammonium sulfate. Crystals were transferred to a cryoprotectant solution of the same reservoir solution containing 10% glycerol for a few seconds and flash-cooled in liquid nitrogen.

Data collection, automated data processing, and refinement. Diffraction data were collected at the SOLEIL synchrotron (Proxima 1). We used AutoPROC (45) for data processing. Images were processed and scaled using XDS (46) and Aimless (47). Phaser (48) was used for molecular replacement. The structure of DENV3 NS5-MTase/AdoMet (PDB code 3P97) was used as the template model. The first ZIKV MTase model was built using ARP/wARP (49), and autoBUSTER (50) was used for structure refinement. Iterative cycles of model building were done with the program COOT (51). The quality of the refined structures was assessed with MOLPROBITY (52). The crystal structure was determined at a resolution of 2.01 Å and belonged to space group P1 with the following unit cell parameters: $a = 37.6$ Å, $b = 64.1$ Å, and $c = 72.0$ Å and $\alpha = 113.1^\circ$, $\beta = 97.8^\circ$, and $\gamma = 92.0^\circ$. Gesamt (53) was used to superimpose the obtained model with DENV3 MTase (PDB code 3P97), ZIKV MTase (PDB code 5KQR), and MODV MTase (PDB code no. 2WA1) and to determine RMSD values. Figures were generated with Chimera.

Accession number(s). The crystal structure of the Zika virus MTase construct generated in this study was deposited in the Protein Data Bank (PDB) under code 5M5B.

ACKNOWLEDGMENTS

We are grateful to Corinne Sallamand, Alexandre Blanjoie, and Théo Guez for technical assistance with the chemical synthesis and purification of the short capped RNA oligonucleotides.

This work was supported by the European program H2020 under the ZIKAlliance project (grant agreement 734548) and the EVAg Research Infrastructure (grant agreement 653316) and by the French research agency ANR (VMTaseIn, grant ANR-ST14-ASTR-0026, and FragVir, grant ANR-13-JS07-0006-01). B.M. is the recipient of a DGA-MRIS scholarship and W.A. the recipient of a scholarship from the “Méditerranée Infection” Foundation.

REFERENCES

- Walker WL, Lindsey NP, Lehman JA, Krow-Lucal ER, Rabe IB, Hills SL, Martin SW, Fischer M, Staples JE. 2016. Zika virus disease cases—50 States and the District of Columbia, January 1–July 31, 2016. *MMWR Morb Mortal Wkly Rep* 65:983–986. <https://doi.org/10.15585/mmwr.mm6536e5>.
- Musso D, Roche C, Robin E, Nhan T, Teissier A, Cao-Lorme VM. 2015. Potential sexual transmission of Zika virus. *Emerg Infect Dis* 21:359–361. <https://doi.org/10.3201/eid2102.141363>.
- Ventura CV, Maia M, Bravo-Filho V, Góis AL, Belfort R. 2016. Zika virus in Brazil and macular atrophy in a child with microcephaly. *Lancet* 387:228. [https://doi.org/10.1016/S0140-6736\(16\)00006-4](https://doi.org/10.1016/S0140-6736(16)00006-4).
- Cao-Lormeau V-M, Blake A, Mons S, Lastere S, Roche C, Vanhomwegen J, Dub T, Baudouin L, Teissier A, Larre P, Vial A-L, Decam C, Choumet V, Halstead SK, Willison HJ, Musset L, Manuguerra J-C, Despres P, Fournier E, Mallet H-P, Musso D, Fontanet A, Neil J, Ghawche F. 2016. Guillain-Barre syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet* 387:1531–1539. [https://doi.org/10.1016/S0140-6736\(16\)00562-6](https://doi.org/10.1016/S0140-6736(16)00562-6).
- Decroly E, Ferron F, Lescar J, Canard B. 2011. Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol* 10: 51–65. <https://doi.org/10.1038/nrmicro2675>.
- Egloff MP, Benarroch D, Selisko B, Romette JL, Canard B. 2002. An RNA cap (nucleoside-2'-O)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. *EMBO J* 21: 2757–2768. <https://doi.org/10.1093/emboj/21.11.2757>.
- Zhou Y, Ray D, Zhao Y, Dong H, Ren S, Li Z, Guo Y, Bernard KA, Shi P-Y, Li H. 2007. Structure and function of flavivirus NS5 methyltransferase. *J Virol* 81:3891–3903. <https://doi.org/10.1128/JVI.02704-06>.
- Muthukrishnan S, Both GW, Furuichi Y, Shatkin AJ. 1975. 5'-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation. *Nature* 255:33–37. <https://doi.org/10.1038/255033a0>.
- Züst R, Cervantes-Barragan L, Habjan M, Maier R, Neuman BW, Ziebuhr J, Szretter KJ, Baker SC, Barchet W, Diamond MS, Siddell SG, Ludewig B, Thiel V. 2011. Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat Immunol* 12:137–143. <https://doi.org/10.1038/ni.1979>.
- Schuberth-Wagner C, Ludwig J, Bruder AK, Herzner AM, Zillinger T, Goldeck M, Schmidt T, Schmid-Burgk JL, Kerber R, Wolter S, Stümpel JP, Roth A, Bartok E, Drosten C, Coch C, Hornung V, Barchet W, Kümmerer BM, Hartmann G, Schlee M. 2015. A conserved histidine in the RNA sensor RIG-I controls immune tolerance to 2'-O-methylated self RNA. *Immunity* 43:41–52. <https://doi.org/10.1016/j.immuni.2015.06.015>.
- Devarkar SC, Wang C, Miller MT, Ramanathan A, Jiang F, Khan AG, Patel SS, 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 U S A* 113:596–601. <https://doi.org/10.1073/pnas.1515152113>.
- Daffis S, Szretter KJ, Schriewer J, Li J, Youn S, Errett J, Lin T-Y, Schneller S, Züst R, Dong H, Thiel V, Sen GC, Fensterl V, Klimstra WB, Pierson TC, Buller RM, Gale M, Shi P-Y, Diamond MS. 2010. 2'-O-methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* 468:452–456. <https://doi.org/10.1038/nature09489>.
- Kumar P, Sweeney TR, Skabkin MA, Skabkina OV, Hellen CUT, Pestova TV. 2014. 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* 42:3228–3245. <https://doi.org/10.1093/nar/gkt1321>.
- Dong H, Chang DC, Xie X, Toh YX, Chung KY, Zou G, Lescar J, Lim SP, Shi PY. 2010. Biochemical and genetic characterization of dengue virus methyltransferase. *Virology* 405:568–578. <https://doi.org/10.1016/j.virol.2010.06.039>.
- Li S-H, Dong H, Li X-F, Xie X, Zhao H, Deng Y-Q, Wang X-Y, Ye Q, Zhu S-Y, Wang H-J, Zhang B, Leng Q-B, Zuest R, Qin E-D, Qin C-F, Shi P-Y. 2013. Rational design of a flavivirus vaccine by abolishing viral RNA 2'-O methylation. *J Virol* 87:5812–5819. <https://doi.org/10.1128/JVI.02806-12>.
- Züst R, Dong H, Li X-F, Chang DC, Zhang B, Balakrishnan T, Toh Y-X, Jiang T, Li S-H, Deng Y-Q, Ellis BR, Ellis EM, Poidinger M, Zolezzi F, Qin C-F, Shi P-Y, Fink K. 2013. Rational design of a live attenuated dengue vaccine: 2'-O-methyltransferase mutants are highly attenuated and immunogenic in mice and macaques. *PLoS Pathog* 9:e1003521. <https://doi.org/10.1371/journal.ppat.1003521>.
- Ferron F, Decroly E, Selisko B, Canard B. 2012. The viral RNA capping machinery as a target for antiviral drugs. *Antiviral Res* 96:21–31. <https://doi.org/10.1016/j.antiviral.2012.07.007>.
- Lim SP, Noble CG, Shi P-Y. 2015. The dengue virus NS5 protein as a target

- for drug discovery. *Antiviral Res* 119:57–67. <https://doi.org/10.1016/j.antiviral.2015.04.010>.
19. Ray D, Shah A, Tilgner M, Guo Y, Zhao Y, Dong H, Deas TS, Zhou Y, Li H, Shi P-Y. 2006. West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O methylations by nonstructural protein 5. *J Virol* 80:8362–8370. <https://doi.org/10.1128/JVI.00814-06>.
 20. Barral K, Sallamand C, Petzold C, Coutard B, Collet A, Thillier Y, Zimmermann J, Vasseur JJ, Canard B, Rohayem J, Debart F, Decroly E. 2013. Development of specific dengue virus 2'-O- and N7-methyltransferase assays for antiviral drug screening. *Antiviral Res* 99:292–300. <https://doi.org/10.1016/j.antiviral.2013.06.001>.
 21. Dong H, Chang DC, Hua MHC, Lim SP, Chionh YH, Hia F, Lee YH, Kukkaro P, Lok S-M, Dedon PC, Shi P-Y. 2012. 2'-O methylation of internal adenosine by flavivirus NS5 methyltransferase. *PLoS Pathog* 8:e1002642. <https://doi.org/10.1371/journal.ppat.1002642>.
 22. Coloma J, Jain R, Rajashankar KR, Garcia-Sastre A, Aggarwal AK. 2016. Structures of NS5 methyltransferase from Zika virus. *Cell Rep* 16:3097–3102. <https://doi.org/10.1016/j.celrep.2016.08.091>.
 23. Bollati M, Alvarez K, Assenberg R, Baronti C, Canard B, Cook S, Coutard B, Decroly E, de Lamballerie X, Gould EA, Grard G, Grimes JM, Hilgenfeld R, Jansson AM, Malet H, Mancini EJ, Mastrangelo E, Mattevi A, Milani M, Moureau G, Neyts J, Owens RJ, Ren J, Selisko B, Speroni S, Steuber H, Stuart DI, Unge T, Bolognesi M. 2010. Structure and functionality in flavivirus NS-proteins: perspectives for drug design. *Antiviral Res* 87:125–148. <https://doi.org/10.1016/j.antiviral.2009.11.009>.
 24. Yap LJ, Luo D, Chung KY, Lim SP, Bodenreider C, Noble C, Shi P-Y, Lescar J. 2010. Crystal structure of the dengue virus methyltransferase bound to a 5'-capped octameric RNA. *PLoS One* 5:e12836. <https://doi.org/10.1371/journal.pone.0012836>.
 25. Zhao Y, Soh TS, Lim SP, Chung KY, Swaminathan K, Vasudevan SG, Shi P-Y, Lescar J, Luo D. 2015. Molecular basis for specific viral RNA recognition and 2'-O-ribose methylation by the dengue virus nonstructural protein 5 (NS5). *Proc Natl Acad Sci U S A* 112:14834–14839. <https://doi.org/10.1073/pnas.1514978112>.
 26. Milani M, Mastrangelo E, Bollati M, Selisko B, Decroly E, Bouvet M, Canard B, Bolognesi M. 2009. Flaviviral methyltransferase/RNA interaction: structural basis for enzyme inhibition. *Antiviral Res* 83:28–34. <https://doi.org/10.1016/j.antiviral.2009.03.001>.
 27. Lim SP, Sonntag LS, Noble C, Nilar SH, Ng RH, Zou G, Monaghan P, Chung KY, Dong H, Liu B, Bodenreider C, Lee G, Ding M, Chan WL, Wang G, Jian YL, Chao AT, Lescar J, Yin Z, Vedananda TR, Keller TH, Shi PY. 2011. Small molecule inhibitors that selectively block dengue virus methyltransferase. *J Biol Chem* 286:6233–6240. <https://doi.org/10.1074/jbc.M110.179184>.
 28. Brecher M, Chen H, Liu B, Banavali NK, Jones SA, Zhang J, Li Z, Kramer LD, Li H. 2015. Novel broad spectrum inhibitors targeting the flavivirus methyltransferase. *PLoS One* 10:e0130062. <https://doi.org/10.1371/journal.pone.0130062>.
 29. Idrus S, Tambunan U, Zubaidi AA. 2012. Designing cyclopentapeptide inhibitor as potential antiviral drug for dengue virus ns5 methyltransferase. *Bioinformation* 8:348–352. <https://doi.org/10.6026/97320630008348>.
 30. Coutard B, Decroly E, Li C, Sharff A, Lescar J, Bricogne G, Barral K. 2014. Assessment of dengue virus helicase and methyltransferase as targets for fragment-based drug discovery. *Antiviral Res* 106:61–70. <https://doi.org/10.1016/j.antiviral.2014.03.013>.
 31. Benmansour F, Trist I, Coutard B, Decroly E, Querat G, Brancale A, Barral K. 2017. Discovery of novel dengue virus NS5 methyltransferase non-nucleoside inhibitors by fragment-based drug design. *Eur J Med Chem* 125:865–880. <https://doi.org/10.1016/j.ejmech.2016.10.007>.
 32. Jansson AM, Jakobsson E, Johansson P, Lantze V, Coutard B, De Lamballerie X, Unge T, Jones TA. 2009. Structure of the methyltransferase domain from the Modoc virus, a flavivirus with no known vector. *Acta Crystallogr Sect D Biol Crystallogr* 65:796–803. <https://doi.org/10.1107/S0907444909017260>.
 33. Selisko B, Peyrane FF, Canard B, Alvarez K, Decroly E. 2010. Biochemical characterization of the (nucleoside-2'-O)-methyltransferase activity of dengue virus protein NS5 using purified capped RNA oligonucleotides 7MeGpppACn and GpppACn. *J Gen Virol* 91:112–121. <https://doi.org/10.1099/vir.0.015511-0>.
 34. Egloff MP, Decroly E, Malet H, Selisko B, Benarroch D, Ferron F, Canard B. 2007. 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* 372:723–736. <https://doi.org/10.1016/j.jmb.2007.07.005>.
 35. Erlanson DA, Fesik SW, Hubbard RE, Jahnke W, Jhoti H. 2016. Twenty years on: the impact of fragments on drug discovery. *Nat Rev Drug Discov* 15:605–619. <https://doi.org/10.1038/nrd.2016.109>.
 36. Klema VJ, Ye M, Hindupur A, Teramoto T, Gottipati K, Padmanabhan R, Choi KH. 2016. Dengue virus nonstructural protein 5 (NS5) assembles into a dimer with a unique methyltransferase and polymerase interface. *PLoS Pathog* 12:e1005451. <https://doi.org/10.1371/journal.ppat.1005451>.
 37. Zhao Y, Soh TS, Zheng J, Chan KWK, Phoo WW, Lee CC, Tay MYF, Swaminathan K, Cornvik TC, Lim SP, Shi PY, Lescar J, Vasudevan SG, Luo D. 2015. A crystal structure of the dengue virus NS5 protein reveals a novel inter-domain interface essential for protein flexibility and virus replication. *PLoS Pathog* 11:e1004682. <https://doi.org/10.1371/journal.ppat.1004682>.
 38. Zhao Y, Soh TS, Ki Chan KW, Yin Fung SS, Swaminathan K, Lim SP, Shi P-Y, Huber T, Lescar J, Luo D, Vasudevan SG. 2015. Flexibility of NS5 methyltransferase-polymerase linker region is essential for dengue virus replication. *J Virol* 89:10717–10721. <https://doi.org/10.1128/JVI.01239-15>.
 39. Lu G, Gong P. 2013. Crystal structure of the full-length Japanese encephalitis virus NS5 reveals a conserved methyltransferase-polymerase interface. *PLoS Pathog* 9:e1003549. <https://doi.org/10.1371/journal.ppat.1003549>.
 40. Potosopon S, Priet S, Collet A, Decroly E, Canard B, Selisko B. 2014. The methyltransferase domain of dengue virus protein NS5 ensures efficient RNA synthesis initiation and elongation by the polymerase domain. *Nucleic Acids Res* 42:11642–11656. <https://doi.org/10.1093/nar/gku666>.
 41. Brecher M, Chen H, Li Z, Banavali NK, Jones SA, Zhang J, Kramer LD, Li H. 2015. Identification and characterization of novel broad-spectrum inhibitors of the flavivirus methyltransferase. *ACS Infect Dis* 1:340–349. <https://doi.org/10.1021/acsinfecdis.5b00070>.
 42. Lavergne T, Bertrand JR, Vasseur JJ, Debart F. 2008. A base-labile group for 2'-OH protection of ribonucleosides: a major challenge for RNA synthesis. *Chemistry* 14:9135–9138. <https://doi.org/10.1002/chem.200801392>.
 43. Zlatev I, Lavergne T, Debart F, Vasseur JJ, Manoharan M, Morvan F. 2010. Efficient solid-phase chemical synthesis of 5'-triphosphates of DNA, RNA, and their analogues. *Org Lett* 12:2190–2193. <https://doi.org/10.1021/ol1004214>.
 44. Thillier Y, Decroly E, Morvan F, Canard B, Vasseur J-J, Debart F. 2012. Synthesis of 5' cap-0 and cap-1 RNAs using solid-phase chemistry coupled with enzymatic methylation by human (guanine-N7)-methyltransferase. *RNA* 18:856–868. <https://doi.org/10.1261/rna.030932.111>.
 45. Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, Womack T, Bricogne G. 2011. Data processing and analysis with the autoPROC toolbox. *Acta Crystallogr Sect D Biol Crystallogr* 67:293–302. <https://doi.org/10.1107/S0907444911007773>.
 46. Kabsch W. 2010. Xds. *Acta Crystallogr Sect D Biol Crystallogr* 66:125–132. <https://doi.org/10.1107/S0907444909047337>.
 47. Evans PR, Murshudov GN. 2013. How good are my data and what is the resolution? *Acta Crystallogr Sect D Biol Crystallogr* 69:1204–1214. <https://doi.org/10.1107/S0907444913000061>.
 48. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* 40:658–674. <https://doi.org/10.1107/S0021889807021206>.
 49. Langer G, Cohen SX, Lamzin VS, Perrakis A. 2008. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3:1171–1179. <https://doi.org/10.1038/nprot.2008.91>.
 50. Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart OS, Vonrhein C, Womack TO. 2011. BUSTER version 2.11.2. Global Phasing Ltd, Cambridge, United Kingdom.
 51. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr Sect D Biol Crystallogr* 66:486–501. <https://doi.org/10.1107/S0907444910007493>.
 52. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, Snoeyink J, Richardson JS, Richardson DC. 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35(Web Server issue):W375–W383.
 53. Krissinel E. 2012. Enhanced fold recognition using efficient short fragment clustering. *J Mol Biochem* 1:76–85.

ANNEXE 6 – Article 2

Titre : Toward the identification of viral cap-methyltransferase inhibitors by fluorescence screening assay.

Auteurs : Aouadi W, Eydoux C, Coutard B, Martin B, Debart F, Vasseur JJ, Contreras JM, Morice C, Quérat G, Jung ML, Canard B, Guillemot JC, Decroly E.

Référence : Antiviral Res. 2017 Aug;144:330-339. doi: 10.1016/j.antiviral.2017.06.021. Epub 2017 Jul 1.

Deux coronavirus (CoV) hautement pathogènes associés à un syndrome respiratoire sévère ont émergé au début du siècle. Le CoV responsable du syndrome respiratoire aigu sévère (SARS-CoV) s'est propagé dans le sud de la Chine en 2003 faisant 8000 cas en quelques mois. En 2012, le CoV responsable du syndrome respiratoire du Moyen Orient (MERS-CoV) a émergé dans la Péninsule Arabique et l'épidémie est toujours en cours avec un fort taux de mortalité. Les CoVs sont depuis considérés comme une menace sanitaire importante, notamment du fait qu'il n'existe actuellement ni vaccin ni traitement contre ces virus. C'est pourquoi de nouvelles molécules inhibitrices doivent être identifiées afin de développer des traitements antiviraux limitant la réplication des CoVs. Dans cette étude, nous nous sommes concentrés sur la protéine nsp14 qui joue un rôle clé dans la réplication virale puisque cette protéine méthyle la coiffe de l'ARN en position N7 de la guanine (N7 méthyltransférase, N7 MTase). Nous avons développé un dosage enzymatique à haut débit de l'activité N7 MTase basé sur le principe HTRF® (Homogenous Time Resolved Fluorescence) afin de cribler des banques de composés (2000 composés au total) sur nsp14 de SARS-CoV. L'inhibition de vingt composés intéressants contre l'activité nsp14 de SARS-CoV a été évaluée grâce à la détermination de leur IC50 et leur spécificité a été analysée en comparant les résultats obtenus sur les activités MTases de flavivirus et de la N7 MTase humaine. Nos résultats ont révélé trois classes de composés : 1) des molécules non spécifiques inhibant différentes MTases ainsi que la polymérase du virus de la dengue, 2) des molécules ciblant les MTases cellulaires et virales, 3) des inhibiteurs visant plus spécifiquement une ou plusieurs MTase(s) virale(s) mais également de la N7 MTase humaine. Ces composés offrent une première base moléculaire vers le développement d'inhibiteurs spécifiques des MTases virales.

Cette étude a été réalisée dans le cadre du projet ANR/DGA Astrid VMTaseIn (ANR-ST14-ASTR-0026) dans lequel j'ai été impliqué pour la caractérisation de composés sur l'activité MTase de SUDV. J'ai participé aux réunions de travail afin de réfléchir sur les résultats obtenus sur les différents virus et aux discussions concernant cette étude en particulier. J'ai également participé au processus de révision du manuscrit.



Toward the identification of viral cap-methyltransferase inhibitors by fluorescence screening assay



Wahiba Aouadi ^{a,1}, Cécilia Eydoux ^{a,1}, Bruno Coutard ^a, Baptiste Martin ^a,
 Françoise Debart ^b, Jean Jacques Vasseur ^b, Jean Marie Contreras ^c, Christophe Morice ^c,
 Gilles Quérat ^d, Marie-Louise Jung ^c, Bruno Canard ^a, Jean-Claude Guillemot ^{a,1},
 Etienne Decroly ^{a,*}

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

^b IBMM, CNRS, Université Montpellier, ENSCM, Campus Triolet, Place E. Bataillon, 34095, Montpellier Cedex 05, France

^c Prestwick Chemical, 67400, Illkirch, Strasbourg, France

^d UMR "Emergence des Pathologies Virales" (EPV: Aix-Marseille Université - IRD 190 - Inserm 1207 - EHESP), Marseille, France

ARTICLE INFO

Article history:

Received 22 May 2017

Received in revised form

28 June 2017

Accepted 29 June 2017

Available online 1 July 2017

Keywords:

Methyltransferase

Coronavirus

Flavivirus

Inhibitor

HTRF

Antiviral

ABSTRACT

Two highly pathogenic human coronaviruses associated with severe respiratory syndromes emerged since the beginning of the century. The severe acute respiratory syndrome SARS-coronavirus (CoV) spread first in southern China in 2003 with about 8000 infected cases in few months. Then in 2012, the Middle East respiratory syndrome (MERS-CoV) emerged from the Arabian Peninsula giving a still ongoing epidemic associated to a high fatality rate. CoVs are thus considered a major health threat. This is especially true as no vaccine nor specific therapeutic are available against either SARS- or MERS-CoV. Therefore, new drugs need to be identified in order to develop antiviral treatments limiting CoV replication. In this study, we focus on the nsp14 protein, which plays a key role in virus replication as it methylates the RNA cap structure at the N7 position of the guanine. We developed a high-throughput N7-MTase assay based on Homogenous Time Resolved Fluorescence (HTRF[®]) and screened chemical libraries (2000 compounds) on the SARS-CoV nsp14. 20 compounds inhibiting the SARS-CoV nsp14 were further evaluated by IC₅₀ determination and their specificity was assessed toward flavivirus- and human cap N7-MTases. Our results reveal three classes of compounds: 1) molecules inhibiting several MTases as well as the dengue virus polymerase activity unspecifically, 2) pan MTases inhibitors targeting both viral and cellular MTases, and 3) inhibitors targeting one viral MTase more specifically showing however activity against the human cap N7-MTase. These compounds provide a first basis towards the development of more specific inhibitors of viral methyltransferases.

© 2017 Published by Elsevier B.V.

1. Introduction

The severe acute respiratory syndrome coronavirus (SARS-CoV) spread in China in 2003 and was responsible for worldwide outbreak causing over 8000 infected people with a fatality rate around 10% (de Wit et al., 2016). In 2012, a novel human coronavirus named Middle East Respiratory Syndrome Coronavirus (MERS-CoV) emerged in the Arabian Peninsula with secondary infection cases reported in Europe, America, Asia and Africa. MERS-

CoV infection resulted in more than 1728 confirmed infected patients till April 2016 with a fatality rate of 36% (de Wit et al., 2016).

CoVs are enveloped viruses possessing a single stranded positive sense RNA genome of approximately 29.7 kb in length (Pan et al., 2008; Van Boheemen et al., 2012). The SARS-CoV initiates its replication cycle after binding to the angiotensin-converting enzyme 2 (ACE2) (Kuhn et al., 2004). After endocytosis, the viral RNA genome is released into the cellular host cytoplasm and translated into two large polyproteins (de Wit et al., 2016; Snijder et al., 2016). The pp1a and pp1ab polyproteins are next cleaved by viral proteases into 11 and 16 non-structural proteins respectively (nsp1 to 16). These proteins form a large replication transcription complex (RTC) which is associated to host cellular

* Corresponding author.

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

¹ These authors contributed equally to this work.

proteins (Van Hemert et al., 2008). CoV RTC ensures the replication of the viral genome and the transcription of genomic and sub-genomic mRNA. These viral RNAs are polyadenylated at their 3' end and protected by a 5' cap structure, which is synthesized by viral enzymes. The CoV capping pathway is thought to follow four sequential steps implicating several nsps: I) the 5'-3' helicase/NTPase nsp13 hydrolyses first the phosphate γ from the nascent 5'-triphosphorylated RNA (Ivanov and Ziebuhr, 2004); II) a GMP molecule is supposedly transferred to the diphosphorylated RNA by a still unknown guanylyltransferase forming a cap (GpppN) structure; III) the cap is then methylated at N7 position of the guanosine by nsp14 in the presence of methyl donor S-adenosyl-L-methionine (SAM) yielding to a cap-0 ($^7\text{mGpppN}$) and S-adenosyl-L-homocysteine (SAH) by-product (Chen et al., 2009); IV) nsp10/nsp16 complex methylates at the ribose 2'OH group of the first transcribed RNA leading to the conversion of the cap-0 ($^7\text{mGpppN}$) into the cap-1 ($^7\text{mGpppN}_{2\text{om}}$) (Bouvet et al., 2010; Chen et al., 2011). *In vitro* assays have deciphered the mechanisms driving the RNA cap methylation in SARS- and MERS-CoV. It follows an obligatory order in which N7-methylation by nsp14 is a pre-requisite for 2'O-methylation by the nsp10/nsp16 complex (Aouadi et al., 2017; Bouvet et al., 2014, 2010).

The guanine N7-MTase activity embedded in the C-terminal domain of SARS-nsp14 has been discovered by yeast trans-complementation assay (Chen et al., 2009). In addition, the N-terminus moiety of nsp14 contains a DEDDh exonuclease (ExoN) domain (Minskaia et al., 2006). The two domains communicate functionally, as truncation experiments showed that the N-terminal region of nsp14 is required for the N7-MTase activity (Chen et al., 2009). Both N7-MTase and ExoN activities have been confirmed by *in vitro* assay showing that the association of nsp10 to nsp14 stimulated >35 fold the ExoN activity while the N7-MTase activity does not depend on the nsp10-nsp14 interaction (Bouvet et al., 2012, 2010; Decroly et al., 2011).

The N7- and 2'O- methylations of the viral mRNA cap are key events for the viral infection. Indeed reverse genetic experiments revealed first that the N7- methylation of cap structures is essential for the synthesis of viral proteins (Case et al., 2016). This observation is corroborated by former biochemical data showing that the N7-methyl guanosine of cap structures is recognized by the eukaryotic translation initiation factor 4E (eIF4E) and participates in the initiation of viral mRNA translation into proteins (Case et al., 2016; Cougot et al., 2004). Accordingly, inhibitors blocking nsp14 N7-MTase activity have been identified by yeast based screening assay on SARS-CoV, and induced a potent antiviral effect demonstrating that nsp14 MTase activity is an attractive antiviral target (Sun et al., 2014). Whereas N7-MTase mutants are replication defective, 2'O-MTase mutants show limited effect on virus replication in cell culture but have an attenuated phenotype in animal models (Li et al., 2013; Menachery et al., 2014; Zhang et al., 2014; Züst et al., 2013). The molecular basis of this attenuated phenotype was recently elucidated: incompletely-capped RNAs have been shown to be recognized by immune sensors such as RIG-I and MDA-5, which trigger innate immunity pathways (Decroly et al., 2012; Schuberth-Wagner et al., 2015; Wu et al., 2013). In turn, RIG-I or MDA-5 induces signalling cascades yielding to the expression of cytokines and type I interferon inducing an antiviral state in neighboring cells. Among the interferon-stimulated genes (ISG), IFIT1 also participates to the restriction of viral replication by sequestering mis-capped viral RNAs (Pichlmair et al., 2011). Thus cap structure is now considered as a kind of "marker of self" and it is currently admitted that 2'O-MTase inhibitors might help virus clearance by stimulation of the immune response (Decroly et al., 2012; Ferron et al., 2012; Züst et al., 2011).

In this work we first developed an HTRF MTase assay in order to

identify compounds inhibiting the N7-MTase activity of SARS-CoV nsp14. Using this system, we screened a library composed of 2000 compounds containing 1280 FDA approved molecules (Prestwick Chemical Library[®]), 320 natural products and 400 pyridazine-derived compounds. The inhibitory effect of the 20 best compounds was confirmed by a radioactive filter-binding assay, and refined by IC₅₀ values determination on SARS nsp14 and human RNA N7-MTase (hRNMT). In addition, the specificity of each compound was further evaluated using the CoV 2'O-MTase (nsp10/nsp16) and the MTases of Dengue and West-Nile flaviviruses as well as the hRNMT involved in the capping of cellular RNAs.

2. Materials and methods

2.1. Description of the libraries

The Prestwick Chemical Library[®] is a unique collection of 1280 small molecules, mostly approved drugs (FDA, EMA and other agencies) selected for their high chemical and pharmacological diversity as well as for their known bioavailability and safety in humans. The Prestwick Phytochemical Library is a collection of 320 natural products, mostly derived from plants, assembled by medicinal chemists and rich in diverse chemotypes, thus realistic for follow-up chemistry. The Prestwick Pyridazine Library is a collection of 400 innovative pyridazine and pyridazone derivatives, based on a series of carefully selected new original scaffolds. All molecules have been designed to ensure optimal diversity, and are suitable for diverse chemical modification.

2.2. Expression and purification of recombinant proteins

The viral MTases (SARS-CoV nsp14, MERS-CoV nsp14, DENV-3 NS5-MTase, West Nile virus NS5-MTase, DENV-2 NS5-RdRp and human RNA N7-methyltransferase (RNMT) coding sequences were cloned in fusion with a N-terminus hexa-histidine tag in Gateway[®] plasmids (pDest14 or pDest17, Life technologies). The proteins were expressed in *E. coli* cells and purified following previously described protocols (Aouadi et al., 2017; Bouvet et al., 2010; Ma et al., 2015; Milhas et al., 2016; Peyrane et al., 2007; Selisko et al., 2006). MERS-CoV nsp14 was produced and purified as follows: the protein were expressed in Arctic Express *E. coli* strain (Agilent) at 13 °C during 24 h after addition of 50 μM IPTG. The bacteria were pelleted (13,000 \times g, 4 °C, 5 min), and lysed by sonication in appropriate buffer (50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5% glycerol, 20 mM imidazole, 5 mM β -mercaptoethanol, 1 mM PMSF, 10 $\mu\text{g}/\text{ml}$ DNase-I and 0.25 mg/ml lysozyme). After clarification (50,000 \times g, 30 min at 4 °C) the recombinant protein was purified by immobilized metal affinity chromatography (IMAC) on 1 ml of HisPur[™] Cobalt Resin (Thermo Scientific) followed by size exclusion chromatography on a Superdex 200 (GE healthcare) equilibrated in (30 mM HEPES pH (7.5), 300 mM NaCl and 5% glycerol). The purified proteins were analyzed on SDS-PAGE gels after Coomassie blue staining (Fig. 1 and Fig. S1).

2.3. HTRF screen on SARS CoV nsp14 MTase

For the initial screening, 400 nl of each compound resuspended in 100% DMSO at 1 mM were dispensed in the reaction volume (5% final DMSO, 50 μM compound) using a Mosquito Crystal pipetting robot platform (TTP labtech). DMSO, sinefungin (20 μM) and SAH (2 μM) were used as controls.

The enzymatic reaction was performed in 8 μl . 3 μl of mix (Buffer + SARS nsp14 at 5 nM final concentration) were added in the assay wells, containing previously dispensed inhibitors (0.4 μl), using a Biomek NX MC pipetting robot (Beckman). The reaction

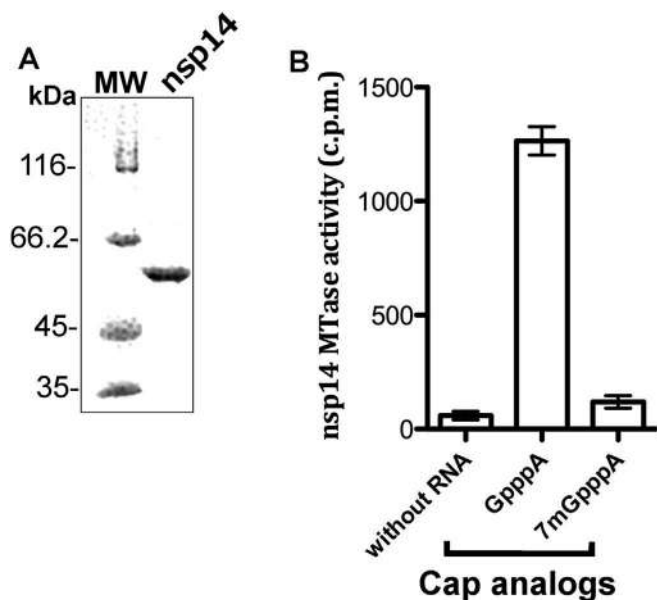


Fig. 1. Characterization of the recombinant SARS-CoV nsp14. A) 2.5 μ g of SARS-CoV nsp14 on a 10% SDS-PAGE gel stained by Coomassie blue. MW, molecular size markers. B) Bar graph representing the N7-MTase activity of SARS-nsp14 determined by monitoring the transfer of tritiated methyl from [3 H] SAM onto GpppA and 7 mGpppA. The tritiated RNA was quantified by DEAE filter binding assay (FBA).

buffer is composed of (Tris-HCl 40 mM (pH 8.0) and DTT 0.5 mM final concentrations). Reaction was started with 4.6 μ l of a mix containing both GpppA (NEB ref S1406) and SAM at 8 μ M and 2 μ M final concentrations respectively and was incubated 20 min at 30 $^{\circ}$ C.

For the detection of released SAH, the EPIgenousTM methyltransferase kit was purchased from CisBio Bioassays and reconstituted according to the supplier protocols. HTRF assays were performed in white 384 Well Small VolumeTM HiBase Polystyrene Microplates (Greiner) with a final volume of 20 μ l. 2 μ l of EPIgenous detection buffer one and 2 μ l of NaCl 5 M were then added in order to stop the reaction. After 10 min of incubation at room temperature, detection reagents were added: 4 μ l of a 1/16 dilution of SAH conjugated to the fluorescent dye d2 in a first place then 4 μ l of Anti-SAH antibody-Lumi4-Tb at a 1/150 dilution. HTRF signals were measured, after a 1 h final incubation at room temperature, using a PHERAstarFS (BMG Labtech) with an excitation filter at 337 nm and fluorescence wavelength measurements at 620 and 665 nm, an integration delay of 60 μ s and an integration time of 400 μ s. Results were analyzed with a two-wavelength signal ratio: [intensity (665 nm)/intensity (620 nm)]* 10^4 (HTRF Ratio). The normalized HTRF ratio is calculated as follow: [(compound signal) - (min signal)]/[(max signal) - (min signal)] * 100, where 'max signal' is the signal ratio without protein and 'min signal' the signal ratio with DMSO in place of compound.

The Z' factor is calculated using the following equation: $Z' = 1 - [3(\text{SD of max}) + 3(\text{SD of min})]/[(\text{mean max signal}) - (\text{mean min signal})]$, where SD is the standard deviation. For IC_{50} measurements, values were normalized and fitted with Prism (GraphPad software) using the following equation: $Y = 100/(1 + ((X/\text{IC}_{50})^{\text{Hill slope}}))$.

The initial screening assay was performed once and the hits were then confirmed by determination of IC_{50} (HTRF) in quadruplicates. IC_{50} (HTRF) is defined as the compound concentration for which the SAH release is decreased by 50%.

2.4. MTase filter binding assay (FBA)

The inhibitor specificity assays were carried out in reaction mixture [40 mM Tris-HCl (pH 8.0), 1 mM DTT, 1 mM MgCl_2 , 2 μ M SAM, and 0.33 μ M ^3H -SAM (Perkin Elmer)] in the presence of 0.7 μ M GpppAC₄ or ^7m GpppAC₄ synthetic RNAs. Purified SARS-nsp14 (10 nM), MERS-nsp14 (250 nM), SARS-nsp10/nsp16 (1 μ M), MERS-nsp10/nsp16 (500 nM), hRNMT (20 nM), DENV-MT (500 nM) and WNV NS5 (500 nM). The enzymes were mixed first with 50 μ M inhibitors before the addition of RNA substrate and SAM and then incubated at 30 $^{\circ}$ C. The final concentration of DMSO was 5%, and control reactions were performed in the presence of 5% DMSO.

Reaction mixtures were stopped after 30 min by their 10-fold dilution in ice-cold 100 μ M SAH. Samples were transferred to Diethylaminoethyl cellulose filters (DEAE) (Perkin Elmer) using a Filtermat Harvester apparatus (Packard Instruments). The unincorporated ^3H SAM was removed from the filter by several washing with 0.01 M ammonium formate pH (8.0), H₂O, then absolute ethanol, before drying of the DEAE filters. The filters were incubated with BetaplateScint (Wallac) scintillation fluid before quantification of ^3H methylation transferred onto RNA substrates using a Wallac 1450 MicroBetaTriLux liquid scintillation counter in counts per minute (cpm).

The inhibition specificity of DV-2 RdRp was performed using 100 nM enzyme, 200 μ M ATP, 4 μ M [^3H] ATP and 100 nM polyU substrate into the buffer reaction [50 mM HEPES, 10 mM DTT, 5 mM MgCl_2 , 2 mM MnCl_2]. The reaction was incubated during 10 min at 30 $^{\circ}$ C then stopped by adding 100 mM EDTA final concentration. The reaction mixture was then transferred onto DEAE filter as described above. The unincorporated [^3H]ATP during the RNA synthesis was removed from the filter with 0.2 M ammonium formate as described above. The IC_{50} values of inhibitors on SARS-nsp14 and RNMT were determined using 0.35 μ M GpppAUUAU RNA mimicking the 5' extremity of SARS-CoV genome and 0.35 μ M GpppAC₄ for hRNMT. The IC_{50} values were fitted as described above.

3. Results

3.1. Expression of active SARS-CoV nsp14 N7 MTase

SARS-CoV encodes the bi-functional enzyme nsp14, which contains an N-terminal ExoN domain followed by a N7-MTase domain involved in RNA capping. The recombinant protein was purified by immobilized metal affinity chromatography (IMAC) followed by size exclusion chromatography. The purified nsp14 is detected as a single band migrating at the expected molecular mass (Fig. 1A, \approx 60 kDa) on Coomassie blue stained SDS-PAGE gels. The purified nsp14 transfers [^3H] methyl onto a GpppA cap analogue whereas no MTase activity is detected onto ^7m GpppA (Fig. 1B) indicating that nsp14 harbors an N7-MTase activity that can be detected using cost effective commercial cap analogues.

3.2. Optimization of HTRF assay for high-throughput screening

We miniaturized the EPIgenous MTase assay based on Homogenous Time Resolved Fluorescence (HTRF) (Degorce et al., 2009) to screen a library of 2000 compounds. The principle of the HTRF assay, shown in Fig. S2, is to follow the MTase activity by indirectly detecting SAH, the by-product of the methyl transfer reaction. Indeed, in the detection system, the released SAH competes with fluorescent SAH-d2, which binds to an anti-SAH antibody coupled to a cryptate fluorophore (see Material and Methods). Thus, increase of SAH leads to decrease of the HTRF signal (Fig. S3D). The experimental conditions for the enzymatic assay were optimized

and adapted for automated screening assay. Briefly, I) the enzyme concentration was set up at 5 nM of nsp14 as titration curves indicated that 80% of the maximal activity after 20 min incubation period is obtained at this enzyme concentration (Fig. S3A); II) GpppA substrate concentration was adjusted at 8 μ M to robustly and reproducibly detect N7-MTase activity (Fig. S3B); III) SAM concentration was adjusted to 2 μ M in order to ensure the optimal conversion rate of SAM to SAH according to the manufacture recommendations (\approx 4–11%, Fig. S3C/D). The SAH detection conditions were optimized and the SAH-d2 and anti-SAH antibody were diluted 1/16 and 1/150 respectively.

Fig. 2A shows the time course experiment of nsp14 MTase activity indicating that using our experimental optimized conditions the N7-methylation reaction reached the saturation phase after a 30 min incubation period. The enzymatic reactions were thus incubated during 20 min in the screening. Sinefungin, which has already been reported to inhibit the SARS-CoV N7-MTase ($IC_{50} \approx$ 500 nM) (Bouvet et al., 2010), was used as positive control for MTase inhibition. The inhibition curve of nsp14 MTase using the optimized HTRF assay revealed an inhibition in a similar range, with IC_{50} (HTRF) value of 110 nM (Fig. 2B). We also assessed that the addition of DMSO up to 5% did not impact the MTase reactions and/or the HTRF quantification of released SAH. The signal variation was estimated to 2–2.8%. The average of Z' value was 0.69.

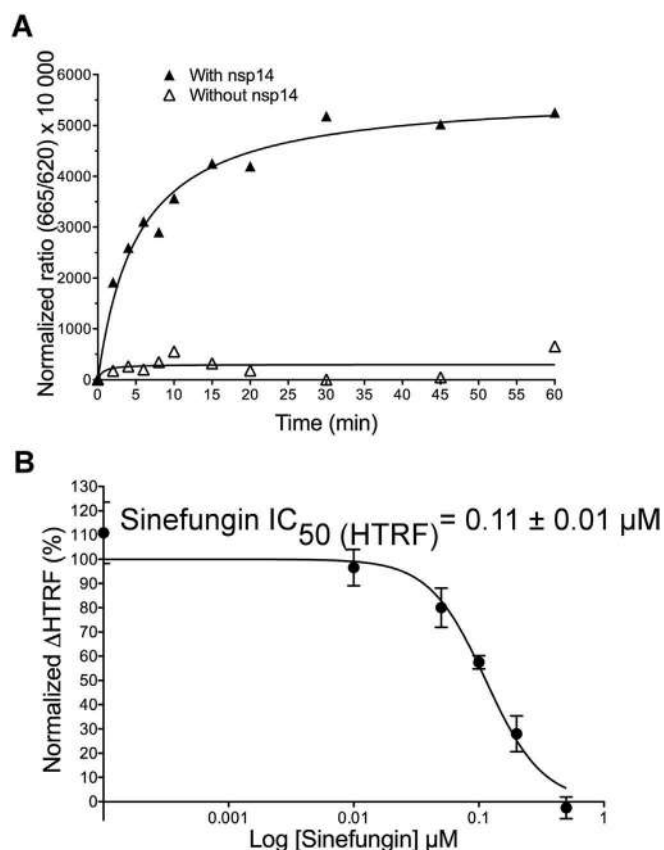


Fig. 2. HTRF assay validation using SARS-nsp14. A) Time course analysis of SARS-CoV nsp14 MTase activity by HTRF assay: 0 or 5 nM of nsp14 (empty or full triangles, respectively) are incubated with 8 μ M GpppA and 2 μ M SAM at 30 °C for increasing time periods. The reaction by-product (SAH) was quantified by measuring the ratio of emission and excitation fluorescence at 665 and 620 nm, on a PolarStar reader as described in materials and methods. B) Sinefungin IC_{50} (HTRF) determination. Increasing concentrations of sinefungin (0–0.5 μ M) were incubated with 5 nM nsp14, 2 μ M SAM, 8 μ M GpppA at 30 °C during 20 min and the SAH by-product was detected as described in panel A. The IC_{50} (HTRF) was then calculated using graphPad Prism equation ($n = 3$; Mean value \pm SD).

3.3. HTRF screening of inhibitors against SARS-CoV nsp14 N7-MTase activity

The chemical library containing 1280 FDA approved molecules, 320 natural substrates and 400 pyridazine derived compounds was used for initial screening at 50 μ M of each compound and the inhibition was determined using HTRF assay. Compounds showing inhibition of the MTase activity higher than 42%, which corresponds to increase of the HTRF signal $>20\%$ (Fig. S4), were first selected (41 compounds, Fig. 3). 21 compounds were discarded due to signal interference with the HTRF assay or by visual inspection. The 20 remaining compounds (hit rate 0.1%, Table 1) were selected for IC_{50} (HTRF) determination by HTRF. The results presented in Table 1 indicate that the IC_{50} (HTRF) vary from 0.16 to >400 μ M, among which 4 have an IC_{50} (HTRF) higher than 75 μ M. IC_{50} (HTRF) for 1159, 597, 194 were not determined due to intrinsic fluorescence interferences at high concentration. Thus, 13 compounds show IC_{50} (HTRF) ranging from 0.16 to 75 μ M.

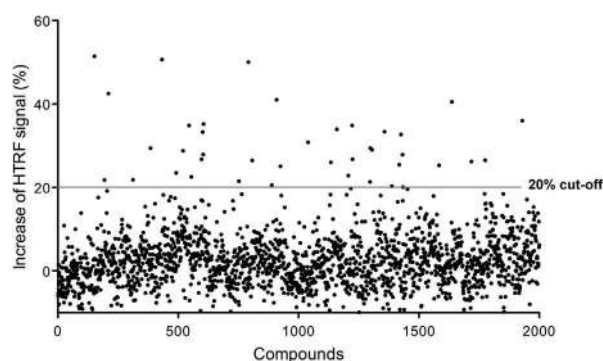


Fig. 3. Distribution of inhibition percentage values obtained by HTRF screening on Prestwick Chemical libraries. The 2000 compounds were screened at 50 μ M, 5% DMSO on SARS CoV nsp14 MTase activity using HTRF assay (5 nM nsp14, 8 μ M GpppA, 2 μ M SAM, SAH-d2 (1/16) and anti-SAH (1/150)). The grey line corresponds to an increase of 20% of the HTRF signal that was used as a threshold for the selection of the primary hits.

Table 1

Screening results and IC_{50} (HTRF) of selected compounds on the SARS-nsp14 by HTRF assay.

Compounds	HTRF signal increase at 50 μ M (%)	IC_{50} (HTRF) (μ M)
152	51.4	171.5 \pm 16
792	50	5.8 \pm 2
210	42.5	17.1 \pm 5.6
909	41	>400
545	34.8	>400
1159	33.9	Not available
602	33.3	73.6 \pm 29
1426	32.7	0.16 \pm 0.02
1040	30.8	5.5 \pm 1
1298	29.4	2 \pm 0.6
1305	29	3.3 \pm 0.7
604	27.9	57.2 \pm 15
597	26.8	Not available
1418	25.4	18.9 \pm 3.8
925	25	5 \pm 0.6
1207	22.8	73 \pm 9.3
312	21.8	38.6 \pm 5
194	21.8	Not available
752	21.5	35.5 \pm 7.3
1434	20	>400
Sinefungin	96.9	0.112 \pm 0.01

Not available: IC_{50} (HTRF) not accurate due to fluorescence issues at high compound concentrations.

3.4. Hit confirmation by radioactive MTase assay on SARS-CoV nsp14

In order to validate the inhibition observed by HTRF, we next performed a radioactive MTase assay (FBA) using freshly solubilized compounds. In this experiment, the MTase activity was analyzed by measuring the [³H] radiolabeled methyl transferred onto a longer synthetic RNA substrate (GpppAC₄). This capped RNA was previously evidenced as a better substrate for nsp14 N7-MTase than the cap analogue GpppA (Bouvet et al., 2010). The low inhibition of the

4 compounds showing IC₅₀ (HTRF) > 75 μM by HTRF (1434, 152, 545, 909), was confirmed by the FBA assay. In addition the fluorescent compounds 194, 597 poorly inhibits the MTase activity in this FBA assay. Finally, eleven compounds display a potent inhibition (>80%) at 50 μM concentration using the radioactive assay (Table 2). Whereas we observed some discrepancies between HTRF and FBA assay for a couple of compounds (210 and 909), which could be explained by the difference in RNA substrate length used in both assays; we observed a good correlation between both assays.

Table 2
Inhibition specificity of compounds at 50 μM on MTases and DENV-2 polymerase activities using DEAE-filter binding assay.

Compounds	Inhibition percentage at 50 μM (%)							
	SARS nsp14	MERS nsp14	SARS nsp10/16	MERS nsp10/16	RNMT	DENV-MT	WNV NS5	DV-2 Rdrp
Sinefungin	98.4	73.5	98.4	77	98.2	98.3	98.4	11
925	96.9	94.2	88.3	98	96	98.9	98.5	99
1426	95.3	38.1	89.5	87.7	97	85	94.1	72
1298	93.6	46.2	64.9	76.8	97	44.9	48.2	98
1305	92.2	37.4	69.5	83.6	97.8	60.3	58.9	85
1159	83	50.8	72.5	8.2	94.7	37.6	30.2	72
1040	90.5	85.6	92.9	94.5	98.6	97.8	66.3	13
792	81.3	62	94	85.5	96.5	19.1	39.6	47
312	93.7	63.5	63.3	9.5	66.7	33	18.5	26
1418	85.9	24.4	33.2	57.5	82.7	30	32	44
1207	93.5	42.1	38.7	13.1	42.5	14.6	33.1	14
604	93.7	39.3	34.2	7.6	95	12.2	22.7	6
752	37.3	38.1	13.7	0	0	28.3	0	8
1434	36.4	34.9	20.3	0	0	37	0	13
602	24.34	33.5	1	0	0	25.8	0	4
194	22.3	4.9	31.6	0	0	32.6	12.8	8
152	19.8	16.7	6.8	13.7	6	3	23.7	0
597	10.9	0	0	0	0	4	15.7	4
545	8.5	36.9	0	0	0	9.4	11.5	11
909	7.3	13.1	13.4	5.9	0	2.4	9.6	0
210	0.63	27.4	36.1	12.7	88.6	15.4	32.2	14

The compounds are clustered and framed according to their inhibition specificity from wide range to specific inhibitors. ND: Not determined.

3.5. Assessment of the inhibitors specificity using other viral and cellular MTases

The specificity of the 20 compounds was then analyzed by measuring their inhibition profile on other viral and cellular MTases (for more details see workflow in Fig. S5). We thus produced and purified the SARS- and MERS-CoV nsp14 and the nsp10/nsp16 complex (2'O-MTase); the full length NS5 of the West Nile (WNV) and NS5-MTase domain of Dengue 3 (DENV-3) viruses and the human RNA N7-MTase (hRNMT) (Fig. S1) in order to compare their inhibition profile. The inhibition of each compound (50 μM) was then determined by FBA MTase activity measurements. In addition, we also tested whether the DENV-2-RdRp activity is inhibited in order to address the compound specificity. The results summarized in Table 2 show that among the 11 best SARS-CoV nsp14 inhibitors confirmed by FBA (Table 5), 5 (925, 1426, 1298, 1305, 1159) show large spectrum inhibition as they inhibit all tested MTases as well as the DENV-2 RdRp. Table 2 also reveals that sinefungin (used as positive control) and compound 1040 exhibit a broad spectrum MTase inhibition profile with no inhibition detected on the RdRp activity of DENV NS5. This broad-spectrum inhibition observed with sinefungin has already been reported and can be related to its structural homology with the SAM/SAH (Dong et al., 2010). Interestingly, compound 792 inhibits mainly the CoV N7- and 2'O-MTase as well as the hRNMT but barely interferes with DENV-3 or WNV NS5 MTase activity. Finally, the four last compounds (312, 1418, 1207, 604) show more specific inhibition profile on MERS- and SARS-CoV nsp14 although they inhibit the hRNMT.

To further characterize the eleven most potent nsp14 N7-MTase inhibitors, a dose response curve was performed with the compounds showing percentage inhibition higher or equal to 40% against SARS-CoV nsp14 and the hRNMT. After pre-incubation with increasing concentrations of inhibitors, the MTase activity was followed by FBA. The IC_{50} of compounds on either nsp14 or hRNMT activities ranged from 19 nM to 23 μM and 50 nM to 5.5 μM , respectively, confirming the observed inhibition potency of these molecules (Table 3). In addition, the IC_{50} values determined in this assay corroborate with those performed by HTRF assay (Table 1). The broad spectrum MTase inhibitor (1040) shows IC_{50} values of 4.5 and 5.1 μM on SARS nsp14 and human hRNMT respectively and four compounds targeting more specifically SARS, MERS and hRNMT have IC_{50} values between 10 and 24 μM , but IC_{50} values observed with for hRNMT are lower in each case.

Altogether these results suggest that structural homology between viral and cellular MTases might impair the discovery of inhibitors targeting specifically the CoV N7-MTase (see discussion about homology). Nevertheless, we also observed that the less active inhibitors on SARS-CoV nsp14 barely inhibit the human

Table 3

Comparison of the IC_{50} value of 11 most active compounds on SARS-nsp14 and human RNMT activities using GpppAUAU and GpppAC₄ substrates respectively.

Compounds	IC_{50} SARS-nsp14 (μM)	IC_{50} RNMT (μM)
Sinefungin	0.36 \pm 0.08	<0.05
925	1.5 \pm 0.11	1 \pm 0.1
1426	0.019	ND
1298	0.74 \pm 0.1	0.32 \pm 0.27
1305	1.3 \pm 0.2	0.4 \pm 0.05
1040	4.5 \pm 0.35	5.1 \pm 2.28
792	31 \pm 3	5.5 \pm 1.92
1159	2 \pm 0.44	<0.05
312	12.3 \pm 1	4.84 \pm 1.6
1418	23.7 \pm 7	10 \pm 3
1207	18.4 \pm 2	ND
604	11 \pm 1	1.28 \pm 0.16

ND: Not Determined.

hRNMT suggesting the possibility to identify more specific MTase inhibitors (Table 2). A structure activity relationship (SAR) study is required in order to gain in activity and specificity.

3.6. Inhibition of SARS-CoV nsp14 by compound 312 analogs already present in the library

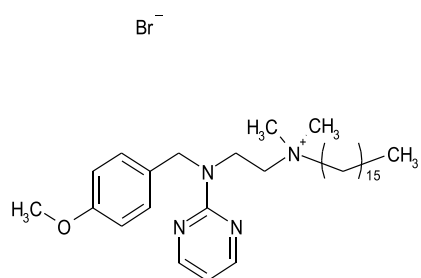
Following the confirmation of compound 312, analogs were

Table 4

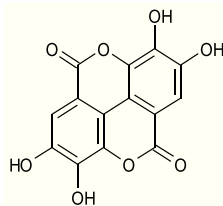
Evaluation of compound 312 analogs. IC_{50} determined by FBA on SARS-CoV nsp14 and RNMT MTase activities using GpppAC₄ substrate.

Compounds	Structure	IC_{50} (μM) on	
		SARS nsp14	RNMT
312		7.3 \pm 0.4	14.3 \pm 1
457		6.2 \pm 0.4	8.8 \pm 0.6
278		9.1 \pm 0.7	16.8 \pm 2
1264		13.5 \pm 2	13.7 \pm 0.8
133		No inhibition	365.2 \pm 186
510		No inhibition	159 \pm 36
706		No inhibition	226.8 \pm 29
503		No inhibition	102.2 \pm 21

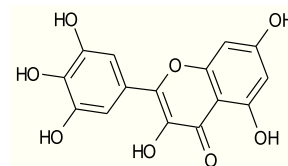
Table 5
Structure of the most active compounds.



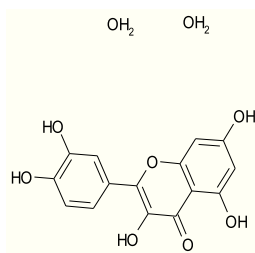
Compound 925



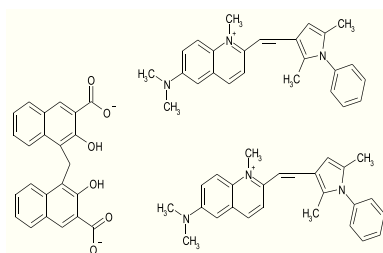
Compound 1426



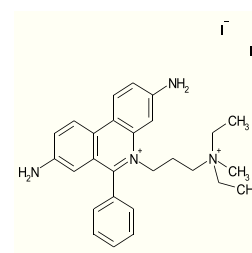
Compound 1298



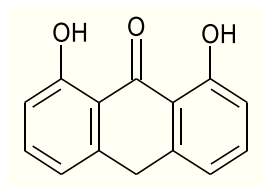
Compound 1305



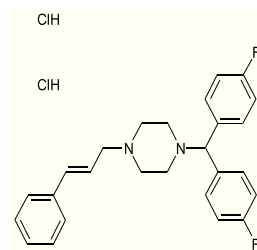
Compound 1040



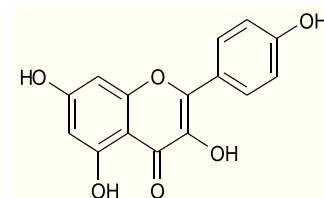
Compound 792



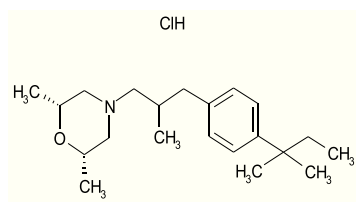
Compound 1159



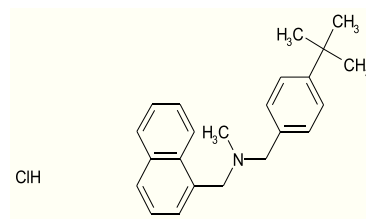
Compound 312



Compound 1418



Compound 1207



Compound 604

searched in the tested libraries. 7 analogs were selected and tested for inhibition of SARS nsp14 and hRNMT MTase activities by FBA (Table 4). All the 7 analogs bear the benzhydryl-piperazine group characteristic to this chemical series in the right-hand side of the structure. Four compounds lacking the aromatic moiety in the left part of the molecule show no inhibition of SARS-CoV nsp14 MTase. In contrast, 3 analogs carrying an aromatic terminal ring in the left-hand side inhibit the SARS-CoV N7-MTase activity with IC₅₀ between 6.2 and 14 μM (457, 278 and 1264). The substitution of the two phenyl rings in the benzhydryl-piperazine moiety (fluorine or chlorine groups) have no impact on the inhibition profile. These compounds were initially not selected from the original screening but their inhibition was close to the selection cut-off (42% inhibition at 50 μM). The compounds inhibited the hRNMT and SARS-CoV nsp14 to similar levels. Although they did not bring additional specificity compared to compound 312, it might be worth sampling a larger chemical space within this family.

4. Discussion

The identification of molecules inhibiting viral enzymes essential for virus replication is a key point to combat emerging viruses. Methylations of viral mRNA cap structures are an essential step in the virus life-cycle (Case et al., 2016; Cougot et al., 2004) preventing their identification as foreign RNA markers by the host cell (Diamond, 2014; Li et al., 2013; Züst et al., 2013). Thus, enzymes such as MTases involved in viral RNA capping are now considered as promising antiviral targets (Ferron et al., 2012). Here, we developed a high-throughput MTase screening assay based on HTRF technology that possesses advantageous features. The HTRF assay allows the minimization of non-specific fluorescence emissions by introducing a time delay (50–150 μs) between the initial light excitation and fluorescence measurement (Degorce et al., 2009). This time delay limits the interference of fluorescent compounds in the screening assay. Indeed, only 37 out of 2000 compounds interfered with the detection system among which were considered as 17 false positive. In addition, a non-fluorescent enzymatic assay (FBA) was used to confirm the hits obtained from the primary screening. The initial screening was performed using a commercial cap analogue substrate (GpppA), as the availability of high amount of capped RNA substrates is a limiting step. This screening allowed the identification of 20 compounds inhibiting at least 42% of the nsp14 N7-MTase activity. The inhibitory effect (below 25 μM, Table 3) of 11 of these compounds was confirmed using a radioactive filter-binding assay performed with longer synthetic RNA substrates (GpppAC₄). In contrast, 9 compounds barely inhibited the MTase activity measured with GpppAC₄ substrate. Thus, it is likely that the use of GpppA as substrate may favor the identification of inhibitors that would be less efficient on longer RNA substrates. The weaker inhibition effect observed using longer RNA substrates might suggest that these compounds interfere with the RNA binding site of the enzyme.

One important issue in the drug discovery process is to identify compounds inhibiting specifically the viral enzymes but barely active on the cognate cellular enzymes in order to limit their cytotoxicity or off-target effects. To determine the compound specificity at the early stage of the screening process, we compared the inhibition activity of the 20 compounds on other viral MTases (SARS- and MERS-CoV 2′O MTase, DEN-V and WNV N7/2′O MTases) and a cellular MTase hRNMT involved in the methylation of cap structures. In addition, we also tested these compounds on the Dengue virus polymerase in order to control the inhibition specificity on the RdRp activity and eliminate molecules showing unspecific inhibition or frequent hitters. This strategy revealed that 5 out of the 11 best hits identified in the nsp14 N7-MTase activity-

screening campaign potentially inhibited all MTases tested as well as the Dengue virus polymerase.

We also observed that most inhibitors identified in this comparative screening potentially inhibit the human RNA N7-MTase hRNMT (12 over 20 compounds). These results might be related to structural and functional homologies between CoV N7-MTase and the human hRNMT limiting the possibility to identify inhibitors targeting specifically the CoV N7-MTase. However, we would like to propose that the inhibitors identified here might well be substrate mimicks and bind to conserved substrate pockets since the protein folding of hRNMT and nsp14 are significantly different, the latter deviating from classical Rossmann-fold MTases. It is also noteworthy that this work was performed with a highly active truncated construct of hRNMT, which is used at very low concentration (20 nM) for IC₅₀ determination. Additionally, it is also possible that as our truncated hRNMT is less selective as it lacks the hRNMT activating Mini-protein (RAM) binding domain and the RAM protein. Indeed it was previously demonstrated that RAM stabilizes the structure and positioning of the hRNMT lobe and increases the specific recruitment of SAM (Varshney et al., 2016). This high inhibition of hRNMT MTase activity may be overcome in the presence of its allosteric activator RAM, which may allow to stringent the specificity of the compounds.

Out of nsp14, the comparative screening using various MTases also reveals that several compounds show a large spectrum of MTase inhibition without any significant activity detected on the Dengue virus polymerase used as a control (compounds 1040 and the positive control sinefungin). The pleiotropic inhibition observed on MTases activities suggests that despite the weak primary sequence conservation, these enzymes do have a common structural organization (Rossmann fold characteristic of most of viral and cellular MTases) and share similar enzymatic properties. This observation is consistent with the common structural organization, with a conserved SAM, RNA and/or cap binding sites, shared by those MTases. The identification of broad-spectrum viral MTases inhibitors may offer a starting point to develop new antiviral compounds. We considered that these molecules could be interesting as we expected that the SAR process will give the opportunity to increase both specificity and inhibition properties. Accordingly, a SAR study previously performed on the flavivirus MTases has demonstrated that chemical analogs of SAH, showed improved activity and selectivity against DENV MTase, but not on the related human enzymes (Lim et al., 2011). The crystal structure revealed that the most active SAH derivative harbors substitutions accommodated in a specific cavity conserved in Flaviviral MTases. It would be thus relevant to initiate the structural analysis of viral MTases in complex with the best compounds in order to improve selectivity.

5. Conclusion

Viral N7- and 2′O-MTase is now considered as promising antiviral target as their inhibition is supposed to impair mRNA translation into viral proteins and to induce the stimulation of the innate immune system of the infected host. Thus it is currently accepted that MTase inhibitors might help virus clearance by stimulation of the immune response (Decroly et al., 2012; Ferron et al., 2012; Züst et al., 2011). In this work we identified 20 molecules inhibiting the N7-MTase activity of SARS-CoV nsp14 from chemical libraries composed of 2000 compounds using an HTRF based inhibition assay. Eleven compounds demonstrated a potent inhibition (>80%) at 50 μM concentration using the radioactive assay. The specificity of each compound was further evaluated using a radioactive assay on CoV 2′O-MTase (nsp10/nsp16) and the MTases of Dengue and West-Nile viruses as well as the human RNA N7-MTase (hRNMT)

involved in the capping of cellular RNAs. This comparative analysis revealed that the selected compounds cluster in 3 classes: 1) barely specific inhibitors, 2) pan-MTases inhibitors inhibiting viral and cellular MTases, 3) inhibitors targeting more specifically one viral MTase but also active against the hRNMT. These latter series of inhibitors may represent a good starting point for the development of antivirals targeting the CoV MTases.

Acknowledgments

We are grateful to Eric Trinquet and Thomas Roux for helpful discussions on the HTRF assay. We thank Alexandre Blanjoie for HPLC purification of high amount of GpppAC₄ substrates.

This work was supported by the French research agency ANR [VMTaseIn, grant ANR-ST14-ASTR-0026], and SARS-RNA-SPA, [grant ANR-12BSV3-0007] and the PACA regional fundings. BM is the recipient of a DGA-MRIS scholarship and WA the recipient of a scholarship from the "Méditerranée Infection" Foundation and Fondation pour la Recherche Médicale (FRM: FDT20160434967).

Appendix A. Supplementary data

Supplementary data related to this article can be found at <http://dx.doi.org/10.1016/j.antiviral.2017.06.021>.

References

- Aouadi, W., Blanjoie, A., Vasseur, J., Canard, B., Decroly, E., 2017. Binding of the methyl donor S-Adenosyl-L-Methionine to Middle East respiratory syndrome coronavirus 2'-O-methyltransferase nsp16 promotes recruitment of the allosteric activator nsp10. *J. Virol.* 91, 1–18.
- Bouvet, M., Debarnot, C., Imbert, I., Selisko, B., Snijder, E.J., Canard, B., Decroly, E., 2010. In vitro reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS Pathog.* 6, e1000863. <http://dx.doi.org/10.1371/journal.ppat.1000863>.
- Bouvet, M., Imbert, I., Subissi, L., Gluais, L., Canard, B., Decroly, E., 2012. RNA 3'-end mismatch excision by the severe acute respiratory syndrome coronavirus nonstructural protein nsp10/nsp14 exoribonuclease complex. *PNAS* 109, 9372–9377. <http://dx.doi.org/10.1073/pnas.1201130109/-DCSupplemental>. www.pnas.org/cgi/doi/10.1073/pnas.1201130109.
- Bouvet, M., Lugari, A., Posthuma, C.C., Zevenhoven, J.C., Bernard, S., Betzi, S., Imbert, I., Canard, B., Guillemot, J.C., Lécine, P., Pfefferle, S., Drosten, C., Snijder, E.J., Decroly, E., Morelli, X., 2014. Coronavirus Nsp10, a critical co-factor for activation of multiple replicative enzymes. *J. Biol. Chem.* 289, 25783–25796. <http://dx.doi.org/10.1074/jbc.M114.577353>.
- Case, J.B., Ashbrook, A.W., Dermody, T.S., Denison, M.R., 2016. Mutagenesis of S-Adenosyl-L-Methionine-Binding residues in coronavirus nsp14 N7-methyltransferase demonstrates differing requirements for genome translation and resistance to innate immunity. *J. Virol.* 90, 7248–7256. <http://dx.doi.org/10.1128/JVI.00542-16>.
- Chen, Y., Cai, H., Pan, J., Xiang, N., Tien, P., Ahola, T., Guo, D., 2009. Functional screen reveals SARS coronavirus nonstructural protein nsp14 as a novel cap N7 methyltransferase. *PNAS* 106, 3484–3489. <http://dx.doi.org/10.1073/pnas.0808790106>.
- Chen, Y., Su, C., Ke, M., Jin, X., Xu, L., Zhang, Z., Wu, A., Sun, Y., Yang, Z., Tien, P., Ahola, T., Liang, Y., Liu, X., Guo, D., 2011. Biochemical and structural insights into the mechanisms of SARS coronavirus RNA ribose 2'-O-methylation by nsp16/nsp10 protein complex. *PLoS Pathog.* 7, e1002294. <http://dx.doi.org/10.1371/journal.ppat.1002294>.
- Cougot, N., Van Dijk, E., Babajko, S., Séraphin, B., 2004. Cap-tabolism. *Trends Biochem. Sci.* 29, 436–444. <http://dx.doi.org/10.1016/j.tibs.2004.06.008>.
- de Wit, E., van Doremalen, N., Falzarano, D., Munster, V.J., 2016. SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* 14, 523–534. <http://dx.doi.org/10.1038/nrmicro.2016.81>.
- Decroly, E., Debarnot, C., Ferron, F., Bouvet, M., Coutard, B., Imbert, I., Gluais, L., Papageorgiou, N., Sharif, A., Bricogne, G., Ortiz-Lombardia, M., Lescar, J., Canard, B., 2011. Crystal structure and functional analysis of the SARS-coronavirus RNA cap 2'-O-methyltransferase nsp10/nsp16 complex. *PLoS Pathog.* 7, e1002059. <http://dx.doi.org/10.1371/journal.ppat.1002059>.
- Decroly, E., Ferron, F., Lescar, J., Canard, B., 2012. Conventional and unconventional mechanisms for capping viral mRNA. *Nat. Rev. Microbiol.* 10, 51–65. <http://dx.doi.org/10.1038/nrmicro2675>.
- Degorce, F., Card, A., Soh, S., Trinquet, E., Knapiak, G.P., Xie, B., 2009. HTRF: a technology tailored for drug discovery - a review of theoretical aspects and recent applications. *Curr. Chem. Genomics* 3, 22–32. <http://dx.doi.org/10.2174/1875397300903010022>.
- Diamond, M.S., 2014. IFIT1: a dual sensor and effector molecule that detects non-2'-O methylated viral RNA and inhibits its translation. *Cytokine Growth Factor Rev.* 25, 543–550. <http://dx.doi.org/10.1016/j.cytogfr.2014.05.002>.
- Dong, H., Liu, L., Zou, G., Zhao, Y., Li, Z., Lim, S.P., Shi, P.Y., Li, H., 2010. Structural and functional analyses of a conserved hydrophobic pocket of flavivirus methyltransferase. *J. Biol. Chem.* 285, 32586–32595. <http://dx.doi.org/10.1074/jbc.M110.129197>.
- Ferron, F., Decroly, E., Selisko, B., Canard, B., 2012. The viral RNA capping machinery as a target for antiviral drugs. *Antivir. Res.* 96, 21–31. <http://dx.doi.org/10.1016/j.antiviral.2012.07.007>.
- Ivanov, K.A., Ziebuhr, J., 2004. Human coronavirus 229E nonstructural protein 13: characterization of duplex-unwinding, nucleoside triphosphatase, and RNA 5'-triphosphatase activities. *J. Virol.* 78, 7833–7838. <http://dx.doi.org/10.1128/JVI.78.14.7833-7838.2004>.
- 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. <http://dx.doi.org/10.1007/s00018-004-4242-5>.
- Li, S.-H., Dong, H., Li, X.-F., Xie, X., Zhao, H., Deng, Y.-Q., Wang, X.-Y., Ye, Q., Zhu, S.-Y., Wang, H.-J., Zhang, B., Leng, Q.-B., Zuest, R., Qin, E.-D., Qin, C.-F., Shi, P.-Y., 2013. Rational design of a flavivirus vaccine by abolishing viral RNA 2'-O methylation. *J. Virol.* 87, 5812–5819. <http://dx.doi.org/10.1128/JVI.02806-12>.
- Lim, S.P., Sonntag, L.S., Noble, C., Nilar, S.H., Ng, R.H., Zou, G., Monaghan, P., Chung, K.Y., Dong, H., Liu, B., Bodenreider, C., Lee, G., Ding, M., Chan, W.L., Wang, G., Jian, Y.L., Chao, A.T., Lescar, J., Yin, Z., Vedananda, T.R., Keller, T.H., Shi, P.Y., 2011. Small molecule inhibitors that selectively block dengue virus methyltransferase. *J. Biol. Chem.* 286, 6233–6240. <http://dx.doi.org/10.1074/jbc.M110.179184>.
- Ma, Y., Wu, L., Shaw, N., Gao, Y., Wang, J., Sun, Y., Lou, Z., Yan, L., Zhang, R., Rao, Z., 2015. Structural basis and functional analysis of the SARS coronavirus nsp14-nsp10 complex. *PNAS* 112, 9436–9441. <http://dx.doi.org/10.1073/pnas.1508686112>.
- Menachery Jr., V.D., B.L.Y., Josset, L., Gralinski, L.E., Scobey, T., Agnihothram, S., Katze, M.G., Baric, R.S., 2014. Attenuation and restoration of severe acute respiratory syndrome coronavirus mutant lacking 2'-O-methyltransferase activity. *J. Virol.* 88, 4251–4264. <http://dx.doi.org/10.1128/JVI.03571-13>.
- Milhas, S., Raux, B., Betzi, S., Dervieux, C., Roche, P., Restouin, A., Basse, M.J., Rebuffet, E., Lugari, A., Badol, M., Kashyap, R., Lissitzky, J.C., Eydoux, C., Hamon, V., Gourdel, M.E., Combes, S., Zimmermann, P., Aurrand-Lions, M., Roux, T., Rogers, C., M?ller, S., Knapp, S., Trinquet, E., Collette, Y., Guillemot, J.C., Morelli, X., 2016. Protein-protein interaction inhibition (2P2I)-oriented chemical library accelerates hit discovery. *ACS Chem. Biol.* 11, 2140–2148. <http://dx.doi.org/10.1021/acschembio.6b00286>.
- Minskaia, E., Hertzog, T., Gorbalyena, A.E., Campanacci, V., Cambillau, C., Canard, B., Ziebuhr, J., 2006. Discovery of an RNA virus 3'->5' exoribonuclease that is critically involved in coronavirus RNA synthesis. *PNAS* 103, 5108–5113. <http://dx.doi.org/10.1073/pnas.0508200103>.
- Pan, J., Peng, X., Gao, Y., Li, Z., Lu, X., Chen, Y., Ishaq, M., Liu, D., DeDiego, M.L., Enjuanes, L., Guo, D., 2008. Genome-wide analysis of protein-protein interactions and involvement of viral proteins in SARS-CoV replication. *PLoS ONE* 3. <http://dx.doi.org/10.1371/journal.pone.0003299>.
- Peyrane, F., Selisko, B., Decroly, E., Vasseur, J.J., Benarroch, D., Canard, B., Alvarez, K., 2007. High-yield production of short GpppA- and 7MeGpppA-capped RNAs and HPLC-monitoring of methyltransfer reactions at the guanine-N7 and adenosine-2'O positions. *Nucleic Acids Res.* 35. <http://dx.doi.org/10.1093/nar/gkl1119>.
- Pichlmair, A., Lassnig, C., Eberle, C.-A., Górná, M.W., Baumann, C.L., Burkard, T.R., Bürckstümmer, T., Stefanovic, A., Krieger, S., Bennett, K.L., Rülcke, T., Weber, F., Colinge, J., Müller, M., Superti-Furga, G., 2011. IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. *Nat. Immunol.* 12, 624–630. <http://dx.doi.org/10.1038/ni.2048>.
- Schuberth-Wagner, C., Ludwig, J., Bruder, A.K., Herzner, A.M., Zillinger, T., Goelck, M., Schmidt, T., Schmid-Burgk, J.L., Kerber, R., Wolter, S., Stümpel, J.P., Roth, A., Bartok, E., Drosten, C., Coch, C., Hornung, V., Barchet, W., Kümmerer, B.M., Hartmann, G., Schlee, M., 2015. A conserved histidine in the RNA sensor RIG-I controls immune tolerance to N1-2'O-methylated self RNA. *Immunity* 43, 41–51. <http://dx.doi.org/10.1016/j.immuni.2015.06.015>.
- Selisko, B., Dutartre, H., Guillemot, J.C., Debarnot, C., Benarroch, D., Khromykh, A., Desprès, P., Egloff, M.P., Canard, B., 2006. Comparative mechanistic studies of de novo RNA synthesis by flavivirus RNA-dependent RNA polymerases. *Virology* 351, 145–158. <http://dx.doi.org/10.1016/j.virol.2006.03.026>.
- Snijder, E.J., Decroly, E., Ziebuhr, J., 2016. Chapter three – the nonstructural proteins directing coronavirus RNA synthesis and processing. In: *Advances in Virus Research*, pp. 59–126. <http://dx.doi.org/10.1016/bs.avir.2016.08.008>.
- Sun, Y., Wang, Z., Tao, J., Wang, Y., Wu, A., Yang, Z., Wang, K., Shi, L., Chen, Y., Guo, D., 2014. Yeast-based assays for the high-throughput screening of inhibitors of coronavirus RNA cap guanine-N7-methyltransferase. *Antivir. Res.* 104, 156–164. <http://dx.doi.org/10.1016/j.antiviral.2014.02.002>.
- Van Boheemen, S., Graaf, M. De, Lauber, C., Bestebroer, T.M., Raj, V.S., Zaki, A.M., Osterhaus, A., Haagmans, B.L., Gorbalyena, A., Snijder, E., Fouchier, R., 2012. Genomic characterization of newly discovered coronavirus associated with acute respiratory distress syndrome in humans. *mBio* 3. <http://dx.doi.org/10.1128/mBio.00473-12.e00473-12>.
- Van Hemert, M.J., Van Den Worm, S.H.E., Knoop, K., Mommaas, A.M., Gorbalyena, A.E., Snijder, E.J., 2008. SARS-coronavirus replication/transcription complexes are membrane-protected and need a host factor for activity in vitro. *PLoS Pathog.* 4. <http://dx.doi.org/10.1371/journal.ppat.1000054>.
- Varshney, D., Petit, A.-P., Bueren-Calabuig, J.A., Jansen, C., Fletcher, D.A., Pegg, M.,

- Weidlich, S., Scullion, P., Pisiakov, A.V., Cowling, V.H., 2016. Molecular basis of RNA guanine-7 methyltransferase (RNMT) activation by RAM. *NAR* 10, gkw637. <http://dx.doi.org/10.1093/NAR/GKW637>.
- Wu, B., Peisley, A., Richards, C., Yao, H., Zeng, X., Lin, C., Chu, F., Walz, T., Hur, S., 2013. Structural basis for dsRNA recognition, filament formation, and antiviral signal activation by MDA5. *Cell* 152, 276–289. <http://dx.doi.org/10.1016/j.cell.2012.11.048>.
- Zhang, Y., Wei, Y., Zhang, X., Cai, H., Niewiesk, S., Li, J., 2014. Rational design of human metapneumovirus live attenuated vaccine candidates by inhibiting viral mRNA cap methyltransferase. *J. Virol.* 88, 11411–11429. <http://dx.doi.org/10.1128/jvi.00876-14>.
- Züst, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B.W., Ziebuhr, J., Szretter, K.J., Baker, S.C., Barchet, W., Diamond, M.S., Siddell, S.G., Ludewig, B., Thiel, V., 2011. Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat. Immunol.* 12, 137–143. <http://dx.doi.org/10.1038/ni.1979>.
- Züst, R., Dong, H., Li, X.F., Chang, D.C., Zhang, B., Balakrishnan, T., Toh, Y.X., Jiang, T., Li, S.H., Deng, Y.Q., Ellis, B.R., Ellis, E.M., Poidinger, M., Zolezzi, F., Qin, C.F., Shi, P.Y., Fink, K., 2013. Rational design of a live attenuated Dengue Vaccine: 2'-O-methyltransferase mutants are highly attenuated and immunogenic in mice and macaques. *PLoS Pathog.* 9, e1003521. <http://dx.doi.org/10.1371/journal.ppat.1003521>.

Résumé

Les virus appartenant à l'ordre des *Mononegavirales* possèdent une « large » protéine L, responsable du cycle réplication/transcription et de maturation des ARNs. Six domaines conservés portent les différentes activités de cette protéine dont la région VI, qui contiendrait le site catalytique d'une activité méthyltransférase (MTase) de la coiffe. La coiffe est une structure chimique constituée d'une guanosine méthylée en position N7 reliée à l'extrémité 5' des ARNm par une liaison 5'-5' triphosphate. Une seconde méthylation est également présente en position 2'O du ribose du premier nucléotide de l'extrémité 5' de l'ARNm. Ces méthylations ont un rôle critique chez les virus car elles permettent la traduction efficace des ARNm mais permettent également aux ARNs viraux d'échapper à leur détection par l'immunité innée de l'hôte. Ainsi, la caractérisation de ce domaine chez le virus Ebola serait un point clé pour une meilleure compréhension de la réplication des filovirus et un pas vers l'élaboration d'une nouvelle stratégie thérapeutique. Nous avons donc produit et purifié le domaine MTase du virus Soudan (SUDV) afin de caractériser son activité. Il a été démontré que le domaine C-terminal de la protéine L joue un rôle dans le recrutement de l'ARN et est donc crucial pour l'activité MTase. En utilisant des ARNs synthétiques portant différents types de coiffe, nous avons pu identifier une activité A-2'O MTase interne tout à fait originale. Bien que la fonction d'une telle méthylation ne soit pas identifiée, nous avons proposé que ces modifications épitranscriptomiques sont impliquées dans d'autres mécanismes d'échappement à l'immunité innée, dans la régulation du cycle réplication/transcription ou dans l'encapsidation médiée par la nucléoprotéine. Le domaine MTase de SUDV est également capable de méthyle les positions N7 et 2'O de la coiffe mais une caractérisation plus approfondie est nécessaire pour mieux comprendre le processus de synthèse de la coiffe. Enfin, nous avons identifié des molécules inhibant l'activité MTase des filovirus. Les résultats d'une analyse biochimique plus poussée de ces inhibitions permettront d'initier le développement d'une nouvelle stratégie antivirale contre le virus Ebola.

Abstract

In the *Mononegavirales* order, viruses encode a large protein (L), which is responsible for replication/transcription and RNA modifications. This protein harbours six conserved domains accountable of these different activities. Among these domains, the conserved region VI (CRVI) has been predicted to support cap-methyltransferase (MTase) activity. The cap consists in a N7-methylated guanosine linked to the first nucleotide at the mRNA 5'-end by a 5'-5' triphosphate bond. This structure can also be methylated at the 2'O position of N1 ribose. These methylations play a critical role in virus life cycle as N7 methylation triggers efficient viral RNA translation and 2'O methylation hampers the detection of viral RNA by the host innate immunity. Thus, the characterization of this domain in Ebola virus is a key point to understand replication of mononegaviruses and design new antiviral strategies. We produced the MTase domain of *Sudan ebolavirus* (SUDV) to characterize its MTase activity. We demonstrated that the protruding C-terminal domain is essential for MTase activity as this domain is a key for the RNA recognition. Using synthetic short RNAs holding different cap structures, we discovered that SUDV MTase harbours an unconventional A-2'O MTase activity. Although the function of such methylations is not yet determined, we hypothesized that such epitranscriptomic RNA modifications may be involved in host immunity escape mechanisms or may regulate translation or RNA packaging by the nucleoprotein. Besides this, the MTase domain is able to methylate the cap structure at N7 and 2'O positions but further characterization would be necessary to fully understand the cap synthesis. Finally, we identified compounds limiting the Ebola virus MTase activity. Further biochemistry and compounds characterization results will thus pave the way towards the development of an innovative antiviral strategy.