

**THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LILLE**

École Doctorale Biologie –Santé de Lille

Pour l'obtention du grade de :

**Docteur de L'UNIVERSITÉ DE LILLE**

**Thèse soutenue le 12 juillet 2018**

**CARACTÉRISATION DE LA PROTÉINE S DU  
CORONAVIRUS HUMAIN 229E**

Centre d'Infection & d'Immunité de Lille (CIIL), Inserm U1019, CNRS UMR 8204,  
Institut Pasteur de Lille, Équipe Virologie Moléculaire & Cellulaire

Directrice de Thèse : **Dr Anne Goffard**

Présentée et soutenue publiquement par :

**Ariane Bonnin**

Membres du Jury :

**Président**

**Pr Gilles Duverlie**

**Rapporteur**

**Dr Sophie Le Poder**

**Rapporteur**

**Dr Olivier Terrier**

**Examineur**

**Pr Fabrice Allain**

**Examineur**

**Pr Benoit Foligné**

**Directrice de thèse**

**Dr. Anne Goffard**



## Remerciements

Je remercie le Dr Jean Dubuisson de m'avoir accueillie dans son équipe pour ma thèse de pharmacie et mon doctorat en science. Tes conseils scientifiques m'ont été et me sont encore très utiles. Je remercie également le Dr Olivier Terrier et le Dr Sophie Le Poder d'avoir accepté d'être rapporteurs de ma thèse, ainsi que le Pr Gilles Duverlie, le Pr Benoît Foligne et le Pr Fabrice Allain d'avoir accepté de faire partie de mon jury. Je tiens également à remercier le Dr Sandrine Duvet, qui faisait parti de mon comité de thèse avec Olivier Terrier et qui m'a soutenu.

Je remercie le Dr Anne Goffard, ma directrice de thèse, pour m'avoir prise comme étudiante et pour tous ses conseils. Merci de m'avoir encouragé quand j'en avais besoin, et de m'avoir permis d'enseigner la virologie. J'ai vraiment apprécié !

Je remercie aussi le Dr Sandrine Belouzard pour tout le temps qu'elle m'a consacré et pour sa patience. Merci également pour tes gâteaux, qui sont décidément très bon !

Un grand merci à Anabelle et Adeline (la team andouillette), qui m'ont aidé à survivre au doctorat. Votre soutien a été indispensable ! Je ne compte plus les fous rires ou discussions sérieuses qu'on a eu. Finalement, entre les bains de boue, les « réunions » au McEwans et Porto, on n'a pas eu le temps de s'ennuyer !

Merci à tout le labo : aux mange-tôt et aux mange-tard. A Nathalie et ses habits roses et à paillettes, à Karin, qui rassemble le labo à St Omer ou autour d'un ticket de loto, et à Muriel, qui me tient compagnie le matin. Merci à Yves pour toutes tes blagues et pour tes conseils. Avec tes Stan Smith et ton Kway, tu resteras la personne qui lance les modes ! Merci à Laurence pour son soutien et sa gentillesse. Merci aussi à Cécile-Marie pour ta bienveillance (je pense à toi quand je vois des chouettes maintenant). Merci à Lydia, jeune padawan des westerns-blot, pour ton sourire perpétuel. Ca nous ramène un peu de soleil de Californie ! Merci à Czes pour son humour. Merci aussi à Soph pour sa bonne humeur. Ton humour est parfois bizarre, mais on s'habitue !! Merci à Claire, qui est attentive aux autres. Merci aux doctorants du labo. Maliki, qui semble avoir un radar à nourriture, Najet, qui nous fait du super bon thé à la menthe sur la plaque chauffante, Juliette, qui fait de bons gâteaux et dernièrement Kevin. Merci aussi à ceux qui sont partis du labo (ou de l'étage), et particulièrement à Laura et Jeanne.

Je tiens aussi à remercier les « voisins » du labo d'à côté, Alex (qui a souvent faim), Jonathan, Sophie, Carine, Stephanie, Christine et tous les autres pour leur bonne humeur. Promis, je vais arrêter de surveiller combien de cafés vous buvez. J'ai aussi une pensée pour Béatrice, avec qui j'ai partagé quelques TPs. Courage, tu as presque fini !

Merci à Christel pour tous les cinés/restos/discussions qu'on a eu depuis le début de l'internat, déjà 5 ans ! Rendez-vous à Toulouse ☺

Merci aussi à ma colocataire Alice, pour sa bonne humeur perpétuelle, son soutien, et toutes ses recettes italiennes (et n'oublions pas le prosecco et l'apéro italien !).

Un grand merci à mes amis, particulièrement à Seb. Tu es toujours là, et je ne t'ai jamais remercié pour ta bonne humeur et ta disponibilité, donc merci pour ces 15 dernières années ☺ (et les 70 à venir !).

Merci aussi aux « vieux », Elise, Jess et Pauline. On a réussi à tous être dans des villes différentes, mais ça ne change rien ! Un grand merci à Evelyne pour son soutien, ainsi qu'à Mion, et à Milou pour les nombreuses images qu'elle m'a envoyée ☺

Je tiens également à remercier ma famille et particulièrement mes parents, mon frère Manu et mes soeurs, Sophie et Claire, qui me soutiennent quoi qu'il arrive et qui m'encouragent. Je ne serais pas arrivée jusque là sans vous !

## TABLE DES MATIERES

---

Liste des illustrations .....	1
Liste des tableaux.....	2
Résumé .....	3
Abstract.....	5
Abréviations .....	7
1 Introduction.....	9
1.1 Généralités sur les coronavirus .....	9
1.1.1 Historique.....	9
1.1.2 Epidémiologie .....	12
1.1.3 Infection par les coronavirus.....	18
1.2 Biologie des coronavirus.....	25
1.2.1 Structure et morphologie de la particule virale .....	25
1.2.2 Organisation du génome .....	26
1.2.3 Protéines structurales .....	27
1.2.4 Protéines non structurales .....	29
1.2.5 Protéines accessoires.....	32
1.3 Cycle viral .....	35
1.3.1 Entrée .....	35
1.3.2 Réplication du génome viral et synthèse des protéines virales.....	37
1.3.3 Assemblage et exocytose.....	40
1.4 Protéine S .....	41
1.4.1 Généralités .....	41
1.4.2 Domaine S1 : Attachement et reconnaissance antigénique .....	42
1.4.3 Domaine S2 : Fusion virale.....	49
1.5 N-glycosylation.....	57
1.5.1 Généralités et mécanisme .....	57
1.5.2 Rôle de la N-glycosylation .....	61
1.6 Objectifs .....	71
2 Matériel et méthodes.....	73

2.1	Cellules .....	73
2.2	Anticorps .....	73
2.3	Inhibiteurs et produits chimiques .....	73
2.4	Plasmides .....	74
2.5	Séquences et alignements.....	74
2.6	Transfection .....	75
2.7	Transduction .....	75
2.8	Western-blot.....	75
2.9	Clivage et biotinylation .....	76
2.10	Co-immunoprécipitation protéine S-récepteur soluble.....	77
2.10.1	Production du récepteur soluble .....	77
2.10.2	Co-immunoprécipitation .....	77
2.11	Production de pseudo-particules et tests d'infection.....	78
2.12	Test d'infection en surface .....	78
2.13	Analyse statistique .....	79
3	Partie I : Etude du clivage protéolytique de la protéine S du coronavirus humain 229E .....	81
3.1	Résultats .....	81
3.1.1	Modélisation de la protéine S du coronavirus humain 229E .....	81
3.1.2	Conservation des sites de clivage de la protéine S.....	82
3.1.3	Rôle de la région S2' dans l'entrée du virus HCoV-229E .....	83
3.1.4	Infection en surface médiée par la trypsine et TMPRSS2.....	85
3.1.5	Clivage de la protéine S du HCoV-229E par la trypsine.....	89
3.2	Discussion.....	91
4	Partie II : Etude du rôle de la N-glycosylation dans les fonctions de la protéine S du coronavirus humain 229E .....	97
4.1	Résultats .....	97
4.1.1	Détermination des séquons N-glycosylés.....	97
4.1.2	Rôle de la N-glycosylation du domaine S1 dans l'entrée virale.....	99
4.1.3	Etude du rôle de la N-glycosylation lors du repliement des protéines mutantes.....	102
4.1.4	Etude du rôle de la N-glycosylation lors de l'interaction de la protéine S avec le récepteur humain APN.....	103

4.1.5 Etude du rôle de la N-glycosylation pour les fonctions de fusion de la protéine S .....	104
4.1.6 Neutralisation.....	106
4.2 Discussion.....	111
5 Partie III : Discussion générale et perspectives .....	119
6 Bibliographie.....	125
7 Annexes .....	145





## LISTE DES ILLUSTRATIONS

---

Figure 1 : Distribution mondiale des coronavirus humains .....	13
Figure 2 : Transmission intra- et inter-espèces des coronavirus humains .....	15
Figure 3: Structure générale d'un coronavirus .....	255
Figure 4: Topologie des protéines structurales présentes sur l'enveloppe des coronavirus .....	277
Figure 5: Organisation du génome de HCoV-229E. ....	322
Figure 6: Schéma général du cycle viral des coronavirus.....	399
Figure 7 : Schéma fonctionnel et représentation 3D de la protéine S des coronavirus .....	455
Figure 8 : Schéma du mécanisme de fusion virale avec la cellule hôte.....	51
Figure 9: Représentation de l'entrée viral des coronavirus .....	555
Figure 10: Biosynthèse des N-glycanes.....	588
Figure 11: Mécanisme de repliement dans le RE des protéines N-glycosylées. ....	60
Figure 12: Organisation du bouclier de glycanes de la protéine S du HCoV-NL63 par cryo-microscopie électronique et spectrométrie de masse .....	688
Figure 13: Modélisation de la protéine S du HCoV-229E .....	82
Figure 14: Infection par des particules pseudotypées avec la protéine S du HCoV-229E... ..	844
Figure 15: Schéma fonctionnel de la protéine S et séquences des régions S1/S2 et S2' de la protéine S et des différents mutants .....	866
Figure 16: Infection en surface des 229Epp médiée par la trypsine et par TMPRSS2.....	888
Figure 17: Clivage de la protéine S du HCoV-229E par la trypsine .....	899
Figure 18 : Analyse des pseudoparticules incorporant la protéine S du HCoV-229E ou un des mutants .....	100
Figure 19 : Rôle des N-glycanes dans le repliement de la protéine S. ....	102
Figure 20 : Etude de l'interaction de la protéine S avec le récepteur soluble shAPN .....	1044
Figure 21 : Infection en surface des 229Epp médiée TMPRSS2 et test de fusion à la surface cellulaire .....	1055
Figure 22 : Clivage de la protéine S du HCoV-229E.....	1066
Figure 23 : Spécificité des sérum E et F .....	1077
Figure 24 : Neutralisation de la protéine S du HCoV-229E par les sérums E et F .....	11010
Figure 25 : Modélisation de la protéine S du HCoV-229E .....	1144

# LISTE DES TABLEAUX

---

Tableau 1 : Genres, espèces, spectre d'hôtes et récepteurs des coronavirus abordés dans l'introduction .....	10
Tableau 2: Table récapitulative des séquons N-glycosylés .....	98
Tableau 3: Séquences des primers. ....	145
Tableau 4 : Numéro d'accession d'alphacoronavirus et betacoronavirus .....	147
Tableau 5: Analyses statistique des pentes .....	148

## RESUME

---

Le coronavirus humain 229E (HCoV-229E) est responsable de rhumes mais peut entraîner de graves complications respiratoires chez les personnes âgées ou atteintes d'une maladie chronique. Les coronavirus sont des virus enveloppés avec un génome à ARN positif simple brin. Trois protéines virales sont ancrées dans l'enveloppe virale : la protéine spike (S), la protéine de membrane (M) et la protéine d'enveloppe (E). Les protéines M et E sont impliquées dans l'assemblage viral et la sécrétion. La protéine S s'assemble en trimères à la surface des virions et joue un rôle-clé dans l'entrée du virus dans sa cellule-cible. Elle est constituée de deux domaines, le domaine S1 responsable de la liaison du virus à son récepteur et le domaine S2 responsable de la fusion de l'enveloppe virale avec une membrane cellulaire. La fusion est activée par des protéases cellulaires par clivage de la protéine S. Dans un premier temps nous avons caractérisé ce mécanisme. Pour cela, nous avons d'abord cloné la protéine S d'un isolat circulant de HCoV-229E. Nous avons analysé le clivage protéolytique de la protéine S par des sérine-protéases de type trypsine conduisant au processus de fusion à l'aide de particules pseudotypées rétrovirales. Les résidus arginine, sites potentiels de reconnaissance par les protéases et présents au niveau de la jonction S1/S2 ou de la région S2' ont été mutés individuellement (R565N, R679N, R683N ou R687N) afin d'étudier leur rôle lors de l'activation de la fusion. Contrairement à d'autres coronavirus, l'activation permettant la fusion de HCoV-229E semble être un processus en une seule étape. En effet, seule la mutation R683N inhibe l'infection médiée par des sérine-protéases et le clivage à l'interface S1/S2 ne semble pas être un pré-requis.

Les protéines S de coronavirus sont fortement N-glycosylées et constituent la principale cible des anticorps neutralisants. Nous avons analysé le rôle de la N-glycosylation du domaine S1 dans les mécanismes d'entrée et dans la neutralisation par des anticorps. L'analyse de la séquence de la protéine clonée montre la présence de 33 sites potentiels de N-glycosylation, dont 18 dans le domaine S1 qui ont été numérotés de N1 à N18. Ces 18 sites de N-glycosylation ont été abolis individuellement par mutagenèse dirigée. L'effet des mutations sur l'infectiosité virale a été évalué en utilisant des particules pseudotypées rétrovirales. L'infectiosité des mutants N6, N7 ou N9 est diminuée tandis que deux mutants N12 et N15 montrent une augmentation de l'infectiosité. Nous n'avons détecté aucune différence d'interaction de ces mutants avec une forme soluble du récepteur, l'aminopeptidase N (APN). Des expériences d'activation de la fusion virale à la surface cellulaire par la trypsine suggèrent que les glycanes présents aux positions 6, 7 et 9 sont impliquées dans la fusion virale, cependant nous n'avons détecté aucune différence de clivage de ces mutants par la trypsine. Pour le mutant N17 uniquement, la diminution partielle de l'infectiosité pourrait s'expliquer par une diminution de l'incorporation de la protéine S dans les pseudoparticules,

due au mauvais repliement de la protéine, comme le montre le profil du mutant en western blot en conditions réductrices ou non.

Nous avons ensuite évalué si les N-glycanes pouvaient moduler la reconnaissance de la protéine S par des anticorps neutralisants. Des pseudoparticules contenant les différents mutants ont été produites et utilisées pour infecter des cellules en présence d'anticorps neutralisants. Nos données montrent que les mutants N4, N10, N11, N12, N15, N16, N17, N18 réduisent la sensibilité des pseudoparticules à la neutralisation des anticorps. Dans ensemble, nos résultats suggèrent que les N-glycanes de la protéine S jouent un rôle important dans l'entrée virale et modulent la reconnaissance de la protéine par des anticorps neutralisants.

## ABSTRACT

---

The human coronavirus 229E (HCoV-229E) is a causative agent of common colds and can lead to severe respiratory complications in elderly persons and those with underlying disease. Coronavirus are enveloped viruses with a single stranded, positive-sense RNA genome. Three viral proteins are anchored in the viral envelope : the spike (S) protein, the membrane (M) protein and the envelope (E) protein. The M and E proteins are involved in viral assembly and secretion. The spike proteins assemble into trimers at the surface of the virions and play a key role in the early steps of viral infection. The spike protein comprised two domains, the S1 domain responsible for receptor binding and the S2 domain responsible for fusion of the viral envelope with the host cell membrane. Coronavirus fusion is activated by the proteolytic processing of the spike protein. First, we characterized the proteolytic processing of the HCoV-229E spike protein by trypsin-like serine-proteases. To do so, we first cloned the spike protein of a circulating isolate of HCoV-229E. To investigate the role of the S1/S2 junction and the specific role of the 3 arginine residues located in the S2' region in the proteolytic activation of HCoV-229E spike protein, the arginine residues present at these positions were mutated individually (R565N, R679N, R683N or R687N). Our results show that unlike other coronaviruses, HCoV-229E fusion activation appears to be a one step process. Indeed, the cleavage of the S1/S2 interface does not seem to be a pre-requisite, and the fusion activation strongly relies on the S2' region, with R683 acting as the cleavage site.

The spike protein is highly N-glycosylated and is the main target of neutralizing antibodies. We analysed the role of S1 domain N-glycosylation in the entry functions of the S protein and in neutralization by antibodies. Analysis of the sequence of the cloned protein shows the presence of 33 potential N-glycosylation sites, 18 being located in the S1 domain (numbered from N1 to N18). We mutated the 18 N-glycosylation sites of S1 individually by site-directed mutagenesis and studied the effect of the mutations using retroviral pseudotyped particles. Infectivity of the spike proteins with mutation either at the N6, N7 or N9 glycosylation site was strongly impaired. We did not detect any difference of interaction of these mutants with the soluble form of the receptor, the aminopeptidase N (APN). Results obtained by inducing the fusion of pseudoparticles at the cell surface with trypsin suggest that N-glycans located at the position N6, N7 and N9 are involved in viral fusion. However, the proteolytic processing of the protein required for fusion activation does not seem to be affected. Two mutants N12 and N15 show an increase of infectivity. Mutation of the N-glycosylation site N17 induces a partial decrease in infectivity. Indeed a decrease of spike protein incorporation into pseudoparticles was observed likely due to misfolding of the protein as shown by the profile of the mutant in western blot under reducing and non-reducing conditions. We next assessed if N-glycans can modulate the recognition of the spike protein by neutralizing antibodies. Pseudoparticles

harbouring the different mutants were produced and used to infect cells in presence or absence of neutralizing antibodies. Our data demonstrate that mutants N4, N10, N11, N12, N15, N16, N17, N18 reduce the sensitivity of pseudoparticules to antibody neutralization. Taken together our results suggest that N-glycans of the S protein play an important role in viral entry and modulate the recognition of the protein by neutralizing antibodies.

## ABBREVIATIONS

---

<b>Ac</b>	Anticorps
<b>ACE2</b>	Enzyme de conversion de l'angiotensine 2
<b>APN</b>	Aminopeptidase N
<b>ARN</b>	Acide ribonucléique
<b>ADN</b>	Acide desoxyribonucléique
<b>CCoV</b>	<i>Canine coronavirus</i>
<b>CEACAM1</b>	<i>Carcinoembryonic Antigen Cell Adhesion Molecule 1</i>
<b>ERGIC</b>	<i>ER-Golgi intermediate compartment</i>
<b>DC-SIGN</b>	<i>Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin</i>
<b>DENV</b>	<i>Dengue virus</i>
<b>DPP4</b>	Dipeptidyl peptidase-4
<b>FP</b>	<i>Fusion peptide</i>
<b>FCoV</b>	<i>Feline coronavirus</i>
<b>HAT</b>	<i>Human airway trypsin</i>
<b>HCoV-229E</b>	<i>Human coronavirus 229E</i>
<b>HCV</b>	<i>Hepatitis C virus</i>
<b>HIV</b>	<i>Human immunodeficiency virus</i>
<b>HKU1</b>	<i>Hong Kong University 1</i>
<b>HCoV-NL63</b>	<i>Human coronavirus NL63</i>
<b>HCoV-OC43</b>	<i>Human coronavirus OC43</i>
<b>HR</b>	<i>Heptad repeat</i>
<b>IBV</b>	<i>Infectious bronchitis virus</i>
<b>IFITM</b>	<i>Interferon-inducible Transmembrane Proteins</i>
<b>MERS-CoV</b>	<i>Middle East respiratory syndrome coronavirus</i>

<b>MHV</b>	<i>Mouse Hepatitis virus</i>
<b>Nsp</b>	<i>Non structural proteins</i>
<b>ORF</b>	<i>Open Reading Frame</i>
<b>PEDV</b>	<i>Porcine epidemic diarrhea virus</i>
<b>PDCoV</b>	<i>Porcine deltacoronavirus</i>
<b>PrCoV</b>	<i>Porcine respiratory coronavirus</i>
<b>SARS-CoV</b>	<i>Severe acute respiratory syndrome coronavirus</i>
<b>RBD</b>	<i>Receptor Binding Domain</i>
<b>RBM</b>	<i>Receptor Binding Motif</i>
<b>RE</b>	<i>Réticulum endoplasmique</i>
<b>RTC</b>	<i>Replicase-Transcription complex</i>
<b>SP</b>	<i>Signal peptide</i>
<b>TEM</b>	<i>Microdomaines enrichis en tétraspanines</i>
<b>TGEV</b>	<i>Transmissible gastro-enteritis coronavirus</i>
<b>TMPRSS2</b>	<i>Transmembrane protease, serine subfamily member 2</i>
<b>TM</b>	<i>Transmembrane domain</i>
<b>TTSP</b>	<i>Type II transmembrane serine protease</i>



# 1 INTRODUCTION

---

## 1.1 GENERALITES SUR LES CORONAVIRUS

### 1.1.1 Historique

Les coronavirus font partie de l'ordre des *Nidovirales*, de la famille des *Coronaviridae* et de la sous-famille des *Coronavirinae* qui regroupe 4 genres : *alphacoronavirus*, *betacoronavirus*, *gammacoronavirus* et *deltacoronavirus*. Les coronavirus infectent une grande diversité d'hôtes, incluant de nombreuses espèces d'oiseaux et de mammifères, dont l'Homme (Monto, 1974). Les *alphacoronavirus* et *betacoronavirus* circulent parmi les mammifères, les *gammacoronavirus* et les *deltacoronavirus* infectent les oiseaux et les mammifères (Tableau 1) (Masters, 2006).

Parmi les animaux, les coronavirus sont responsables de pathologies variées. Le *gammacoronavirus Infectious Bronchitis Virus* (IBV) est responsable d'infections respiratoires chez le poulet, et a un impact important pour l'industrie avicole (Cook et al., 2012). Le *betacoronavirus mouse hepatitis virus* (MHV) infecte les souris et peut provoquer des hépatites ou des encéphalites selon la souche virale (Weiss and Leibowitz, 2011). De nombreux coronavirus infectant les animaux sont responsables d'infections gastro-intestinales. Parmi ceux-ci, l'*alphacoronavirus porcine epidemic diarrhea coronavirus* (PEDV) a été à l'origine de la mort de nombreux porcelets dans les élevages porcins en Amérique et en Asie en 2013 (Lau and Chan, 2015). Au sein des *betacoronavirus*, on distingue 4 lignages : le lignage A auquel appartiennent les coronavirus humains HKU1 et OC43, le lignage B auquel appartient le SARS-CoV (pour *Severe Acute Respiratory Syndrome Coronavirus*), le lignage C auquel appartient le MERS-CoV (*Middle East Respiratory Syndrome Coronavirus*) et le lignage D auquel appartiennent les coronavirus de chauve-souris HKU4 et HKU5, proches du MERS-CoV.

**Tableau 1** : Genres, espèces, spectre d'hôtes et récepteurs des coronavirus abordés dans l'introduction

Genres	Espèce					
	Lignage	Isolat	Spectre d'hôtes	Récepteurs		
<i>Alphacoronavirus</i>	Coronavirus humain 229E		Homme	Aminopeptidase N humain (hAPN)		
	Coronavirus humain NL63		Homme	Enzyme de conversion de l'angiotensine (ACE2)		
	Virus de la diarrhée épidémique porcine (PEDV)		Porc	Aminopeptidase N porcine (pAPN)		
	Alphacoronavirus 1 dont		Virus de la gastroentérite transmissible (TGEV)	Porc	Aminopeptidase N porcine (pAPN)	
			Coronavirus félin (FCoV)	Chat	Aminopeptidase N félin (fAPN)	
		Coronavirus canin (CCoV)	Chien	Aminopeptidase N canin		
		Coronavirus respiratoire porcine (PrCoV)	Porc	Aminopeptidase N porcine (pAPN)		
<i>Betacoronavirus</i>	A	Betacoronavirus 1	Coronavirus humain OC43	Homme	Acide N-acetyl-9-O-acetylneuraminique (Neu 5,9 Ac2)	
			Bovine Coronavirus (BCoV)	Boeuf	Acide N-acetyl-9-O-acetylneuraminique (Neu 5,9 Ac2)	
	Coronavirus humain HKU1		Homme	Inconnu		
	Murine coronavirus	Murine hepatitis virus (MHV)	Souris	Carcinoembryonic Antigen Cell Adhesion Molecule 1 (CEACAM1)		
	B	Severe Acute Respiratory Syndrome (SARS) related coronavirus dont :	Human SARS-CoV	Homme	Enzyme de conversion de l'angiotensine (ACE2)	
		Coronavirus de chauve-souris Rhinolophus	Chauve-souris	Enzyme de conversion de l'angiotensine (ACE2)		

	C	Middle East Respiratory Syndrome-related coronavirus	Middle East Respiratory Syndrome Coronavirus (MERS-CoV)	Homme	Dipeptidyl peptidase-4 (DPP4)
	D	Bat coronavirus HKU4 (BtCoV-HKU4)		Chauve-souris	Dipeptidyl peptidase-4 (DPP4)
		Bat coronavirus HKU5 (BtCoV-HKU5)		Chauve-souris	Inconnu
<i>Gammacoronavirus</i>		Avian coronavirus	Infectious Bronchitis Virus (IBV)	Poulet	Inconnu
<i>Deltacoronavirus</i>		Coronavirus HKU15	Porcine deltacoronavirus PDCoV	Porc	Inconnu

Les premiers coronavirus humains 229E (HCoV-229E) et OC43 (HCoV-OC43) furent isolés dans les années 1960 et sont classés respectivement parmi les *alphacoronavirus* et *betacoronavirus*. En particulier, HCoV-229E a été découvert au Royaume-Uni dans des lavages nasals de personnes ayant un rhume (Hamre and Procknow, 1966). Parce que ces virus étaient peu pathogènes, cette famille de virus suscita peu d'intérêt en médecine humaine jusque dans les années 2000, ce qui explique le peu de données actuelles malgré l'ancienneté de leur découverte.

En 2002, un premier coronavirus hautement pathogène émergea dans la population humaine, le *betacoronavirus* SARS-CoV. Il fut responsable d'une épidémie de syndromes respiratoires aigus sévères qui débuta en Chine avant de se propager rapidement à travers le monde (Peiris et al., 2004). Malgré des mesures de quarantaine efficaces qui permirent de contenir l'épidémie, environ 8000 personnes furent infectées avec un taux de mortalité de 10% environ (WHO, 2004). A la suite de l'épidémie de SARS-CoV, les *betacoronavirus* humains NL63 (HCoV-NL63) et HKU1 ont été respectivement découverts en 2004 et en 2005 et sont responsables d'infections peu sévères (van der Hoek et al., 2004; Woo et al., 2005). De nouveau en 2012, un coronavirus fut isolé chez un patient atteint d'une pneumonie sévère (Zaki et al., 2012). Ce *betacoronavirus*, apparu en Arabie Saoudite, fut nommé *Middle East Respiratory Syndrome Coronavirus* (MERS-CoV). A ce jour, le MERS-

CoV est responsable d'une épidémie qui est toujours en cours, mais restreinte à la péninsule arabique, avec un nombre total de plus de 2000 cas recensés et un taux de mortalité d'environ 35% (WHO, 2018).

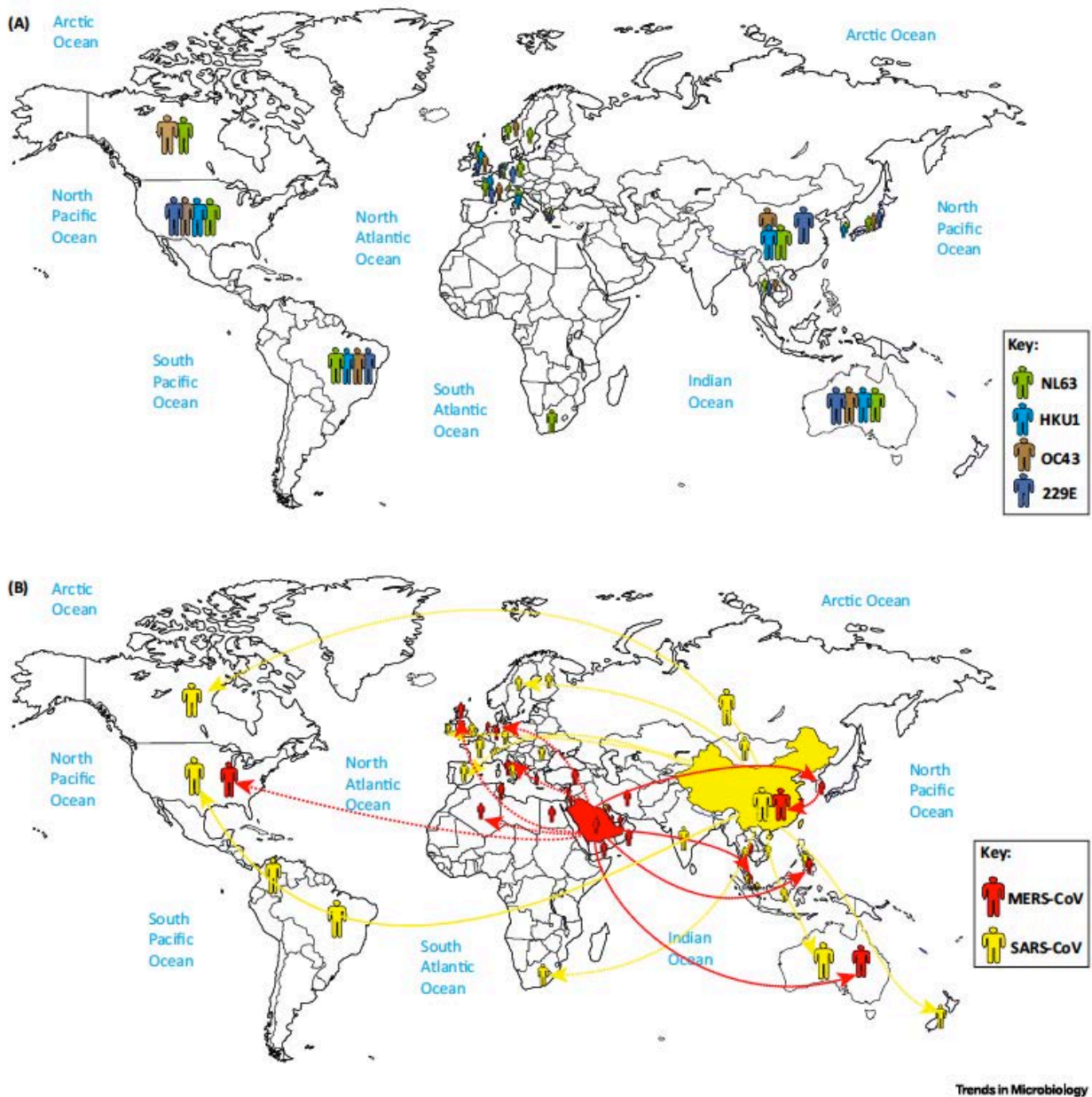
On distingue cliniquement les coronavirus humains faiblement pathogènes, dont font partie HCoV-229E, HCoV-NL63, HKU1 et OC43 et deux coronavirus humains hautement pathogènes, le MERS-CoV et le SARS-CoV.

### **1.1.2 Epidémiologie**

#### **Distribution géographique des coronavirus**

Les coronavirus ont une distribution mondiale. HCoV-229E et HCoV-OC43 sont identifiés depuis les années 1960 et ont été isolés sur tous les continents. Les coronavirus humains isolés plus récemment (HCoV-NL63 et HKU1) ont été décrits dans de nombreux pays. Des cas d'infections par HKU1 en dehors de l'Asie ont été détectés à New Haven aux Etats-Unis chez des enfants, ainsi qu'en Australie, en France, et au Brésil, indiquant une circulation mondiale du virus. HCoV-NL63 a été initialement isolé aux Pays-Bas fin 2004, puis a été isolé dans de nombreux pays (Figure 1A) (Su et al., 2016). L'épidémie causée par le SARS-CoV a débuté dans la province de Guangdong dans le sud de la Chine en novembre 2002, puis s'est répandue à d'autres pays en Asie, en Amérique du Nord et en Europe (37 pays/régions au total) pendant 9 mois (WHO, 2017). Le SARS-CoV ne circule plus actuellement (Figure 1B). Les pays autour de la péninsule arabique sont reconnus comme endémiques pour le MERS-CoV et l'Arabie Saoudite est le pays qui a signalé le plus de cas. Depuis sa découverte en 2012, 27 pays ont déclaré des cas d'infection par le MERS-CoV dont l'origine est souvent liée à un voyage dans la péninsule arabique. En 2015, l'exportation d'un cas d'infection en Corée a été à l'origine d'une épidémie secondaire de 186 cas, soulignant la capacité de ce coronavirus à disséminer rapidement dans une population naïve (Figure 1B) (de Wit et al., 2016).

## Réservoir et origine des coronavirus

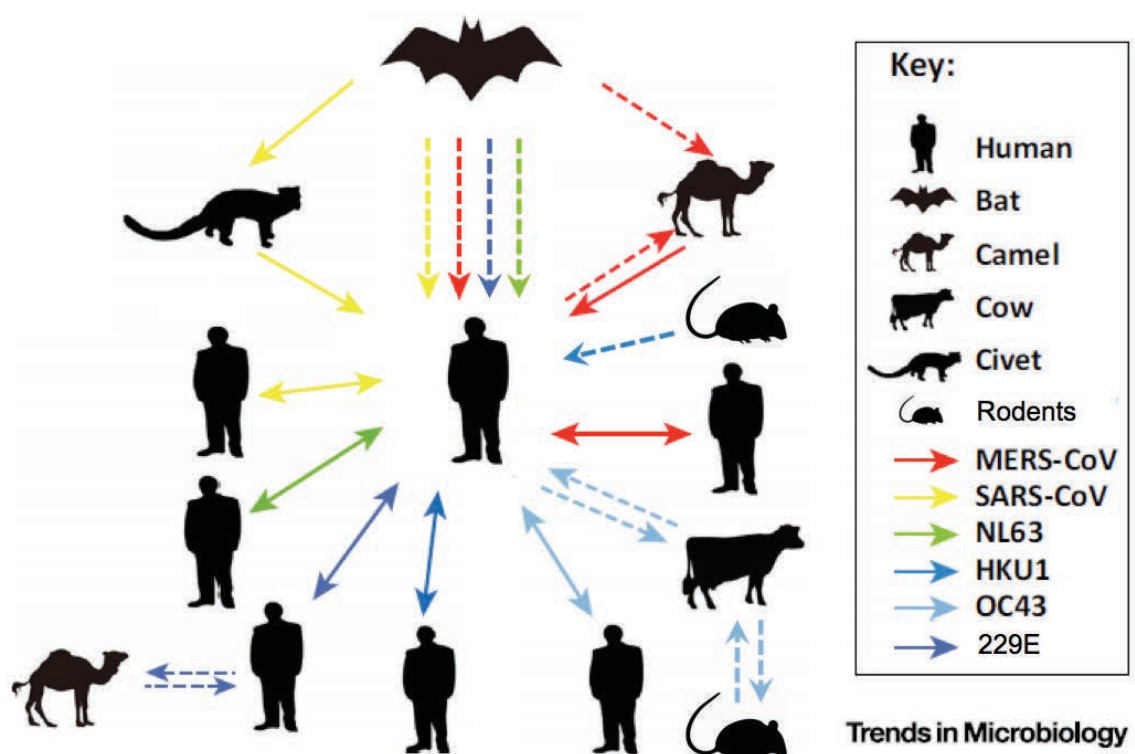


**Figure 1 : Distribution mondiale des coronavirus humains** (A) Les couleurs vertes, bleues, brunes et violettes représentent la distribution globale des coronavirus humains NL63, HKU1, OC43 et 229E respectivement. (B) Les couleurs rouges et jaunes représentent la distribution globale du MERS-CoV et du SARS-CoV respectivement (schéma issu de Su et al, 2016).

Les *alpha-* et *betacoronavirus* semblent avoir comme origine les chauve-souris, sauf certains *betacoronavirus* qui pourraient avoir émergé à partir de rongeurs (Lau et al., 2015a) tandis que les *delta-* et *gammacoronavirus* auraient pour origine les oiseaux (Woo et al., 2012). Il semble que les 4 genres de coronavirus aient divergé d'un ancêtre commun il y a 10 000 ans puis auraient co-évolué avec leurs hôtes (Woo et al., 2012). La date d'émergence des HCoV-NL63 et HCoV-229E dans la population humaine a été estimée entre 500 à 800 ans et de 200 ans respectivement. En ce qui concerne HCoV-OC43, il semble avoir émergé à partir d'un franchissement de barrière d'espèce du coronavirus bovin (BCoV) il y a 120 ans (Figure 2). HCoV-HKU1 semble avoir émergé dans les années 1950 (Chan et al., 2013 ; Forni et al., 2017).

Il a été montré que le SARS-CoV et le MERS-CoV ont des origines zoonotiques (Figure 2). L'introduction de ces virus dans la population humaine a probablement eu lieu *via* des espèces intermédiaires, mais cela reste débattu en ce qui concerne le SARS-CoV (Guan et al., 2003 ; Sharif-Yakan and Kanj, 2014). Le SARS-CoV appartient à l'espèce *SARS-related coronavirus* qui comprend aussi de nombreux virus de chauve-souris. En 2003, des virus proches du SARS-CoV ont été identifiés chez les chauve-souris, cependant ils n'étaient pas capables d'utiliser l'enzyme de conversion de l'angiotensine 2 (ACE2) humaine comme récepteur. Les *SARS-related coronavirus* ont été identifiés chez au moins 11 espèces de chauve-souris en Asie, en Afrique et en Europe (Lau et al., 2005; Li et al., 2005b; Pfefferle et al., 2009). Il a été suggéré que la civette palmée (*Paguma larvata*) puisse être un hôte intermédiaire dans lequel le virus se soit adapté pour finalement infecter les humains. Les *SARS-related coronavirus* de chauve-souris diffèrent notamment du SARS-CoV humain car ce sont des virus incapables d'utiliser le même récepteur cellulaire ACE2 (*angiotensin-converting enzyme 2*). De plus, il a été montré que l'ORF8 du SARS-CoV a une délétion de 29 nucléotides comparé à l'ORF8 des *SARS-related coronavirus* de chauve-souris (Lau et al., 2015b). Une étude récente suggère que les chauve-souris *Horseshoe* chinoises de la famille des *Rhinolophidae* pourraient être le réservoir naturel du SARS-CoV (Ge et al., 2013 ; Poon et al., 2005). Dans cette étude, deux *SARS-related coronavirus* ont été identifiés à partir de

prélèvements fécaux de chauve-souris, dont la séquence est la plus proche du virus humain jamais rapportée. Il a été montré qu'un des isolats (WIV1) était capable de reconnaître le récepteur ACE2 humain et de se répliquer dans certaines lignées de cellules humaines, ce qui suggère que ce virus puisse infecter l'Homme sans adaptation. La pathogénicité de cette souche n'a cependant pas été établie (Ge et al., 2013; Poon et al., 2005; Wu et al., 2016).



**Figure 2 : Transmission intra- et inter-espèces des coronavirus humains.** Les flèches rouges, jaunes, vertes, bleues, marrons et violettes représentent la transmission du MERS-CoV, SARS-CoV, HCoV-NL63, HCoV-HKU1, HCoV-OC43, et HCoV-229E respectivement, entre les chauves-souris, les dromadaires, les vaches, les humains et les civettes palmées (voir légende). Les flèches pleines représentent les transmissions confirmées entre deux espèces et les flèches en pointillé représentent les transmissions suspectées (Figure adaptée de Su et al, 2016).

Le MERS-CoV a probablement émergé à partir de coronavirus de chauve-souris (Annan et al., 2013 ; Lu et al., 2013 ; Memish et al., 2013). En effet, après séquençage du génome du MERS-CoV, les virus de chauves-souris BtCoV-HKU4 et BtCoV-HKU5 se sont révélés les plus proches phylogénétiquement même si ces virus ne sont pas des ancêtres directs (Wang et al., 2014). D'autre part, des études menées à Oman, en Arabie Saoudite, au Qatar, dans les Emirats Arabes Unis et en Jordanie ont montré que les dromadaires étaient séropositifs pour les anticorps neutralisants dirigés contre le MERS-CoV (Reusken et al., 2013). Des anticorps neutralisants ont ensuite été retrouvés chez les dromadaires du Moyen-Orient originaire d'Afrique, incluant l'Egypte, le Kenya, le Nigeria, l'Ethiopie, la Tunisie, la Somalie et le Soudan (Chan et al., 2015). Ces données suggèrent qu'il y a eu une première transmission d'un virus de chauves-souris vers les dromadaires et qu'après adaptation il ait émergé dans la population humaine. Aujourd'hui les virus circulant chez les dromadaires et chez l'Homme sont très proches suggérant que le dromadaire est un réservoir du virus. Cependant, de nombreux cas confirmés d'infection par le MERS-CoV n'ont pas d'historique de contact avec les dromadaires (Samara and Abdoun, 2014), même s'ils sont aujourd'hui considérés comme responsables de la réintroduction continue du virus dans la population humaine.

Les coronavirus humains 229E, OC43, NL63 et HKU1 sont bien adaptés à l'Homme et aucun de ces virus ne semble être maintenu à l'aide d'un réservoir animal. Un coronavirus très proche de HCoV-229E a cependant été identifié chez le dromadaire (Corman et al., 2016; Sabir et al., 2016). De manière générale, il est probable que la plupart des coronavirus humains viennent des chauves-souris (Drexler et al., 2014; Menachery et al., 2017). Les épidémies apparues au cours des 15 dernières années soulignent la capacité des coronavirus à franchir les barrières d'espèces pour causer des pathologies graves chez l'Homme et montre que cette famille de virus constitue un réservoir de pathogènes émergents (Corman et al., 2018; Sabir et al., 2016; Su et al., 2016).



## Transmission

Les coronavirus se transmettent principalement par aérosols de liquides biologiques et par contamination des surfaces. Ils infectent les voies aériennes humaines en entrant dans les cellules de l'épithélium respiratoire par le pôle apical. Les nouveaux virions sont également excrétés à partir du pôle apical, facilitant la dissémination du virus par la toux et les éternuements (Jonsdottir and Dijkman, 2016). Les coronavirus humains faiblement pathogènes (HCoV-NL63, HCoV-HKU1, HCoV-OC43 et HCoV-229E) restent détectables pendant 2 semaines dans les voies respiratoires (Elden et al., 2004), tandis que l'ARN du SARS-CoV est détectable jusqu'à un mois dans les sécrétions respiratoires, les urines et selles des patients après le début des signes cliniques (Chan et al., 2004). Le MERS-CoV semble persister dans les échantillons respiratoires jusqu'à 30 jours après infection (Poissy et al., 2014).

Des études ont montré qu'il existait une transmission interhumaine du MERS-CoV et du SARS-CoV, particulièrement en milieu hospitalier (Chowell et al., 2015; Guery et al., 2013; Hui et al., 2015; WHO, 2016). Le nombre de cas de MERS-CoV semble être limité, sauf dans le cas de l'épidémie de mai et juin 2015 en Corée du Sud, où la plupart des cas étaient secondaires voir tertiaires (WHO, 2015a). Des études portant sur la transmission et le potentiel épidémique du MERS-CoV ont montré que le taux de reproduction de base ( $R_0$  = nombre de nouveaux cas infectieux qu'un individu peut produire quand il est introduit dans une espèce susceptible) des patients infectés par le MERS-CoV est de 0,6 à 0,69 (Bauch and Oraby, 2013; Breban et al., 2013). Ce  $R_0 < 1$  suggère que le MERS-CoV n'a pas encore un potentiel pandémique. D'autres études ont suggéré que le  $R_0$  peut atteindre 0,8 à 1,3 en l'absence de mesure de contrôle de l'infection (Cauchemez et al., 2014)

## Prévalence et incidence

Il est couramment admis que les infections par les coronavirus ont lieu majoritairement entre la fin de l'hiver et le début du printemps, mais ces infections

ont aussi été rapportées à d'autres périodes de l'année (Gaunt et al., 2010; Vabret et al., 2009; Zhang et al., 2018). De manière générale, des études ont montré que les coronavirus humains faiblement pathogènes sont responsables de 15 à 30% des infections respiratoires chaque année (Fehr and Perlman, 2015; Mahony et al., 2011; Walsh et al., 2013), et de 5% de toutes les infections respiratoires aiguës d'enfants hospitalisés pendant l'hiver (Vabret et al., 2008).

### **1.1.3 Infection par les coronavirus**

#### **1.1.3.1 Clinique**

##### **Période d'incubation**

La période d'incubation des infections à HCoV-NL63, HCoV-HKU1, HCoV-OC43 et HCoV-229E est de 2 à 5 jours en moyenne tandis qu'elle est de 2 à 13 jours en moyenne pour le MERS-CoV et le SARS-CoV. Dans le cas du HCoV-229E, les symptômes induits par l'infection durent en moyenne 7 jours, mais peuvent parfois dépasser 18 jours (Bradburne et al., 1967; Falsey and Walsh, 2006; Su et al., 2016).

##### **Symptômes**

Les infections respiratoires induites par les coronavirus peuvent être hautes (rhinites, laryngites) ou basses (bronchites, bronchiolites ou pneumopathies) chez les humains, et ne sont cliniquement pas distinguables des infections respiratoires causées par d'autres pathogènes comme les rhinovirus ou le virus de la grippe A (Su et al., 2016). Des adultes volontaires en bonne santé furent infectés par HCoV-229E et il a été montré que 50% développent un rhume, les 50% restants ne développant pas de symptômes. L'infection est caractérisée par un malaise, des céphalées, une rhinorrhée, des frissons et des maux de gorge. Une petite portion des patients (10 à 20%) ont aussi montré de la toux et une fièvre (Bradburne et al., 1967). Les infections à HCoV-NL63, HCoV-HKU1 et HCoV-OC43 sont caractérisées par les mêmes symptômes que l'infection par le HCoV-229E. Une tachypnée, une hypoxie et des laryngites obstructives sont également fréquemment observées lors des infections à HCoV-NL63 (van der Hoek et al., 2006; Sahar Abdul-Rasool, 2010). Une

forte incidence de convulsions fébriles a été rapportée parmi les enfants infectés par HCoV-HKU1 et un cas de méningite a également été décrit (Lau et al., 2006; Woo et al., 2009). La plupart des infections à HCoV-HKU1 restent généralement bénignes, mais deux décès de patients atteints d'une pneumonie à HCoV-HKU1 ont été rapportés (Woo et al., 2005).

Les manifestations cliniques des infections par le SARS-CoV ou MERS-CoV sont plus sévères que les infections par les coronavirus faiblement pathogènes. Les patients infectés par le SARS-CoV présentaient initialement une fièvre, des myalgies, des maux de tête, un malaise et des frissons, suivis par une toux non productive, une dyspnée et une détresse respiratoire généralement 5 à 7 jours après le début des symptômes, ce qui peut entraîner la mort. Dans certains cas, une infection du foie, des reins et du cerveau a été rapportée (Weiss and Leibowitz, 2011). L'infection du tractus gastro-intestinal a également été décrite et environ un tiers des patients infectés présentaient des symptômes tels que des diarrhées et des vomissements. Les manifestations cliniques de l'infection par MERS-CoV vont d'une infection asymptomatique à une pneumonie sévère avec une détresse respiratoire aiguë et un choc septique (Zumla et al., 2015). L'insuffisance rénale aiguë est la caractéristique la plus frappante de la maladie causée par MERS-CoV, ce qui est jusqu'à présent unique pour les infections à coronavirus humains (Chan et al., 2015; Gao et al., 2016; de Wit et al., 2016). Une étude a rapporté des atteintes rénales dans 58% des cas (Arabi et al., 2014), bien que cette fréquence semble variable (Al-Abdallat et al., 2014; Guery et al., 2013). Les cas mortels d'infections à SARS-CoV et MERS-CoV ont lieu après respectivement 23 et 14 jours en moyenne. Le MERS-CoV semble avoir un taux de mortalité plus important (35%) que le SARS-CoV (10%), ce qui peut être expliqué par une prévalence importante des comorbidités parmi les patients infectés.

### **Terrains particuliers**

Les coronavirus humains faiblement pathogènes (HCoV-NL63, HCoV-HKU1, HCoV-OC43 et HCoV-229E) sont des virus que l'Homme rencontre tôt au cours de

sa vie : tous les enfants sont infectés pendant leur petite enfance, sans qu'il n'y ait nécessairement de symptômes (Prill et al., 2012). Globalement, il semble que l'infection se produise sans distinction de sexe, d'âge ou de lieu (Dijkman and van der Hoek, 2009; Garbino et al., 2006). Certains terrains favorisent les infections par les coronavirus humains comme les âges extrêmes de la vie, l'immunodépression et les pathologies respiratoires (Esper et al., 2005; Falsey and Walsh, 2006; Glezen et al., 2000; Jevšnik et al., 2012). Les pathologies respiratoires font partie des comorbidités jouant un rôle majeur dans la sévérité de l'infection par les coronavirus faiblement pathogènes. Les infections respiratoires virales ont été documentées dans 16% des épisodes aigus de maladies respiratoires tel que la bronchopathie chronique obstructive (BPCO) ou l'asthme. Il a été rapporté que 22,9% des 70 virus associés à des infections respiratoires chez les patients ayant une BPCO étaient dues à HCoV-229E ou HCoV-OC43. Les coronavirus sont aussi les virus les plus fréquemment identifiés lors d'infections respiratoires associées avec une nécessité d'hospitalisation (Beckham et al., 2005; Glezen et al., 2000; Greenberg et al., 2000). Il semble cependant que les coronavirus aient un impact relativement faible sur l'évolution de la mucoviscidose (Filho et al., 2012; Frickmann et al., 2012).

Il a également été montré que les personnes âgées sont plus susceptibles à la maladie provoqué par le SARS-CoV, avec un taux de mortalité supérieur à 50% (Weiss and Leibowitz, 2011). Les personnes atteintes de diabète, d'insuffisance rénale, de pneumopathie chronique ou d'immunodéficiences sont plus vulnérables à l'infection par le MERS-CoV : 75% des patients infectés ont au moins une autre comorbidité et les patients décédés étaient plus susceptibles d'avoir une pathologie chronique préexistante (Gao et al., 2016; de Wit et al., 2016; Yang et al., 2017).

### **1.1.3.2 Diagnostic**

Dans la plupart des cas, le diagnostic de l'infection à coronavirus n'est pas nécessaire, la maladie étant bénigne et se résolvant par elle-même. Toutefois, le diagnostic de la pneumonie virale nécessite une confirmation en laboratoire à cause

des ressemblances entre les symptômes cliniques des différents virus à tropisme respiratoire, comme le virus de la grippe (Falsey and Walsh, 2006). De plus, le diagnostic des infections à MERS-CoV et SARS-CoV est essentiel pour la gestion clinique et le contrôle épidémiologique des infections. De manière générale, le diagnostic viral inclut la culture, la détection rapide d'antigènes, la RT-PCR (*Reverse Transcription Polymerase Chain Reaction*) et les tests sérologiques, la disponibilité et la sensibilité de ces tests variant selon le virus étudié. Les prélèvements biologiques appropriés pour les tests incluent des aspirations rhinopharyngées, des écouvillonnages nasaux et trachéaux, des prélèvements d'expectorations et des prélèvements de lavages broncho-alvéolaires. Dans le cas du SARS-CoV, il est également possible d'effectuer des prélèvements de selles (Pozzetto and Groupe Révir, 2007).

Le diagnostic de l'infection par les coronavirus est effectué essentiellement par RT-PCR, méthode très sensible et spécifique (Chan et al., 2017; Mahony et al., 2011). Des RT-PCR multiplex en temps réel ont été développées qui permettent la détection des 4 coronavirus circulants (HCoV-229E, HCoV-NL63, HCoV-OC43 et HCoV-HKU1). La détection des antigènes viraux intracellulaires peut être réalisée par immunofluorescence indirecte sur frottis de cellules respiratoires. Il est possible d'utiliser le microscope électronique pour rechercher des virions dans des selles de patients, mais c'est une méthode qui n'est pas utilisée pour le diagnostic des infections virales. Le diagnostic sérologique n'est pas utile, les Ac apparaissant tardivement. La culture des coronavirus nécessite des conditions de croissance particulières demandant des installations de recherche pour l'isolement. Elle n'est pas utilisée pour le diagnostic (HCSP, 2013; Pozzetto and Groupe Révir, 2007). Le diagnostic d'une infection par le MERS-CoV nécessite souvent l'analyse de plusieurs échantillons prélevés sur plusieurs jours. C'est pourquoi une enquête doit être menée avant même d'avoir les résultats du laboratoire pour les cas suspects, lorsqu'un patient ayant une infection respiratoire aiguë sévère a voyagé au Moyen-Orient ou a été en contact avec des personnes infectées (WHO, 2015b).

### 1.1.3.3 Principes de la prise en charge et de la prévention

Aucun médicament antiviral ni vaccin n'est actuellement approuvé pour le traitement et la prévention des infections à coronavirus. Le traitement symptomatique ainsi que la prévention des complications restent les stratégies les plus importantes pour la gestion des infections par le SARS-CoV ou le MERS-CoV ainsi que lors de cas sévères d'infections par le HCoV-229E, le HCoV-NL63, le HCoV-HKU1 ou le HCoV-OC43. De nombreux composés peuvent inhiber la réplication et/ou l'entrée de coronavirus en culture cellulaire ou en modèle animal, cependant ces composés n'ont pas toujours prouvé leur efficacité chez l'humain. Dans le cas du SARS-CoV et du MERS-CoV, certains médicaments ont été utilisés de manière empirique ou évalués dans des essais cliniques non randomisés (Zumla et al., 2016). Ainsi des patients ont été traités avec différentes combinaisons de drogues incluant la ribavirine, des corticostéroïdes, l'interféron alpha ou des inhibiteurs de protéases. Cependant, le bénéfice de ces traitements pour les patients infectés par le SARS-CoV ou MERS-CoV n'est pas certain. Parmi les potentiels traitements étudiés, l'interféron de type I (IFN- $\alpha$ ) est capable de moduler la permissivité virale et l'efficacité de réplication en faisant basculer les cellules infectées et voisines dans un état antiviral. L'IFN- $\alpha$  présente une activité antivirale puissante vis-à-vis du HCoV-229E *in vitro* et *in vivo* (Djikman et Van der Hoek, 2009 ; Zumla et al, 2016). Cependant, l'administration intranasale prolongée d'IFN- $\alpha$  à des volontaires infectés par HCoV-229E a donné lieu à une décharge nasale teintée de sang, un effet secondaire qui est peut-être pire que le rhume causé par HCoV-229E (Tyrell et al, 1986). La ribavirine est un analogue nucléosidique de la guanosine agissant sur la réplication virale. Elle possède un large spectre antiviral. Elle semble être active contre le MERS-CoV et le SARS-CoV mais à haute dose, ce qui entraîne des effets secondaires sévères parmi les patients traités. De plus, le tableau clinique n'est pas amélioré significativement. Parmi les traitements testés, le lopinavir est un des plus intéressants. C'est un inhibiteur de protéase utilisé pour traiter les infections par le VIH (virus de l'immunodéficience humaine), habituellement commercialisés avec le ritonavir pour obtenir un effet « boost » (lopinavir-ritonavir). Le lopinavir et/ou le lopinavir-ritonavir ont une activité anti-coronavirus *in vitro* et *in vivo* chez les primates

non humains infectés par le MERS-CoV et dans les essais non randomisés de patients atteints par le SARS-CoV (Zumla et al., 2016).

La prévention de l'infection par les coronavirus est importante puisqu'aucun traitement ni vaccin n'est commercialisé. Il a été montré que l'infection par les coronavirus dont HCoV-229E dans l'enfance n'induit pas d'immunité protectrice (Callow et al., 1990). La réinfection a lieu quand le niveau d'anticorps dirigés contre le virus est faible. Il a été montré que les anticorps anti-HCoV-229E sont détectés jusqu'à un an après l'infection, indiquant que la protection conférée par ces anticorps est courte. Il est donc probable que chaque individu contracte de nombreuses infections par les coronavirus faiblement pathogènes au cours de sa vie. Pour le SARS-CoV et le MERS-CoV, le contact rapproché dans la vie courante ou lors de l'hospitalisation a été impliqué dans les infections interhumaines. Il semble que les patients ayant été infectés par le SARS-CoV gardent des anticorps dirigés contre le *receptor binding domain* de la protéine S pendant au moins 3 ans après l'infection (Cao et al., 2010). Dans le cas d'infections par les coronavirus hautement pathogènes, les patients atteints sont isolés et les agents de santé doivent porter des équipements de protection individuels (EPI) et avoir une hygiène des mains appropriée (WHO, 2015c). Les EPI comprennent une blouse, des gants propres et une protection faciale c'est-à-dire un masque médical plus une visière et des lunettes ou un écran facial. La prévention de l'infection par isolement respiratoire n'est pas nécessaire pour les coronavirus faiblement pathogènes (Falsey et Walsh 2006). Cependant, il est recommandé d'utiliser les mesures de précautions standard à l'hôpital (lavage des mains, matériel à usage unique etc) (Société Française d'Hygiène Hospitalière, 2017).

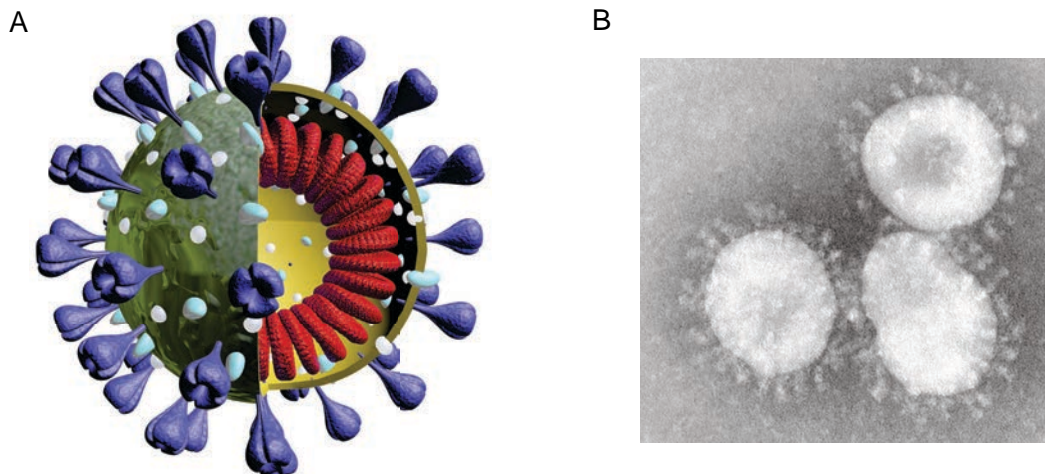




## 1.2 BIOLOGIE DES CORONAVIRUS

### 1.2.1 Structure et morphologie de la particule virale

Les coronavirus sont des virus sphériques enveloppés d'un diamètre de 80 à 120 nm (Masters, 2006). La capsid virale formée par la nucléoprotéine (N) et le génome est contenue dans l'enveloppe et est de symétrie hélicoïdale. A la surface des particules sont enchâssées trois protéines structurales, la protéine de membrane M, la protéine d'enveloppe E et la protéine S. La protéine S, également nommée *spike*, donne cet aspect de couronne en microscopie électronique et le nom de cette famille virale (Figure 3).



**Figure 3: Structure générale d'un coronavirus.** A : Schéma d'un virion. Les trimères de protéine S sont représentés en violet, la protéine M est en bleu, la protéine E en blanc et la nucléocapside en rouge. B : Virions de HCoV-229E infectieux purifiés.

## 1.2.2 Organisation du génome

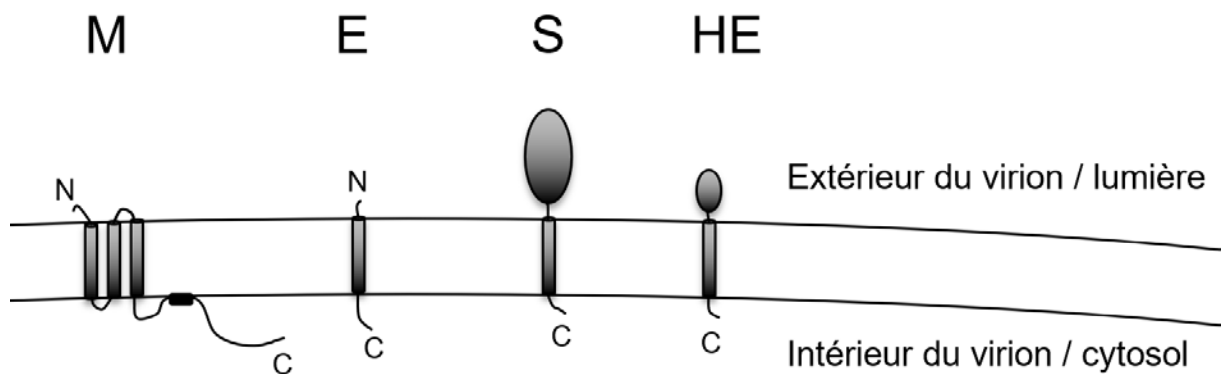
Le génome du HCoV-229E est un ARN simple brin de polarité positive de 27,2 kb (Farsani et al., 2012), qui le classe parmi les plus grands génomes de virus à ARN connus. Le génome possède une extrémité 5' coiffée et une queue polyadénylée (polyA) en 3'. L'extrémité 5' contient également une séquence régulatrice de transcription (*transcriptional regulatory sequence* TRS) leader et une région non traduite (*untranslated region* UTR) contenant plusieurs structures secondaires nécessaires à la réplication et à la transcription. Des structures ARN indispensables à la réplication et à la synthèse d'ARN viral sont également présentes dans l'UTR de l'extrémité 3'. Plusieurs cadres ouverts de lecture sont présents. Deux cadres ouverts de lecture ORF1a et ORF1b (*Open Reading Frame*) chevauchants occupent les deux premiers tiers du génome (Figure 5). Ils codent des polyprotéines clivées en 16 protéines non structurales (nsp) nécessaires à la réplication virale (Ponnusamy et al., 2006). Le dernier tiers du génome code les protéines structurales S, M, N et E et contient le gène codant l'unique protéine accessoire ORF4a du HCoV-229E (Zhang et al., 2014).

L'ordre des gènes, inchangé parmi tous les membres de la famille des coronavirus, est 5'-réplicase-S-E-M-N-3' (Figure 5). Chaque gène est précédé d'une séquence régulatrice de la transcription TRS nécessaire à l'expression de chacun de ces gènes (Fehr and Perlman, 2015). Les séquences TRS sont constituées d'une séquence « core » et de séquences régulatrices en 5' et 3' de cette séquence core. La séquence core du TRS leader en 5' du génome est identique à la séquence core des TRS localisé à l'extrémité 5' des gènes, ce qui joue un rôle important dans la synthèse discontinue des ARNs.

## 1.2.3 Protéines structurales

### 1.2.3.1 Protéine N

La protéine N est une protéine de 50 kDa qui encapside le génome viral pour former la nucléocapside hélicoïdale. La protéine N du TGEV (*Transmissible Gastro-Enteritis Virus*, un *alphacoronavirus* porcin proche du HCoV-229E) semble être phosphorylée (Calvo et al., 2005), bien que le rôle de cette phosphorylation soit encore mal connu. Il a été montré que la protéine N du HCoV-229E peut s'auto-assembler pour former des oligomères (Lo et al., 2013). La nucléocapside est entourée d'une bicouche lipidique dérivée de la cellule hôte, dans laquelle sont enchâssées les trois protéines d'enveloppe S, M et E.



**Figure 4: Topologie des protéines structurales présentes de l'enveloppe des coronavirus.** Toutes les protéines sont sous forme de monomère, mais les protéines S et HE forment respectivement des homotrimères et homodimères. L'extrémité N-terminale est signalée par « N » tandis que l'extrémité C-terminale est signalée par un « C ».

### 1.2.3.2 Protéine S

La protéine S joue un rôle majeur lors de l'entrée virale (Belouzard et al., 2012). C'est une protéine transmembranaire de type I fortement N-glycosylée, de 180 à 200 kDa qui s'assemble en homotrimères à la surface de la particule virale (Delmas and Laude, 1990). Elle possède un long domaine N-terminal et un court domaine C-terminal (Figure 4). Elle joue une double fonction dans l'entrée virale en permettant d'une part la liaison du récepteur cellulaire et d'autre part la fusion de l'enveloppe virale avec les membranes des cellules cibles. Elle a un rôle déterminant pour le tropisme cellulaire et pour la pathogénicité (Hulswit et al., 2016). Le rôle de cette protéine sera développé dans la partie 1.4.

### 1.2.3.3 Protéine M

La protéine de membrane M est la plus abondante de l'enveloppe et joue un rôle majeur dans l'assemblage du virion. Elle est capable d'interagir avec elle-même et toutes les autres protéines structurales : S, E et N (Ujike and Taguchi, 2015). Elle a un poids moléculaire de 25 à 30 kDa (Masters, 2006). L'extrémité N-terminale, également appelée ectodomaine, est localisée à l'extérieur du virion (Figure 4). Toutes les protéines M connues sont glycosylées au niveau de l'ectodomaine. Les protéines M des *alpha-* et *gammacoronavirus* sont N-glycosylées tandis que la plupart des protéines des *betacoronavirus* sont O-glycosylées. Le SARS-CoV, bien qu'étant un *betacoronavirus*, a une protéine M N-glycosylée (Oostra et al., 2006). L'ectodomaine est suivi de 3 passages transmembranaires. L'extrémité C-terminale, constituant environ la moitié de la protéine, se situe dans le virion, et est également nommée endodomaine. Cette topologie N-exo/C-endo de la protéine M est commune à la plupart des coronavirus, à l'exception du TGEV (*Transmissible gastroenteritis coronavirus*). Ce virus, en plus de la topologie N-exo/C-endo, peut adopter une topologie additionnelle C-exo (Risco et al., 1995). L'extrémité C-terminale de la protéine M des coronavirus contient un domaine amphipathique conservé, associé aux membranes et proche du troisième domaine transmembranaire (Masters, 2006).

#### **1.2.3.4 Protéine E**

La petite protéine d'enveloppe E a un rôle dans la sécrétion des virions. C'est une protéine peu abondante de 9 à 10 kDa. Elle a une extrémité N-terminale courte, suivie d'un domaine transmembranaire puis d'une extrémité C-terminale constituant la majorité de la protéine (Figure 4). La protéine E ne possède pas de peptide signal clivable ce qui suggère que c'est une protéine transmembranaire de type II (Martin J B Raamsman, 2000). Cependant il n'y a pas de consensus concernant la topologie de cette protéine qui reste controversée, avec plusieurs modèles rapportés. La protéine E du HCoV-229E a une activité de canal ionique *in vitro* dont le rôle dans le cycle viral reste à démontrer (Masters, 2006; Wilson et al., 2006).

#### **1.2.3.5 Protéine HE**

Certains *betacoronavirus* comme les *betacoronavirus* MHV ou OC43 possèdent une quatrième protéine structurale, l'hémagglutinine-estérase HE, à la surface des virions. Cette protéine transmembranaire de type I de 70 kDa possède un grand ectodomaine, un domaine transmembranaire et un endodomaine court. En tant qu'hémagglutinine, cette protéine peut se lier aux acides sialiques. De plus, elle possède une activité acétyl estérase clivant les groupements acétyl des acides 9-O- ou 4-O-neuraminiques acétylés (Masters, 2006).

### **1.2.4 Protéines non structurales**

Les protéines non structurales des coronavirus sont situées dans les deux premiers tiers du génome, composés de deux ORFs (ORF1a et ORF1b) chevauchantes. Ces deux ORFs sont traduites en deux polyprotéines, pp1a et pp1ab, qui sont elles-même clivées par deux protéases virales permettant la libération des 16 protéines non structurales (nsps) (Figure 6).

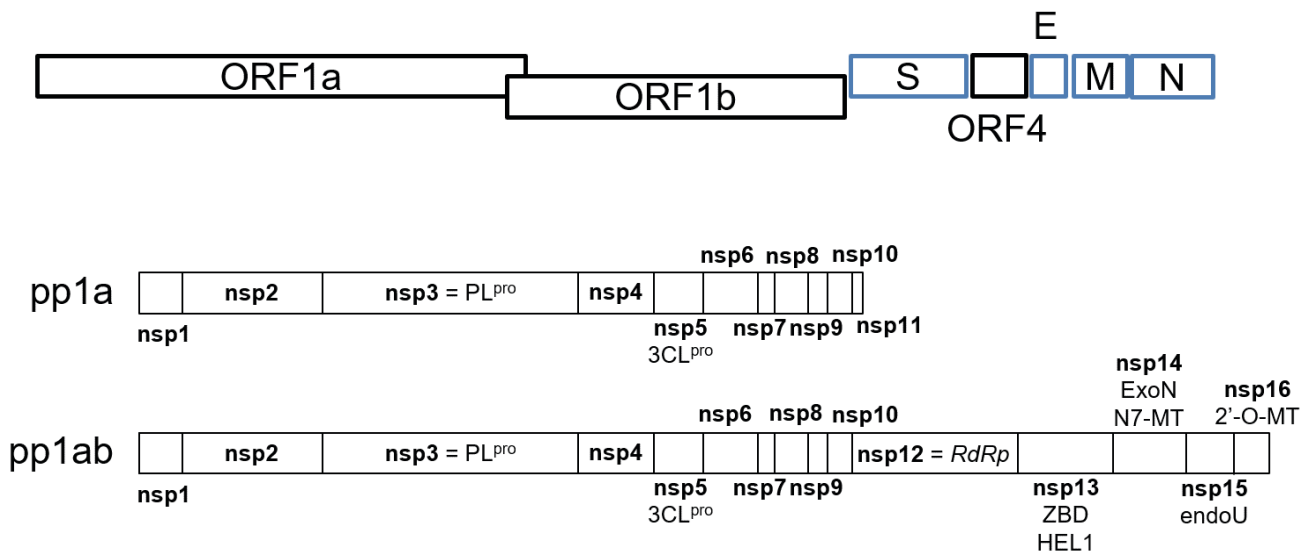
Il semble que **nsp1** ait un rôle dans la protection contre le système immunitaire en neutralisant la réponse immunitaire dépendante à l'interféron. Le rôle de **nsp2** est mal connu. La protéine **nsp3** contient des domaines protéases homologues à la papaine (*papaine-like protease* PL<sup>pro</sup>). Cette protéase permet la libération des nsp1, nsp2 et nsp3 par clivage de la polyprotéine. Elle aurait également un rôle majeur lors de la synthèse de l'ARN des coronavirus et est importante lors de l'ancrage dans le réseau membranaire, tout comme les protéines **nsp6** et **nsp4**. La deuxième protéase virale **nsp5** est une protéase de type sérine qui appartient à la famille de la chymotrypsine (*3C-like proteinase* 3CL<sup>pro</sup> ou *Main protease* M<sup>pro</sup>). Elle permet la libération des nsps 4 à 16 (Fehr and Perlman, 2015).

L'enzyme-clé lors de la réplication est la protéine **nsp12** qui est l'ARN polymérase ARN-dépendante (*RdRp*). Il semble que **nsp8** ait également une activité ARN polymérase ARN-dépendante, en association avec **nsp7**. **Nsp9** peut se lier à de l'ARN et de l'ADN et pourrait protéger l'ARN naissant au cours de la synthèse.

La coiffe présente à l'extrémité 5' des ARN est un facteur critique pour l'initiation de la synthèse protéique et la stabilisation des ARNm. Dans les cellules eucaryotes, la formation de la coiffe des ARNm par les enzymes cellulaires est co-transcriptionnelle et se fait dans le noyau. Trois activités enzymatiques sont essentielles pour la formation de la coiffe : une ARN 5'-triphosphatase, une ARN guanylyltransférase et une ARN méthyltransférase (Shuman, 2000). La réplication des coronavirus étant cytoplasmique, la coiffe est formée grâce à des protéines virales. La protéine **nsp10** fait parti du complexe de méthylation des ARNm et stabiliserait l'activité de **nsp16** (Denison et al., 2011). La protéine **nsp13**, qui a une activité ARN hélicase, est nécessaire pour le déroulement de substrats d'ARN ou d'ADN à partir de leur extrémité 5'. Elle possède une activité ARN 5'-triphosphatase probablement utile à la formation de la coiffe. Deux activités importantes pour la coiffe, l'activité N7-méthyltransférase (N7-MT) et l'activité 2'O-méthyltransférase (2'-O-MT) sont fournies par les **nsps 14 et 16** respectivement. La méthylation de coiffe sert de protection contre la dégradation par les nucléases de l'hôte.

La protéine **nsp15** est une endoribonucléase retrouvée chez tous les Nidovirus mais son rôle lors de la réplication virale est mal connu.

Enfin, la protéine **nsp14**, qui a une activité d'exoribonucléase (ExoN) en plus de son activité N7-méthyltransférase, assure la fidélité de la réplication. Elle joue un rôle critique dans les mécanismes assurant la stabilité génétique du génome viral, ce qui permet probablement aux coronavirus d'assurer la réplication de leur génome exceptionnellement long (Neuman et al., 2014). En effet, la substitution des résidus du site catalytique de l'exoribonucléase du SARS-CoV et le HCoV-229E par des alanines perturbe son activité (Ma et al., 2015; Minskaia et al., 2006). Il a été montré pour le MHV que des mutations de l'exoribonucléase permettaient d'obtenir des virus mutants viables, mais avec un taux de réplication et une synthèse d'ARN défectueuse. Le taux de mutation du virus est alors similaire aux autres virus à ARN qui ne possèdent pas d'exoribonucléase, ce qui montre l'importance de nsp14 pour la fidélité de la réplication (Eckerle et al., 2007). L'inactivation de l'exoribonucléase du SARS-CoV entraîne une diminution de la fidélité de la réplication également *in vivo*, et une diminution de la pathogénèse chez des souris infectées. Cette inactivation de l'exoribonucléase semble être stable *in vivo* (Graham et al., 2012). L'activité de l'exoribonucléase semble conférer une résistance du virus aux antiviraux mutagènes (analogues nucléosidiques), comme il a été montré pour le SARS-CoV, confirmant le rôle de correction des erreurs par nsp14 (Smith et al., 2013). Graepel *et al.* ont ensuite étudié l'impact de l'accumulation de passages *in vitro* de MHV dans lequel l'activité ribonucléase a été perturbée par des mutations. Trois hypothèses avaient été envisagées à terme : l'extinction du virus, une réversion des mutations de la ribonucléase ou une compensation de l'activité de correction des erreurs. Le virus a évolué en réduisant sa susceptibilité aux analogues nucléosides, ce qui est cohérent avec une augmentation de la fidélité de réplication. Cette résistance du virus aux analogues nucléosidique s'explique seulement en partie par des mutations acquises sur nsp12 et nsp14. D'autres protéines contribuent donc probablement à la régulation de la fidélité du MHV (Graepel et al., 2017).



**Figure 5: Organisation du génome de HCoV-229E.** En noir sont représentées les protéines non structurales et en bleu les protéines structurales. En dessous du génome sont représentées les polyprotéines pp1a et pp1ab obtenues par décalage du cadre de lecture. Les protéases nsp3 ( $PL^{pro}$ ) et nsp5 ( $3CL^{pro}$ ) vont permettre le clivage des polyprotéines. Les 16 protéines non structurales (nsp) obtenues sont indiquées ainsi que les domaines importants de la réplicase.  $PL^{pro}$ , papain-like proteinase (nsp3);  $3CL^{pro}$ , 3C-like proteinase (nsp5); RdRp, ARN polymérase ARN dépendante (nsp12); ZBD, zinc-binding domain (nsp13); HEL1, superfamily 1 helicase (nsp13); ExoN, exoribonucléase (nsp14); N7-MT, N7-methyl transférase (nsp14); endoU, uridylate-specific endoribonuclease (nsp15); 2'-O-MT, 2'-O-methyl transférase (nsp16).

### 1.2.5 Protéines accessoires

En plus des protéines structurales et non structurales, des ORFs accessoires spécifiques d'espèces peuvent être localisées entre les locus des gènes des différentes protéines structurales. A ce jour, une seule protéine accessoire a été décrite pour HCoV-229E : **ORF4a**. Le locus de cette protéine est situé entre les protéines structurales S et E, et est conservé parmi les coronavirus. Elle correspond à l'ORF3a du SARS-CoV pour laquelle il a été montré qu'elle pouvait être localisée à la surface cellulaire où il semble qu'elle forme des homotétramères (Lu et al., 2006). Afin de tester la fonction de l'ORF3a en tant que canal ionique, la protéine a été



exprimée dans des ovocytes de *Xenopus*, puis la différence de conductance entre les ovocytes contenant la protéine et les ovocytes contrôles a été évaluée (Plugge et al., 2000). Lu *et al.* ont montré que l'ORF3a formait un canal ionique sensible au potassium, mais sans montrer une sélectivité vis-à-vis du potassium. Il a ensuite été montré *in vitro* à l'aide de siRNA que l'inhibition de l'ORF3a diminuait l'exocytose du virus, sans modifier la réplication (Lu et al., 2006). L'étude *in vitro* de l'ORF3a a montré que cette protéine aurait une fonction pro-apoptotique caspase-dépendante (Law et al., 2005), et l'activité de canal ionique semble être indispensable pour l'apoptose (Chan et al., 2009). L'ORF3a pourrait être une viroporine. La protéine ORF4a du HCoV-229E est tronquée en comparaison de la protéine accessoire du SARS-CoV présente au même locus (Zeng et al., 2004). Elle n'a que peu été étudiée, mais il semblerait qu'elle présente plusieurs caractéristiques similaires à l'ORF3a du SARS-CoV. L'ORF4a du HCoV-229E peut se présenter sous forme d'oligomères, et il semble qu'elle ait un rôle de canal ionique non sélectif et puisse augmenter la perméabilité membranaire, régulant ainsi la production virale (Zhang et al., 2014).



## 1.3 CYCLE VIRAL

### 1.3.1 Entrée

La voie d'entrée des virus est définie par les interactions entre les particules virales et les récepteurs à la surface cellulaire (Belouzard et al., 2012). Après la liaison initiale au récepteur, les virus enveloppés doivent fusionner leur enveloppe avec les membranes de la cellule hôte pour libérer la nucléocapside dans la cellule cible (Figure 6). La fusion d'un virus enveloppé peut avoir lieu à la surface cellulaire ou dans des endosomes après internalisation de la particule virale. Pour les coronavirus, la fusion est dirigée par la protéine S qui subit de larges réarrangements conformationnels au cours de ce processus. L'activation de ces changements de conformation peut être déclenchée par différents stimuli : liaison du récepteur, exposition à un pH acide ou aux protéases cellulaires.

Dans certains cas, la liaison du coronavirus à son récepteur peut suffire à induire le changement de conformation de la protéine S nécessaire à la fusion, comme c'est le cas pour certaines souches du *betacoronavirus* MHV. Dans d'autres cas, l'activation de la fusion du coronavirus est dépendante du pH acide des endosomes de la cellule, ce qui implique une entrée par voie endosomale, comme pour le *gammacoronavirus* IBV. La plupart des coronavirus dépendent des protéases cellulaires pour leur entrée. Selon les coronavirus, ils peuvent utiliser une protéase ubiquitaire comme la furine, des protéases de la surface cellulaire de la famille des *Type II Transmembrane Serine Protease* (TTSP) ou des protéases endosomales telles que les cathepsines L ou B. Le site de fusion lors de l'entrée virale dépend probablement en grande partie des protéases utilisées. Il a été montré que le MHV (souche A59) entre par endocytose par la voie dépendante de la clathrine (Burkard et al., 2014).

HCoV-229E peut utiliser les protéases endosomales lors de l'entrée virale, comme la cathepsine L, mais il est possible que d'autres protéases endosomales

soient impliquées (Kawase et al., 2009). Il a aussi été montré pour le HCoV-229E que la fusion pourrait avoir lieu à la surface cellulaire en présence de protéases de la famille des TTSPs comme la *Transmembrane protease, serine subfamily member 2* (TMPRSS2) ou *Human airway trypsin* (HAT). En effet, l'utilisation de TMPRSS2 et HAT rend la fusion virus/cellule insensible à des inhibiteur de cathepsine L (Bertram et al., 2013). Il semble que cette voie soit plus avantageuse pour le virus que la voie endosomale (Shirato et al., 2016).

De nombreuses protéases utilisées par les coronavirus ainsi que leurs récepteurs sont transmembranaires et exprimés à la surface cellulaire. Il a été montré que la protéine S du MHV était clivée directement après liaison à son récepteur lorsque le virus entre à la surface cellulaire (Matsuyama and Taguchi, 2009). Ce résultat suggère que les protéases et les récepteurs doivent être à proximité l'un de l'autre. Une étude récente a montré que les protéases de la famille des TTSPs (TMPRSS2 et HAT) sont rassemblées au niveau de microdomaines enrichis en tétraspanines (TEMs) (Earnest et al., 2015). Les tétraspanines forment une famille de 34 protéines chez l'Homme et la souris. Elles sont formées de 4 domaines transmembranaires et deux boucles extracellulaires de tailles différentes. Les extrémités N- et C-terminales sont cytoplasmiques. La plupart des tétraspanines sont localisées à la surface cellulaire et il semble qu'elles soient organisées en réseau et plus particulièrement en TEMs. Les TEMs sont des plateformes fonctionnelles ordonnées, composées de tétraspanines, de lipides et de protéines du cytosquelette permettant de moduler l'adhésion, la migration et la différenciation cellulaire. D'autres protéines peuvent en faire partie, comme certains membres de la famille des intégrines et des immunoglobulines, certaines ectopeptidases et metalloprotéases ... (Charrin et al., 2009; van Deventer et al., 2017; Hemler, 2005; Yáñez-Mó et al., 2011). En particulier, la tétraspanine CD81 a été largement étudié pour son rôle lors de l'entrée du virus de l'hépatite C (HCV) et il a été montré que cette tétraspanine était un récepteur du HCV par liaison à la glycoprotéine E2 (Cocquerel et al., 2006).

Il a donc été suggéré que les TEMs pouvaient être le site d'entrée des coronavirus. En effet, en plus des protéases de la surface cellulaires, plusieurs récepteurs de coronavirus, tel que ACE2, APN (aminopeptidase N), CEACAM

(*Carcinoembryonic antigen-related cell adhesion molecule 1*) ou DPP4 (dipeptidyl peptidase-4) sont également majoritairement localisés au niveau des TEMs (Earnest et al., 2015). Les TEMs rassemblent donc les protéases et récepteurs nécessaires à l'entrée virale, et il a été montré que la tétraspanine CD9 jouait un rôle important dans l'infection des coronavirus médiée par les TTSPs (Earnest et al., 2015, 2017). En effet dans des cellules 293β5 exprimant la protéase TMPRSS2 et les récepteurs de HCoV-229E, SARS-CoV et MERS-CoV, l'entrée virale de ces différents virus est inhibée par des anticorps dirigés contre les tétraspanines CD9, CD63 ou CD81. Cependant la surexpression de TMPRSS2 permet de restaurer l'infection. Des lignées cellulaires déficientes en CD9 ou CD81 ont ensuite été générées grâce à la technologie CRISPR/Cas9 pour confirmer les résultats précédents. Aucune différence d'entrée virale n'a été observée en absence de CD81. En ce qui concerne la lignée déficiente en CD9, l'entrée des particules pseudotypées avec la protéine S du MERS-CoV ou de HCoV-229E est inhibée, tandis que celle des pseudoparticules MHV et SARS-CoV n'est pas affectée. L'infection par des pseudoparticules dans ces lignées déficientes en CD9 est aussi restaurée par une surexpression de TMPRSS2. Les auteurs ont montré que le complexe CD9-hDPP4-TMPRSS2 favorise une entrée précoce et donc plus efficace du virus MERS-CoV. Le rôle de CD9 et TMPRSS2 pour l'infection par le MERS-CoV a aussi été confirmé *in vivo* en modèle animal (Earnest et al., 2017). Il est intéressant de noter que les interactions entre le récepteur DPP4 du MERS-CoV et la tétraspanine CD9 sont semblables aux interactions existant entre le récepteur APN du HCoV-229E et CD9 (Zöller, 2008). Les mécanismes d'entrée médiés par la protéine S seront détaillés dans la partie 1.4.

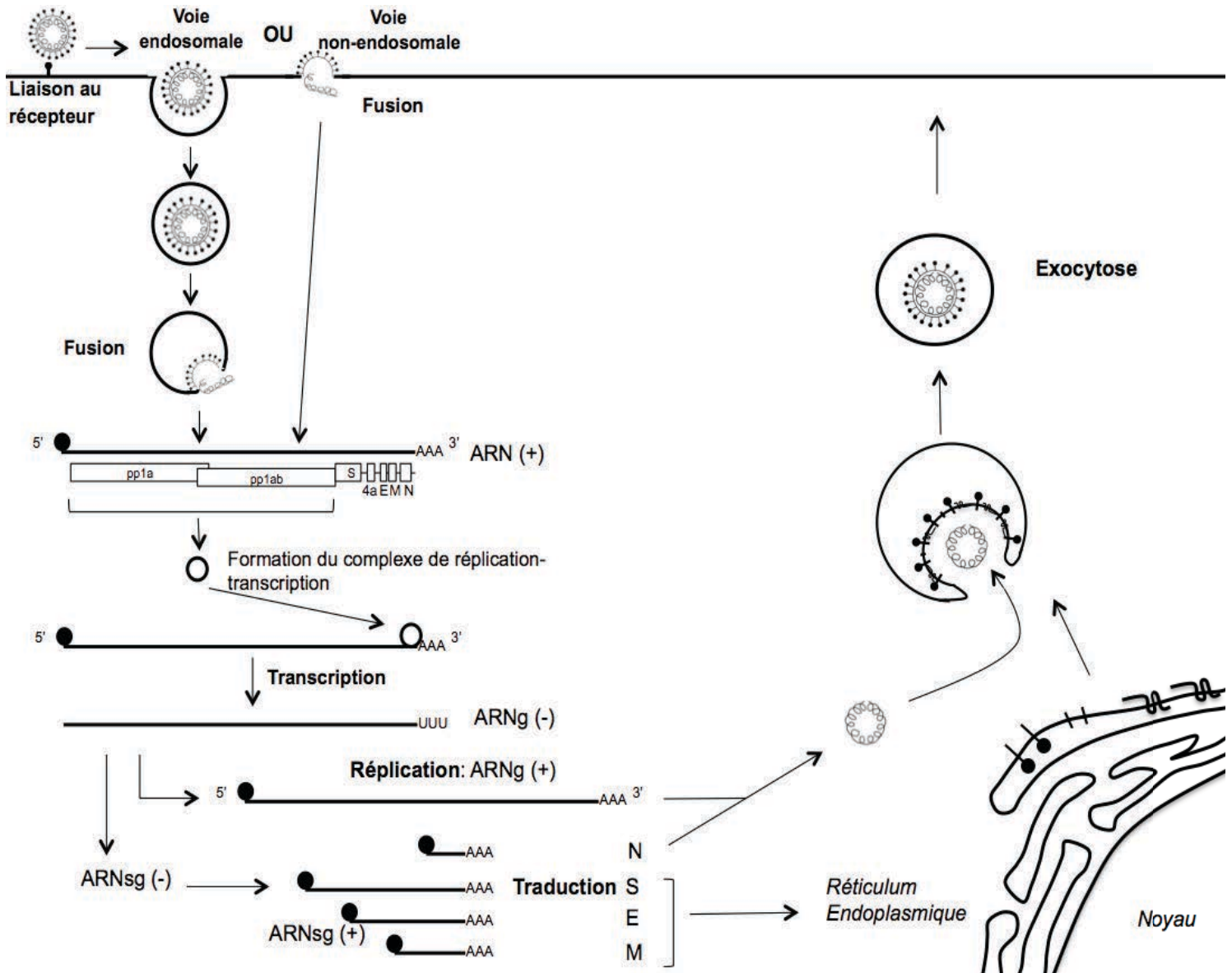
### 1.3.2 Réplication du génome viral et synthèse des protéines virales

Une fois dans le cytoplasme, le génome viral doit être répliqué et traduit afin de produire de nouveaux virions. La première étape de la réplication est la traduction du gène de la réplicase à partir de l'ARN viral. Ce gène est organisé en deux cadres de lecture chevauchants ORF1a et ORF1b (Figure 5) qui sont traduits en deux polyprotéines pp1a et pp1ab par la machinerie cellulaire (Figure 6). La traduction de l'ORF1b est permise par un décalage du cadre de lecture (*frameshift*) de type (-1)

entraînant un prolongement de la polyprotéine pp1a en une polyprotéine plus longue pp1ab. Ce glissement ribosomal du cadre de lecture nécessite une séquence « glissante » de 8 nucléotides et une structure en pseudonoeud situées juste en amont du codon stop de l'ORF1a (Sola et al., 2015). Les deux polyprotéines pp1a et pp1ab subissent ensuite une maturation protéolytique permettant la production de 16 protéines non structurales par les protéases virales nsp3 et nsp5 (Fehr and Perlman, 2015).

Après la maturation des polyprotéines, la plupart des nsps s'assemblent pour former le complexe de réplication-transcription (RTC). Le RTC permet la réplication et la transcription du génome viral. Il a été montré pour le SARS-CoV que le RTC est localisé dans un réseau membranaire issu des membranes du réticulum endoplasmique. L'ARN double brin viral serait localisé dans des vésicules à double membrane (DMV) (Knoops et al., 2008). Parmi les nsps présentes dans le complexe RTC, l'enzyme-clé est la protéine nsp12 qui est l'ARN polymérase ARN-dépendante (*RdRp*). Les protéines non structurales peuvent avoir un rôle soit lors de la synthèse de l'ARN en permettant sa formation ou sa modification, ou lors des interactions entre le virus et l'hôte en protégeant par exemple l'ARN (Neuman et al., 2014; Snijder et al., 2016; Subissi et al., 2012).

Après l'assemblage du RTC, la synthèse d'un ARN viral (-) intermédiaire permet la production d'ARN génomique de pleine longueur. La synthèse d'ARN génomique est nommée réplication. D'autre part, lors de la synthèse d'ARN viral (-), un mécanisme de synthèse discontinue, appelé transcription va permettre la production d'ARNs subgénomique (ARNsg) (Figure 6). La séquence TRS leader en amont du gène va s'hybrider avec la séquence TRS située à l'extrémité 5' du génome. Il semble que la *RdRp* s'arrête à chaque séquence TRS puis continue l'élongation jusqu'à la prochaine séquence TRS ou « saute » pour aller amplifier la séquence leader à l'extrémité 5' du génome. Les différents ARN subgénomiques obtenus codent la protéine accessoire ORF4a et les protéines structurales (Fehr and Perlman, 2015; Sola et al., 2015; Subissi et al., 2012).



**Figure 6: Schéma général du cycle viral des coronavirus.** Le virus entre dans la cellule cible après reconnaissance du récepteur cellulaire. La fusion entre les membranes virale et cellulaire a lieu après endocytose ou directement à la surface. Le génome viral est ensuite relargué dans le cytoplasme où il est traduit par la machinerie cellulaire. Deux polyprotéines (pp1a et pp1ab) sont produites et maturées par les protéases virales nsp3 et nsp5. Les 16 protéines non structurales (nsps) ainsi libérées s'associent pour former le complexe de transcription-réplication (RTC). Le RTC va alors transcrire les ARNm subgénomiques (ARNsg) codant les protéines structurales et la protéine accessoire et également répliquer l'ARN génomique (ARNg). L'ARN génomique est encapsidé par la protéine N et va acquérir son enveloppe après bourgeonnement dans le compartiment ERGIC (Endoplasmic Reticulum-Golgi Intermediate Compartment).

Les coronavirus sont connus pour leur capacité de recombinaison, que ce soit de manière homologue ou non homologue. Cette capacité est liée à la capacité de la *RdRp* de passer d'un brin d'ARN à un autre. La recombinaison joue un rôle majeur dans l'évolution virale (Fehr and Perlman, 2015).

### 1.3.3 Assemblage et exocytose

Après la réplication du génome viral et la synthèse des ARN subgénomiques, les protéines structurales S, M et E sont traduites et insérées dans la membrane du réticulum endoplasmique (RE). Elles migrent le long de la voie de sécrétion jusqu'au compartiment intermédiaire entre l'appareil de Golgi et le réticulum endoplasmique (ERGIC) (Figure 6) où a lieu l'assemblage de la particule virale. La protéine M dirige la majorité des interactions entre les protéines structurales nécessaires à l'assemblage et jouerait un rôle central dans l'organisation de cet assemblage. Ces mécanismes restent cependant encore mal connus, particulièrement le recrutement des génomes au niveau du ERGIC. Le génome est encapsidé par la protéine N et bourgeonne à travers les membranes portant les protéines d'enveloppe, pour former les particules virales. Ces dernières suivent la voie de sécrétion habituelle avant d'être libérées dans le milieu extracellulaire par exocytose (de Haan and Rottier, 2005; Travis R Ruch, 2012). Sur une particule virale, il semble que le ratio entre les différentes protéines structurales pour le SARS-CoV soient de  $1S_3/16M/4N$  à  $1S_3/25M/4N$  (où  $S_3$  correspond à des trimères de protéine S) (Neuman et al., 2006).



## 1.4 PROTEINE S

### 1.4.1 Généralités

La protéine S du HCoV-229E est une protéine transmembranaire de type I. Elle est fortement N-glycosylée : elle possède de 26 à 33 sites potentiels de N-glycosylation selon les souches (Farsani et al., 2012; Raabe et al., 1990; Shirato et al., 2012). La protéine S s'assemble en trimères lors de sa synthèse dans la cellule. Elle est ensuite incorporée dans les particules virales au niveau du site d'assemblage. Il a été montré pour le SARS-CoV qu'une particule virale pouvait contenir de 50 à 100 trimères de protéine S (Neuman et al., 2006).

Sur le plan fonctionnel, la protéine S du HCoV-229E est divisée en deux domaines S1 et S2 (Figure 7A). Toutes les protéines S des différents coronavirus partagent cette organisation en deux domaines, cependant une différence réside dans leur clivage. Certains coronavirus comme le MHV possèdent un site de clivage par la furine entre S1 et S2 tandis que d'autres virus comme le HCoV-229E arborent une protéine non clivée à leur surface. La région S1, située du côté N-terminal, permet la reconnaissance et la liaison à son récepteur cellulaire, l'aminopeptidase N (APN) pour HCoV-229E, et détermine en grande partie le tropisme cellulaire et l'hôte du virus. Le domaine S1 est également un déterminant antigénique majeur (Shirato et al., 2012). La région S2 est située du côté C-terminal et est impliquée dans la fusion de l'enveloppe du virus avec les membranes cellulaires. C'est une protéine de fusion de classe I (Liu et al., 2006).

## 1.4.2 Domaine S1 : Attachement et reconnaissance antigénique

### 1.4.2.1 Structure du domaine S1

Le domaine S1 du HCoV-229E est situé à l'extrémité N-terminale de la protéine S et est constitué de 565 acides aminés. En 2016 et 2017, les structures des protéines S des virus MHV, HKU1 et HCoV-NL63 ont été déterminées en cryomicroscopie électronique (Kirchdoerfer et al., 2016; Ou et al., 2017; Walls et al., 2016b, 2016a). Du point de vue structural, le domaine S1 est constitué de 4 domaines A, B, C et D (Figure 7B). Pour le virus HCoV-NL63, la protéine S contient en plus un domaine 0 en amont du domaine A à l'extrémité N-terminale de la protéine. Ce domaine est présent sur la protéine S de nombreux *alphacoronavirus* (Hulswit et al., 2016), mais pas sur celle du HCoV-229E. La structure de trimères de protéine S des *betacoronavirus* MHV et HKU1 montre que les domaines B de chaque protomère restent proche de l'axe central. Les domaines A sont plus éloignés du centre (Figure 7D). La conformation des trimères de protéine S de l'*alphacoronavirus* NL63 est similaire à celle des *betacoronavirus*, bien que la vue du sommet d'un trimère montre que les domaines A et B de chaque protéine sont regroupés (Figure 7C).

Les domaines A et B contiennent le domaine de liaison au récepteur cellulaire (RBD pour *Receptor Binding Domain*). Les structures du RBD complexé au récepteur ont été résolues pour plusieurs coronavirus dont le SARS-CoV et HCoV-NL63 (Lu et al., 2013; Wu et al., 2009; Yuan et al., 2017). La position du RBD peut varier parmi les coronavirus, celui-ci étant situé dans la première moitié N-terminale du domaine S1 (domaine A) pour le *betacoronavirus* murin MHV (Murine Hepatitis Virus), ou dans le domaine C-terminal pour HCoV-229E, HCoV-NL63 ou le SARS-CoV (domaine B). Au sein du RBD, les résidus qui permettent d'établir des interactions avec des résidus du récepteur constituent le motif de liaison au récepteur (RBM pour *Receptor Binding Motif*).

La position du RBD du HCoV-229E n'est pas connue avec certitude. En effet, plusieurs études ont identifié différentes régions comme étant le RBD. Afin de déterminer les régions de la protéine S impliquées dans l'interaction avec hAPN, diverses méthodes ont été utilisées. Différentes protéines S tronquées, solubles et purifiées ont été exprimées à l'aide d'un système d'expression baculovirus et la liaison de ces protéines avec le récepteur a été déterminée, soit en utilisant une forme soluble du récepteur, soit en exprimant le récepteur à la surface des cellules (Bonavia et al., 2003; Breslin et al., 2003). Ces deux études ont identifié la même région entre les résidus 407 à 547 comme importante pour la liaison du récepteur. L'interaction entre la protéine S du HCoV-229E et hAPN a ensuite été étudiée à l'aide de protéines S chimères dont certains domaines de la protéine S du HCoV-229E ont été remplacés par des domaines correspondant de la protéine S du HCoV-NL63. L'interaction avec le récepteur a été déterminé par cytométrie en flux. Les auteurs ont montré que l'interaction avec le récepteur dépendait de l'extrémité C-terminale du domaine S1 ainsi que des acides aminés 278 à 329 (Hofmann et al., 2006). Récemment, une étude du complexe hAPN – RBD du HCoV-229E par rayon X a été réalisée. La structure cristalline du complexe montre que la liaison se fait par trois boucles (*extended loop*). Ces boucles sont situées entre les résidus 308 – 325 ; 352 – 359 et 404 – 408 (Wong et al., 2017).

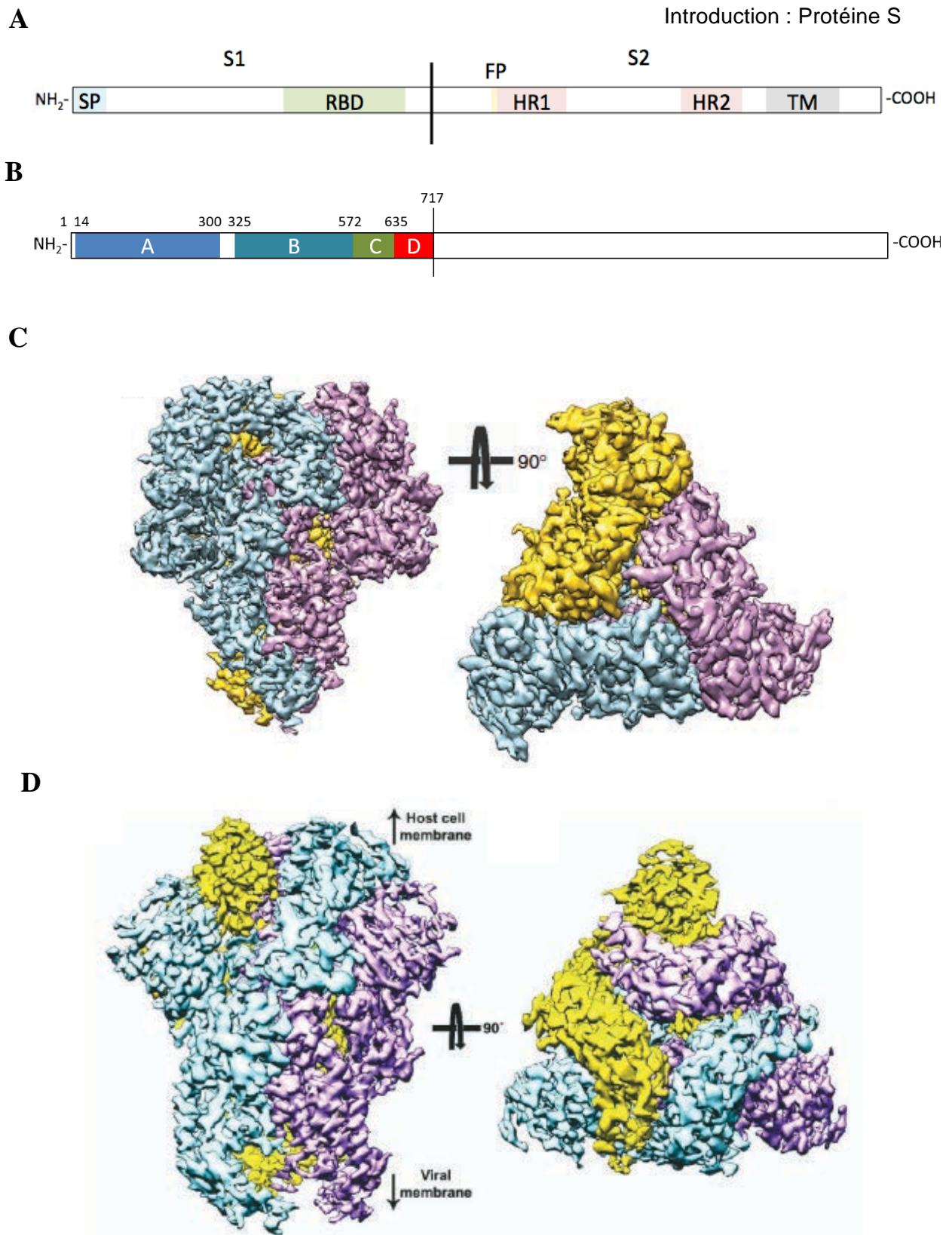
La séquence en acide aminé du domaine S1 est peu conservée, pourtant la structure tertiaire du domaine C-terminal et plus particulièrement du RBD est similaire pour les coronavirus appartenant à un même genre (Li, 2016). Le domaine B contenant le RBD des protéines S des *alphacoronavirus* aurait divergé en une structure «  $\beta$ -sandwich » constituée de deux feuillets de trois brins  $\beta$  tandis que celui des protéines S des *betacoronavirus* aurait divergé en une structure composée d'un feuillet de cinq brins  $\beta$ . Ce type de structure a été décrite pour la protéine S des *betacoronavirus* MHV (Walls et al., 2016b), HKU1 (Kirchdoerfer et al., 2016), SARS-CoV (Li et al., 2005a), et MERS-CoV (Wang et al., 2013a). Il a également été montré pour ces *betacoronavirus* que le domaine A se replie pour former une structure de type «  $\beta$ -sandwich » tandis que les domaines C et D sont riches en feuillets  $\beta$  et proches de la tige S2. Si la structure tridimensionnelle du domaine de liaison des

coronavirus semble conservée, des modifications de structure des motifs de liaison permettraient de reconnaître les récepteurs spécifiques de chaque virus (Chen et al., 2013; Jiang et al., 2012; Peng et al., 2011; Wang et al., 2013a).

#### **1.4.2.2 Attachement viral et récepteurs cellulaires**

L'attachement du virion à la cellule hôte est initié par les interactions entre le RBD de la protéine S et son récepteur, ce qui détermine le spectre d'hôte mais aussi le tropisme tissulaire du coronavirus. Le SARS-CoV et HCoV-NL63 utilisent l'enzyme de conversion de l'angiotensine 2 (ACE2) pour entrer dans les cellules hôtes (Hofmann et al., 2005; Li et al., 2003), MHV entre après liaison à CEACAM1 (Williams et al., 1991), et le MERS-CoV se lie à DPP4 pour entrer dans les cellules (Raj et al., 2013). De nombreux *alphacoronavirus* utilisent l'aminopeptidase N (APN) comme récepteur, dont le coronavirus humain 229E (Tableau 1) (Yeager et al., 1992).

APN (ou CD13) est une glycoprotéine transmembranaire de type II de 150 à 160 kDa. C'est une protéase zinc-dépendante qui clive les acides aminés N-terminaux des peptides actifs biologiquement. Elle est exprimée comme homodimère sur la membrane apicale des cellules épithéliales des tractus respiratoires et entériques, sur les cellules endothéliales, au niveau des jonctions synaptiques des cellules rénales et sur les cellules du système immunitaire (monocytes, cellules dendritiques et granulocytes) (Wong et al., 2012). Elle joue un rôle pivot dans de nombreux processus physiologiques, et ses caractéristiques structurales lui permettent de fonctionner de façon ubiquitaire dans le métabolisme peptidique et d'interagir avec d'autres protéines pour agir sur la motilité cellulaire et l'adhésion (Mina-Osorio, 2008). APN est également une cible potentielle importante pour le traitement de certains cancers (Hitzerd et al., 2014).



**Figure 7 : Schéma fonctionnel et représentation 3D de la protéine S des coronavirus.**  
 A. Schéma général de la protéine S. SP peptide signal, RBD Receptor Binding Domain, FP peptide de fusion, HR1 et HR2 heptad repeat domain 1 et 2, TM domaine transmembranaire. B. Représentation des domaines A, B, C et D de la protéine S du MCoV (selon Walls et al, 2016b). C et D. Structure 3D de trimères de la protéine S de HCoV-NL63 (C) et de MCoV (D) résolue par cryo-microscopie électronique. Les protomères de protéine S sont de couleur bleu, jaune ou rose. Les trimères sont présentés vus de côté (gauche) ou vus d'en haut (droite) (d'après Walls et al, 2016a et b).

L'aminopeptidase N est le récepteur de plusieurs *alphacoronavirus*, comme le coronavirus humain 229E, et les coronavirus animaux TGEV (*Transmissible gastro-enteritis coronavirus*), FCoV (*Feline coronavirus*), CCoV (*Canine coronavirus*) et PRCoV (*Porcine respiratory coronavirus*). Les protéines APN originaires de l'Homme, de la souris, du rat, du lapin, du porc, de la vache, du chat et du chien ont une séquence en acides aminés hautement conservée (70 à 80% d'homologie). Cependant, le virus HCoV-229E n'est capable de lier que l'APN humaine et par exemple, TGEV ne peut lier que l'aminopeptidase porcine. La restriction d'espèce de la liaison à APN a été étudiée. L'introduction dans hAPN d'un séquen encodant un site potentiel de N-glycosylation au niveau de l'acide aminé 291, comme celui présent sur l'APN porcine, ou entre les acides aminés 288 à 290, comme observé sur l'APN murine, abolit son activité de récepteur pour HCoV-229E (Tusell et al., 2007; Wentworth and Holmes, 2001). Concernant les sites d'interactions entre le RBD des coronavirus et APN, il a été montré qu'HCoV-229E, TGEV, FCoV et CCoV étaient capables de se lier à différents sites d'APN présents à l'extérieur des cellules, sans que les fonctions physiologiques du récepteur, dont l'activité enzymatique, ne soient affectées dans la plupart des cas (Chen et al., 2012; Delmas et al., 1994). Comme pour HCoV-229E, l'*alphacoronavirus* respiratoire porcin, PRCoV, utilise également APN comme récepteur, bien que ces coronavirus n'interagissent pas avec les mêmes sites d'APN à cause de différences de conformation de leur RBD. Deux résidus majeurs la tyrosine 528 (Tyr528) et le tryptophane 571 (Trp571) de la protéine S de PRCoV et TGEV ont été identifiés pour les nombreuses liaisons qu'ils forment avec pAPN. Le résidu Tyr528 du RBD se lie avec les résidus glutamine 731 et tryptophane 737 de pAPN. Le résidu Trp571 est proche des résidus histidine 786 et proline 787 de pAPN et se lie à l'asparagine 783 de pAPN. De plus, le N-glycane porté par l'asparagine 736 de pAPN est impliqué dans la liaison du RBD de PRCoV et TGEV lorsque pAPN est exprimé à la surface cellulaire. Ce N-glycane est également conservé dans les protéines APN canines et félines, mais n'est pas présent dans hAPN, ce qui suggère que d'autres résidus interviennent dans la liaison entre le RBD d'HCoV-229E et son récepteur. Le résidu tyrosine n'est pas présent dans le RBD du HCoV-229E mais le résidu tryptophane est conservé, bien que la boucle du RBD le contenant soit plus courte dans HCoV-229E que dans TGEV (Reguera et al., 2012, 2014). Le site de liaison précis d'HCoV-229E avait été

déterminé au niveau des résidus 283-292/295 de hAPN (Chen et al., 2012; Kolb et al., 1997), mais la structure du complexe RBD d'HCoV-229E/APN a récemment été résolue, apportant de nouvelles informations sur les interactions entre le RBD de HCoV-229E et hAPN. Il semble que la protéine S se lie par 3 boucles entre les résidus 242 à 318 d'hAPN. Il a été montré que les résidus Asp287, Tyr289, Val290, Glu291, Ile309 et Leu318 interagissaient particulièrement avec la 1ère boucle du RBD d'HCoV-229E, le résidu Asp315 interagit avec les 2ème et 3ème boucles du RBD et le résidu Lys292 avec la 3ème boucle (Wong et al., 2017). Le fait que HCoV-229E et PRCoV se lient à différents sites sur APN a des conséquences importantes. APN peut se trouver dans une conformation ouverte, intermédiaire ou fermée en fonction de l'accessibilité de son site catalytique. Le passage d'une conformation à l'autre est considéré comme importante pour la catalyse de ses substrats. Il semble que la forme fermée soit la forme protéolytiquement active (Santiago et al., 2017; Wong et al., 2012). Le RBD du HCoV-229E se lie à hAPN lorsqu'il est en conformation fermée tandis que le RBD du PRCoV se lie spécifiquement à pAPN lorsqu'il est en conformation ouverte. Ces différences dans la liaison et la conformation des récepteurs se reflètent dans le fait que la protéine S de PRCoV inhibe l'activité catalytique de APN mais pas celle de HCoV-229E (Breslin et al., 2003; Santiago et al., 2017; Wong et al., 2017).

La raison des interactions avec des peptidases (APN, ACE2, DPP4) comme récepteur cellulaire pour la plupart des coronavirus n'est pas connue. En effet, la liaison des coronavirus à leur récepteur ne suffit pas, la protéine S à la surface du virus doit subir une maturation protéolytique. Pourtant les coronavirus n'utilisent pas l'activité catalytique des peptidases servant de récepteur pour cette maturation, mais entrent après l'action de protéases localisées proches des récepteurs. Il a été montré que la liaison du SARS-CoV à son récepteur ACE2 entraînait son internalisation et une diminution de l'activité enzymatique à la surface de ACE2, ce qui pourrait expliquer en partie la sévérité des infections par le SARS-CoV. Concernant le MERS-CoV, ce type d'interactions entre le virus et les protéines de l'hôte n'a pas été montré (Bosch et al., 2014). L'implication du domaine enzymatique de ces peptidases a également été étudiée lors de l'entrée virale du TGEV à l'aide de son

récepteur pAPN. Des inhibiteurs de l'activité enzymatique de APN ont été testés, et aucun n'a exercé d'effet significatif sur l'infection de TGEV. De plus, la mutation de résidus appartenant au site actif de la métalloprotéase n'avait pas affecté l'entrée virale (Delmas et al., 1994). Il semble donc que l'activité catalytique des peptidases ACE2, APN et DPP4 n'est pas requise pour l'entrée des coronavirus, mais l'interférence des coronavirus avec les fonctions de ces peptidases pourrait en partie expliquer la pathogenèse virale.

### **1.4.2.3 Reconnaissance antigénique**

La neutralisation de l'infection par les anticorps est un mécanisme de protection contre les virus. L'infection peut être inhibée en bloquant la liaison aux récepteurs cellulaires ou en interférant avec la fusion virale. De plus, dans le cas des virus enveloppés, les anticorps peuvent recruter des cellules effectrices ou le complément, permettant ainsi la destruction des cellules infectées ou la lyse des particules virales (Corti and Lanzavecchia, 2013). La protéine S est la protéine antigénique majeure des coronavirus: le domaine S1 contient la plupart des épitopes reconnus par les anticorps neutralisants lors d'une infection (Godet et al., 1994; Hofmann et al., 2004; Sayaka et al., 1991), bien que des sites aient été décrits dans d'autres régions de la protéine S (YangYang et al., 2014). Il a été montré pour le SARS-CoV, le MERS-CoV, TGEV et PEDV que le RBD situé dans le domaine S1 serait la cible la plus importante pour les anticorps neutralisants (Chen et al., 2017; Du et al., 2017; Hofmann et al., 2004; Li et al., 2017a; Reguera et al., 2012; Wang et al., 2018). Le domaine de liaison au récepteur des glycoprotéines virales doit être accessible pour permettre l'entrée virale, c'est pourquoi c'est une cible majeure des anticorps. Plus particulièrement, certaines structures secondaires comme les boucles (*extended loop*) semblent être particulièrement immunogènes, ce qui a déjà été montré pour l'*alphacoronavirus* TGEV (Wong et al., 2017). A l'inverse du RBD, les domaines permettant la fusion sont peu accessibles à cause du repliement serré des glycoprotéines virales ou de l'exposition trop transitoire durant l'étape de fusion. C'est pourquoi peu d'épitopes sont décrits dans ces régions (Corti and Lanzavecchia, 2013). Des variations dans le domaine S1 de la protéine S tel que la



N-glycosylation pourraient influencer l'antigénicité. Pour le moment, le domaine S1 du HCoV-229E a été déterminé comme important lors de la neutralisation par les anticorps (Shirato et al., 2012). Une étude récente du complexe hAPN – RBD du HCoV-229E a montré que trois boucles présentes dans le RBD étaient responsables de la liaison avec le récepteur. D'autre part au cours de ces cinquante dernières années, c'est-à-dire depuis l'identification de HCoV-229E, il apparaît que ces boucles ont évoluées. Les analyses de séquences ont permis de grouper les RBD en 6 classes différentes et ont montré une évolution au cours du temps avec un remplacement successif des RBD d'une classe par la classe suivante. D'autre part, un anticorps neutralisant dirigé contre le RBD de classe 1 (correspondant au RBD des souches les plus anciennes) n'est pas capable de lier les RBD des 5 autres classes. Ces résultats montrent que les anticorps neutralisants bloquant la liaison au récepteur peuvent être contrecarrés par une augmentation de l'affinité du RBD du virus pour son récepteur, permettant donc un échappement immunitaire du virus. De plus, une affinité de liaison optimale au récepteur n'existe que dans un environnement donné. Des modifications de pression de sélection dues à l'adaptation à une nouvelle espèce, à des changements dans le tropisme tissulaire ou à des différences dans les niveaux d'expression des récepteurs peuvent conduire à des changements dans l'affinité de liaison au récepteur. Cette évolution constante permet d'expliquer pourquoi l'infection par HCoV-229E ne confère pas d'immunité protectrice contre les différentes souches (Wong et al., 2017). Les régions de la protéine S du HCoV-229E précisément impliquées dans la reconnaissance par les anticorps n'ont pas encore été complètement élucidées.

### **1.4.3 Domaine S2 : Fusion virale**

#### **1.4.3.1 Mécanisme de la fusion virale**

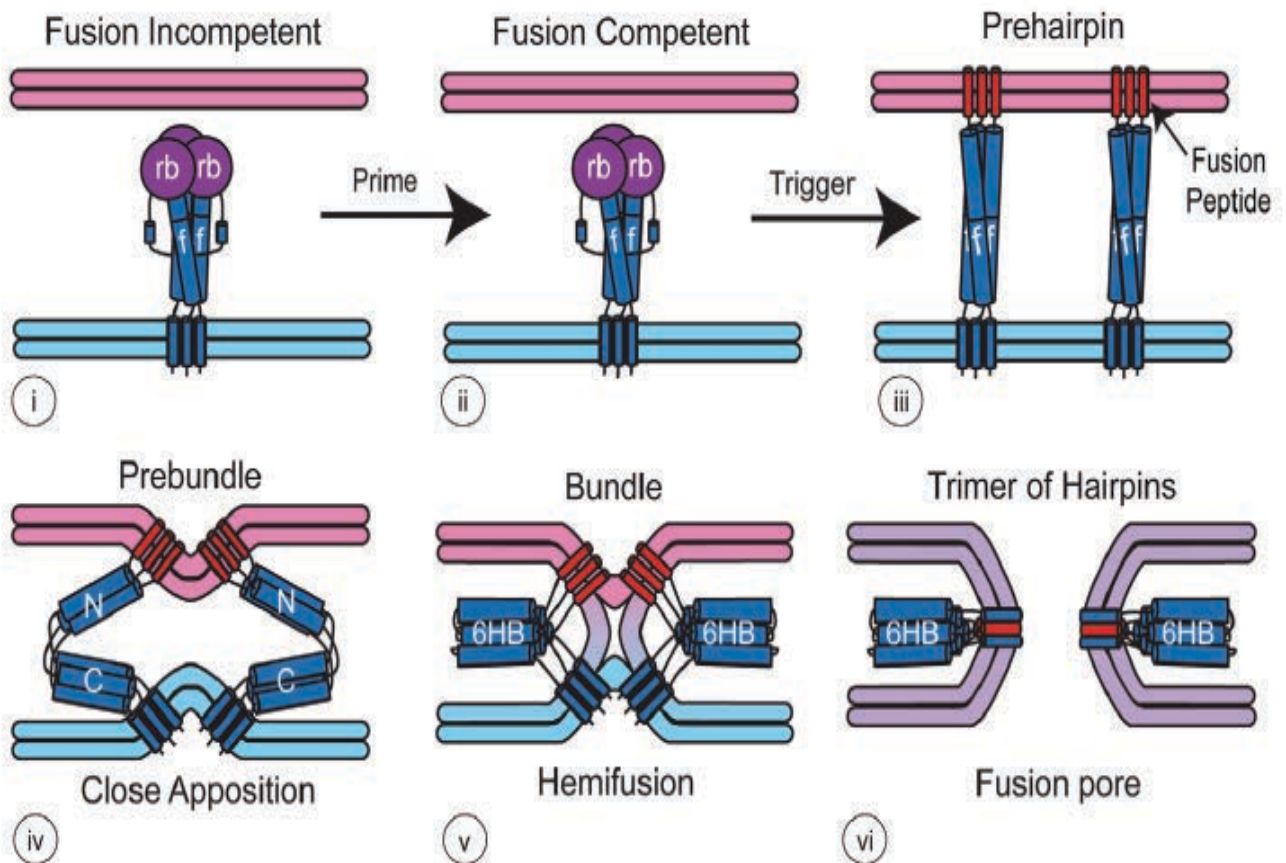
La protéine S du HCoV-229E est une protéine virale de fusion de classe I et est impliquée dans la fusion de l'enveloppe du virus avec les membranes cellulaires (Liu et al., 2006). Le domaine S2 est situé du côté C-terminal de la protéine S et contient les quatre éléments requis pour la fusion des membranes : un peptide de

fusion ou boucle, deux *heptad repeat* HR1 et HR2 et un domaine transmembranaire. Les HR sont constitués d'une répétition de peptides *abcdefgh* où *a* et *d* sont des résidus hydrophobes caractéristiques de la formation de superhélices  $\alpha$  ( *$\alpha$ -helical coiled-coil structure*) qui participent au processus de fusion. La formation d'une structure de type superhélice  $\alpha$  est caractéristique des protéines de fusion de classe I (Heald-Sargent and Gallagher, 2012).

Lors de la fusion entre les enveloppes virales et cellulaires, la protéine S des coronavirus passe par une série de réarrangements conformationnels (Figure 8). Le processus de fusion commence par la liaison au récepteur. D'autres événements peuvent être nécessaires à la fusion, comme un clivage protéolytique par des protéases endosomales ou présentes à l'extérieur de la cellule, ou une acidification du pH (Belouzard et al., 2012; White and Whittaker, 2016). Le premier changement conformationnel permet l'exposition du peptide de fusion, qui va s'insérer dans la membrane de la cellule hôte. La structure adoptée par la protéine est alors une structure étirée dans laquelle les *heptads repeats* sont alignées. Les changements de conformation suivants impliquent le rapprochement de l'extrémité C-terminale de la protéine S vers le peptide de fusion, ce qui entraîne le rapprochement des membranes virales et cellulaires. Une première étape d'hémifusion consiste en la fusion des feuilletts extérieurs des membranes virales et cellulaires. Puis les feuilletts intérieurs des membranes fusionnent pour former un pore reliant le virus à la cellule hôte (Figure 8).

Lors du changement de conformation final, la protéine adopte une structure en épingle à cheveux dans laquelle les HR  $\alpha$ -hélicoïdaux s'assemblent pour former un faisceau d'hélices ou *6-helix coiled-coil bundle*. Dans cette structure finale de la protéine de fusion, les trois HR1 de chaque protéine S du trimère interagissent pour former une structure de type *coiled-coil*, tandis que les trois HR2 encadrent cette structure dans un modèle anti-parallèle. La présence de cette structure a été démontrée pour plusieurs coronavirus, dont HCoV-229E (Gao et al., 2013; Liu et al., 2006; Ma et al., 2005), et est fortement similaire aux formes post-fusion d'autres virus comme le virus de la grippe (Blijleven et al., 2016). La structure finale de fusion de la protéine S est très stable, et est résistante aux protéases après dénaturation ou

après que la fusion des membranes soit achevée (Heald-Sargent and Gallagher, 2012)



**Figure 8 : Schéma du mécanisme de fusion virale avec la cellule hôte.** La fusion est permise par des événements comme la liaison au récepteur, un abaissement du pH et/ou l'activation protéolytique de la protéine S. La protéine est alors compétente pour la fusion. Il y a alors dissociation des domaines S1 (rb, en violet) et S2 (f, en bleu) de la protéine S. Le peptide de fusion (FP, en rouge) exposé va alors s'insérer dans la membrane de la cellule cible. Puis un nouveau changement conformationnel rassemblant les heptad repeat 1 et 2 (HR1 et HR2) de la protéine S rapproche les membranes virales (bleu clair) et cellulaires (en rose), la structure finale de la protéine étant nommée 6-helix coiled-coil bundle (6HB) (schéma issu de White et Whittaker, 2016).

### **1.4.3.2 Clivage de la protéine S**

Le clivage protéolytique est une modification post-traductionnelle irréversible qui entraîne des changements de conformation créant l'apparition ou la disparition de propriétés biologiques. Les protéases sont des enzymes capables de cliver les liaisons peptidiques des protéines (Li et al., 2017b; Steiner, 1998). Elles peuvent intervenir dans de nombreux processus biologiques, comme la maturation des protéines ou la digestion des aliments et sont présentes au niveau des endosomes intracellulaires ou à la surface cellulaire. On distingue les exoprotéases, qui coupent le premier ou le dernier acide aminé d'une chaîne peptidique, et les endoprotéases, qui sont capables de couper dans la chaîne peptidique. Il est possible de regrouper les protéases en fonction de leur nature ou des acides aminés de leur site actif, comme pour les protéases à cystéine ou les protéases à sérine (Polgár, 2013a, 2013b). Comme leur nom l'indique, les protéases à cystéine possèdent une cystéine dans leur site actif tandis que les protéases à sérine possèdent trois acides aminés dans leur site catalytique, dont la sérine. De nombreuses protéases clivent en fonction des acides aminés entourant le site de coupure : elles reconnaissent un motif précis de la chaîne polypeptidique.

Comme mentionné précédemment la plupart des coronavirus ont besoin de protéases cellulaires pour activer la fusion de leur enveloppe virale avec les membranes des cellules cibles. Les coronavirus peuvent utiliser des protéases endosomales, comme les cathepsines (L et B) ou la furine ou des protéases de la surface cellulaire comme TMPRSS2 (Millet and Whittaker, 2015).

Après sa découverte, les mécanismes d'entrée du SARS-CoV ont beaucoup été étudiés (Figure 9). Très rapidement, il a été montré que l'entrée était inhibée par des agents lysomotropiques mais que leur action était indirecte et liée à l'inhibition de la cathepsine L, une protéase endosomale active à pH acide (Matsuyama et al., 2005; Simmons et al., 2004). Il a ensuite été montré que l'entrée virale par la voie endosomale pouvait être court-circuitée lorsque les virus liés à la surface des cellules

étaient exposés à des protéases exogènes telles que la trypsine ou l'élastase (Matsuyama et al., 2005). Pour le SARS-CoV, la voie d'entrée directe par action de la trypsine est 100 à 1000 fois plus efficace que l'entrée virale par voie endosomale par action de la cathepsine (Matsuyama et al., 2005). D'autre part, en modèle de fusion cellule-cellule, la fusion des cellules exprimant la protéine S du SARS-CoV à leur surface peut être induite par un traitement par la trypsine. L'ensemble de ces données a donc rapidement indiqué que la fusion du SARS-CoV était stimulée par un clivage protéolytique. Ces mécanismes ont d'abord été étudiés en utilisant la trypsine, comme protéase modèle. Ainsi il a été montré que la fusion virale a lieu après un clivage séquentiel au niveau de 2 sites de la protéine S : il y a un premier clivage au niveau de l'arginine 667 à la jonction entre S1 et S2 qui semble faciliter un second clivage au niveau de l'arginine 797 située dans une région appelée S2', responsable de l'activation de la fusion (Belouzard et al., 2009). En effet, ce deuxième clivage a lieu au niveau N-terminal d'une région très conservée parmi les coronavirus qui présente les caractéristiques d'un peptide de fusion (Madu et al., 2009). La protéine S du SARS-CoV peut également être clivée par d'autres protéases comme l'élastase qui permet un clivage quelques acides aminés en amont du site S2' du SARS-CoV (Matsuyama et al., 2005), suggérant que la position du clivage en S2' permettant l'activation de la fusion présente une certaine flexibilité.

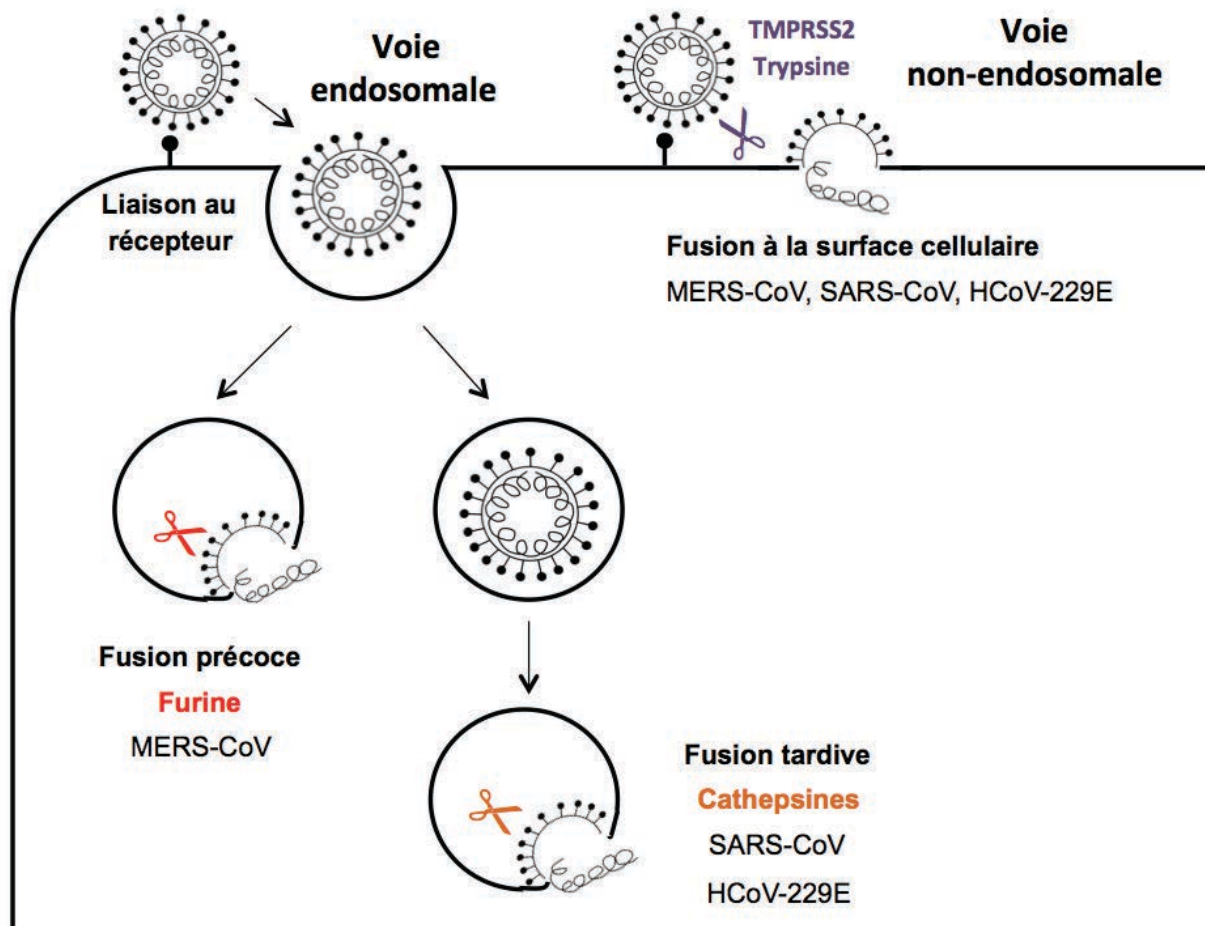
La trypsine utilisée comme outil n'est pas très pertinente du point de vue physiologique. Ainsi des travaux ont déterminé quelles protéases pulmonaires de type sérine-protéase pourraient être impliquées dans l'infection du SARS-CoV. Les meilleures candidates appartiennent à la famille des TTSPs. Parmi ses membres, deux protéines sont capables de cliver la protéine S du SARS-CoV, TMPRSS2 et HAT (ou TMPRSS11d). Il a été montré que HAT clive la protéine S au niveau de la jonction S1/S2, tandis que TMPRSS2 clive la protéine S à de multiples sites majoritairement proches du site S2' mais dont la localisation précise n'avait pas été déterminée. Une étude récente a suggéré que le résidu R667 situé au niveau de la jonction S1/S2 n'était pas nécessaire à l'activation de la protéine S par TMPRSS2 puisque sa mutation n'inhibe pas l'infection médiée par TMPRSS2. Par contre, le résidu R797 situé en S2' est indispensable pour l'activation par TMPRSS2 (Reinke et

al., 2017). Seule TMPRSS2 est capable d'induire l'infection des cellules à la surface cellulaire tandis que HAT pourrait jouer un rôle dans la propagation du virus en facilitant la formation de *syncytia*. De plus, l'activation par TMPRSS2 protège la protéine S de l'inhibition par les IFITM (*Interferon-inducible Transmembrane Proteins*), des protéines cellulaires induites par l'interféron qui inhibent l'entrée cellulaire de plusieurs virus enveloppés. A ce jour, les processus protéolytiques qui permettent à la cathepsine L d'induire la fusion restent inconnus (Bailey et al., 2014; Belouzard et al., 2012; Bertram et al., 2013; Wrensch et al., 2014).

Le rôle du clivage de la région S2' dans l'activation de la fusion des coronavirus a ensuite été confirmé pour plusieurs virus dont MHV-A59 et le MERS-CoV. Les études de la protéine S du MERS-CoV montrent que de nombreuses protéases cellulaires sont capables de l'activer (Figure 9) (Shirato et al., 2013). Il a été montré que la protéine S du MERS-CoV, comme celle du SARS-CoV, subit un clivage séquentiel en S1/S2 puis en S2' (Millet and Whittaker, 2014). A la différence du SARS-CoV, la protéine S du MERS est clivée dans un premier temps lors de l'excrétion du virus par les cellules productrices (région S1/S2) puis dans un deuxième temps lors de l'entrée virale dans de nouvelles cellules-hôte (au niveau de S2'). En particulier, il a été montré que la protéine S du MERS-CoV peut être clivée par la furine ou une cathepsine lors de son entrée par voie endosomale. Le MERS-CoV peut aussi fusionner à la surface cellulaire en présence de TMPRSS2 (Gierer et al., 2013; Shirato et al., 2013).

Le clivage précoce de la protéine S du MERS-CoV à la jonction S1/S2 semble permettre un changement de conformation facilitant l'accès des protéases au niveau S2' de la protéine S. Park *et al.* ont étudié l'importance du clivage S1/S2 dans l'infection par le MERS-CoV en réalisant des mutations de cette jonction. Lorsque le clivage de la protéine S par la furine au niveau de la jonction S1/S2 a lieu avant l'entrée cellulaire, la protéine S est alors sensible aux protéases présentes à la surface cellulaire et la fusion se déroule dans des endosomes précoces. A l'inverse, lorsque la protéine S n'est pas clivée à la jonction S1/S2 avant l'entrée cellulaire, la fusion se déroule au niveau des endosomes tardifs (Park et al., 2016). La présence d'un site de clivage par la furine en S2' pourrait expliquer l'infection extrapulmonaire

observé chez les patients atteints par le MERS-CoV puisque la furine est une protéase ubiquitaire (Burkard et al., 2014; Millet and Whittaker, 2014; Millet et al., 2016).



**Figure 9: Représentation de l'entrée virale des coronavirus.** Après liaison au récepteur de la protéine S du coronavirus, le virus entre par la voie endosomale ou la voie non-endosomale. L'activation protéolytique de la protéine S peut être réalisée par différentes protéases. Au niveau de la voie endosomale, la furine et les cathepsines ont été décrites. Au niveau de la voie non-endosomale, la trypsine a été décrite, ainsi que TMPRSS2. Les coronavirus utilisant ces protéases sont mentionnés en-dessous de celles-ci.

Il existe également quelques études sur l'utilisation des protéases cellulaires par HCoV-229E lors de l'entrée virale (Figure 9). Sa protéine S n'est pas clivée à la surface des virions entre le domaine S1 et S2 durant l'assemblage et l'exocytose des

virions, comme pour le SARS-CoV. Il a d'abord été montré que le virus HCoV-229E dépendait de la cathepsine L pour son entrée dans les cellules cibles (Kawase et al., 2009). La protéase TMPRSS2 peut aussi cliver et activer la protéine S du HCoV-229E permettant ainsi une entrée cellulaire indépendante de la cathepsine L tandis que HAT est seulement capable d'induire la fusion cellule-cellule mais pas l'infection. Ce résultat semble intéressant car TMPRSS2 est exprimée dans les cellules cibles du HCoV-229E (l'épithélium humain respiratoire) et peut donc améliorer la propagation du virus chez les humains (Bertram et al., 2013). De plus, il a été constaté que les souches cliniques d'HCoV-229E étaient moins capables d'utiliser la cathepsine L pour l'entrée cellulaire que des souches virales ayant subies plusieurs passages en culture *in vitro*. Il est donc probable que la voie endosomale soit plus avantageuse *in vitro* qu'*in vivo* (Shirato et al., 2016) et que l'adaptation du virus à la culture cellulaire facilite l'utilisation de la cathepsine L. Les acides aminés nécessaires au clivage par les cathepsines ne sont pas connus, mais cette étude a montré que les mutations des résidus arginine 642 de la souche de laboratoire en méthionine (R642M) et asparagine 714 en lysine (N714K) présentes dans les souches cliniques étaient suffisantes pour diminuer sa capacité à entrer dans les cellules par une voie dépendante de la cathepsine L. D'autre part, après passage d'une souche clinique en culture cellulaire, une seule mutation I577S augmente la capacité de la souche à utiliser la cathepsine L lors de l'entrée. Ces substitutions pourraient modifier la structure conformationnelle de la protéine S, modifiant l'accès de la cathepsine L au site de clivage (Shirato et al., 2016).

Récemment, des résultats similaires ont été obtenus avec les virus HCoV-HKU1 et HCoV-OC43 (Shirato et al., 2017). En effet, les isolats cliniques peuvent être cultivés sur des cellules épithéliales respiratoires et leur entrée est majoritairement médiée par TMPRSS2. Leur adaptation en culture cellulaire favorise leur capacité à utiliser des protéases endosomales. Ceci suggère que l'entrée par la voie endosomale favoriserait l'induction de la réponse immunitaire antivirale de l'hôte comme cela a été observé pour HCoV-229E (Bertram et al., 2013).



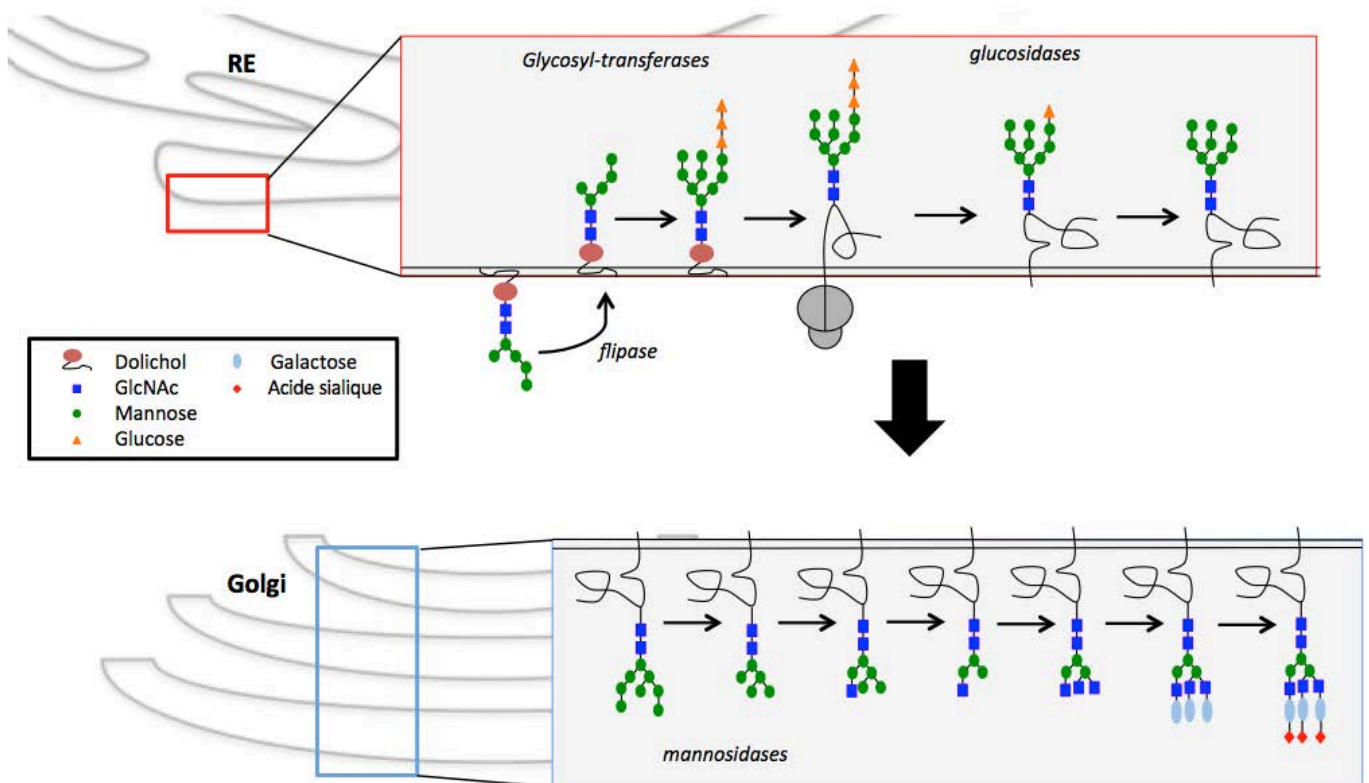
## 1.5 N-GLYCOSYLATION

### 1.5.1 Généralités et mécanisme

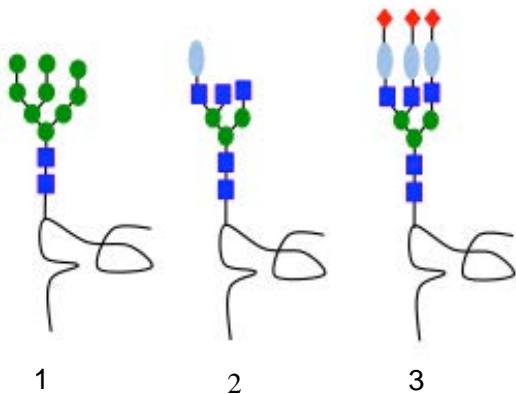
La N-glycosylation a lieu chez tous les Eucaryotes, ainsi que chez les Archaea et les bactéries (Calo et al., 2010; Nothaft and Szymanski, 2010). Les glycanes, ou chaînes oligosaccharidiques, sont des oses liés de façon covalente à des protéines et qui peuvent avoir des structures variables (Ohtsubo and Marth, 2006). Les protéines peuvent être O- ou N-glycosylées, ou former des glycosaminoglycanes, aussi nommés protéoglycanes, selon le site d'ancrage du N-glycane, ainsi que la taille et la nature des chaînes liées à la protéine (Schachter, 2000; Yan and Lennarz, 2005).

La N-glycosylation est une modification co-translationnelle qui consiste en l'attachement d'oligosaccharides sur un atome d'azote appartenant à un résidu asparagine (Asn) d'une protéine (Figure 10A). L'attachement du glycane sur la protéine nécessite la reconnaissance d'une séquence consensus, dans laquelle le résidu Asn appartient à la séquence Asn-X-Ser/Thr (Ser = sérine, Thr = thréonine) où X est n'importe quel acide aminé autre que la proline (Pro). Les sites de N-glycosylation sont également déterminés par la structure de la protéine elle-même et le type cellulaire dans laquelle elle est synthétisée (Helenius and Aebi, 2001). La N-glycosylation débute dans le réticulum endoplasmique par le transfert en bloc d'un précurseur oligosaccharidique préformé (Helenius and Aebi, 2001; Suzuki et al., 1995). Une chaîne oligosaccharidique précurseur est initialement formée dans le cytoplasme. Les 7 premiers oses de cette chaîne sont ajoutés sur le dolichol, un lipide enchâssé dans la membrane du RE. Puis, le complexe bascule dans la lumière du RE à l'aide d'une flipase, où 7 oses sont ajoutés sur la chaîne oligosaccharidique.

A



B



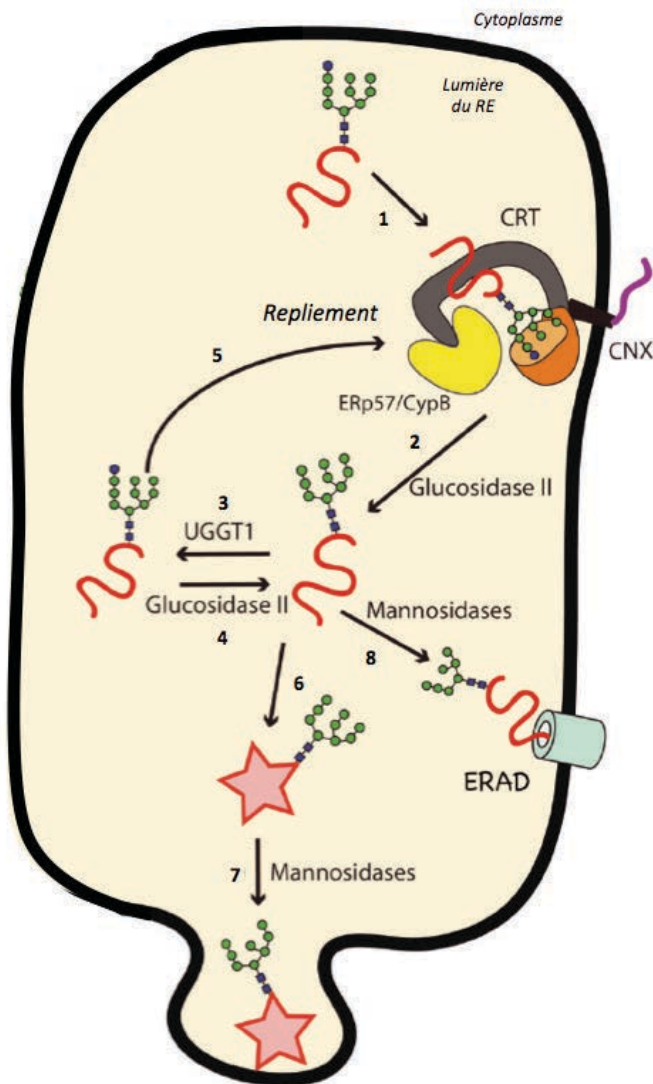
**Figure 10: Biosynthèse des N-glycane.** A. La N-glycosylation est initiée dans le RE. Une chaîne oligosaccharidique précurseur est formée puis transférée sur la glycoprotéine. Le N-glycane est ensuite mûri dans le RE puis dans l'appareil de Golgi. B. Les différents types de structure du N-glycane. 1 : Structure riche en mannose, 2 : structure hybride, 3 : structure complexe

Le précurseur de 14 oses est alors composé de 2 N-acétylglucosamines (GlcNAc), de 9 mannoses (Man) et de 3 glucoses (Glc). Il est ajouté co-translationnellement sur le résidu asparagine de la protéine à modifier. La liaison ainsi formée est une liaison N-glycosidique, d'où le nom de N-glycosylation. La

réaction établissant cette liaison covalente est contrôlée par la glycosyltransférase qui est localisée dans la lumière du réticulum endoplasmique. La chaîne subit ensuite des modifications très ordonnées. Deux glucoses sont séquentiellement enlevés par les glucosidases I et II, créant une protéine monoglycosylée. Ce N-glycane monoglycosylé est un substrat pour la calnexine (CNX) et/ou la calréticuline (CRT), qui sont des lectines chaperonnes du RE (Figure 11). La glucosidase II peut aussi supprimer le dernier glucose, la calnexine et la calréticuline ne peuvent alors plus se lier à la chaîne protéique. Si la protéine est mal repliée, le *UDP-glucose glycoprotein glycosyltransferase* (UGGT) transfère un glucose sur la protéine, permettant une nouvelle liaison aux protéines chaperonnes. Pendant ce processus, ERp57, une protéine disulfide isomérase (PDI) appartenant au RE peut moduler le repliement des glycoprotéines en aidant à la formation et l'isomérisation de liaisons disulfures. Si la protéine est toujours mal repliée, elle est alors envoyée vers la voie de l'ERAD (*Endoplasmic Reticulum associated protein degradation*) afin d'être dégradée. Le processus de maturation et de contrôle qualité des glycoprotéines de la voie de sécrétion précoce dépend donc de la calnexine, de la calréticuline, de ERp57 et de UGGT pour fonctionner correctement (Lamriben et al., 2016). Lorsque la glycoprotéine est correctement repliée, la maturation de l'arbre glucidique peut commencer. Une mannosidase clive un seul mannose avant que la glycoprotéine ne transite vers l'appareil de Golgi (Figure 10A). Avant l'entrée dans l'appareil de Golgi, toutes les protéines N-glycosylées portent des N-glycanes précurseurs ayant la même structure (2 GlcNAc, 8 Man). La chaîne oligosaccharidique est ensuite maturée par clivage et addition d'oses de manière séquentielle dans les différents compartiments de l'appareil de Golgi (cis, médian, trans, réseau trans-golgien) où sont localisées les enzymes modifiant le N-glycane.

On distingue 3 types de structure de la chaîne oligosaccharidique selon les modifications subies dans l'appareil de Golgi : structure riche en mannose, structure hybride et structure complexe (Figure 10B). Pour les N-glycanes riche en mannose, seuls des mannoses peuvent être clivés mais aucun ose n'est ajouté à la chaîne oligosaccharidique. Les deux derniers types de N-glycanes ont une structure de base commune : l'oligosaccharide de base est coupé par la mannosidase I et est glycosylé

par la GlcNAc transférase. L'oligosaccharide est ensuite modifié par la mannosidase II. Toutes les glycoprotéines qui atteignent ce stade sont alors résistantes à l'hydrolyse de N-glycanes par l'endoglycosidase H qui est utilisée expérimentalement pour déterminer si une glycoprotéine contient une forte teneur en mannose ou des oligosaccharides complexes. De multiples GlcNAc transférases sont capables de transférer différents oses sur le N-glycane après l'action de la mannosidase II, déterminant ainsi la structure du N-glycane complexe. La diversité des N-glycanes formés est très importante, mais rend la détermination de la structure des glycoprotéines difficile.



**Figure 11: Mécanisme de repliement dans le RE des protéines N-glycosylées.** La protéine arborant un seul glucose (cercle bleu) est un substrat pour la calnexine (CNX) et la calréticuline (CRT) (1). L'élimination du dernier glucose par la glucosidase II libère le substrat de la CNX/CRT (2). Les glycoprotéines mal repliées peuvent être de nouveau glycosylées par UGGT (UDP-glucose glycoprotéine glycosyltransférase) (3) afin de pouvoir se lier de nouveau à la CNX/CRT, ou être modifiées par la glucosidase II (4). Les protéines correctement repliées (étoile rouge) (6) sont modifiées par la mannosidase et dirigées vers l'appareil de Golgi (7). Les protéines mal repliées sont prises en charge par une mannosidase et dirigées vers l'ERAD (Endoplasmic Reticulum associated protein degradation) (8) afin d'être dégradées (figure adaptée de Lamriben et al, 2016).

## 1.5.2 Rôle de la N-glycosylation

### 1.5.2.1 Dans les cellules

Les N-glycanes portés par les glycoprotéines ont plusieurs fonctions qui passent le plus souvent par leur liaison à des lectines, c'est-à-dire des protéines ayant une affinité pour certains motifs glucidiques. Les N-glycanes peuvent participer au repliement des glycoprotéines qui les portent, permettant l'acquisition de leurs structures tridimensionnelles. Ce processus de contrôle qualité fait intervenir les protéines chaperonnes calnexine et calréticuline. Une fois les protéines formées et repliées, les N-glycanes peuvent être supprimés sans effets majeurs sur la conformation de la protéine (Helenius and Aebi, 2001; Lamriben et al., 2016). Après la maturation des N-glycanes portés par la glycoprotéine lors du passage dans la voie de sécrétion (constituée du RE puis de l'appareil de Golgi), de nombreuses glycoprotéines sont localisées à la surface cellulaire, où ces N-glycanes sont généralement bien exposés. Ils participent en particulier à l'adhésion cellulaire en fixant par des liaisons faibles des lectines membranaires appartenant à d'autres cellules, ou des lectines possédant plusieurs sites de liaison pouvant assurer un pontage entre différentes glycoprotéines. Les glycoprotéines participent également à d'autres fonctions biologiques, comme le trafic intracellulaire, la signalisation cellulaire, la réponse immunitaire etc (Ohtsubo and Marth, 2006).

Les protéines glycosylées de la surface cellulaire peuvent être utilisées comme récepteur, co-récepteur ou facteur d'attachement par de nombreux virus. La N-glycosylation est très conservée chez les mammifères, bien qu'il y ait des variations spécifiques aux différentes espèces. Ces variations déterminent la susceptibilité de certains organismes hôtes à des pathogènes dont les virus, impactant, par exemple, le tropisme cellulaire, la spécificité d'hôte, l'adaptation, la transmission inter-espèce et la pathogénèse des virus (Gagneux and Varki, 1999). Par exemple, il a été montré que le virus influenza de type A entre dans les cellules à l'aide d'acides sialiques. L'hémagglutinine HA se lie préférentiellement aux N-glycanes sialylés cellulaires possédant une liaison de type  $\alpha 2,6$  ou  $\alpha 2,3$  entre l'acide

sialique terminal et le galactose respectivement pour les virus de mammifères et les virus aviaires (Dietrich et al., 2017).

### **1.5.2.2 N-glycosylation des protéines virales**

Les virus possèdent à leur surface des protéines glycosylées grâce à la machinerie cellulaire, ce qui leur permet de profiter de plusieurs avantages. Les N-glycanes ont une fonction différente selon leur localisation : un N-glycane proche d'un épitope peut permettre l'échappement immunitaire en perturbant la reconnaissance par les anticorps. Un N-glycane situé près d'un site de clivage peut entraîner une diminution de la virulence ou modifier l'affinité au récepteur s'il est proche d'un site de liaison au récepteur.

#### 1.5.2.2.1 Repliement des glycoprotéines virales

Le repliement correct des protéines virales et le trafic intracellulaire sont permis par l'interaction entre les protéines chaperonnes et les N-glycanes. Il semble que tous les virus étudiés jusqu'à présent utilisent le cycle calnexine/calréticuline (Larkin and Imperiali, 2011; Vigerust and Shepherd, 2007). L'hémagglutinine HA du virus influenza a beaucoup été étudiée. Il a été montré que les N-glycanes présents sur celle-ci se lient aux protéines chaperonnes calnexine, calréticuline et ERp57 permettant ainsi la régulation du repliement de la glycoprotéine HA (Daniels et al., 2003). Il a été montré pour le virus de l'hépatite C que deux N-glycanes présents sur les glycoprotéines d'enveloppe E2 sont impliqués dans le repliement de la protéine (Goffard et al., 2005). La N-glycosylation peut également assurer la stabilisation des glycoprotéines, comme il a été montré pour le virus de la dengue. En effet, la glycoprotéine non structurale NS1, qui est importante pour la réplication virale et lors de la réponse immunitaire, possède deux sites de N-glycosylation, N130 et N207. Il a été montré qu'au moins un des deux sites de N-glycosylation était indispensable pour la production de virus viable. Ces N-glycanes n'affectent pas la dimérisation de

la protéine NS1 mais semblent impliqués dans la stabilisation des dimères. En particulier, les dimères de NS1 sont plus sensibles à la chaleur donc plus rapidement dégradés lorsque ces N-glycanes sont supprimés. De plus, la plupart des protéines NS1 secrétées forment des hexamères. Lorsqu'un ou les deux sites de N-glycosylation sont mutés, la quantité d'hexamères diminue. Ces résultats indiquent que les N-glycanes sont importants pour la sécrétion de cette protéine et la stabilité des hexamères (Yap et al., 2017).

#### 1.5.2.2.2 Entrée virale

La N-glycosylation des protéines de surfaces virales joue un rôle dans les interactions avec le récepteur cellulaire ainsi que dans l'infectiosité du virion. Il a été montré pour le virus de l'immunodéficience humaine (VIH) que la mutation de 6 des 25 sites de N-glycosylation présents sur la protéine d'enveloppe gp120 entraînait une perte significative d'infectiosité du virus, ce qui a été confirmé dans une étude plus récente pour certains de ces N-glycanes (Mathys and Balzarini, 2015; Wang et al., 2013b). L'étude des glycoprotéines d'enveloppe E1 et E2 du virus de l'hépatite C a permis de montrer l'importance de sites de N-glycosylation présents sur la glycoprotéine E2 pour l'infection virale (Goffard et al., 2005; Helle et al., 2010). Le virus de la dengue (DENV) infecte les cellules dendritiques, les monocytes, les cellules endothéliales et les cellules des ganglions lymphatiques grâce aux lectines de type C, notamment le *Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin* (DC-SIGN). La protéine d'enveloppe E de DENV possède deux sites de N-glycosylation N67 et N153. Il a été montré le récepteur DC-SIGN interagissait directement avec le N-glycane N67 de la glycoprotéine E. De plus, il semble que la présence de N-glycanes riches en mannose sur la glycoprotéine d'enveloppe soit impliqués dans la liaison à son récepteur DC-SIGN (Yap et al., 2017). A l'opposé, l'unique site de N-glycosylation N154 de la glycoprotéine E du virus Zika (ZIKV) semble être défavorable à l'attachement du virus sur des cellules de moustique (Fontes-Garfias et al., 2017).

L'importance de la N-glycosylation lors de l'entrée virale a été étudiée pour le SARS-CoV. La protéine S du SARS-CoV peut lier des lectines de type C (DC-SIGN et L-SIGN) cependant leur rôle en tant que facteur d'attachement ou récepteur reste controversé. Des analyses de mutagenèse ont identifié 7 sites de N-glycosylation (N109, N118, N119, N158, N227, N589 and N699) de la protéine S du SARS-CoV comme étant critiques pour l'entrée médiée par DC-SIGN et L-SIGN. Cependant, la mutation de plusieurs sites de N-glycosylation n'a pas d'effet de synergie sur l'entrée virale (Han et al., 2007). Une étude précédente avait identifié seulement deux sites de N-glycosylation (N330 et N357) importants pour l'interaction avec DC-SIGN (Shih et al., 2006). Ces résultats contradictoires s'expliquent par des méthodologies différentes. Une lectine sérique, *mannose-binding lectin* MBL, est également étudiée pour son rôle lors de l'infection par le SARS-CoV. La lectine MBL joue un rôle important dans l'immunité innée en agissant comme une opsonine ou en activant le complément. La MBL peut se lier sélectivement au SARS-CoV, inhibant l'infection des cellules permissives à des concentrations inférieures à celles observées dans le sérum d'individus en bonne santé. Il a été montré que la mutation d'un seul site de N-glycosylation N330 de la protéine S abolit spécifiquement l'interaction avec la MBL (Zhou et al., 2010).

#### 1.5.2.2.3 Clivage protéolytique

Le rôle de la N-glycosylation dans la modulation de l'activation de protéines virales de surface par clivage protéolytique a peu été étudié.

Pour être compétente pour la fusion, l'hémagglutinine HA de du virus influenza synthétisée sous la forme d'un précurseur HA0, doit être clivée entre les deux domaines HA1 et HA2. Ce clivage joue un rôle important dans la pathogénèse des virus de la grippe. Pour les influenza virus H5N2, Il a été montré que des souches peu virulentes avaient gagné en virulence (souche peu virulente A/chicken/Pennsylvania/1/83, souches virulentes A/chicken/Pennsylvania/1370/83) avec la perte d'un site de N-glycosylation. Ainsi la présence d'un N-glycane à proximité du site de clivage empêcherait l'accès des protéases à l'hémagglutinine



virale HA (Deshpande et al., 1987; Kawaoka et al., 1984). Une étude récente sur la glycoprotéine HA d'une autre souche d'influenzavirus H5N1 (A/Mallard/Huadong/S/2005) a identifié le chevauchement de deux sites potentiels de N-glycosylation (nommés 10 et 11 : NNST) au niveau du site de clivage de la protéine. Une seule de ces arginines peut porter un N-glycane à cause de l'encombrement stérique. Le site 10 avait été dans un premier temps décrit comme N-glycosylé (Zhang et al., 2015). Il a ensuite été montré que le site 11 était le site réellement N-glycosylé et que la mutation de celui-ci impactait le clivage de l'hémagglutinine HA (Yin et al., 2017).

Concernant les coronavirus, en se basant sur la structure de la protéine S de PDCoV, Xiong *et al.* ont suggéré que les sites de N-glycosylation N652 et N661 pourraient limiter l'accès des protéases à la région S2' de la protéine et jouer un rôle dans l'activation de la fusion. Ces hypothèses doivent cependant être confirmées (Xiong et al., 2017). Le clivage de la protéine S du SARS-CoV a également été étudié. Il a été montré que lors de la digestion de la protéine S, les fragments générés par TMPRSS2 et la trypsine avaient un poids moléculaire différents, cependant le traitement de ces fragments par la PNGase permettait d'obtenir des fragments de taille similaire. Ces résultats suggèrent que TMPRSS2 et la trypsine clivent la protéine au même endroit mais sur des protéines modifiées différemment par des N-glycanes. Il est possible que ce résultat soit dû à une différence de localisation entre TMPRSS2 et la trypsine. La trypsine agit quasi exclusivement à la surface cellulaire, tandis que TMPRSS2 peut cliver la protéine S précocement dans les cellules, avant que les N-glycanes de la protéine S ne soient complètement matures (Reinke et al., 2017). Récemment, le rôle des N-glycanes du domaine S1 lors de la fusion virale du *gammacoronavirus Infectious Bronchitis Virus* (IBV), souche Beaudette, a été étudié. La protéine S de ce coronavirus possède 29 sites potentiels de N-glycosylation, et les 13 sites situés à l'extrémité N-terminale ont été analysés. La chromatographie en phase liquide haute performance couplée à de la spectrométrie de masse a permis d'identifier 8 sites effectivement N-glycosylés : N212, N237, N247, N276, N513, N591, N1051 et N1074. Le rôle des 13 sites potentiels de N-glycosylation lors de la fusion cellule/cellule a été étudié par

mutagénèse dirigée. Il a été montré que les sites N212 et N276 étaient importants pour la fusion cellulaire, potentiellement en modulant le clivage protéolytique de la protéine S. Le site N212 se trouve dans le RBD de la protéine S tandis que N276 est proche du RBD, et ces deux sites sont effectivement N-glycosylés. Il est intéressant de noter que les auteurs ont choisi d'abolir la N-glycosylation en mutant l'asparagine (N) en acide aspartique (D). Pour le site N283, ils ont pu montrer que la mutation N283D avait un impact sur l'infectivité et la réplication virale, tandis que la mutation de l'asparagine en glutamine n'avait aucun impact. Ce résultat montre que c'est le résidu asparagine qui est important pour l'infectivité et la réplication virale et pas le N-glycane N283 (Zheng et al., 2018). Aucune donnée n'est actuellement disponible sur le rôle des N-glycanes dans le clivage de la protéine S de HCoV-229E.

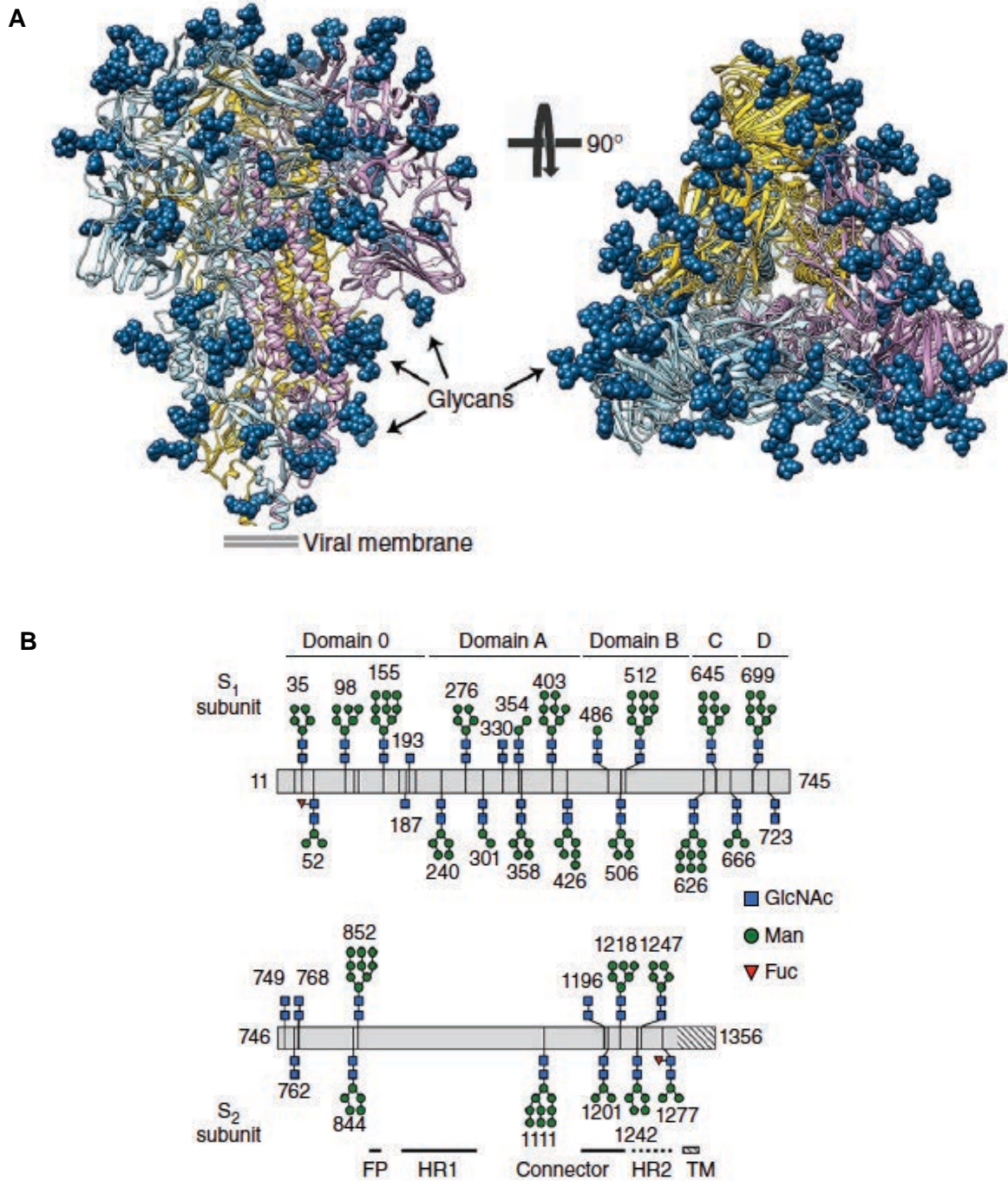
#### 1.5.2.2.4 Reconnaissance par les anticorps

La neutralisation par les anticorps est un moyen important de défense contre les pathogènes. La N-glycosylation est importante pour le repliement des protéines, influençant ainsi potentiellement l'accessibilité des épitopes impliqués dans la reconnaissance par les anticorps. Les N-glycanes eux-mêmes peuvent être reconnus par certains anticorps neutralisants. Les virus peuvent s'adapter à leur environnement et plus particulièrement à la pression immunitaire en modulant les N-glycanes présents à leur surface (Raska et al., 2014).

La protéine d'enveloppe (Env) du virus de l'immunodéficience humaine (VIH) a beaucoup été étudiée. La protéine Env est un trimère d'hétérodimères gp41-gp120, eux-mêmes produits par clivage protéolytique de la protéine précurseur gp160. La forte densité de N-glycanes présents à la surface de la protéine Env du VIH sert de bouclier de N-glycanes : les anticorps neutralisants ciblant ce bouclier doivent avoir une plasticité importante et tolérer la variabilité virale (Behrens and Crispin, 2017). Une équipe a étudié la N-glycosylation de la glycoprotéine d'enveloppe Env du VIH, qui possède 25 sites potentiels de N-glycosylation. La mutation individuelle de 4 N-glycanes (N197, N301, N442 et N625) a rendu le virus plus sensible à la

neutralisation par les anticorps. Il est possible que l'abolition de la N-glycosylation entraîne la suppression du bouclier de N-glycanes et/ou un changement de conformation suffisant pour permettre un meilleur accès des anticorps aux épitopes. A l'inverse, la mutation individuelle de 3 N-glycanes (N142, N355 et N463) a entraîné la résistance du virus à la neutralisation par les anticorps, ce qui suggère que ces N-glycanes sont reconnus par les anticorps (Wang et al., 2013b). Le même type d'étude a été mené pour le virus de l'hépatite C (HCV). Il a été montré qu'au moins 5 N-glycanes situés sur la glycoprotéine d'enveloppe E2 de HCV diminuaient la sensibilité aux anticorps neutralisants, en cachant des épitopes (Helle et al., 2010). La délétion des N-glycanes présents sur E1 et E2 permettraient une amélioration de l'immunogénicité de HCV (Ren et al., 2015).

A partir des études de structures de protéines de coronavirus, il est possible de formuler des hypothèses concernant la relation entre N-glycosylation et neutralisation par les anticorps. De nombreux sites de N-glycosylations sont conservés sur les séquences des protéines S de HCoV-NL63 et de PDCoV. Ces données suggèrent que la pression immunitaire est similaire chez les hôtes respectifs de ces *alpha*- et *deltacoronavirus*. Les régions cachées par ce bouclier de N-glycanes semblent donc importantes pour les fonctions de la protéine S (Walls et al., 2016a; Xiong et al., 2017). Le bouclier de N-glycanes permettrait donc un échappement immunitaire, comme cela a été décrit pour le HIV. Pour HCoV-229E, l'importance du domaine S1 lors de la reconnaissance par les anticorps a été confirmée. Il a été montré que des souches cliniques de HCoV-229E n'avaient pas le même profil antigénique que la souche de laboratoire VR-740, probablement à cause des différences entre les domaines S1 (Shirato et al., 2012). L'implication de la N-glycosylation de la protéine S du HCoV-229E lors de la reconnaissance par les anticorps n'a pas encore été déterminée.



**Figure 12: Organisation du bouclier de glycanes de la protéine S du HCoV-NL63 par cryo-microscopie électronique et spectrométrie de masse.** A. Des trimères de protéine S sont montrés vus de côté (A, à gauche) et depuis le dessus (A, à droite). Les glycanes sont représentés par des sphères bleu foncé. B. Schéma des types de N-glycanes. La structure des glycanes détectés à chaque site est représentée. FP peptide fusion, HR1 et 2 heptad-repeat domain 1 et 2, TM domaine transmembranaire (les stries indiquent les régions non étudiées), GlcNAc N-acetylglucosamine, Man mannose, Fuc fucose (Figure issue de Walls et al, 2016).

#### 1.5.2.2.5 Description des N-glycanes présents sur les protéines S des coronavirus.

Les N-glycanes présents sur la protéine S du SARS-CoV ont été caractérisés par chromatographie en phase liquide à haute performance associée à de la spectrométrie de masse. La protéine S du SARS-CoV possède vingt-trois sites potentiels de N-glycosylation. Les N-glycanes décrits étaient à 30% à haute teneur en mannose, à 28% des N-glycanes hybrides et à 42% des N-glycanes complexes à deux, trois ou quatre ramifications (Ritchie et al., 2010).

Les trimères de protéine S de HCoV-NL63 ont récemment été étudiés par cryo-microscopie électronique et spectrométrie de masse. Sur les 39 sites potentiels de N-glycosylation identifiés, il a été montré que 34 étaient effectivement N-glycosylés. L'identification des N-glycanes répartis à la surface des trimères montre que 102 N-glycanes couvrent la surface d'un trimère (Figure 12), masquant la surface des protéines, ce qui suggère que l'accès pour de potentiels anticorps neutralisants puissent être limité (Walls et al., 2016a). Du point de vue phylogénétique, les virus HCoV-229E et HCoV-NL63 sont très proches, ainsi les protéines S partagent 56% d'identité. Le domaine 0 est absent de la protéine S du virus HCoV-229E, ce domaine présent dans la protéine S du virus HCoV-NL63 et d'autres *alphacoronavirus* pourrait être impliqué dans la liaison d'acides sialiques. La même équipe a également étudié la protéine S du *deltacoronavirus* porcine PDCoV par cryo-microscopie électronique et les N-glycanes présents sur la protéine S ont été identifiés (Xiong et al., 2017). Il semble que la conformation de la protéine S du PDCoV soit similaire à celle de la protéine S d'*alphacoronavirus*. Vingt-six sites potentiels de N-glycosylation ont été confirmés, masquant la surface de la protéine de la même manière que pour HCoV-NL63. Il est intéressant de constater que de nombreux N-glycanes identifiés sur la protéine S de PDCoV sont situés au même endroit sur la protéine ou occupent le même espace que les N-glycanes identifiés sur la protéine S du HCoV-NL63. Actuellement, la structure de la protéine S du HCoV-229E n'a pas été résolue.



## 1.6 OBJECTIFS

Le coronavirus humain 229E, bien que connu depuis plus de 50 ans, n'a suscité que peu d'intérêt jusqu'à la découverte des coronavirus hautement pathogènes SARS-CoV et MERS-CoV. L'étude des coronavirus à l'aide de modèles animaux restant difficile (Gretebeck and Subbarao, 2015; Lassnig et al., 2005), l'étude des mécanismes moléculaires *in vitro* nous apporte une meilleure compréhension de la physiologie des infections liées à cette famille virale.

L'objectif de cette thèse est d'étudier le clivage protéolytique de la protéine S du HCoV-229E qui permet d'activer la fusion virale et le rôle de la N-glycosylation dans les fonctions de la protéine S.

Pour la protéine S du SARS-CoV, il a été montré que l'activation de la fusion nécessitait une activation protéolytique en deux étapes, un premier clivage de la jonction S1/S2 facilite un deuxième clivage dans la région S2' primordial pour médier la fusion. La protéine S du HCoV-229E n'est pas clivée à la surface de la particule virale comme celle du SARS-CoV. Nous avons donc testé si la fusion du virus HCoV-229E était activée par un mécanisme similaire en utilisant un système de particules rétrovirales pseudotypées avec la protéine S de HCoV-229E. Cet outil a été utilisé pour étudier spécifiquement l'entrée de plusieurs virus, dont le virus de l'hépatite C et les coronavirus (Bartosch et al., 2003; Corver et al., 2007).

Selon les souches, la protéine S de HCoV-229E possède de 26 à 33 sites potentiels de N-glycosylation. La protéine S est donc fortement N-glycosylée. Le rôle de la N-glycosylation dans les fonctions de la protéine S n'a été que peu étudié. Afin d'étudier leur rôle dans les fonctions d'entrée de la protéine S, nous avons employé une stratégie de mutagenèse dirigée et étudié les effets de ces mutations à l'aide de particules rétrovirales pseudotypées avec la protéine S de HCoV-229E. Nous avons

aussi étudié le rôle de ces N-glycanes dans la reconnaissance de la protéine par des anticorps neutralisants. Afin de pouvoir étudier la neutralisation du virus par des sérums de patients, nous avons cloné la séquence de la protéine d'un virus de patient. En effet depuis l'isolement de la souche VR-740 dans les années 1960, le virus a évolué et il a été montré que la souche VR-740 n'est pas neutralisée par des anticorps dirigés contre des protéines S de souches cliniques. Inversement, les anticorps dirigés contre les souches clinique sont inefficaces contre VR-740 (Shirato et al., 2012).

Ces résultats participeront à une meilleure connaissance du coronavirus respiratoire humain 229E, dans le but de développer des moyens de lutter contre les coronavirus.



## 2 MATERIEL ET METHODES

---

### 2.1 CELLULES

Les cellules embryonnaires de rein humain (HEK293T) et cellules de carcinome hépatocellulaire humain (Huh-7) ont été maintenues à 37°C et 5 % de CO<sub>2</sub> dans du DMEM (*Dulbecco's Modified Eagle's medium*) contenant 10 % de SVF (Sérum de Veau Foetal).

### 2.2 ANTICORPS

Aucun anticorps monoclonal dirigé contre la protéine S du HCoV-229E n'étant actuellement disponible, toutes les protéines ont un tag VSVG à l'extrémité C-terminale afin de permettre leur détection. Les anticorps monoclonaux P5D4 (anti-VSVG) et R187 (anti-capside du *murine leukemia virus* MLV; ATCC CRL1912) ont été produits au laboratoire.

### 2.3 INHIBITEURS ET PRODUITS CHIMIQUES

La chlorométhylketone (Decanoyl-Arg-Val-Lys-Arg-chlorométhylketone) a été utilisé pour inhiber la furine (Enzo Life Science) à une concentration de 5 µM. Le chlorure d'ammonium NH<sub>4</sub>Cl a été commandé chez Sigma-Aldrich et a été utilisé à des concentrations de 25 mM. Dans tous les cas, les cellules ont été traitées de 30 min avant l'infection jusqu'à 30 minutes après l'infection.

## 2.4 PLASMIDES

La séquence de la protéine S du HCoV-229E (France/Lille/2014) a été clonée à partir d'un prélèvement respiratoire provenant d'un patient infecté par HCoV-229E en 2014 et séquencée (Genoscreen). Les sites potentiels de N-glycosylation et la position du peptide signal ont été prédits à l'aide du logiciel NetNGly 1.0 du *Center for Biological Sequence Analysis* (<http://www.cbs.dtu.dk/services/>). La séquence d'ADN codant pour la protéine S (France/Lille/2014) a ensuite été synthétisée par GeneArt afin d'optimiser l'utilisation des codons et insérée dans le vecteur pHCMV entre les enzymes BglIII et NheI. Les différentes mutations ont été réalisées par PCR à l'aide du mélange Q5® *High-Fidelity DNA Polymerase* (New England Biolabs). Afin d'abolir les sites potentiels de N-glycosylation, le codon de l'asparagine a été remplacé par celui codant pour une glutamine. Afin d'abolir les sites potentiels de clivage, les codons des arginines (565, 679, 683 et 687) ont été remplacés par des asparagines. Des sites de reconnaissance par la furine ont aussi été introduits en position 565 et 687 : la séquence  ${}_{559}\text{IIAVQPR}_{565}$  a été remplacée par  ${}_{559}\text{RRRRKRR}_{565}$  et  ${}_{683}\text{RVAGR}_{687}$  par  ${}_{683}\text{RRSRR}_{687}$  (pour les primers, voir tableau 1 en annexe). Chaque plasmide obtenu a été vérifié par séquençage (Genoscreen).

La séquence codante de la protéine aminopeptidase N des acides aminés 32 à 967 a été fusionnée à la séquence codant pour un peptide signal (MIHSVFLLMFLLTPTESY) afin de permettre la sécrétion de la protéine. La séquence d'ADN a ensuite été clonée dans un vecteur d'expression pCEP4 entre les enzymes NotI et XhoI permettant l'ajout à l'extrémité C-terminale d'un tag avec 6 histidines pour permettre la purification ainsi qu'un tag V5 pour permettre la détection. Le plasmide obtenu a été vérifié par séquençage (Genoscreen).

## 2.5 SEQUENCES ET ALIGNEMENTS

Les alignements de séquences des sites S2' et S1/S2 de la protéine S de chaque coronavirus ont été réalisés avec Clustal Omega v1.2.4 (EMBL-EBI). Les

séquences utilisées et leur identifiant GenBank sont présentées dans l'annexe, tableau 2.

## 2.6 TRANSFECTION

Les cellules HEK293T ont été transfectées à l'aide de l'agent de transfection TurboFect (ThermoFischer Scientific) avec un ratio d'1 µg d'ADN pour 3 µl d'agent de transfection dans un volume final de 100 µl de milieu Opti-MEM selon les recommandations du fabricant.

## 2.7 TRANSDUCTION

Pour produire les vecteurs lentiviraux exprimant TMPRSS2, des cellules HEK293Tensemencées en plaque 6 puits ont été co-transfectées avec trois plasmides : le pTRIP-TMPRSS2 (500ng) et les vecteurs d'expression pour gag-pol (400ng) et VSVG (100ng). 72h post-transfection, le surnageant a été récolté, filtré puis conservé à -80°C. Des cellules Huh-7ensemencées en plaque 6 puits ont été transduites avec 300 µl/puits de surnageant, le lendemain elles ont été réparties en plaque 96 puits et infectées avec les pseudoparticules 48h après transduction.

## 2.8 WESTERN-BLOT

Les protéines ont été séparées par électrophorèse sur un gel de SDS-PAGE (sodium dodecyl sulfate polyacrylamide), composé d'un gel de concentration à 5 % (5 % d'acrylamide (ratio acrylamide/bisacrylamide 37,5/1), 125 mM de Tris-HCl (trishydroxyméthylaminométhane) pH 6,8 , 0,1 % de SDS, 0,1 % d'APS (ammonium persulfate), 0,1 % de TEMED (N,N,N',N'-tetramethylethylenediamine)) et d'un gel de séparation à 10 % (10% d'acrylamide, 375 mM de Tris-HCl pH 8,8 , 0,1 % de SDS, 0,1% d'APS, 0,1 % de TEMED). La migration s'est effectuée à 70 volts pendant 20 minutes puis à 200 volts pendant 1h, dans du tampon de migration (25 mM de Tris, 194 mM de glycine, 1% de SDS). Les protéines ont ensuite été transférées sur une

membrane de nitrocellulose pendant 1h30 à 100 volts, dans du tampon de transfert (47,9 mM de Tris, 38,6 mM de glycine, 20 % d'éthanol absolu, 0,04% de SDS). Les sites d'interactions non spécifiques de la membrane ont été bloqués par une solution de 5 % de lait écrémé, 0,1 % de NP40 (Nonidet-P40) dans du PBS pendant 1 h à température ambiante. La protéine S et ses mutants ont été immunodétectés par un anticorps primaire monoclonal anti-VSVG, dans la solution de blocage, par incubation une nuit à 4°C sous agitation. La membrane a été lavée trois fois par un tampon de lavage (0,1 % NP40 dans du PBS) pendant 10 min. La membrane a ensuite été incubée avec un anticorps secondaire de chèvre anti-IgG de souris couplé à la HRP (horseradish peroxidase) (au 1/10000) pendant 45 min à température ambiante sous agitation. La membrane a été lavée trois fois par le tampon de lavage pendant 10 min. La révélation a été réalisée à l'aide d'un substrat chimioluminescent (enhanced chemiluminescent, ECL) provenant du kit de révélation « Pierce ECL Western Blotting Substrate » et à l'aide d'un film autoradiographique.

## 2.9 CLIVAGE ET BIOTINYLATION

Des cellules HEK293Tensemencées en plaques de 6 puits recouvertes de poly-D-lysine ont été transfectées avec 1 µg par puits de plasmide exprimant la protéine S ou ses mutants. A 48h post-transfection, les cellules ont été rincées avec du DMEM chaud sans SVF, puis traitées par 0 ou 2 µg/ml de trypsine par puits pendant 20 min à 37°C. Les cellules ont ensuite été rincées par du PBS froid, sur glace. Le marquage des protéines en surface des cellules a été réalisé en incubant les cellules avec une solution de 300 µg/ml de biotine clivable et non perméable (EZ link™sulfo-NHS-SS-Biotin) dans du PBS froid deux fois 15 min à 4°C. La biotine non liée a été bloquée par deux lavages de 10 min à 4°C avec une solution à 50 mM de glycine dans du PBS. Après rinçage par du PBS, les cellules ont été lysées par un tampon B1 (50 mM de Tris-HCl pH 7,5, 100 mM de NaCl, 2 mM d'EDTA, 1 % de triton X-100, 0,1 % de SDS), supplémenté par des inhibiteurs de protéase (Roche). Après centrifugation à 14000 rpm pendant 10 min à 4°C, le surnageant a été récupéré, et les protéines biotinylées ont été précipitées par des billes de streptavidine-agarose sous agitation à 4°C pendant 12 h. Les billes ont ensuite été

lavées trois fois avec le tampon B1, deux fois avec le tampon B2 (50 mM de Tris-HCl pH 7,5, 100 mM de NaCl, 2 mM d'EDTA, 0,1 % de triton X-100, 0,5 % de SDS, 0,5 % de deoxycholate), deux fois avec le tampon B3 (50 mM de Tris-HCl pH 7,5, 500 mM de NaCl, 2 mM d'EDTA, 0,1 % de triton X-100) et une fois avec le tampon B4 (50 mM de Tris-HCl pH 7,5, 100 mM de NaCl, 2 mM d'EDTA). Les billes ont été re-suspendues dans le tampon de charge Laemmli (pour une concentration finale de 1 % SDS, 10 % glycérol, 3 mM de Tris-HCl pH 6,8, 50 mM de dithiothréitol (DTT) et de bleu de bromophénol), chauffées 10 min à 70 °C. Après centrifugation 1 min à 14000 G, les échantillons ont été déposés sur un gel 10% SDS-PAGE (sodium dodecyl sulfate-polyacrylamide) pour une analyse par western-blot.

## **2.10 CO-IMMUNOPRECIPITATION PROTEINE S-RECEPTEUR SOLUBLE**

### **2.10.1 Production du récepteur soluble**

Pour faciliter l'étude des interactions entre hAPN et la protéine S, une forme soluble de hAPN (shAPN) a été produite (Breslin et al. 2003). Le plasmide contenant la séquence codant pour shAPN a été transfectée dans des cellules HEK293T. Le surnageant a été récolté, centrifugé 5 min à 1000 rpm pour séparer les débris puis filtré avec un filtre 0,22 µm. Les surnageants aliquotés ont été conservés à -80°C.

### **2.10.2 Co-immunoprécipitation**

Des cellules HEK293T ensemencées en plaques de 6 puits recouvertes de poly-D-lysine ont été transfectées avec 1 µg par puits de plasmide exprimant la protéine S ou ses mutants. A 48h post-transfection, les cellules ont été rincées par du PBS froid, sur glace, puis lysées par 250 µl de tampon de lyse (1 % de triton X-100 dans du PBS froid, supplémenté par des inhibiteurs de protéases (Roche)). Les lysats ont été incubés pendant 1h à 4°C avec du récepteur soluble aminopeptidase N (shAPN) puis avec billes de protéine G, préalablement incubées avec un anticorps anti-VSVG, pendant 3h à 4°C. Les billes ont ensuite été lavées trois fois avec le tampon de lyse. Les billes ont été re-suspendues dans le tampon de charge Laemmli, et chauffées 10 min à 70 °C. Après centrifugation 1 min à 14000 G, les

échantillons ont été déposés sur un gel 10% SDS-PAGE pour une analyse par western-blot.

## 2.11 PRODUCTION DE PSEUDO-PARTICULES ET TESTS D'INFECTION

Pour générer des pseudo-particules de HCoV-229E (229Epp), des cellules HEK293T ont été co-transfectées avec des vecteurs d'expression codant les composants viraux : la glycoprotéine S ou un de ses mutants (300 ng), les protéines de capsid et polymérase (pTG-gag-pol, 300 ng) du *murine leukemia virus* (MLV) et un minigénome contenant un gène rapporteur de la *Firefly-luciférase* (pTG-luc, 400 ng). Les cellules ont été incubées 72h à 33°C et le surnageant contenant les particules a été récolté et filtré sur des membranes avec des pores de 0,45 µm. Des cellules Huh-7 préalablement réparties en plaques 96 puits ont ensuite été infectées et lysées 48h plus tard. L'activité luciférase a été détectée à l'aide du Luciferase Assay Kit (Promega).

Pour vérifier la production de pseudoparticules dans les surnageants, les pseudoparticules ont été concentrées par ultracentrifugation à 175000g à 4°C pendant 3h sur un coussin de sucrose 30 %. Chaque culot a été re-suspendu avec du tampon de charge Laemmli et analysé en western-blot.

## 2.12 TEST D'INFECTION EN SURFACE

Des cellules Huh-7 ont été réparties en plaques de 48 puits 24h avant infection. Les cellules ont été pré-incubées 1h en présence ou absence de 25 mM de NH<sub>4</sub>Cl à 37°C pour inhiber l'entrée virale par la voie endosomale, puis transférées sur de la glace. Des 229Epp produites comme décrit précédemment ont été liées à la surface cellulaire pendant 2h à 4°C dans du DMEM/hepes (DMEM, 0,2% de BSA, 25 mM d'hépes) en absence ou en présence de 25 mM de NH<sub>4</sub>Cl. Les cellules ont ensuite été incubées 5 min à 37°C dans du DMEM/hépes contenant 2 µg/ml de trypsine en présence ou en absence de 25 mM de NH<sub>4</sub>Cl. Puis les cellules ont été incubées

pendant 30 min supplémentaires dans du milieu complet contenant ou non 25 mM de  $\text{NH}_4\text{Cl}$ , puis le milieu a été remplacé. L'activité luciférase a été détectée à l'aide du Luciferase Assay Kit (Promega) 48h après infection.

## 2.13 ANALYSE STATISTIQUE

Les données expérimentales ont été analysées et les graphiques réalisés à l'aide de Prism (GraphPad). Pour l'analyse de significativité statistique, des tests de one-way ANOVA test ont été utilisés. Pour décrire la significativité de la valeur p, nous avons utilisé les codes suivants : non significatif (n.s.), pour  $P > 0,05$ ; significatif (\*) pour  $P \leq 0,05$ ; fortement significatif (\*\*) pour  $P \leq 0,01$ ; très fortement significatif (\*\*\*) pour  $P \leq 0,001$ .

Pour les analyses statistiques de la neutralisation par les anticorps, le logiciel SAS (SAS v9.4, SAS Institute). Les données sont présentées sous forme de moyenne  $\pm$  écart-type estimé ou médiane [intervalle interquartile] pour les variables continues, et sous forme de proportions pour les variables catégorielles. L'évolution du ratio de chaque mutant (moyenne de 3 réplicats, divisée par la moyenne des 3 réplicats de la dilution 0 pour le mutant et le sérum considérés) a été étudiée par un modèle de régression linéaire. Le modèle incluait la dilution, le type de mutant, le type de sérum et leur interaction comme facteurs prédictifs indépendants à effets fixes, et l'expérience comme facteur à effets aléatoires. La relation entre la réponse ln-transformée et la dilution était linéaire. Les hypothèses sous-jacentes de la régression ont été vérifiées graphiquement à l'aide des résidus et un test de Shapiro-Wilk a été effectué pour vérifier la normalité de la distribution des résidus standardisés. Pour toutes les analyses, le risque de première espèce était fixé à 5%.





### **3 PARTIE I : ÉTUDE DU CLIVAGE PROTEOLYTIQUE DE LA PROTEINE S DU CORONAVIRUS HUMAIN 229E**

---

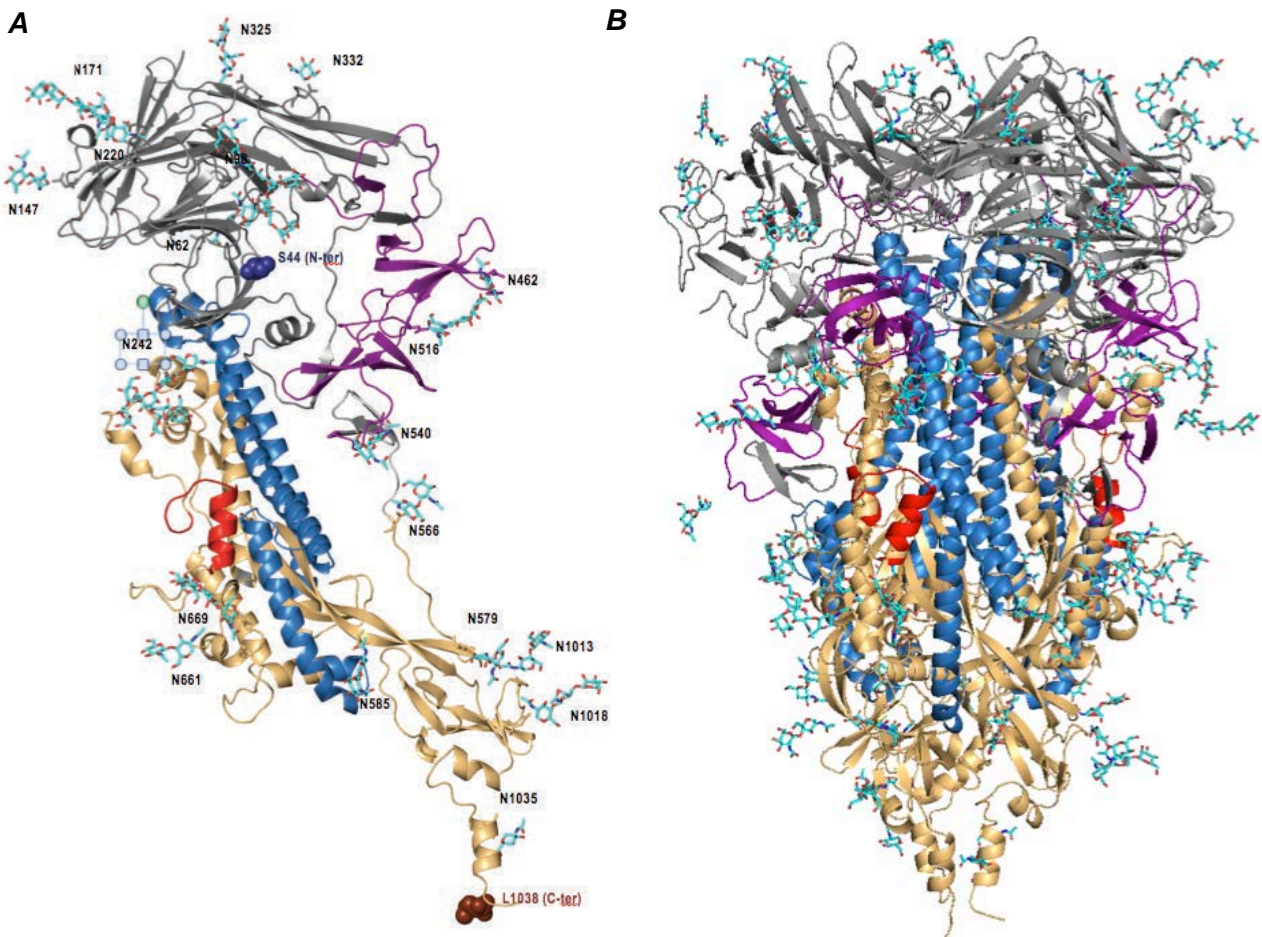
#### **3.1 RESULTATS**

##### **3.1.1 Modélisation de la protéine S du coronavirus humain 229E**

Pour initier nos travaux sur la protéine S du virus HCoV-229E, nous avons d'abord cloné la séquence de la protéine S d'un patient infecté par HCoV-229E et hospitalisé à Lille, France en 2014. La séquence codon-optimisée de la protéine S a été insérée dans un vecteur d'expression pCDNA3.1(+) et fusionnée avec une séquence codant un tag VSVG à l'extrémité C-terminale afin de faciliter la détection de la protéine par immunoblot. La séquence de l'isolat français de la protéine S montre une homologie de 99% avec les isolats de la protéine S de Niigata, 0349 et USA/932-72/1993 et seulement 94% d'homologie avec la souche de laboratoire VR-740 (Figure 14A).

La modélisation de la protéine S du HCoV-229E a été réalisée en collaboration avec l'équipe du Dr Petrescu (Figure 13). Cette modélisation a été construite en utilisant la structure résolue par cryo-microscopie électronique de la protéine S du HCoV-NL63 (Walls et al., 2016a). Les protéines S des HCoV-229E et HCoV-NL63 sont proches, avec une homologie de 56%. Les acides aminés 44 à 1045 sont inclus dans la structure. La région C-terminale incluant HR2 et le domaine transmembranaire ne sont pas inclus. Il a été montré par cryo-microscopie électronique et spectrométrie de masse pour HCoV-NL63 que 34 des 39 sites potentiels de N-glycosylation présents sur la protéine S étaient effectivement glycosylés (Walls et al., 2016a). Dans la protéine S France/Lille/2014, nous avons identifié 33 sites potentiels de N-glycosylation. Les 24 sites de N-glycosylation conservés entre la protéine S du HCoV-229E et HCoV-NL63 ont été représentés sur la modélisation. Les sept sites potentiels de N-glycosylation présents uniquement sur

la protéine S du HCoV-229E ne sont pas représentés. Cette modélisation montre que comme pour les coronavirus HCoV-NL63 et PDCoV, qu'une partie importante de la surface des trimères de protéine S est recouverte de N-glycanes.



**Figure 13: Modélisation de la protéine S du HCoV-229E.** A Modélisation d'un monomère de protéine S. B. Modélisation d'un trimère de protéine S. Gris: domaine S1, or : domaine S2, violet : domaine de liaison au récepteur, rouge : peptide de fusion, bleu : HR1.

### 3.1.2 Conservation des sites de clivage de la protéine S

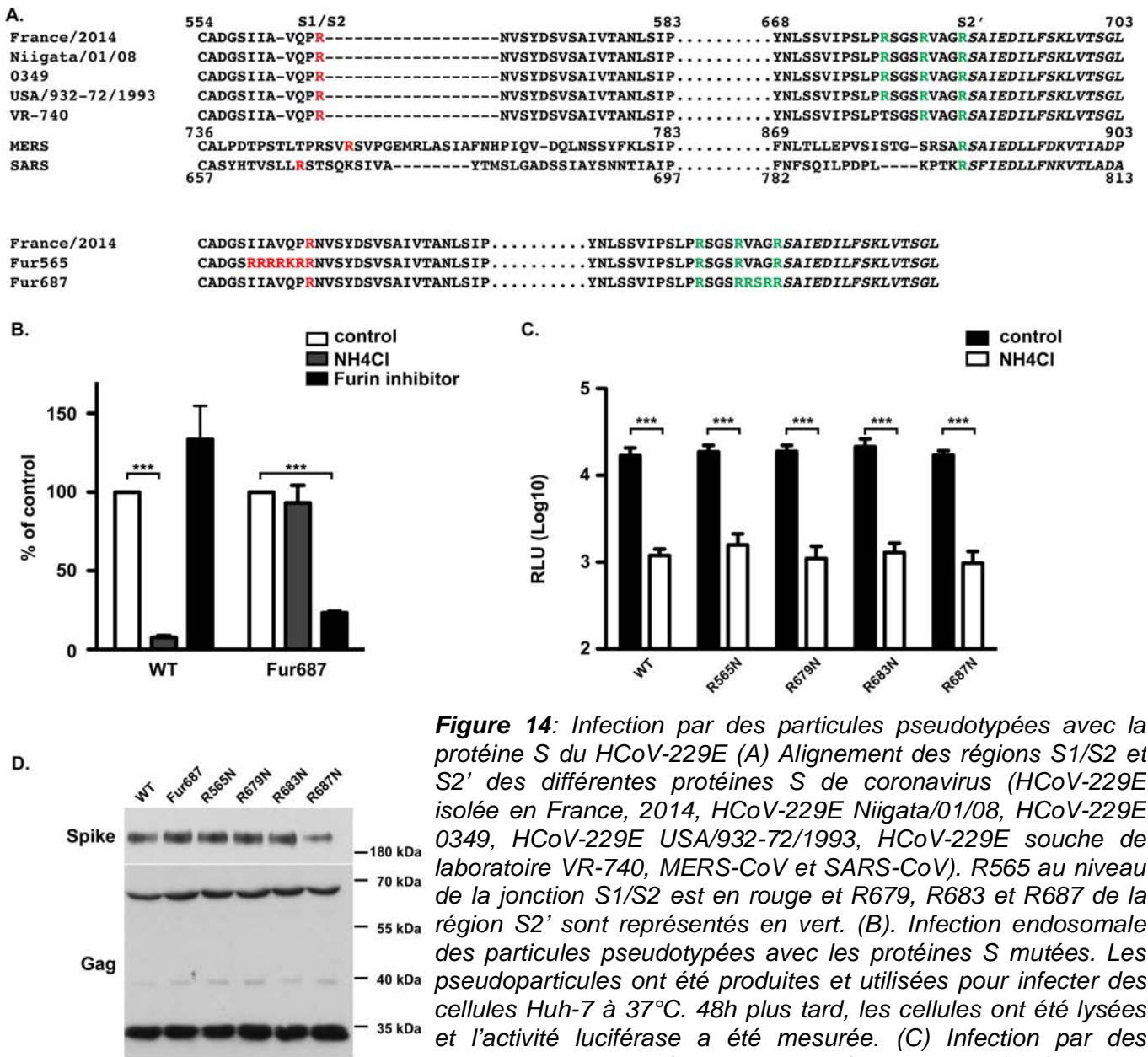
La protéine S du virus HCoV-229E est susceptible de subir un processus protéolytique similaire à celui du SARS-CoV lors de l'activation de la fusion par la trypsine ou d'autres protéases à sérine comme TMPRSS2 avec un premier clivage

de la jonction S1/S2 suivi d'un deuxième clivage S2'. Ces protéases clivent les protéines au niveau de résidus basiques. Nous avons d'abord analysé les résidus présents au niveau des jonctions S1/S2 et dans la région S2' de la protéine S de HCoV-229E (Figure 14A). Des séquences d'*alphacoronavirus* et de *betacoronavirus* humains ont été alignées.

L'alignement de la région S1/S2 montre qu'un site de clivage par la furine est présent sur le MERS-CoV dans cette région. La protéine S du virus HCoV-229E quant à elle possède un unique résidu basique (R565) comme le SARS-CoV. L'arginine 565 (Figure 14A) est immédiatement suivie d'un site potentiel de N-glycosylation très conservé parmi les différentes souches de HCoV-229E. L'alignement de la région S2' montre que le résidu arginine en position 687 sur la séquence d'HCoV-229E (France/Lille/2014) est fortement conservé sur toutes les séquences. Cette arginine est localisée à l'extrémité N-terminale d'un peptide très conservé ayant les caractéristiques d'un peptide de fusion (Madu et al., 2009). Dans la région S2', HCoV-229E possède deux arginines supplémentaires en position 679 et 683. Seule la souche de laboratoire VR-740 ne contient pas d'arginine à la position 679.

### 3.1.3 Rôle de la région S2' dans l'entrée du virus HCoV-229E

Nous avons étudié si le clivage S1/S2 de la protéine S du virus HCoV-229E était nécessaire à l'activation de la fusion. Nous avons aussi déterminé quelle arginine de la région S2' était reconnue par les protéases à sérine pour activer la fusion. Pour étudier l'entrée du virus HCoV-229E, nous avons utilisé un modèle de particules rétrovirales pseudotypées avec la protéine S du virus HCoV-229E. Ces particules sont constituées d'une nucléocapside rétrovirale (*Murine Leukemia Virus*) ayant incorporé un minigénome contenant un gène rapporteur, la luciférase, enveloppées d'une bicouche lipidique sur laquelle est enchâssée la protéine S de HCoV-229E.



**Figure 14:** Infection par des particules pseudotypées avec la protéine S du HCoV-229E (A) Alignement des régions S1/S2 et S2' des différentes protéines S de coronavirus (HCoV-229E isolée en France, 2014, HCoV-229E Niigata/01/08, HCoV-229E 0349, HCoV-229E USA/932-72/1993, HCoV-229E souche de laboratoire VR-740, MERS-CoV et SARS-CoV). R565 au niveau de la jonction S1/S2 est en rouge et R679, R683 et R687 de la région S2' sont représentés en vert. (B) Infection endosomale des particules pseudotypées avec les protéines S mutées. Les pseudoparticules ont été produites et utilisées pour infecter des cellules Huh-7 à 37°C. 48h plus tard, les cellules ont été lysées et l'activité luciférase a été mesurée. (C) Infection par des particules pseudotypées avec la protéine sauvage (WT) ou avec la protéine S dans laquelle un site de clivage par la furine a été introduit dans la région S2' (Fur687). L'infection a été réalisée en présence ou absence de 25 mM de NH<sub>4</sub>Cl ou de 5 μM d'inhibiteur de furine. Les résultats sont présentés en pourcentage d'infection en absence de drogue et correspondent à la moyenne de trois expériences indépendantes. Les barres d'erreur représentent l'erreur standard des moyennes (SEM). (D) Analyse de la formation des pseudoparticules. Les pseudoparticules ont été produites avec les différentes protéines S mutées et concentrées sur coussin de sucrose, puis les pseudoparticules ont été lysées et la production a été analysée par western-blot avec un anticorps anti-VSVG afin de détecter la protéine S et un anticorps anti-gag.

Dans un premier temps, pour confirmer le rôle de la région S2' de la protéine S lors de l'entrée, nous avons inséré un site de clivage par la furine (Fur687) dans la protéine S, à la position N-terminale du peptide de fusion.

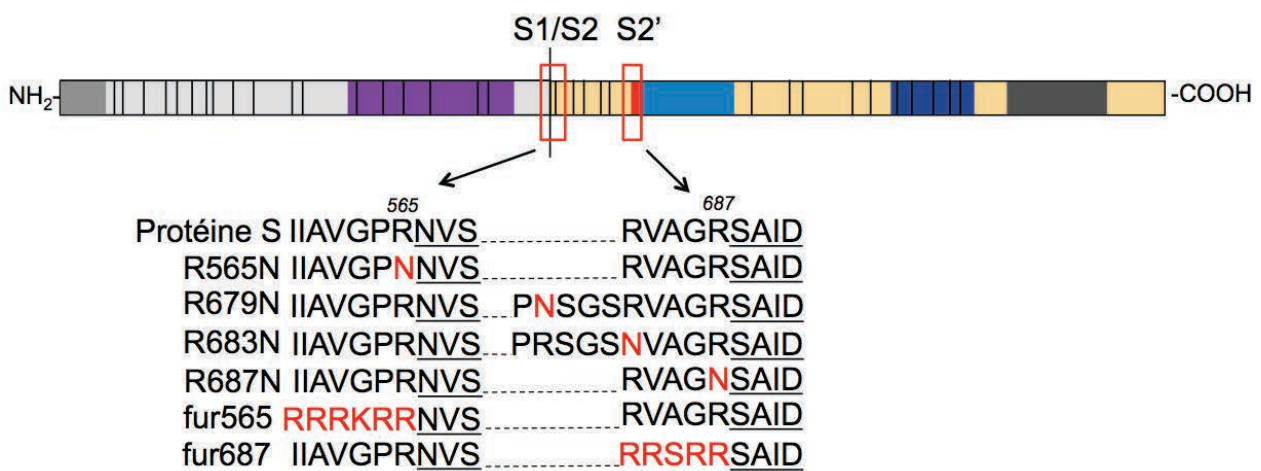
Afin de confirmer la fonctionnalité du site de reconnaissance par la furine, les cellules Huh-7 ont été infectées par des particules pseudotypées avec la protéine sauvage ou avec le mutant Fur687 (Figure 14B). La neutralisation du pH endosomal par le traitement par le chlorure d'ammonium inhibe l'infection des pseudoparticules de HCoV-229E. Ce résultat montre que HCoV-229E infecte les cellules Huh-7 par voie endosomale et utilise des protéases actives à bas pH pour médier sa fusion. Comme montré dans la figure 14B, les virions pseudotypés avec la protéine S contenant un site de clivage par la furine à la position S2' sont insensibles à la neutralisation du pH endosomal, mais sont affectés par un inhibiteur de la furine convertase.

Nous avons montré que la protéine S contenant un site de clivage par la furine ne dépend pas des cathepsines endosomales lors de l'entrée, mais nécessite la furine. Ce résultat confirme l'implication de la région S2' lors de l'activation de la fusion du HCoV-229E, comme observé pour les autres coronavirus (Burkard et al., 2014).

### **3.1.4 Infection en surface médiée par la trypsine et TMPRSS2**

Le rôle de la jonction S1/S2 et le rôle spécifique des 3 résidus arginine localisés dans la région S2' (R679, R683 et R687) lors de l'activation protéolytique de la protéine S du HCoV-229E par les sérines protéases de type trypsine a été étudié. Les résidus arginines présents en S1/S2 et en S2' ont été mutés individuellement (R565N, R679N, R683N et R687N) (Figure 15). Nous avons d'abord regardé l'effet de ces mutations lorsque le virus entre par la voie endosomale en utilisant les protéases de type cathepsine présentes dans les cellules. En parallèle, la production de particules a été analysée en western-blot. La production de particules

pseudotypées avec les mutants R565N, R679N, R683N et R687N est similaire aux particules pseudotypées avec la protéine S (Figure 14D). L'incorporation des glycoprotéines dans les 229Epp est similaire quelque soit le mutant. Les particules pseudotypées incorporant les mutants R565N, R679N, R683N et R687N ont des niveaux d'infectiosités comparables à celles incorporant la protéine S (Figure 14C). D'autre part, nous avons confirmé que ces différents mutants utilisaient bien la voie endosomale, puisque leur infection est sensible à la neutralisation du pH endosomal par le chlorure d'ammonium.



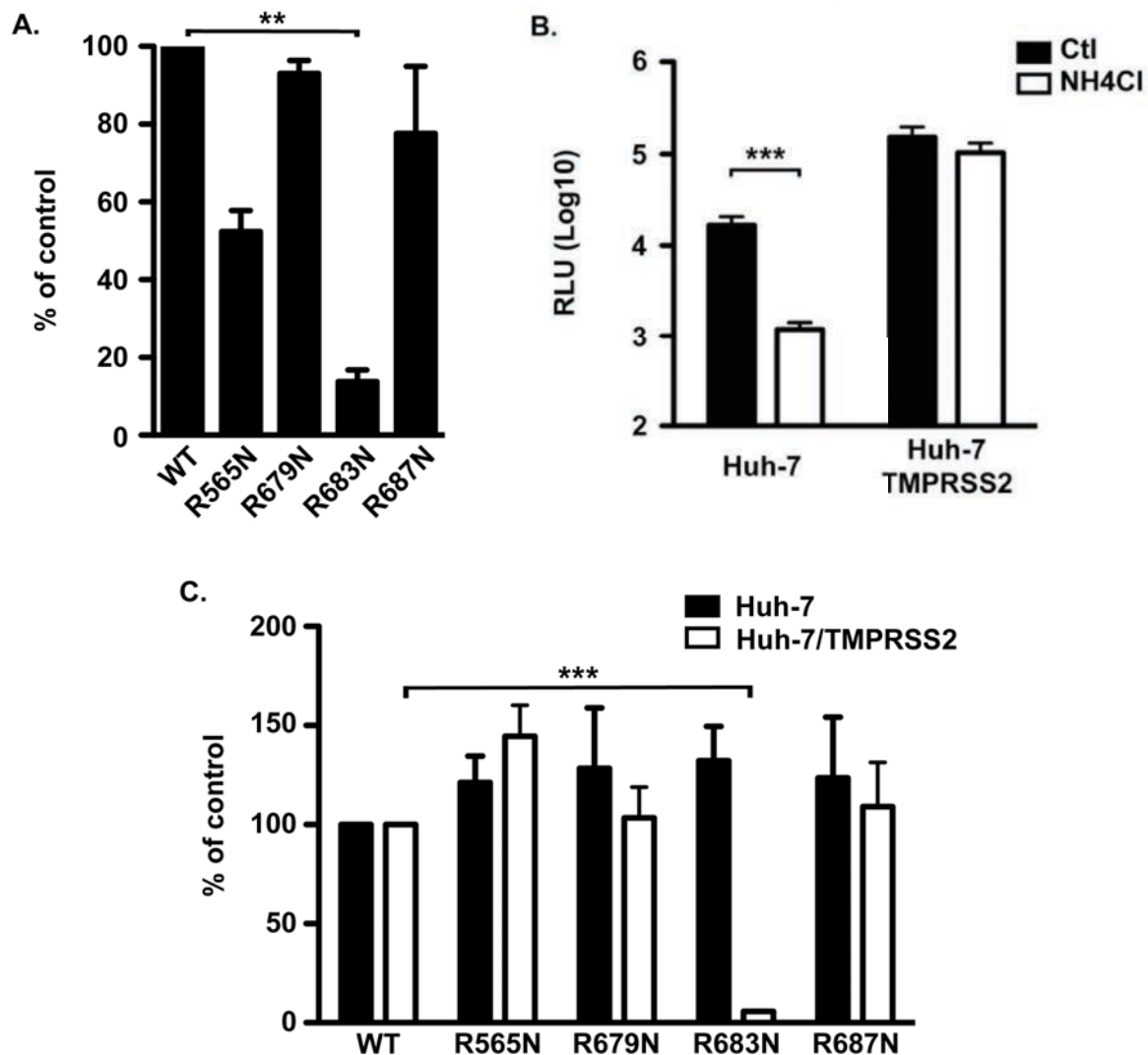
**Figure 15: Schéma fonctionnel de la protéine S et séquences des régions S1/S2 et S2' de la protéine S et des différents mutants.** Les sites potentiels de N-glycosylation sont représentés par un trait noir court, et le site potentiel de clivage S1/S2 est représenté par un grand trait noir. Gris clair: domaine S1, or: domaine S2, violet: domaine de liaison au récepteur, rouge: peptide de fusion, bleu clair: HR1, bleu foncé: HR2, domaine transmembranaire: gris foncé. Les arginines (R) et lysine (K) potentiellement impliquées dans le clivage sont représentées respectivement en rouge et orange. Les sites potentiels de N-glycosylation sont soulignés.

*In vivo*, il est considéré que les coronavirus respiratoires utilisent une voie d'entrée dépendante des TTSP. Pour étudier cette voie d'entrée, nous avons d'abord utilisé la trypsine comme protéase modèle. Pour cela, les particules pseudotypées avec les différents mutants arginines des régions S1/S2 et S2' ont d'abord été liées à la surface des cellules à 4°C et la fusion a été induite par un court traitement (5 min) avec de la trypsine (Figure 16A). Pour inhiber l'infection par la voie endosomale, les cellules ont été pré-traitées avec du chlorure d'ammonium afin de neutraliser le pH

endosomal. Pour le mutant R683N, le traitement par la trypsine n'a pas permis de rétablir une infection similaire à celle de la protéine sauvage, suggérant que la protéolyse par la trypsine a lieu au niveau du résidu R683 (Figure 16A). La mutation de R679 et R687 n'a pas d'effet sur l'infection médiée par la trypsine. L'infection induite par la trypsine lorsque l'arginine présente à la jonction S1/S2 est mutée (R565N) est diminuée par rapport à celle de la protéine sauvage, bien que ce ne soit pas statistiquement significatif, suggérant qu'un clivage dans cette région puisse être nécessaire pour l'activation de la fusion. Cependant ce clivage ne semble pas indispensable puisque l'inhibition de l'infection n'est que partielle.

Afin de confirmer ces résultats, des cellules Huh-7 ont été transduites par un vecteur lentiviral exprimant la protéase TMPRSS2, physiologiquement plus pertinente que la trypsine (Figure 16B). L'infection par les 229Epp des cellules exprimant TMPRSS2 est supérieure à celle des cellules contrôles (Figure 16B). L'infection dans les cellules exprimant TMPRSS2 n'est plus sensible au traitement par le chlorure d'ammonium confirmant que les 229Epp n'utilisent pas la voie endosomale dans ces cellules mais la protéase TMPRSS2 exprimée à la surface des cellules. La mutation individuelle du résidu basique de la jonction S1/S2 (R565) et des deux résidus de la région S2' (R679 et R687) n'affecte pas l'infection médiée par TMPRSS2. Au contraire, la mutation R683N inhibe complètement l'infection médiée par TMPRSS2, de la même manière qu'avec la trypsine, ce qui indique que l'arginine 683 est un résidu majeur pour l'entrée médiée par TMPRSS2 (Figure 16C).

La mutation de l'arginine présente à la jonction S1/S2 (R565) n'a aucun effet sur l'infection médiée par TMPRSS2 tandis que la mutation de R683 inhibe les infections médiées par la trypsine ou TMPRSS2. Ces résultats montrent que contrairement au SARS-CoV ou au MERS-CoV, la protéine S de HCoV-229E ne semble pas subir un clivage séquentiel lors de l'activation de la fusion par des protéases à sérine. En effet, l'activation semble dépendre uniquement du clivage dans la région S2' au niveau de l'arginine 683.

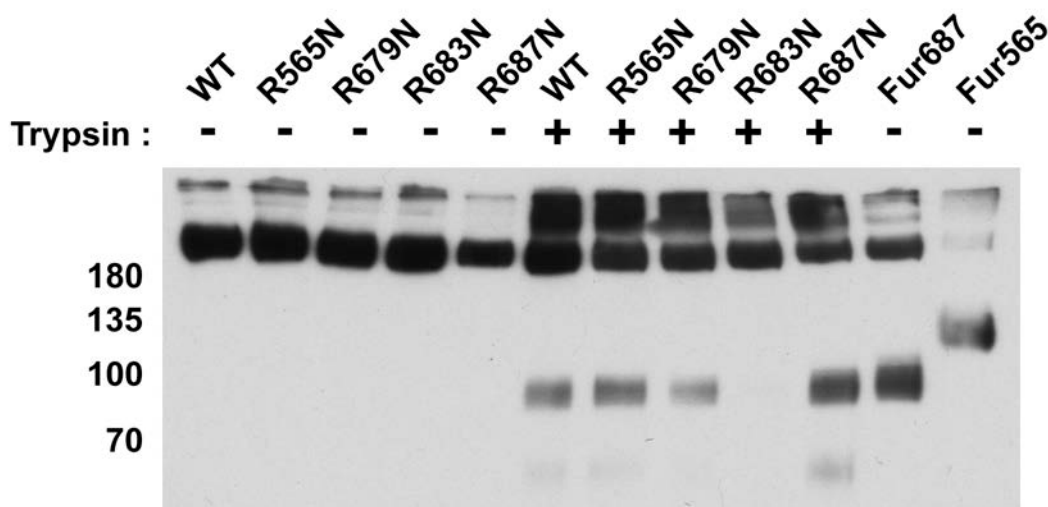


**Figure 16: Infection en surface des 229Epp médiée par la trypsine et par TMPRSS2.** (A) Infection médiée par la trypsine à la surface cellulaire. Les particules pseudotypées avec les différentes protéines mutées ont été liées à la surface cellulaire à 4°C en présence de 25 mM de NH<sub>4</sub>Cl et ont été traitées pendant 5 min par 2 µg/ml de trypsine pour induire la fusion. 48h après, les cellules ont été lysées et l'activité luciférase a été mesurée. (B) Les cellules Huh-7 ont été transduites ou non par un lentivirus exprimant TMPRSS2. Puis les cellules ont été infectées en présence ou en absence de 25 mM NH<sub>4</sub>Cl. Les résultats sont présentés en relative light units (RLU) et correspondent à la moyenne de trois expériences indépendantes. (C) Les cellules Huh-7 transduites ou non par un lentivirus exprimant TMPRSS2 ont été infectées par les particules pseudotypées avec les différents mutants de la protéine S. Sauf mention contraire, les résultats sont présentés en pourcentage d'infection observée pour les particules pseudotypées avec la protéine sauvage et correspondent à la moyenne de trois expériences indépendantes. Les barres d'erreur représentent l'erreur standard des moyennes (SEM).



### 3.1.5 Clivage de la protéine S du HCoV-229E par la trypsine

Pour confirmer que ce clivage médié par les sérine protéases de type trypsine a lieu au niveau du résidu R683, des cellules HEK293T exprimant de façon transitoire la protéine S ou différentes protéines mutantes ont été traitées par la trypsine ou non, puis les protéines présentes à la surface cellulaire ont été biotinylées à 4°C. Après la précipitation des protéines biotinylées avec des billes de streptavidine, les produits de clivage ont été analysés par western-blot. Des mutants contenant un site de clivage par la furine entre S1 et S2 (Fur565) et au niveau de S2' (Fur687) ont été utilisés comme contrôles afin de déterminer la taille des produits de clivage correspondant à ces deux positions (Figure 17).



**Figure 17: Clivage de la protéine S du HCoV-229E par la trypsine.** La protéine S et ses mutants ont été exprimés en cellules HEK293T. 48h après, les cellules ont été traitées ou non par 2 µg/ml de trypsine pendant 1h à 37°C. Les protéines ont été biotinylées à 4°C puis lysées. Les protéines biotinylées ont été précipitées par des billes de streptavidine puis analysées par western-blot à l'aide d'un Ac anti-VSVG détectant l'extrémité C-terminale de la protéine.

En absence de trypsine et de toute modification post-transductionnelle, la taille attendue de la protéine S est d'environ 130 kDa. Etant donné que cette protéine est très fortement N-glycosylée, la taille observée est d'environ 180 kDa en absence de clivage. Le clivage de la protéine S sauvage par la trypsine montre l'apparition d'une bande supplémentaire de faible intensité de 85 kDa et des agrégats de haut poids moléculaire qui sont aussi observés en absence de trypsine mais en moindre quantité. En théorie, toujours en absence de modification post-traductionnelle, un clivage en S1/S2 à la position 565 devrait générer deux produits de clivage de 61 kDa (domaine S1) et 67 kDa (domaine S2). Un clivage en S2' en position 687 devrait générer deux produits de clivage de 75 et 53 kDa (fragment S2'). L'anticorps utilisé reconnaît le tag positionné à l'extrémité C-terminale de la protéine, ainsi les mutants Fur565 et Fur687 permettent de détecter respectivement les fragments S2 et S2' et de déterminer leur taille apparente. Pour le mutant Fur565 une bande d'environ 110-120 kDa est détectée et pour le mutant Fur687, une bande d'environ 85 kDa est détectée. Ainsi le clivage par la trypsine de la protéine S observé semble correspondre à un clivage au niveau de la région S2', au contraire des autres coronavirus pour lesquels la trypsine induit majoritairement un clivage à la jonction S1/S2 (Belouzard et al., 2009; Millet and Whittaker, 2014).

La mutation de R683 abolit le clivage induit par la trypsine confirmant la maturation protéolytique de la protéine S à cette position par la trypsine.

L'ensemble de nos résultats montrent que l'activation de la fusion du virus HCoV-229E par les protéases à sérine comme TMPRSS2 et la trypsine ne dépend pas d'un clivage séquentiel de la protéine S mais d'un seul événement de clivage après l'arginine 683.

## 3.2 DISCUSSION

Le clivage protéolytique de la protéine S permettant d'activer la fusion a été décrit pour plusieurs coronavirus (Belouzard et al., 2012). Afin d'étudier le clivage protéolytique du HCoV-229E, des alignements de séquences de coronavirus des régions S1/S2 et S2' ont été réalisés. La jonction S1/S2 est située au niveau du résidu basique R565. Un site potentiel de N-glycosylation N566, très proche de la jonction S1/S2, est conservé parmi les différentes souches de HCoV-229E. Dans la région S2', un résidu basique (R687 pour HCoV-229E) localisé à l'extrémité N-terminale du peptide de fusion est très conservé parmi les coronavirus. L'arginine localisée à la même position dans la protéine S du SARS-CoV est un site majeur de reconnaissance par la trypsine. La protéine S de HCoV-229E contient deux arginines (R679 et R683) supplémentaires en amont de l'arginine 687. L'arginine 679 est présente dans les souches cliniques du HCoV-229E tandis que la souche de laboratoire VR-740 contient une thréonine. L'arginine 683 est très conservée parmi les souches HCoV-229E. La conservation des arginines suggère un rôle important de ces résidus dans les fonctions de la protéine S et plus particulièrement lors du clivage protéolytique. Actuellement, le ou les sites de clivage de la protéine S du HCoV-229E ne sont pas connus (Bertram et al., 2013; Kawase et al., 2009). Nous avons donc essayé de déterminer si la protéine S était clivée en S1/S2 et si l'arginine R565 était impliquée dans ce clivage. De plus, l'arginine impliquée dans le clivage en S2' a été déterminée.

Nous n'avons pas observé de clivage de la protéine S par la trypsine au niveau de la jonction S1/S2, ce qui suggère que soit le clivage de cette jonction n'est pas nécessaire soit que ce clivage a lieu mais que le fragment S2 est rapidement traité au niveau de S2'. Etant donné que nous n'observons aucune accumulation de produit de clivage d'une taille plus importante lorsque R683 est muté, ceci est plutôt en défaveur d'un clivage séquentiel de la protéine comme cela a été montré pour la protéine S du SARS-CoV et du MERS-CoV (Belouzard et al., 2009; Millet and Whittaker, 2014). L'introduction d'un site de clivage par la furine dans la région S1/S2

permet un clivage protéolytique de la protéine S, mais ce mutant ne permet pas la production de pseudoparticules et l'étude de son effet sur l'entrée virale. Ce résultat suggère qu'un clivage trop important pourrait être délétère pour l'infection virale. L'infection médiée par la trypsine des pseudoparticules incorporant une protéine S avec la mutation R565N est inférieure à celle incorporant la protéine S sauvage mais l'infection n'est pas complètement inhibée. La mutation de cette arginine n'a pas d'impact sur l'infection médiée par TMPRSS2, une protéase plus pertinente du point de vue physiologique. Ce résultat suggère que l'arginine 565 n'est pas indispensable à l'activation de la fusion médiée par la trypsine ou TMPRSS2. Le clivage S1/S2 n'est donc pas primordial pour la protéine S du HCoV-229E, à l'inverse du SARS-CoV pour lequel l'arginine située à la jonction S1/S2 (R667) a un rôle majeur sur l'infection médiée à la surface cellulaire par la trypsine. Nous avons identifié un site potentiel de N-glycosylation proche de la jonction S1/S2, le N-glycane présent à cette position pourrait moduler l'accès des protéases à l'arginine 565.

L'infection des cellules à la surface cellulaire en induisant la fusion du virus par la trypsine est jusqu'à trois fois supérieure au niveau d'infection par la voie endosomale (non montré). L'utilisation de TMPRSS2 par les 229Epp permet une infection à la surface jusqu'à six fois supérieure au niveau d'infection par la voie endosomale (Figure 16B). Ces résultats sont concordants avec des observations précédentes sur le SARS-CoV qui suggèrent que l'entrée à la surface cellulaire est bien plus efficace que par la voie endosomale (Matsuyama et al., 2005). De plus, les travaux de Shirato *et al.* montrent que la capacité du virus HCoV-229E à utiliser une voie d'entrée plutôt qu'une autre diffère selon les souches. En effet, les souches cliniques semblent utiliser la voie endosomale dépendante de la cathepsine moins efficacement que la souche HCoV-229E de laboratoire (souche VR740 de l'*American Type Culture Collection ATCC*) (Shirato et al., 2016). En effet, le passage du virus en cellules qui expriment peu de protéase à leur surface a certainement conduit le virus à s'adapter pour utiliser d'autres protéases cellulaires comme les protéases endosomales. D'autre part, il a été montré que l'inhibition de l'entrée par la voie endosomale par l'inhibition de la cathepsine L est moins efficace pour le virus HCoV-229E que le virus du SARS-CoV (Kawase et al., 2009), ce qui suggère qu'une autre

protéase que la cathepsine L pourrait être impliquée dans l'entrée de HCoV-229E. *In vivo*, la voie principale d'entrée des coronavirus respiratoires est donc probablement médiée par la voie TTSP.

Nous avons mis en évidence le rôle de la région S2' dans l'entrée du virus HCoV-229E. Lorsqu'un site de clivage par la furine est introduit en amont du peptide de fusion, l'entrée n'est alors plus dépendante de la cathepsine L et n'est plus inhibée par le chlorure d'ammonium. L'entrée est alors dépendante de la furine. Ce résultat confirme l'implication de la région S2' pour l'activation de la fusion de HCoV-229E, comme observé pour d'autres coronavirus (Belouzard et al., 2009; Burkard et al., 2014; Millet and Whittaker, 2014; Yamada and Liu, 2009). Il est intéressant de noter que la protéine S incorporée dans les pseudoparticules n'a pas été clivée par la furine dans les cellules productrices, indiquant que cette protéine ne subit pas de maturation durant la sortie des cellules productrices mais pendant l'entrée dans les cellules cibles. Ce résultat suggère que lors de la sortie et de la sécrétion du virus, le site de clivage n'est pas accessible à la furine et que lors de l'entrée, les changements conformationnels induits par la liaison au récepteur pourraient exposer le site de clivage.

L'activation du processus de fusion par la trypsine et les sérine-protéases de type trypsine (particulièrement TMPRSS2 exprimé dans l'épithélium respiratoire humain) semble dépendre fortement du résidu R683 comme site de clivage. L'arginine 683 n'a cependant pas de rôle lors de l'entrée par voie endosomale. En effet, lorsque HCoV-229E est internalisé, le virus est dépendant de la cathepsine L endosomale pour médier la fusion. Mais à la différence des autres protéases, la cathepsine L peut reconnaître plusieurs résidus et les sites précis de protéolyse restent peu clairs. Des résultats similaires ont été rapportés pour le SARS-CoV et le MERS-CoV (Belouzard et al., 2009; Millet and Whittaker, 2014). Ces résultats montrent la redondance des mécanismes d'activation par clivage protéolytique.

L'arginine 679 ne semble pas être impliquée dans l'entrée virale médiée par TMPRSS2 (Shirato et al., 2016), ce que nos résultats semblent confirmer. Cette arginine est présente dans les souches cliniques du HCoV-229E mais pas dans la

souche de laboratoire VR-740. La souche VR-740 isolée dans les années 60 a très probablement subi des modifications liées au passage de la souche en culture cellulaire mais malheureusement la séquence de l'isolat d'origine n'est pas disponible. Il a été montré que les souches cliniques utiliseraient de préférence la voie TTSP (Shirato et al., 2016, 2017), ce qui suggère que la présence du résidu arginine à la position 679 n'est pas impliquée dans cette adaptation.

Il est surprenant que la mutation de R687 n'ait pas d'effet sur l'infection médiée par la trypsine. En effet, R687 semble être localisé à l'extrémité N-terminale du peptide de fusion et l'arginine localisée à la même position dans la protéine S du SARS-CoV est un site majeur de reconnaissance par la trypsine. Il a été montré que la position du site de clivage de la région S2' du SARS-CoV pouvait être décalée par rapport au peptide de fusion mais avec un effet sur l'activation de la fusion et sur l'infection (Belouzard et al., 2010). De plus, même si la localisation du site de clivage présente une flexibilité, certaines positions dans le voisinage du peptide de fusion ne semblent pas accessibles aux protéases. Pour la protéine S du HCoV-229E, le site de clivage par la trypsine n'est pas localisé à l'extrémité N-terminale du peptide de fusion mais est présent 3 résidus en amont. Aucun produit de clivage n'a été observé quand R683 a été muté suggérant que les arginines 679 et 687 n'étaient pas accessibles aux protéases. En effet si ces résidus pouvaient être reconnus comme site de clivage mais sans être fonctionnels pour activer la fusion, la même bande aurait été observée que pour le mutant R683N.

Une étude du clivage de la protéine S de HCoV-229E par la trypsine avait mis en évidence un produit de clivage de 80 à 85 kDa, ce que nous avons également montré (Kawase et al., 2009). En se basant sur le poids moléculaire, les auteurs avaient conclu qu'il s'agissait du produit d'un clivage entre les domaines S1 et S2, mais le site exact de clivage n'avait pas été identifié. Il est délicat d'estimer le site de clivage en se basant uniquement sur le poids moléculaire, en effet la protéine S est fortement N-glycosylée, ce qui modifie considérablement le poids moléculaire de la protéine. Le produit de clivage détecté par Kawase *et al.* correspondait probablement à un clivage en S2' de la protéine (Kawase et al., 2009). Une autre étude a montré que la protéine S était clivée par TMPRSS2 et par la trypsine en quatre produits de

clivage identiques : 120, 65, 50 et 47 kDa (Bertram et al., 2013). La différence de poids moléculaire et de nombre de produits de clivage entre cette étude et la nôtre est probablement due à des conditions expérimentales différentes et surtout à la méthode de détection de la protéine S : nous utilisons un Ac dirigé contre un tag situé à l'extrémité C-terminale de la protéine S tandis que Bertram *et al.* ont utilisé un sérum dirigé contre HCoV-229E. Il est donc possible que le produit de clivage détecté à 120 kDa corresponde au côté N-terminal de la protéine S après clivage en S2', tandis que notre méthode détecte uniquement le côté C-terminal.

Il a été montré que la séquence du site de clivage en amont du peptide de fusion est importante pour le site intracellulaire de fusion des coronavirus (Burkard et al., 2014). En effet, la furine est localisée partiellement à la surface cellulaire et au niveau des endosomes précoces et permet la fusion du virus plus précocément que la cathepsine L dans la voie d'endocytose. La protéine S de HCoV-229E ne possède pas de site de reconnaissance de la furine à ce niveau, il est donc probable que la protéine S fusionne au niveau des endosomes tardifs, comme la protéine S du MHV-A59 (Burkard et al., 2014) lorsque les cellules n'expriment pas TMPRSS2. Nos résultats montrent que TMPRSS2 est capable d'activer la fusion de HCoV-229E en clivant la protéine de façon similaire à la trypsine au niveau de R683, entraînant une entrée plus précoce des 229Epp mais aussi plus efficace.

Tous les résultats que nous avons obtenus, c'est-à-dire l'absence d'effet de la mutation R565N ou l'absence d'observation d'un clivage S1/S2 par la trypsine en western-blot, montrent que la protéine S de HCoV-229E ne dépend pas d'un clivage séquentiel pour l'activation de sa fonction de fusion. Ce qui diffère de ce qui est observé pour d'autres coronavirus comme le SARS-CoV et le MERS-CoV. En effet pour ces deux virus, le clivage par la trypsine induit la formation de fortes quantités de produit de clivage à la jonction S1/S2, alors que le clivage protéolytique à la jonction S2' induit la formation de faibles quantités de produit de clivage (Belouzard et al., 2009; Millet and Whittaker, 2014). A l'inverse, nous n'avons pas détecté de clivage dans la région S1/S2 alors que le clivage protéolytique dans la région S2' est

très efficace. La comparaison de structures tridimensionnelles de protéine S disponibles suggère que la jonction S1/S2 et la région S2' sont exposées à la surface pour HCoV-229E, MERS-CoV et SARS-CoV. Yuan *et al.* ont également montré que le peptide de fusion (ainsi que HR1) sont exposés à la surface des trimères de protéine S du SARS-CoV et du MERS-CoV (Yuan et al., 2017). Il n'est donc pas possible pour le moment d'expliquer la différence de maturation des différentes protéines S par les protéases. Il semble donc que le clivage en S2' du HCoV-229E soit majoritaire alors qu'il est minoritaire pour le MERS-CoV et le SARS-CoV. L'ensemble de nos résultats confirment que le site de clivage immédiatement en amont du peptide de fusion, S2', est essentiel pour l'entrée virale des coronavirus. L'étude de l'activation protéolytique permettra d'apporter des indications sur le tropisme et la pathogénicité de ces virus.



## **4 PARTIE II : ETUDE DU ROLE DE LA N-GLYCOSYLATION DANS LES FONCTIONS DE LA PROTEINE S DU CORONAVIRUS HUMAIN 229E**

---

### **4.1 RESULTATS**

#### **4.1.1 Détermination des séquons N-glycosylés**

Les sites effectivement N-glycosylés ont été déterminés en collaboration avec JM Saliou (Plateforme de Protéomique et Peptides Modifiés, Lille) par chromatographie en phase liquide couplée à de la spectrométrie de masse. Pour cela, c'est la protéine S de la souche HCoV-229E VR-740 qui a été utilisée. Nous n'avons pas pu utiliser la souche France/Lille/2014, cette souche n'ayant pas été isolée. La souche clinique France/Lille/2014 de la protéine S du HCoV-229E possède 33 sites potentiels de N-glycosylation tandis que la souche de laboratoire VR-740 en possède 30. Dans la souche France/Lille/2014, nous avons numéroté les sites potentiels de N1 à N33. Ils sont répartis sur toute la protéine S, à l'exception de HR1, du domaine transmembranaire et de l'extrémité C-terminale. Ces sites ont été numérotés de N1 à N33 (Figure 18). Les sites N1, N5 et N15 ne sont pas présents sur la souche VR-740. L'homologie entre ces deux souches est de 94%. Le virus HCoV-229E VR-740 a été produit et purifié. Une partie du virus purifié a ensuite été traité par la PNGase. En effet l'étude est basée sur le fait qu'un résidu asparagine glycosylé lorsqu'il est traité par la PNGase devient un acide aspartique ce qui entraîne une modification des spectres de peptides obtenus. Le virus lysé a ensuite été déposé sur gel SDS-PAGE. Après coloration au bleu de coomassie du gel, la bande correspondant à la protéine S a été excisée afin de réaliser les analyses de glycosylation. Des traitements par différentes protéases ont été réalisés. L'analyse des spectres obtenus est encore en cours et seuls des résultats partiels sont disponibles. Nous ne nous sommes intéressés pour l'instant qu'au domaine S1.

**Tableau 2: Table récapitulative des séquons N-glycosylés**

Numéro et position des sites potentiels de N-glycosylation			Probabilité de glycosylation	Site de N-glycosylation correspondant dans la protéine S de HCoV-NL63	Glycosylation
N°	France/ Lille/ 2014	VR-740			
N1	Asn 20	absent	Aucune information	absent	
N2	Asn 23	Asn 23	Aucune information	Asn 203	Non
N3	Asn 62	Asn 62	haute	Asn 240	Oui
N4	Asn 98	Asn 98	haute	Asn 276	Oui
N5	Asn 111	absent			
N6	Asn 147	Asn 147	Aucune information	Asn 330	Oui
N7	Asn 171	Asn 171	haute	Asn 354	Oui
N8	Asn 176	Asn 176	haute	Asn 358	Oui
N9	Asn 220	Asn 220	Aucune information	Asn 403	Oui
N10	Asn 242	Asn 243	haute	Asn 426	Oui
N11	Asn 325	Asn 326	haute	Asn 506	Oui
N12	Asn 332	Asn 333	haute	Asn 512	Oui
N13	Asn 438	Asn 440	haute	Asn 626	Oui
N14	Asn 462	Asn 464	haute	Asn 645	Oui
N15	Asn 486	absent		Asn 666	Oui
N16	Asn 516	Asn 518	haute	Asn 699	Oui
N17	Asn 536	Asn 538	haute		
N18	Asn 540	Asn 542	haute	Asn 723	Oui

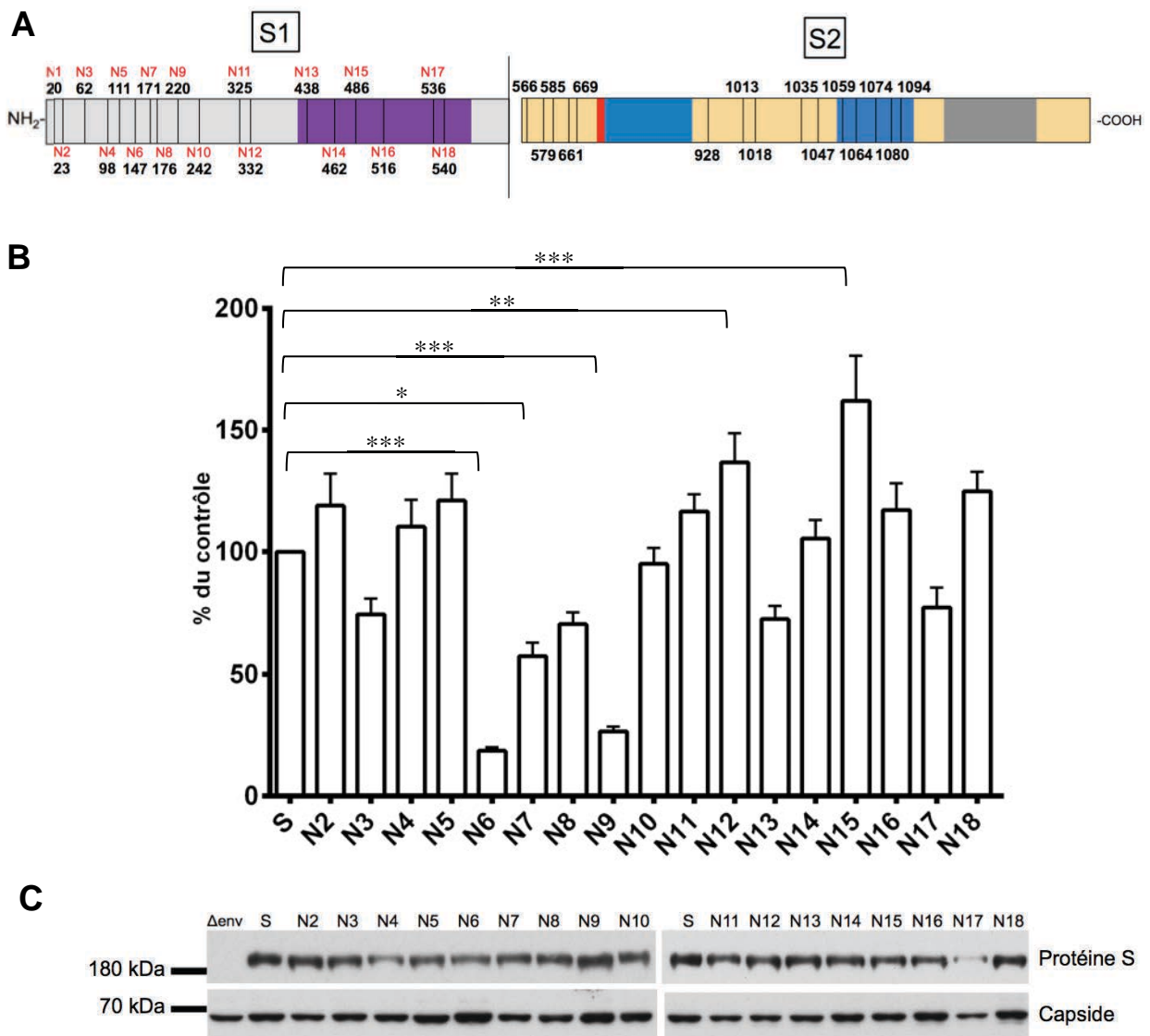
Les résultats que nous avons déjà, montrent que les sites N3, N4, N7, N8, N10, N11, N12, N13, N14, N16, N17 et N18 sont très probablement N-glycosylés (voir tableaux récapitulatifs). Aucune information n'est disponible pour l'instant pour les sites N2, N6 et N9 présents sur la souche VR-740. Comme détaillé en introduction, la structure de la protéine S de HCoV-NL63 a été résolue, or les protéines S de HCoV-229E et HCoV-NL63 sont très proches. Sur les 34 sites de N-

glycosylation présents sur la protéine S du HCoV-NL63, 25 sont strictement conservés sur la protéine S du HCoV-229E. Ainsi les sites N6 et N9 sont conservés sur la séquence de la protéine S de HCoV-NL63 dans laquelle ils sont glycosylés (Walls et al., 2016a). Il est donc probable qu'ils soient effectivement N-glycosylés sur la protéine S du HCoV-229E. Le site correspondant au site N2 n'est pas quant à lui glycosylé. Ces résultats confirment que la protéine S du HCoV-229E est fortement N-glycosylée et montrent que les N-glycanes sont fortement conservés entre les protéines S de HCoV-229E et HCoV-NL63.

#### **4.1.2 Rôle de la N-glycosylation du domaine S1 dans l'entrée virale.**

Etant donné le nombre important de sites potentiels de N-glycosylation, nous sommes concentrés dans un premier temps sur le rôle de la N-glycosylation du domaine S1. Pour cela, des mutants individuels des sites potentiels de N-glycosylation du domaine S1 ont été construits par mutagenèse dirigée en remplaçant le résidu asparagine des séquons par une glutamine. Les mutants ont été nommés en fonction du site muté (par exemple, le mutant N2 correspond au deuxième site potentiel de N-glycosylation muté) (Figure 19A). Seul le premier site potentiel de N-glycosylation n'a pas été muté, ce site n'est pas présent dans la souche VR-740 et de plus, par sa position très proche du site de clivage du peptide signal et du deuxième site de N-glycosylation, il est peu probable que ce site soit N-glycosylé.

Nous avons d'abord regardé l'effet de ces mutations de la protéine S sur l'entrée virale à l'aide de particules pseudotypées avec la protéine S ou des mutants (Figure 19B). Pour cela, des pseudoparticules ont été produites pour chaque mutant et utilisées pour infecter des cellules Huh-7. En parallèle, la production de particules a été analysée en western-blot (Figure 19C).



**Figure 18 : Analyse des pseudoparticules incorporant la protéine S du HCoV-229E ou un des mutants.** A. Schéma fonctionnel de la protéine S. Les sites potentiels de N-glycosylation sont représentés par un trait noir court, leur position en acide aminé en noir et le nom du mutant correspondant en rouge. Le site de clivage S1/S2 est représenté par un grand trait noir. Gris clair: domaine S1, or: domaine S2, violet: domaine de liaison au récepteur, rouge: peptide de fusion, bleu clair: HR1, bleu foncé: HR2, domaine transmembranaire: gris foncé. B. Infectiosité des pseudoparticules incorporant la protéine S ou un des mutants. Des 229Epp ont été produites et utilisées pour infecter des cellules Huh-7. Les résultats sont exprimés en pourcentage de l'infection des particules pseudotypées avec la protéine sauvage. L'expérience a été répétée au moins 3 fois. C. Analyse de la production de pseudoparticules. Les particules ont été concentrées sur coussin de sucrose puis lysées et déposées sur gel SDS-PAGE. En haut: western blot réalisé avec un Ac anti-VSVG reconnaissant la protéine S. En bas: western-blot réalisé avec un Ac anti-gag reconnaissant la protéine de capsid

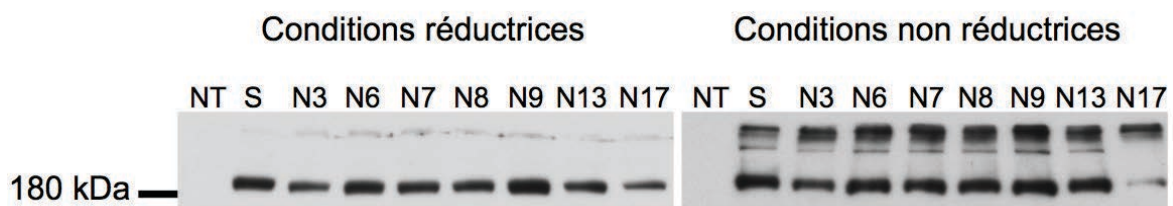
Le niveau de production des particules pseudotypées avec les mutants N2 à N18 est similaire à celui des particules pseudotypées avec la protéine S (Figure 19C). Les profils obtenus avec l'Ac anti-VSVG montrent que les quantités de protéines mutantes incorporées dans les 229Epp étaient comparables à la quantité de protéine S sauvage incorporée, sauf pour le mutant N17 (Figure 19C) qui est moins bien incorporé.

L'analyse de l'infectiosité des 229Epp incorporant les mutants comparés à l'infectiosité de celles incorporant la protéine S sauvage révèle 4 phénotypes : dans le groupe 1 l'infectiosité des mutants n'est pas affectée (N2, N4, N5, N10, N11, N14, N16 et N18). Le groupe 2 contient les mutants dont l'infectiosité est atténuée de 20 à 50% (N3, N7, N8, N13 et N17). Le groupe 3 contient deux mutants (N6 et N9) dont l'infectiosité est très fortement diminuée (respectivement 81,5% et 73,5%). Le groupe 4 contient les mutants dont l'infectiosité est augmentée (N12 et N15). Notre analyse montre que seuls les mutants N6, N7, N9, N12 et N15 sont statistiquement différents de la protéine S. Ces résultats suggèrent que les N-glycanes en position N6 et N9 jouent un rôle majeur dans l'entrée virale.

Nous avons ensuite étudié quelles fonctions de la protéine S étaient dépendantes de la présence de ces N-glycanes. En ce qui concerne le mutant N17, la diminution du niveau d'infectiosité des 229Epp peut être expliquée par une diminution d'incorporation de la protéine S dans les pseudoparticules (Figure 19C) qui pourrait être due à un défaut de repliement de la protéine mutante. Les N-glycanes peuvent également être impliqués dans plusieurs étapes de l'entrée virale. Nous avons donc testé les interactions des différents mutants avec le récepteur APN ou leur rôle dans la fusion de l'enveloppe virale.

### 4.1.3 Etude du rôle de la N-glycosylation lors du repliement des protéines mutantes

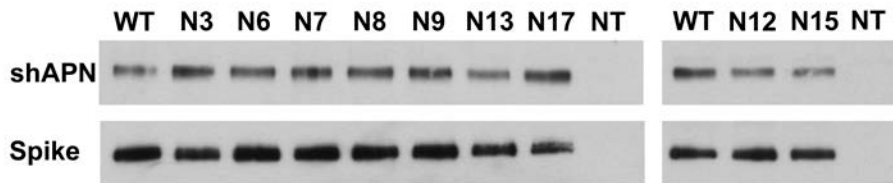
Afin de déterminer si les différents mutants dont l'infectiosité est diminuée étaient correctement repliés, ils ont été exprimés de façon transitoire en cellules HEK293T, puis les cellules ont été lysées et les échantillons ont été analysés en western-blot avec l'anticorps anti-VSVG en conditions réductrices et non réductrices afin de déterminer si ceux-ci étaient correctement repliés. La comparaison du profil de migration des mutants N3, N6, N7, N8, N9, N13 et N17 en conditions réductrices avec celui de la protéine S sauvage montre que les mutants sont détectés au même poids moléculaire (> 180 kDa) que la protéine sauvage (Figure 20). En conditions non dénaturantes, le profil de migration de la protéine S montre deux bandes majoritaires, la première vers 180 kDa est identique à celle observée en conditions réductrices et correspond aux monomères de protéine S. La deuxième bande a un très haut poids moléculaires puisque ces protéines ont peu migré dans le gel, ces formes correspondent probablement à des protéines mal repliées qui constituent des agrégats. Le mutant N17 est majoritairement retenu sous la forme d'agrégats de haut poids moléculaire, en effet la proportion de protéine sous forme de monomères est beaucoup plus faible. Ces résultats suggèrent que le mutant N17 est mal replié et que le N-glycane présent à cette position est impliqué dans le repliement de la protéine S.



**Figure 19 : Rôle des N-glycanes dans le repliement de la protéine S.** Les plasmides codant la protéine S sauvage, ou les mutants de N-glycosylation diminuant l'infectiosité des 229Epp ont été transfectés dans des cellules HEK293T. Après 24h, les cellules ont été lysées en présence d'iodoacétamide puis les protéines ont été analysées par western-blot en conditions réductrices et en conditions non réductrices à l'aide d'un Ac anti-VSVG.

#### **4.1.4 Etude du rôle de la N-glycosylation lors de l'interaction de la protéine S avec le récepteur humain APN**

Le domaine S1 de la protéine S du coronavirus HCoV-229E contient le domaine de liaison du récepteur du virus : l'aminopeptidase N. Plusieurs études en utilisant des techniques différentes ont décrit des domaines de liaison différents (Bonavia et al., 2003; Breslin et al., 2003; Hofmann et al., 2006; Wong et al., 2017). En effet une première étude avait identifié les résidus 407-547 comme le RBD puis une deuxième les résidus 278 à 329. Ces deux études utilisaient différentes formes tronquées de protéine S solubles et analysaient leur liaison soit à une forme soluble du récepteur soit à des cellules exprimant ou non le récepteur. Plus récemment la structure du RBD lié au récepteur a été résolue et trois boucles se liant à APN ont été identifiées comprenant les résidus 308-325, 352-359 et 404-408. Il est possible que les autres régions identifiées puissent aussi moduler l'interaction avec le récepteur. Si on considère l'ensemble des résidus qui ont été décrits dans ces études (278-549), cette région contient 8 sites potentiels de N-glycosylation de N11 à N18. Afin de tester le rôle des N-glycanes dans l'interaction avec le récepteur hAPN (aminopeptidase N humain), des expériences de co-précipitation ont été réalisées. Pour cela des lysats de cellules exprimant la protéine S ou les mutants de N-glycosylation diminuant (N6, N7 et N9) ou augmentant l'infectiosité des 229Epp (N12 et N15) ont été produits et incubés en présence de récepteur soluble APN (shAPN). Les complexes protéine S/récepteur ont été immunoprécipités à l'aide d'un anticorps anti-VSVG (Figure 20). Le récepteur co-immunoprécipité avec la protéine S a ensuite été révélé en western-blot. Le niveau de transfection des différents mutants de N-glycosylation est similaire à celui de la protéine S et le récepteur APN est précipité par tous les mutants de la protéine S. Aucune différence de précipitation du récepteur n'est observée entre les mutants de N-glycosylation et la protéine S, suggérant que les N-glycanes présents sur ces positions ne sont pas impliqués dans l'interaction entre la protéine S et APN.



**Figure 20 : Etude de l'interaction de la protéine S avec le récepteur soluble shAPN.** Les plasmides codant la protéine S sauvage ou les mutants de N-glycosylation diminuant (N3, N6, N7, N8, N9, N13 et N17) ou augmentant significativement l'infectiosité des 229Epp (N12 et N15) ont été transfectés en cellules HEK293T. 24h après la transfection, les cellules ont été lysées et un test de co-précipitation avec la forme soluble du récepteur humain APN (shAPN) a été réalisé. La présence de shAPN dans les précipitats a été révélée en western-blot avec l'Ac anti-V5. L'expression de la protéine S et des mutants a été détectée dans les lysats cellulaires.

#### 4.1.5 Etude du rôle de la N-glycosylation pour les fonctions de fusion de la protéine S

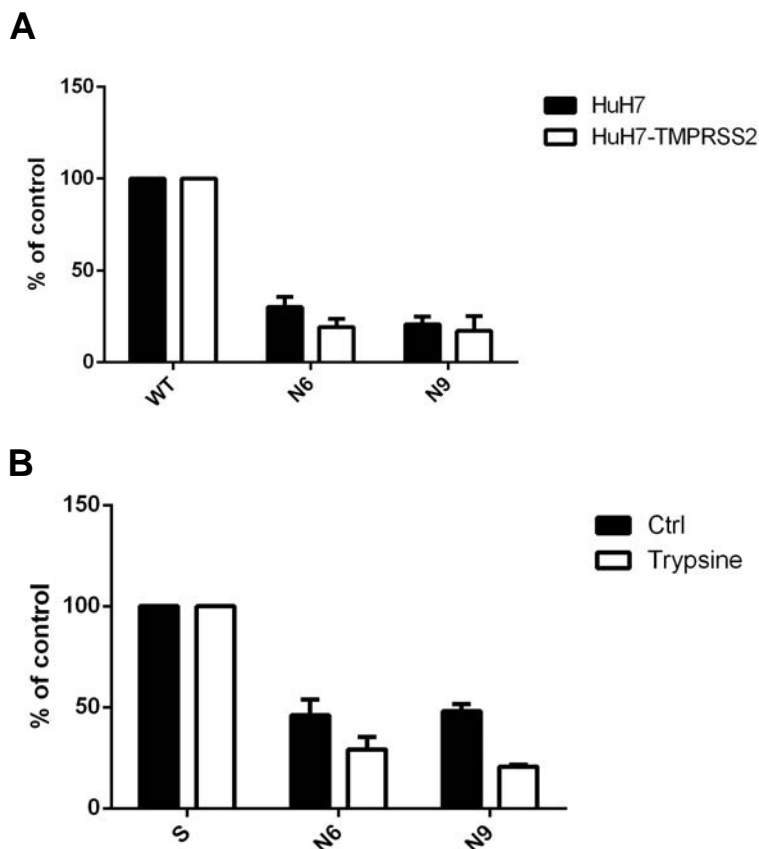
##### 4.1.5.1 Infection médiée par la trypsine et TMPRSS2

Comme décrit précédemment, les coronavirus respiratoires utilisent de préférence une voie d'entrée dépendante des TTSP *in vivo*. Nous avons donc testé si la mutation des sites potentiels de N-glycosylation affectait aussi la voie d'entrée dépendante de TMPRSS2. Pour cela des cellules Huh-7 ont été transduites avec un vecteur lentiviral permettant l'expression de la protéase puis infectées avec des particules pseudotypées avec la protéine sauvage ou les deux mutants. Quelque soit la voie utilisée, endosomale ou à la surface cellulaire, l'entrée des deux mutants est diminuée de façon importante (Figure 21A). Ces résultats suggèrent que la fusion du virus puisse être affectée. En effet il a déjà été observé pour le virus IBV que des mutations de site de N-glycosylation dans le domaine S1 inhibaient le clivage de la protéine spike et donc la fusion de l'enveloppe virale avec la membrane cellulaire (Zheng et al., 2018). Nous avons donc réalisé un test de fusion dans lequel les pseudoparticules sont liées à la surface cellulaire à 4°C et la fusion est induite par un



court traitement par la trypsine (Figure 21B). Le traitement par la trypsine ne permet pas de rétablir une fusion similaire à celle observée avec la protéine sauvage.

Nos résultats suggèrent ces N-glycanes (N6 et N9) sont importants pour l'entrée virale, quelque soit la voie utilisée, endosomale ou médiée par TMPRSS2. Il est probable que N6 et N9 affectent la fusion virale.

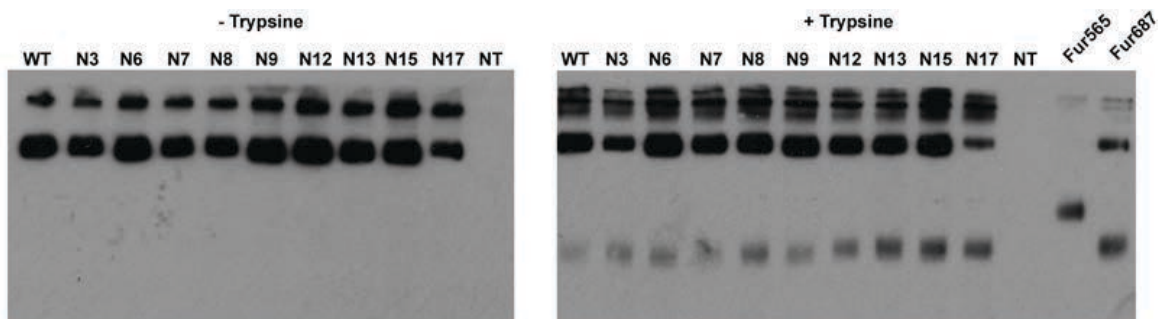


**Figure 21 : Infection en surface des 229Epp médiée TMPRSS2 et test de fusion à la surface cellulaire.** (A) Des cellules contrôles ou transduites avec un vecteur lentiral permettant l'expression de TMPRSS2 ont été infectées avec les particules pseudotypées avec la protéine sauvage ou les mutants N6 et N9. (B) Les particules pseudotypées avec les différentes protéines mutées ont été liées à la surface cellulaire à 4°C en présence de 25 mM de  $\text{NH}_4\text{Cl}$  et ont été traitées pendant 5 min par 2  $\mu\text{g/ml}$  de trypsine pour induire la fusion. 48h après, les cellules ont été lysées et l'activité luciférase a été mesurée. Les résultats sont présentés en pourcentage d'infection observée pour les particules pseudotypées avec la protéine sauvage et correspondent à la moyenne de deux expériences indépendantes. Les barres d'erreur représentent l'erreur standard des moyennes (SEM).

#### 4.1.5.2 Clivage par la trypsine

Nous avons ensuite étudié le clivage de la protéine S par la trypsine. Pour cela, la protéine sauvage et les différents mutants de N-glycosylation exprimés à la surface des cellules ont été traités ou non par de la trypsine puis biotinylés comme

précédemment décrit. Les protéines ont été analysées en western blot à l'aide d'un anticorps anti-VSVG (Figure 22). Le niveau d'expression à la surface des cellules des différents mutants est similaire à celui de la protéine S (Figure 23, gel de gauche). Le clivage de la protéine S sauvage par la trypsine montre une bande de faible intensité de 85 kDa, une bande plus forte de plus de 180 kDa et des agrégats de haut poids moléculaire. Comme montré précédemment (partie 3), la protéine S est clivée dans la région S2'. Les mutants de N-glycosylation montrent un clivage similaire à celui de la protéine S.



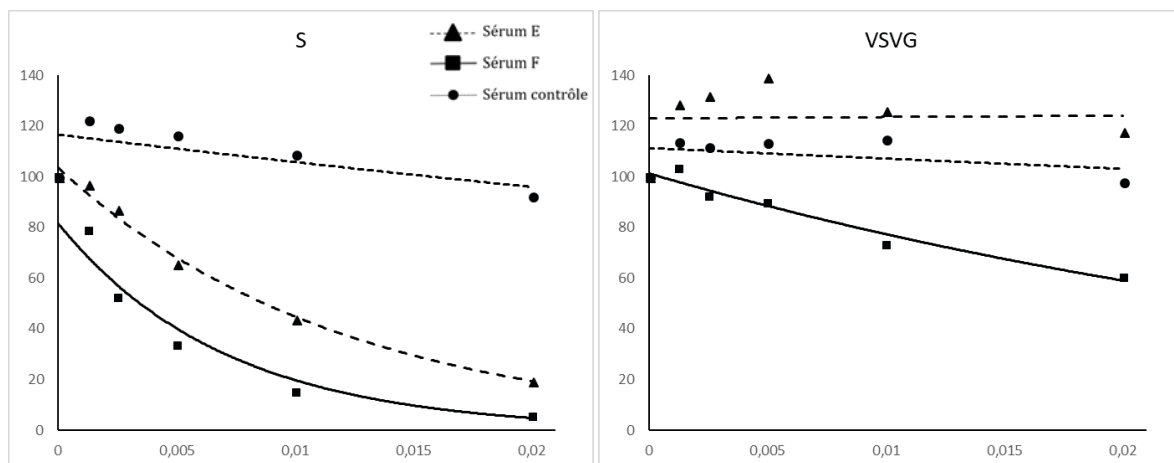
**Figure 22 : Clivage de la protéine S du HCoV-229E.** La protéine S et ses mutants de N-glycosylation ont été produits en cellules HEK293T. Les cellules ont été traitées ou non par 2  $\mu$ g de trypsine puis les protéines ont été biotinylées puis lysées. Après avoir incubé les lysats une nuit en présence de billes de streptavidine afin de ne récupérer que les protéines biotinylées, les billes ont été lavées puis la protéine S a été détectée à l'aide d'un anticorps anti-VSVG après western-blot.

#### 4.1.6 Neutralisation

Nous avons testé le rôle des N-glycanes dans la reconnaissance du domaine S1 par des anticorps neutralisants. Deux sérums (E et F) collectés chez des patients hospitalisés et infectés par HCoV-229E et un sérum (Ctrl) prélevé chez un patient hospitalisé ne présentant pas d'infection HCoV-229E au moment du prélèvement ont été testés pour leur capacité de neutralisation des pseudoparticules 229Epp. Des pseudoparticules incorporant la protéine G du virus de la stomatite vésiculaire ont été utilisées pour vérifier la spécificité des sérums (Figure 23). Les pseudoparticules 229Epp ou VSVpp ont été incubées en présence de quantités décroissantes de

sérums E, F et Ctrl. Les deux sérums montrent une inhibition dose-dépendante de l'infection par les pseudoparticules HCoV-229E mais n'ont aucun effet sur l'infection par les pseudoparticules VSV. De même le sérum contrôle n'a aucun effet sur les infections 229Epp et VSVpp. La spécificité de la neutralisation des 229Epp par les anticorps sériques a été confirmée. Les anticorps présents dans les sérums E et F neutralisent efficacement l'infection par rapport au contrôle. Il est à noter que le sérum F neutralise beaucoup plus efficacement les pseudoparticules que le sérum E. En effet, pour atteindre 50% d'inhibition il faut diluer le sérum F au 1/295 tandis que le sérum E doit être dilué seulement au 1/114.

L'effet de la mutation des sites potentiels de N-glycosylation du domaine S1 sur la sensibilité des 229Epp à la neutralisation par des anticorps sériques a ensuite été évalué (Figure 24). Chaque mutant incorporé dans des 229Epp a été testé avec un sérum contrôle, avec le sérum E ou le sérum F. Pour tous les mutants, l'absence de neutralisation par le sérum contrôle a été validée. Les résultats sont présentés en pourcentage d'infection observée pour les particules pseudotypées avec la protéine sauvage ou un mutant pour une dilution de sérum donnée. Une courbe de décroissance de la moyenne de trois expériences indépendantes a été obtenue pour

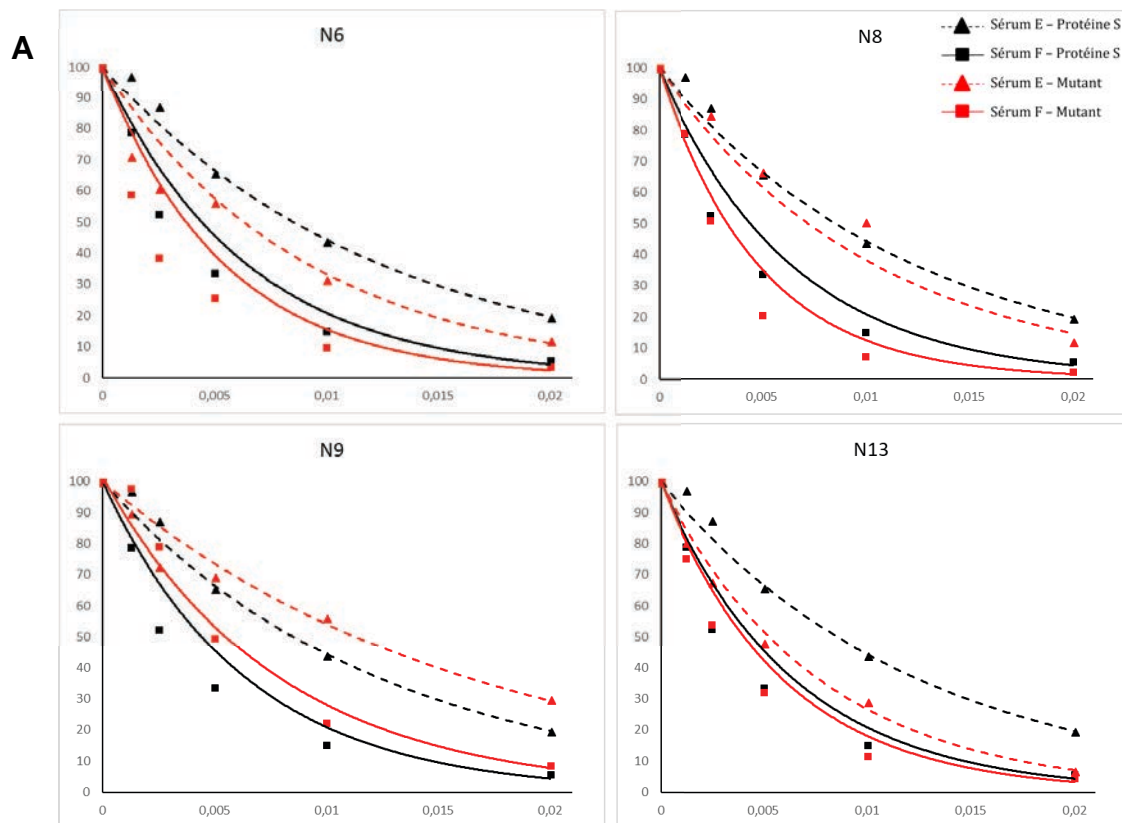


**Figure 23 : Spécificité des sérum E et F.** Un test de neutralisation a été réalisé en incubant des pseudoparticules incorporant la protéine S du HCoV-229E ou la protéine G du VSV avec des quantités décroissantes de sérum pendant 1h. Des cellules Huh7 ont ensuite été infectées avec les mélanges pendant 3h puis rincées. La quantité de luciférase a été mesurée après 48h. Cette expérience a été répétée 2 fois.

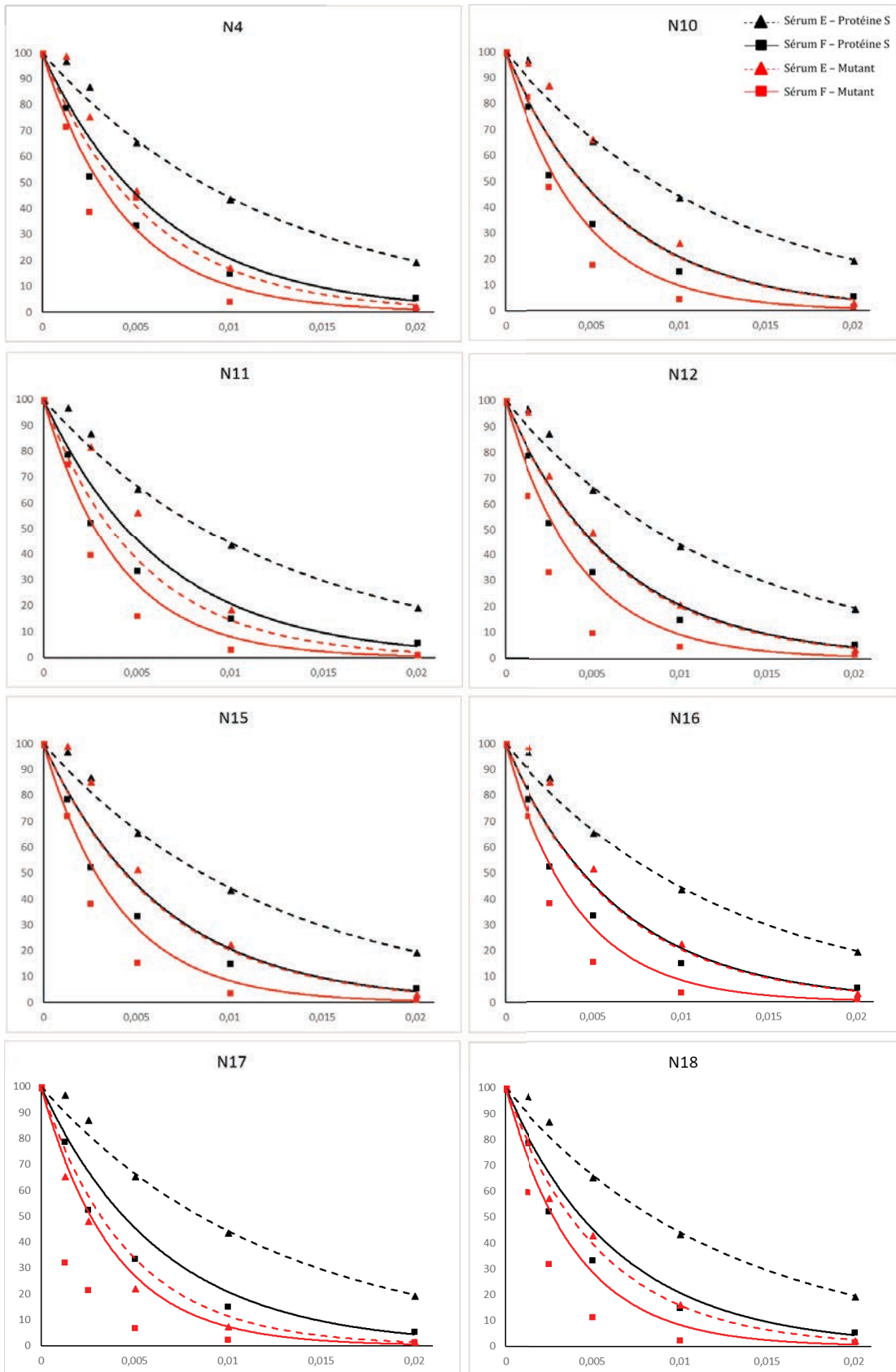
chaque couple protéine S/sérum. La pente a été estimée pour chaque courbe puis une analyse statistique a été réalisée pour déterminer quel mutant présentait une différence significative de neutralisation par rapport à la protéine sauvage (voir résultats en annexe). Il semble que la mutation des sites de N-glycosylation N6, N8, N9 et N13 n'affecte pas la neutralisation, que ce soit par le sérum E ou le sérum F.

Pour les pseudoparticules incorporant les mutants N4, N10, N11, N12, N15, N16, N17 et N18 l'infection est neutralisée plus efficacement que pour les 299Epp incorporant la protéine sauvage quelque soit le sérum utilisé.

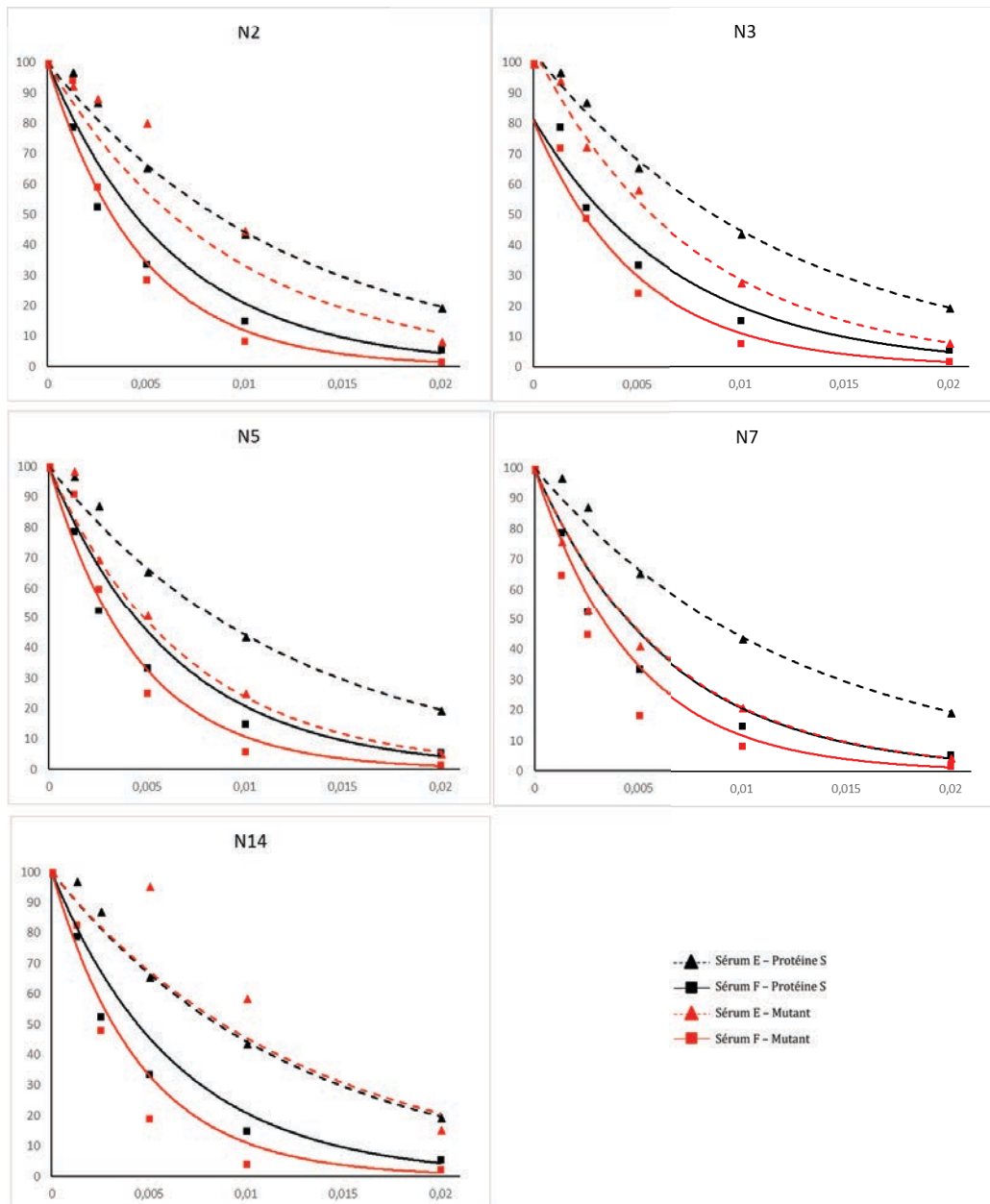
Pour les cinq mutants restants N2, N3, N5, N7 et N14, on observe une neutralisation plus efficace avec le sérum F cependant aucune différence n'est observée avec le sérum E. Cette différence peut être en partie due au fait que le sérum F est plus neutralisant que le sérum E. Ces résultats suggèrent que les N-glycanes en position 2, 3, 5, 7 et 14 joueraient un rôle moins important pour protéger la protéine des anticorps neutralisants. Il est intéressant de constater que les N-glycanes situés dans la région C-terminale de S1 modulent tous fortement la neutralisation par les anticorps, ce qui avait déjà été montré précédemment (Shirato et al., 2012).



**B**



C



**Figure 24 : Neutralisation de la protéine S du HCoV-229E par les sérums E et F.** Un test de neutralisation a été réalisé en incubant des pseudoparticules incorporant la protéine S du HCoV-229E avec des quantités décroissantes de sérum pendant 1h. Des cellules Huh-7 ont ensuite été infectées avec les mélanges pendant 3h puis rincées. La quantité de luciférase a été mesurée après 48h. Les pseudoparticules incorporant les mutants pour lesquels la neutralisation n'est pas modifiée sont représentés en A. Ceux pour lesquels l'infection est mieux neutralisée que pour la protéine sauvage quelque soit le sérum sont représentés en B et ceux pour lesquels l'infection est mieux neutralisée uniquement pour le sérum F sont représentés en C. L'axe des abscisses correspond à la dilution de sérum et l'axe des ordonnées correspond au pourcentage d'infection pour une dilution de sérum donnée par rapport au contrôle sans sérum.

## 4.2 DISCUSSION

De nombreuses protéines d'enveloppe virale utilisent la machinerie cellulaire afin d'être N-glycosylées, ce qui peut avoir un impact sur l'entrée virale ou la reconnaissance par les anticorps neutralisants. Afin de pouvoir étudier la neutralisation du virus par des sérums de patients, nous avons cloné la séquence de la protéine d'un virus de patient. En effet depuis l'isolement de la souche VR-740 dans les années 1960, le virus a évolué et il a été montré que la souche VR-740 n'est pas neutralisée par des anticorps dirigés contre des protéines S de souches cliniques. Inversement, les anticorps dirigés contre les souches cliniques sont inefficaces contre VR-740 (Shirato et al., 2012).

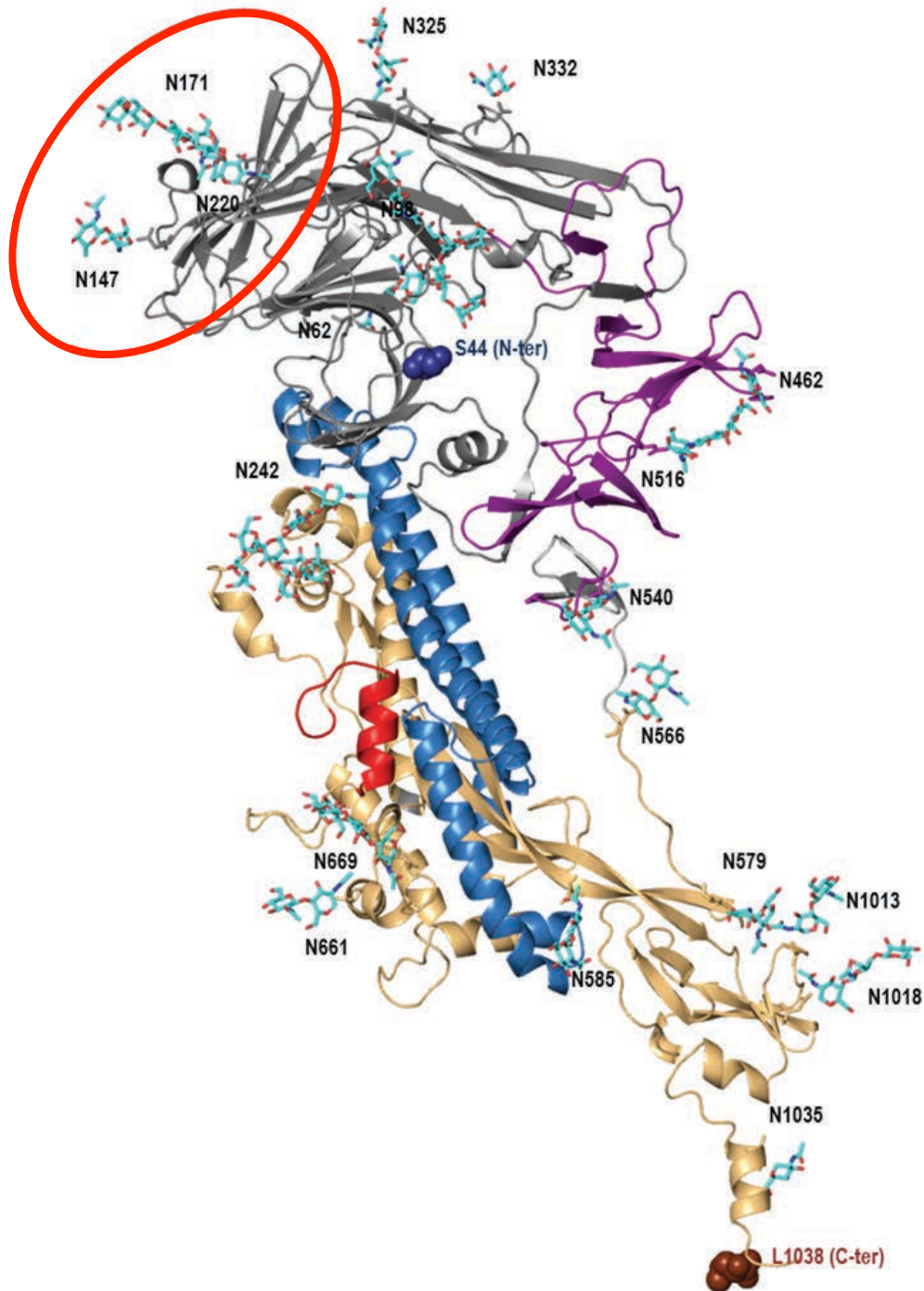
La protéine S du HCoV-229E (France/Lille/2014) a été isolée et 33 sites potentiels de N-glycosylation ont été identifiés. Nous avons initié des analyses afin de déterminer quels sont les sites qui sont effectivement N-glycosylés en utilisant la souche de laboratoire VR-740. Cette souche est très proche de notre souche clinique, avec une homologie de 94%. Ces analyses sont encore en cours mais les premières analyses des spectres obtenus suggèrent que parmi les 18 sites potentiels de N-glycosylation dans le domaine S1, 12 sont glycosylés. Les sites N1, N5 et N15 sont absents de la souche VR-740 et aucune information n'est pour l'instant disponible pour les 3 derniers sites N2, N6 et N9. Il est intéressant de noter que les sites potentiels de N-glycosylation N1 et N5 de la souche clinique ne sont présents ni sur la souche de laboratoire du HCoV-229E ni sur la protéine S du HCoV-NL63, ce qui suggère que ces sites ne sont pas indispensables et donc ne sont probablement pas N-glycosylés sur la souche France/Lille/2014. D'ailleurs nous n'avons observé aucun effet de la mutation N5 que ce soit pour les fonctions d'entrée de la protéine S ou dans les tests de neutralisation de l'infection. Les sites de N-glycosylation N6 et N9 sont strictement conservés sur la protéine S du HCoV-NL63, et sont effectivement N-glycosylés (Walls et al., 2016a). Il est donc probable qu'ils soient également glycosylés sur la souche France/Lille/2014 du HCoV-229E, d'autant plus que la mutation de ces deux sites a un effet fonctionnel très important. Nous n'avons aucune information pour le site N2 car nous n'avons obtenu aucun spectre

peptidique correspondant à cette région. Ce site est situé à l'extrémité N-terminale de la protéine (Asn 23), le site correspondant dans la protéine de HCoV-NL63 (Asn 203) n'est pas glycosylé. Il est possible que ce site soit N-glycosylé dans HCoV-229E et pas dans HCoV-NL63 car la localisation dans la protéine du site est probablement différente. Dans la protéine S de HCoV-229E, ce site est à l'extrémité N-terminale tandis que la protéine S de HCoV-NL63 possède en amont le domaine 0. Les sites effectivement N-glycosylés de la protéine S de PDCoV ont aussi été déterminés (Xiong et al., 2017), et les auteurs ont montré que 26 des 27 sites étaient réellement N-glycosylés. La différence entre le nombre de N-glycanes sur la protéine S de NL63 et du PDCoV peut s'expliquer par la présence de 6 N-glycanes sur le domaine 0 uniquement présent sur la protéine S de NL63. Il semble que le nombre de N-glycanes soit donc similaire pour les *alphacoronavirus* NL63 ou HCoV-229E ou le *deltacoronavirus* PDCoV (Walls et al., 2016a; Xiong et al., 2017).

Le rôle de la N-glycosylation lors de l'entrée virale a été étudié à l'aide de particules pseudotypées avec la protéine S ou un mutant de N-glycosylation, dans lequel l'asparagine a été muté en glutamine. Deux sites de N-glycosylation jouent un rôle majeur dans l'entrée du virus qui est diminuée de plus de 70% quand ils sont mutés, ce sont les N-glycanes en position 6 et 9. Le N-glycane en position 7 module lui aussi l'infectiosité puisque sa mutation entraîne une diminution de l'infection de 45%. Les mutants N6, N7 et N9, correspondent respectivement aux acides aminés N147, N171 et N220. Leur position sur un trimère de protéine S est importante. En effet d'après la modélisation de la protéine S du virus HCoV-229E réalisée par le Dr Petrescu, ces sites sont assez regroupés et semblent être accessibles sur le dessus du trimère de protéines S (Figure 26). Malheureusement pour les sites N6 et N9 nous n'avons pas confirmé leur N-glycosylation, mais nous pensons que ces sites sont très certainement N-glycosylés puisqu'on observe que la plupart des sites potentiels de N-glycosylation des protéines S de HCoV-NL63 et PDCoV sont effectivement N-glycosylés. D'autant plus que les deux sites correspondant à ces sites dans la protéine S de HCoV-NL63 sont effectivement glycosylés. Nos expériences d'induction de la fusion à la surface cellulaire, suggèrent que les sites N6 et N9 sont importants pour la fusion de l'enveloppe virale avec la membrane



cellulaire. De part leur localisation relativement éloignée de la machinerie de fusion (domaine S2), ces résultats sont surprenants. Cependant, une étude récente sur le rôle des N-glycanes du domaine S1 lors de la fusion virale de IBV, souche Beaudette, a permis de distinguer deux N-glycanes, N212 et N276, important pour le clivage protéolytique de la protéine S et pour la fusion virale. Ces deux N-glycanes sont situés respectivement dans le RBD de la protéine S, et très proche du RBD (Zheng et al., 2018). En effet, le RBD d'IBV est situé à l'extrémité N-terminale de S1 tandis que le RBD du HCoV-229E est situé à l'extrémité C-terminale de S1. Ce résultat montre que des N-glycanes éloignés du site de clivage de la protéine S peuvent moduler le clivage et la fusion virale. En se basant sur la structure de la protéine S de PDCoV qu'ils ont déterminé, Xiong *et al.* ont émis l'hypothèse que deux sites de glycosylations (N652 et N661) présents dans le domaine S2 pourraient limiter l'accès des protéases à la région S2' de la protéine et jouer un rôle dans l'activation de la fusion. Ces sites de glycosylations correspondraient aux sites potentiels de N-glycosylation N22 et N23 de la protéine S du HCoV-229E. Il a été montré pour le virus de la grippe que la présence d'un site de N-glycosylation à proximité du site de clivage de la glycoprotéine HA0 modulait l'accès des protéases à la protéine virale (Deshpande et al., 1987; Kawaoka et al., 1984; Yin et al., 2017). Cependant contrairement à ce qui a été observé avec IBV, nos tests de clivage des mutants N6 et N9 par la trypsine ne montrent pas un défaut de clivage de ces mutants, ce qui suggère que c'est une autre étape du processus de fusion qui est affectée. Il est possible la fusion virale soit modifiée par un changement de conformation de la protéine S à cause de l'absence de certains N-glycanes. Cependant nous ne pouvons exclure que ces les deux résidus N147 et N220 soient nécessaires à la fusion virale indépendamment de leur N-glycosylation.



**Figure 25 : Modélisation de la protéine S du HCoV-229E.** Modélisation d'un monomère de protéine S. Gris : domaine S1, or : domaine S2, violet : domaine de liaison au récepteur, rouge : peptide de fusion, bleu : HR1. En rouge sont entourés les sites de glycosylation N6 (N147), N7 (N171) et N9 (N220).

Il est connu que la N-glycosylation peut avoir un rôle majeur lors du repliement des protéines virales (Daniels et al., 2003; Goffard et al., 2005). Nous avons donc étudié cet aspect. Nos résultats montrent que les particules pseudotypées avec le mutant N17 semblent être moins infectieuses qu'avec la protéine sauvage, bien que ce ne soit pas significatif. Or ce mutant est mal incorporé dans les pseudoparticules. L'étude de ce mutant en conditions dénaturantes ou non montre que la mutation de ce site de N-glycosylation entraîne un défaut de repliement de la protéine S. Ainsi mutés individuellement, seule la mutation N17 a un effet sur le repliement de la protéine, mais il est probable que d'autres N-glycanes jouent un rôle important pour le repliement de la protéine. En effet, nous avons observé que la mutation simultanée de plusieurs sites potentiels de N-glycosylation (par groupe de 3 ou 4) avait un effet majeur sur le repliement de la protéine (non montré).

La mutation de deux sites potentiels de N-glycosylation N12 et N15 correspondant aux acides aminés N332 et N486 augmente l'infectiosité de la protéine S. Le site de N-glycosylation N15, situé dans le RBD semble également être situé vers l'extérieur des trimères, sur le côté. Il a été montré pour le VIH que les N-glycanes présents sur la protéine d'enveloppe gp120 avaient un rôle majeur lors de l'entrée virale : en effet, la mutation de certains d'entre eux entraîne une perte significative d'infectiosité du virus (Mathys and Balzarini, 2015; Wang et al., 2013b) mais l'augmentation de l'infectiosité provoquée par la mutation d'un site de N-glycosylation d'une glycoprotéine a déjà été décrite pour le HIV (Wang et al., 2013b). Ce phénomène a également été décrit pour la protéine E2 du HCV, mais dans le cas, l'augmentation de l'infectiosité n'était pas significative (Helle et al., 2010). Il est possible que la suppression des N-glycanes N12 et N15 modifie la conformation de la protéine S et permette une liaison plus aisée entre les boucles du RBD du HCoV-229E et son récepteur aminopeptidase, comme il avait été suggéré pour HIV et HCV (Helle et al., 2010; Wang et al., 2013b). En effet, les N-glycanes peuvent être également important pour l'interaction avec les récepteurs ou co-récepteurs. C'est le cas notamment pour des virus qui utilisent des lectines comme récepteur, c'est ce qui a été montré pour la protéine d'enveloppe E de DENV. En effet, le récepteur DC-SIGN de ce virus interagit directement avec le N-glycane porté par l'asparagine 67 de la protéine E. Il a été montré pour le VIH que la mutation sites de N-glycosylation présents sur la protéine d'enveloppe gp120 diminuait l'infection de cellules

CD4/CCR5/CXCR4 (Mathys and Balzarini, 2015). En tenant compte des études de localisation du RBD de HCoV-229E (Bonavia et al., 2003; Breslin et al., 2003; Hofmann et al., 2006; Wong et al., 2017), 8 N-glycanes (de N11 correspondant à N325 à N18 correspondant à N540) pourraient potentiellement moduler la liaison au récepteur hAPN. Dans le cas du HCoV-229E, nos résultats montrent qu'aucun N-glycane présent sur la protéine S ne semble jouer un rôle important dans l'interaction avec son récepteur aminopeptidase N. Cependant la technique que nous avons utilisée n'est probablement pas assez sensible pour détecter un effet modulateur de nos mutations sur l'interaction de la protéine avec le récepteur. Une perte partielle d'interaction est probablement difficile à visualiser avec notre technique.

L'importance des N-glycanes présents à la surface des virus lors de la neutralisation par les anticorps a été montrée pour plusieurs virus, dont HCV et VIH. Nos résultats montrent que la plupart des N-glycanes présents dans le domaine S1 du HCoV-229E masquent des épitopes reconnus par les anticorps de patients. En particulier, la présence des N-glycanes N4, N10, N11, N12, N15, N16, N17 et N18 diminuent fortement la neutralisation par les anticorps. Ces résultats suggèrent que la forte densité de N-glycanes présents sur la protéine S sert de bouclier, comme il a été décrit pour la protéine Env du VIH, et permet un échappement immunitaire du virus (Behrens and Crispin, 2017; Wang et al., 2013b). L'abolition de la N-glycosylation pourrait alors créer des failles dans ce bouclier, ou modifier légèrement la conformation de la protéine S ou des trimères, rendant accessible les épitopes présents sous les N-glycanes. Nos résultats suggèrent également que les anticorps présents dans le sérum E sont moins diversifiés que ceux présents dans le sérum F. En effet, la mutation de 5 sites potentiels de N-glycosylation a permis la neutralisation de la protéine S par les anticorps présents dans le sérum F mais pas dans le sérum E. Il est probable que la région C-terminale de S1 masquée par le bouclier de N-glycane pour le HCoV-229E soit similaire à d'autres coronavirus, comme HCoV-NL63 et PDCoV. Il avait été suggéré que la pression immunitaire est similaire chez les hôtes respectifs des *alpha*- et *deltacoronavirus* et que les régions cachées par ce bouclier de N-glycanes soient importantes pour les fonctions de la

protéine S (Walls et al., 2016a; Xiong et al., 2017), ce que nos résultats semblent confirmer.



## 5 PARTIE III : DISCUSSION GENERALE ET PERSPECTIVES

---

La glycoprotéine d'enveloppe S des coronavirus joue un rôle majeur lors de l'entrée virale. Elle se lie au récepteur cellulaire et dirige la fusion entre les membranes virale et cellulaire. Son exposition à la surface des virus fait aussi de cette glycoprotéine une cible majeure pour les anticorps neutralisants.

La protéine S provenant d'un patient infecté par HCoV-229E a été isolée et 33 sites potentiels de N-glycosylation ont été identifiés. Nous avons déterminé qu'au moins 26 sites de N-glycosylation étaient effectivement N-glycosylés. Sur les 39 sites potentiels de N-glycosylation présents sur la protéine S de NL63, 34 sont effectivement glycosylés et sur les 27 sites potentiels de la protéine S du PDCoV, 26 sont effectivement glycosylés (Walls et al., 2016a; Xiong et al., 2017). Peu de glycoprotéines présentes à la surface des virus ont autant de sites potentiels de N-glycosylation que la protéine S des coronavirus. La conservation d'un nombre important de glycanes sur les différentes protéines S de coronavirus montre qu'il existe un avantage évolutif des N-glycanes, que ce soit lors du cycle viral ou lors des interactions hôte-pathogène. Le rôle de la N-glycosylation du domaine S1 lors de l'entrée virale et de la neutralisation par les anticorps a été évalué en mutant des sites de N-glycosylation présent sur la protéine S à la surface de pseudoparticules.

L'un des mutants qui diminuait l'infectivité des 229Epp de manière non significative, N17, semble être important pour le repliement de la protéine. Nous avons aussi vérifié si les autres sites de N-glycosylation présents dans le domaine S1 modulaient le repliement de la protéine S, mais aucun autre site que N17 ne semble intervenir lorsqu'il est muté individuellement. Il est un peu surprenant que seul un site de N-glycosylation soit important pour le repliement de la protéine, compte tenu du nombre total de sites identifiés. En effet, sur les 11 sites potentiels de la glycoprotéine E2 d'HCV, 2 N-glycanes sont importants pour le repliement de la protéine (Goffard et al., 2005). Des résultats préliminaires concernant les sites de N-

glycosylations du domaine S2 montrent que ces glycanes ne semblent pas non plus être impliqués dans le repliement de la protéine, que ce soit lorsque la protéine S possède une unique mutation ou lorsque deux sites potentiels de N-glycosylation sont mutés (résultats non montrés). Il est possible que lors du repliement de la protéine S par les protéines chaperonnes, la calnexine et la calréticuline puissent se lier à plusieurs glycanes, ce qui expliquerait que la mutation d'un seul glycanes n'entraîne généralement pas de défaut de repliement. Ce mécanisme expliquerait également pourquoi la délétion de 3 à 4 sites de N-glycosylation dans la même protéine entraîne un défaut de repliement de la protéine S. Il a été montré que la calnexine était importante pour que la protéine S du SARS-CoV soit correctement repliée et puisse être incorporée dans les pseudoparticules lors de leur production (Fukushi et al., 2012). Il serait intéressant de tester les interactions entre les différents mutants de N-glycosylation et la calnexine et/ou calréticuline par co-immunoprécipitation, afin de déterminer les N-glycanes réellement importants pour le repliement de la protéine S.

La mutation des sites potentiels de N-glycosylation N6, N7 et N9 diminue l'entrée virale de manière significative. L'étude de la fusion virale nous a permis de montrer que ces mutants ne modifiaient pas le clivage protéolytique de la protéine S, mais qu'ils devraient avoir un rôle lors de la fusion. Il est possible que l'absence de ces N-glycanes provoque un changement de conformation de la protéine S. Des résultats préliminaires sur les sites de N-glycosylation du domaine S2 indiquent que d'autres glycanes pourraient avoir un rôle lors de la fusion virale, et en particulier des glycanes présents avant (N26, N27, N28) et dans l'*heptad repeat 2* (N33) (résultats non montrés). Ces résultats sont moins étonnants que l'implication de N6 et N9 dans la fusion virale, qui sont éloignés de la machinerie permettant la fusion. Cependant, N6 (N147) est à deux acides aminés d'une cystéine C145 qui pourrait être impliqué dans un pont disulfure et N9 (N220) est proche de C229, ce qui pourrait expliquer un changement de conformation suffisant de la protéine S pour diminuer la fusion virale. En effet, il avait été montré pour la glycoprotéine gp120 de HIV que la présence de N-glycanes proche de cystéines impliquées dans des ponts disulphures était importante pour l'infectivité du virus (Mathys and Balzarini, 2015). Le rôle de ces



mutants lors de la fusion virale devra être confirmé en réalisant des expériences de fusion cellule-cellule. Une étude sur la N-glycosylation et les ponts disulfures pourrait également apporter des informations leur importance lors de la fusion de la protéine S. D'autre part, la protéine S est assemblée en trimères à la surface de la particule virale, des expériences pour vérifier que les mutants N6 et N9 forment des trimères ont été entreprises, cependant des mises au point sont encore en cours.

Le clivage protéolytique de notre souche clinique de la protéine S du HCoV-229E a été caractérisé. Nous avons montré que contrairement à la plupart des coronavirus (Belouzard et al., 2009; Millet and Whittaker, 2014; Yamada and Liu, 2009), la protéine S du HCoV-229E ne subissait pas un clivage séquentiel dans la région S1/S2 puis en S2'. Nous n'avons pas pu mettre en évidence de clivage dans la région S1/S2, mais il est possible que nos outils ne nous permettent pas la détection d'un produit de clivage qui serait très transitoire. Nous avons montré que la protéine S du HCoV-229E était clivée dans la région S2' et que le résidu important pour le clivage par la trypsine ou TMPRSS2 dans la région S2' était R683. Il serait intéressant de confirmer la protéase utilisée par HCoV-229E lors de l'entrée par voie endosomale et de déterminer les résidus importants pour l'entrée du virus par cette voie. Il avait été montré par Shirato *et al.* que la HCoV-229E pouvait utiliser la cathepsine L lors de l'entrée par voie endosomale mais pas la cathepsine B (Shirato et al., 2016). De plus, il avait été suggéré qu'une autre protéase que la cathepsine L intervienne lors de l'entrée endosomale. L'étude des protéases et résidus impliqués dans l'activation protéolytique permettra d'apporter des nouvelles possibilités thérapeutiques. Lors de nos études, nous avons utilisé les cellules Huh-7 comme modèle cellulaire. Ce sont des cellules issues d'hépatocarcinome qui répliquent très bien les coronavirus mais les cellules hépatiques ne constituent pas une cible principale du virus HCoV-229E. Il serait donc intéressant de confirmer les résultats que nous avons obtenu dans des épithéliums respiratoires primaires.

La glycosylation peut avoir un rôle important lors de l'interaction avec le récepteur comme il a été montré pour DENV (Yap et al., 2017). C'est pourquoi nous avons testé si les mutations des sites potentiels de N-glycosylation qui affectent l'infectiosité des pseudoparticules, notamment pour les mutants N6 et N9 modifiaient la liaison d'une forme soluble du récepteur. Aucun des 8 N-glycanes présents dans

le RBD ne semble jouer un rôle dans l'interaction avec son récepteur aminopeptidase N, cependant notre méthode n'est pas quantitative. Nos résultats indiquent que les N-glycanes présents dans le RBD auraient plutôt un rôle lors de la neutralisation par les anticorps. De manière intéressante, la mutation des sites de N-glycosylation N12 et N15 augmentent significativement l'entrée virale. Ces N-glycanes sont situés dans le RBD de la protéine S, il est possible que ces mutants se lient plus efficacement au récepteur mais que nous ne l'ayons pas détecté. Par contre, nous avons montré que les mutations des sites de N-glycosylation N12 et N15 augmentent la sensibilité des particules pseudotypées avec la protéine S aux anticorps neutralisants. Ce résultat suggère donc qu'il existe une balance entre la diminution de l'infectivité et le masquage d'épitopes lié à la présence de ces N-glycanes. La N-glycosylation semble aussi avoir un double rôle dans le cas du virus Nipah : il a été montré que la glycosylation de la protéine de fusion F permet d'augmenter la résistance à la neutralisation par les anticorps mais cause une diminution de la fusion des membranes et de l'entrée cellulaire (Aguilar et al., 2006).

Nous avons montré que sur les 18 sites potentiels de N-glycosylation du domaine S1, seul 5 N-glycanes ont un rôle lors du cycle viral et plus particulièrement lors de l'entrée. Il est donc possible que les autres N-glycanes soient majoritairement impliqués dans les relations hôte-pathogène. Nous avons confirmé que la plupart des glycanes présents dans le domaine S1 du HCoV-229E masquent des épitopes reconnus par les anticorps de patients. Sur les 8 N-glycanes présents dans le RBD, nous avons montré que N11, N12, N15, N16, N17 et N18, soit 6 d'entre eux masquaient des épitopes. Ces résultats suggèrent que la forte densité de glycanes présents sur la protéine S sert de bouclier, comme il a été décrit pour la protéine gp120 du HIV, qui a une quantité de glycanes similaire à la protéine S et qui a beaucoup été étudiée (Mascola and Montefiori, 2003; Wei et al., 2003). L'acquisition et le réarrangement des glycanes présents à la surface du virus permet une évolution constante du virus et un échappement à la neutralisation par les anticorps. Cette évolution constante des glycanes à la surface des glycoprotéines explique pourquoi une grande partie des N-glycanes de la protéine S participe à la protection contre la neutralisation. Il faudrait déterminer si les N-glycanes présents dans le domaine S2 sont également impliqués dans la protection de la protéine S contre la neutralisation par les anticorps.

L'ensemble de notre étude s'est appuyée sur les pseudoparticules comme outil pour l'étude de l'entrée. Il faudrait valider ces résultats avec un modèle de génétique inverse, dans lequel toutes les protéines du HCoV-229E seraient exprimées. On pourrait alors tester les différents mutants de la protéine S, et valider le clivage protéolytique de la protéine S ainsi que le rôle des mutants de N-glycosylation.



## 6 BIBLIOGRAPHIE

---

- Aguilar, H.C., Matreyek, K.A., Filone, C.M., Hashimi, S.T., Levroney, E.L., Negrete, O.A., Bertolotti-Ciarlet, A., Choi, D.Y., McHardy, I., Fulcher, J.A., et al. (2006). N-Glycans on Nipah Virus Fusion Protein Protect against Neutralization but Reduce Membrane Fusion and Viral Entry. *J. Virol.* *80*, 4878–4889.
- Al-Abdallat, M.M., Payne, D.C., Alqasrawi, S., Rha, B., Tohme, R.A., Abedi, G.R., Nsour, M.A., Iblan, I., Jarour, N., Farag, N.H., et al. (2014). Hospital-Associated Outbreak of Middle East Respiratory Syndrome Coronavirus: A Serologic, Epidemiologic, and Clinical Description. *Clin. Infect. Dis. Off. Publ. Infect. Dis. Soc. Am.* *59*, 1225–1233.
- Annan, A., Baldwin, H.J., Corman, V.M., Klose, S.M., Owusu, M., Nkrumah, E.E., Badu, E.K., Anti, P., Agbenyega, O., Meyer, B., et al. (2013). Human Betacoronavirus 2c EMC/2012–related Viruses in Bats, Ghana and Europe. *Emerg. Infect. Dis.* *19*, 456–459.
- Arabi, Y.M., Arifi, A.A., Balkhy, H.H., Najm, H., Aldawood, A.S., Ghabashi, A., Hawa, H., Alothman, A., Khaldi, A., and Al Raiy, B. (2014). Clinical Course and Outcomes of Critically Ill Patients With Middle East Respiratory Syndrome Coronavirus Infection. *Ann. Intern. Med.* *160*, 389–397.
- Bailey, C.C., Zhong, G., Huang, I.-C., and Farzan, M. (2014). IFITM-Family Proteins: The Cell's First Line of Antiviral Defense. *Annu. Rev. Virol.* *1*, 261–283.
- Bartosch, B., Dubuisson, J., and Cosset, F.-L. (2003). Infectious Hepatitis C Virus Pseudo-particles Containing Functional E1–E2 Envelope Protein Complexes. *J. Exp. Med.* *197*, 633–642.
- Bauch, C.T., and Oraby, T. (2013). Assessing the pandemic potential of MERS-CoV. *The Lancet* *382*, 662–664.
- Beckham, J.D., Cadena, A., Lin, J., Piedra, P.A., Glezen, W.P., Greenberg, S.B., and Atmar, R.L. (2005). Respiratory viral infections in patients with chronic, obstructive pulmonary disease. *J. Infect.* *50*, 322–330.
- Behrens, A.-J., and Crispin, M. (2017). Structural principles controlling HIV envelope glycosylation. *Curr. Opin. Struct. Biol.* *44*, 125–133.
- Belouzard, S., Chu, V.C., and Whittaker, G.R. (2009). Activation of the SARS coronavirus spike protein via sequential proteolytic cleavage at two distinct sites. *Proc. Natl. Acad. Sci. U. S. A.* *106*, 5871–5876.
- Belouzard, S., Madu, I., and Whittaker, G.R. (2010). Elastase-mediated Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein at Discrete Sites within the S2 Domain. *J. Biol. Chem.* *285*, 22758–22763.
- Belouzard, S., Millet, J.K., Licitra, B.N., and Whittaker, G.R. (2012). Mechanisms of coronavirus cell entry mediated by the viral spike protein. *Viruses* *4*, 1011–1033.

- Bertram, S., Dijkman, R., Habjan, M., Heurich, A., Gierer, S., Glowacka, I., Welsch, K., Winkler, M., Schneider, H., Hofmann-Winkler, H., et al. (2013). TMPRSS2 activates the human coronavirus 229E for cathepsin-independent host cell entry and is expressed in viral target cells in the respiratory epithelium. *J. Virol.* *87*, 6150–6160.
- Blijleven, J.S., Boonstra, S., Onck, P.R., van der Giessen, E., and van Oijen, A.M. (2016). Mechanisms of influenza viral membrane fusion. *Rhomboid Superfamily Dev. Dis. Fusion* *60*, 78–88.
- Bonavia, A., Zelus, B.D., Wentworth, D.E., Talbot, P.J., and Holmes, K.V. (2003). Identification of a receptor-binding domain of the spike glycoprotein of human coronavirus HCoV-229E. *J. Virol.* *77*, 2530–2538.
- Bosch, B.J., Smits, S.L., and Haagmans, B.L. (2014). Membrane ectopeptidases targeted by human coronaviruses. *Curr. Opin. Virol.* *0*, 55–60.
- Bradburne, A.F., Bynoe, M.L., and Tyrrell, D.A. (1967). Effects of a “new” human respiratory virus in volunteers. *Br. Med. J.* *3*, 767–769.
- Breban, R., Riou, J., and Fontanet, A. (2013). Interhuman transmissibility of Middle East respiratory syndrome coronavirus: estimation of pandemic risk. *The Lancet* *382*, 694–699.
- Breslin, J.J., Mørk, I., Smith, M.K., Vogel, L.K., Hemmila, E.M., Bonavia, A., Talbot, P.J., Sjöström, H., Norén, O., and Holmes, K.V. (2003). Human coronavirus 229E: receptor binding domain and neutralization by soluble receptor at 37 degrees C. *J. Virol.* *77*, 4435–4438.
- Burkard, C., Verheije, M.H., Wicht, O., van Kasteren, S.I., van Kuppeveld, F.J., Haagmans, B.L., Pelkmans, L., Rottier, P.J.M., Bosch, B.J., and de Haan, C.A.M. (2014). Coronavirus Cell Entry Occurs through the Endo-/Lysosomal Pathway in a Proteolysis-Dependent Manner. *PLoS Pathog.* *10*, e1004502.
- Callow, K.A., Parry, H.F., Sergeant, M., and Tyrrell, D.A. (1990). The time course of the immune response to experimental coronavirus infection of man. *Epidemiol. Infect.* *105*, 435–446.
- Calo, D., Kaminski, L., and Eichler, J. (2010). Protein glycosylation in Archaea: Sweet and extreme. *Glycobiology* *20*, 1065–1076.
- Calvo, E., Escors, D., López, J.A., González, J.M., Álvarez, A., Arza, E., and Enjuanes, L. (2005). Phosphorylation and subcellular localization of transmissible gastroenteritis virus nucleocapsid protein in infected cells. *J. Gen. Virol.* *86*, 2255–2267.
- Cao, Z., Liu, L., Du, L., Zhang, C., Jiang, S., Li, T., and He, Y. (2010). Potent and persistent antibody responses against the receptor-binding domain of SARS-CoV spike protein in recovered patients. *Virol. J.* *7*, 299–299.
- Cauchemez, S., Fraser, C., Kerkhove, M.D.V., Donnelly, C.A., Riley, S., Rambaut, A., Enouf, V., Werf, S. van der, and Ferguson, N.M. (2014). Middle East respiratory

- syndrome coronavirus: quantification of the extent of the epidemic, surveillance biases, and transmissibility. *Lancet Infect. Dis.* *14*, 50–56.
- Chan, C.-M., Tsoi, H., Chan, W.-M., Zhai, S., Wong, C.-O., Yao, X., Chan, W.-Y., Tsui, S.K.-W., and Chan, H.Y.E. (2009). The ion channel activity of the SARS-coronavirus 3a protein is linked to its pro-apoptotic function. *Int. J. Biochem. Cell Biol.* *41*, 2232–2239.
- Chan, J.F.-W., To, K.K.-W., Tse, H., Jin, D.-Y., and Yuen, K.-Y. (2013). Interspecies transmission and emergence of novel viruses: lessons from bats and birds. *Trends Microbiol.* *21*, 544–555.
- Chan, J.F.W., Lau, S.K.P., To, K.K.W., Cheng, V.C.C., Woo, P.C.Y., and Yuen, K.-Y. (2015). Middle East Respiratory Syndrome Coronavirus: Another Zoonotic Betacoronavirus Causing SARS-Like Disease. *Clin. Microbiol. Rev.* *28*, 465–522.
- Chan, J.F.W., Sridhar, S., Yip, C.C.Y., Lau, S.K.P., and Woo, P.C.Y. (2017). The role of laboratory diagnostics in emerging viral infections: the example of the Middle East respiratory syndrome epidemic. *J. Microbiol.* *55*, 172–182.
- Chan, K.H., Poon, L.L.L.M., Cheng, V.C.C., Guan, Y., Hung, I.F.N., Kong, J., Yam, L.Y.C., Seto, W.H., Yuen, K.Y., and Peiris, J.S.M. (2004). Detection of SARS Coronavirus in Patients with Suspected SARS. *Emerg. Infect. Dis.* *10*, 294–299.
- Charrin, S., le Naour, F., Silvie, O., Milhiet, P.-E., Boucheix, C., and Rubinstein, E. (2009). Lateral organization of membrane proteins: tetraspanins spin their web. *Biochem. J.* *420*, 133.
- Chen, L., Lin, Y.-L., Peng, G., and Li, F. (2012). Structural basis for multifunctional roles of mammalian aminopeptidase N. *Proc. Natl. Acad. Sci. U. S. A.* *109*, 17966–17971.
- Chen, Y., Rajashankar, K.R., Yang, Y., Agnihothram, S.S., Liu, C., Lin, Y.-L., Baric, R.S., and Li, F. (2013). Crystal Structure of the Receptor-Binding Domain from Newly Emerged Middle East Respiratory Syndrome Coronavirus. *J. Virol.* *87*, 10777–10783.
- Chen, Y., Lu, S., Jia, H., Deng, Y., Zhou, J., Huang, B., Yu, Y., Lan, J., Wang, W., Lou, Y., et al. (2017). A novel neutralizing monoclonal antibody targeting the N-terminal domain of the MERS-CoV spike protein. *Emerg. Microbes Infect.* *6*, e37.
- Chowell, G., Abdirizak, F., Lee, S., Lee, J., Jung, E., Nishiura, H., and Viboud, C. (2015). Transmission characteristics of MERS and SARS in the healthcare setting: a comparative study. *BMC Med.* *13*, 210.
- Cocquerel, L., Voisset, C., and Dubuisson, J. (2006). Hepatitis C virus entry: potential receptors and their biological functions. *J. Gen. Virol.* *87*, 1075–1084.
- Cook, J.K.A., Jackwood, M., and Jones, R.C. (2012). The long view: 40 years of infectious bronchitis research. *Avian Pathol.* *41*, 239–250.
- Corman, V.M., Eckerle, I., Memish, Z.A., Liljander, A.M., Dijkman, R., Jonsdottir, H., Juma Ngeiywa, K.J.Z., Kamau, E., Younan, M., Al Masri, M., et al. (2016). Link of a

- ubiquitous human coronavirus to dromedary camels. *Proc. Natl. Acad. Sci. U. S. A.* *113*, 9864–9869.
- Corman, V.M., Muth, D., Niemeyer, D., and Drosten, C. (2018). Chapter Eight - Hosts and Sources of Endemic Human Coronaviruses. In *Advances in Virus Research*, M. Kielian, T.C. Mettenleiter, and M.J. Roossinck, eds. (Academic Press), pp. 163–188.
- Corti, D., and Lanzavecchia, A. (2013). Broadly Neutralizing Antiviral Antibodies. *Annu. Rev. Immunol.* *31*, 705–742.
- Corver, J., Broer, R., van Kasteren, P., and Spaan, W. (2007). GxxxG Motif of Severe Acute Respiratory Syndrome Coronavirus Spike Glycoprotein Transmembrane Domain Is Not Involved in Trimerization and Is Not Important for Entry. *J. Virol.* *81*, 8352–8355.
- Daniels, R., Kurowski, B., Johnson, A.E., and Hebert, D.N. (2003). N-Linked Glycans Direct the Cotranslational Folding Pathway of Influenza Hemagglutinin. *Mol. Cell* *11*, 79–90.
- Delmas, B., and Laude, H. (1990). Assembly of coronavirus spike protein into trimers and its role in epitope expression. *J. Virol.* *64*, 5367–5375.
- Delmas, B., Gelfi, J., Kut, E., Sjöström, H., Noren, O., and Laude, H. (1994). Determinants essential for the transmissible gastroenteritis virus-receptor interaction reside within a domain of aminopeptidase-N that is distinct from the enzymatic site. *J. Virol.* *68*, 5216–5224.
- Denison, M.R., Graham, R.L., Donaldson, E.F., Eckerle, L.D., and Baric, R.S. (2011). Coronaviruses: An RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol.* *8*, 270–279.
- Deshpande, K.L., Fried, V.A., Ando, M., and Webster, R.G. (1987). Glycosylation affects cleavage of an H5N2 influenza virus hemagglutinin and regulates virulence. *Proc. Natl. Acad. Sci. U. S. A.* *84*, 36–40.
- van Deventer, S.J., Dunlock, V.-M.E., and van Sriel, A.B. (2017). Molecular interactions shaping the tetraspanin web. *Biochem. Soc. Trans.* *45*, 741.
- Dietrich, M.H., Harprecht, C., and Stehle, T. (2017). The bulky and the sweet: How neutralizing antibodies and glycan receptors compete for virus binding. *Protein Sci.* n/a–n/a.
- Dijkman, R., and van der Hoek, L. (2009). Human Coronaviruses 229E and NL63: Close Yet Still So Far. *J. Formos. Med. Assoc.* *108*, 270–279.
- Drexler, J.F., Corman, V.M., and Drosten, C. (2014). Ecology, evolution and classification of bat coronaviruses in the aftermath of SARS. *Antiviral Res.* *101*, 45–56.
- Du, L., Yang, Y., Zhou, Y., Lu, L., Li, F., and Jiang, S. (2017). MERS-CoV spike protein: a key target for antivirals. *Expert Opin. Ther. Targets* *21*, 131–143.



- Earnest, J.T., Hantak, M.P., Park, J.-E., and Gallagher, T. (2015). Coronavirus and Influenza Virus Proteolytic Priming Takes Place in Tetraspanin-Enriched Membrane Microdomains. *J. Virol.* *89*, 6093–6104.
- Earnest, J.T., Hantak, M.P., Li, K., McCray, P.B., Perlman, S., and Gallagher, T. (2017). The tetraspanin CD9 facilitates MERS-coronavirus entry by scaffolding host cell receptors and proteases. *PLoS Pathog.* *13*, e1006546.
- Eckerle, L.D., Lu, X., Sperry, S.M., Choi, L., and Denison, M.R. (2007). High Fidelity of Murine Hepatitis Virus Replication Is Decreased in nsp14 Exoribonuclease Mutants. *J. Virol.* *81*, 12135–12144.
- Elden, V., R, L.J., M, A., M, A., van Alphen, F., Hendriksen, K.A.W., Hoepelman, A.I.M., Kraaij, V., J, M.G., Oosterheert, J.-J., et al. (2004). Frequent Detection of Human Coronaviruses in Clinical Specimens from Patients with Respiratory Tract Infection by Use of a Novel Real-Time Reverse-Transcriptase Polymerase Chain Reaction. *J. Infect. Dis.* *652*–657.
- Esper, F., Weibel, C., Ferguson, D., Landry, M.L., and Kahn, J.S. (2005). Evidence of a Novel Human Coronavirus That Is Associated with Respiratory Tract Disease in Infants and Young Children. *J. Infect. Dis.* *191*, 492–498.
- Falsey, A.R., and Walsh, E.E. (2006). Viral Pneumonia in Older Adults. *Clin. Infect. Dis.* *42*, 518–524.
- Farsani, S.M.J., Dijkman, R., Jebbink, M.F., Goossens, H., Ieven, M., Deijs, M., Molenkamp, R., and van der Hoek, L. (2012). The first complete genome sequences of clinical isolates of human coronavirus 229E. *Virus Genes* *45*, 433–439.
- Fehr, A.R., and Perlman, S. (2015). Coronaviruses: An Overview of Their Replication and Pathogenesis. *Methods Mol. Biol. Clifton NJ* *1282*, 1–23.
- Filho, da S., Ferreira, L.V.R., Zerbinati, R.M., Tateno, A.F., Boas, L.V., Almeida, D., Buarque, M., Levi, J.E., Drexler, J.F., Drosten, C., et al. (2012). The Differential Clinical Impact of Human Coronavirus Species in Children With Cystic Fibrosis. *J. Infect. Dis.* *206*, 384–388.
- Fontes-Garfias, C.R., Shan, C., Luo, H., Muruato, A.E., Medeiros, D.B., Mays, E., Xie, X., Zou, J., Roundy, C.M., Wakamiya, M., et al. (2017). Functional analysis of glycosylation of Zika virus envelope protein. *Cell Rep.* *21*, 1180–1190.
- Forni, D., Cagliani, R., Clerici, M., and Sironi, M. (2017). Molecular Evolution of Human Coronavirus Genomes. *Trends Microbiol.*
- Frickmann, H., Jungblut, S., Hirche, T.O., Groß, U., Kuhns, M., and Zautner, A.E. (2012). Spectrum of viral infections in patients with cystic fibrosis. *Eur. J. Microbiol. Immunol.* *2*, 161–175.
- Fukushi, M., Yoshinaka, Y., Matsuoka, Y., Hatakeyama, S., Ishizaka, Y., Kirikae, T., Sasazuki, T., and Miyoshi-Akiyama, T. (2012). Monitoring of S protein maturation in the endoplasmic reticulum by calnexin is important for the infectivity of severe acute

respiratory syndrome coronavirus. *J. Virol.* *86*, 11745–11753.

Gagneux, P., and Varki, A. (1999). Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* *9*, 747–755.

Gao, H., Yao, H., Yang, S., and Li, L. (2016). From SARS to MERS: evidence and speculation. *Front. Med.* *10*, 377–382.

Gao, J., Lu, G., Qi, J., Li, Y., Wu, Y., Deng, Y., Geng, H., Li, H., Wang, Q., Xiao, H., et al. (2013). Structure of the fusion core and inhibition of fusion by a heptad repeat peptide derived from the S protein of Middle East respiratory syndrome coronavirus. *J. Virol.* *87*, 13134–13140.

Garbino, J., Crespo, S., Aubert, J.-D., Rochat, T., Ninet, B., Deffernez, C., Wunderli, W., Pache, J.-C., Soccac, P.M., and Kaiser, L. (2006). A Prospective Hospital-Based Study of the Clinical Impact of Non-Severe Acute Respiratory Syndrome (Non-SARS)-Related Human Coronavirus Infection. *Clin. Infect. Dis.* *43*, 1009–1015.

Gaunt, E.R., Hardie, A., Claas, E.C.J., Simmonds, P., and Templeton, K.E. (2010). Epidemiology and Clinical Presentations of the Four Human Coronaviruses 229E, HKU1, NL63, and OC43 Detected over 3 Years Using a Novel Multiplex Real-Time PCR Method. *J. Clin. Microbiol.* *48*, 2940–2947.

Ge, X.-Y., Li, J.-L., Yang, X.-L., Chmura, A.A., Zhu, G., Epstein, J.H., Mazet, J.K., Hu, B., Zhang, W., Peng, C., et al. (2013). Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* *503*, 535–538.

Gierer, S., Bertram, S., Kaup, F., Wrensch, F., Heurich, A., Krämer-Kühl, A., Welsch, K., Winkler, M., Meyer, B., Drosten, C., et al. (2013). The spike protein of the emerging betacoronavirus EMC uses a novel coronavirus receptor for entry, can be activated by TMPRSS2, and is targeted by neutralizing antibodies. *J. Virol.* *87*, 5502–5511.

Glezen, W.P., Greenberg, S.B., Atmar, R.L., Piedra, P.A., and Couch, R.B. (2000). Impact of Respiratory Virus Infections on Persons With Chronic Underlying Conditions. *JAMA* *283*, 499–505.

Godet, M., Grosclaude, J., Delmas, B., and Laude, H. (1994). Major receptor-binding and neutralization determinants are located within the same domain of the transmissible gastroenteritis virus (coronavirus) spike protein. *J. Virol.* *68*, 8008–8016.

Goffard, A., Callens, N., Bartosch, B., Wychowski, C., Cosset, F.-L., Montpellier, C., and Dubuisson, J. (2005). Role of N-linked glycans in the functions of hepatitis C virus envelope glycoproteins. *J. Virol.* *79*, 8400–8409.

Graepel, K.W., Lu, X., Case, J.B., Sexton, N.R., Smith, E.C., and Denison, M.R. (2017). Proofreading-Deficient Coronaviruses Adapt for Increased Fitness over Long-Term Passage without Reversion of Exoribonuclease-Inactivating Mutations. *MBio* *8*, e01503-17.

- Graham, R.L., Becker, M.M., Eckerle, L.D., Bolles, M., Denison, M.R., and Baric, R.S. (2012). A live, impaired-fidelity coronavirus vaccine protects in an aged, immunocompromised mouse model of lethal disease. *Nat. Med.* 18, 1820.
- Greenberg, S.B., Allen, M., Wilson, J., and Atmar, R.L. (2000). Respiratory Viral Infections in Adults With and Without Chronic Obstructive Pulmonary Disease. *Am. J. Respir. Crit. Care Med.* 162, 167–173.
- Gretebeck, L.M., and Subbarao, K. (2015). Animal models for SARS and MERS coronaviruses. *Curr. Opin. Virol.* 13, 123–129.
- Guan, Y., Zheng, B.J., He, Y.Q., Liu, X.L., Zhuang, Z.X., Cheung, C.L., Luo, S.W., Li, P.H., Zhang, L.J., Guan, Y.J., et al. (2003). Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* 302, 276–278.
- Guery, B., Poissy, J., el Mansouf, L., Séjourné, C., Ettahar, N., Lemaire, X., Vuotto, F., Goffard, A., Behillil, S., Enouf, V., et al. (2013). Clinical features and viral diagnosis of two cases of infection with Middle East Respiratory Syndrome coronavirus: a report of nosocomial transmission. *The Lancet* 381, 2265–2272.
- de Haan, C.A.M., and Rottier, P.J.M. (2005). Molecular interactions in the assembly of coronaviruses. *Adv. Virus Res.* 64, 165–230.
- Hamre, D., and Procknow, J.J. (1966). A New Virus Isolated from the Human Respiratory Tract. *Exp. Biol. Med.* 121, 190–193.
- Han, D.P., Lohani, M., and Cho, M.W. (2007). Specific asparagine-linked glycosylation sites are critical for DC-SIGN- and L-SIGN-mediated severe acute respiratory syndrome coronavirus entry. *J. Virol.* 81, 12029–12039.
- HCSP (2013). Règles d'hygiène pour la prise en charge des patients suspects d'infections dues au nouveau coronavirus (Paris: Haut Conseil de la Santé Publique).
- Heald-Sargent, T., and Gallagher, T. (2012). Ready, Set, Fuse! The Coronavirus Spike Protein and Acquisition of Fusion Competence. *Viruses* 4, 557–580.
- Helenius, A., and Aebi, M. (2001). Intracellular Functions of N-Linked Glycans. *Science* VOL 291.
- Helle, F., Vieyres, G., Elkrief, L., Popescu, C.-I., Wychowski, C., Descamps, V., Castelain, S., Roingard, P., Duverlie, G., and Dubuisson, J. (2010). Role of N-Linked Glycans in the Functions of Hepatitis C Virus Envelope Proteins Incorporated into Infectious Virions. *J. Virol.* 84, 11905–11915.
- Hemler, M.E. (2005). Tetraspanin functions and associated microdomains. *Nat. Rev. Mol. Cell Biol.* 6, 801.
- Hitzerd, S.M., Verbrugge, S.E., Ossenkuppele, G., Jansen, G., and Peters, G.J. (2014). Positioning of aminopeptidase inhibitors in next generation cancer therapy. *Amino Acids* 46, 793–808.

- van der Hoek, L., Pyrc, K., Jebbink, M.F., Vermeulen-Oost, W., Berkhout, R.J.M., Wolthers, K.C., Wertheim-van Dillen, P.M.E., Kaandorp, J., Spaargaren, J., and Berkhout, B. (2004). Identification of a new human coronavirus. *Nat. Med.* *10*, 368–373.
- van der Hoek, L., Pyrc, K., and Berkhout, B. (2006). Human coronavirus NL63, a new respiratory virus. *FEMS Microbiol. Rev.* *30*, 760–773.
- Hofmann, H., Hattermann, K., Marzi, A., Gramberg, T., Geier, M., Krumbiegel, M., Kuate, S., Überla, K., Niedrig, M., and Pöhlmann, S. (2004). S Protein of Severe Acute Respiratory Syndrome-Associated Coronavirus Mediates Entry into Hepatoma Cell Lines and Is Targeted by Neutralizing Antibodies in Infected Patients. *J. Virol.* *78*, 6134–6142.
- Hofmann, H., Pyrc, K., van der Hoek, L., Geier, M., Berkhout, B., and Pöhlmann, S. (2005). Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 7988–7993.
- Hofmann, H., Simmons, G., Rennekamp, A.J., Chaipan, C., Gramberg, T., Heck, E., Geier, M., Wegele, A., Marzi, A., Bates, P., et al. (2006). Highly conserved regions within the spike proteins of human coronaviruses 229E and NL63 determine recognition of their respective cellular receptors. *J. Virol.* *80*, 8639–8652.
- Hui, D.S., Perlman, S., and Zumla, A. (2015). Spread of MERS to South Korea and China. *Lancet Respir. Med.* *3*, 509–510.
- Hulswit, R.J.G., de Haan, C.A.M., and Bosch, B.-J. (2016). Chapter Two - Coronavirus Spike Protein and Tropism Changes. In *Advances in Virus Research*, J. Ziebuhr, ed. (Academic Press), pp. 29–57.
- Jevšnik, M., Uršič, T., Žigon, N., Lusa, L., Krivec, U., and Petrovec, M. (2012). Coronavirus infections in hospitalized pediatric patients with acute respiratory tract disease. *BMC Infect. Dis.* *12*, 365–365.
- Jiang, S., Lu, L., Liu, Q., Xu, W., and Du, L. (2012). Receptor-binding domains of spike proteins of emerging or re-emerging viruses as targets for development of antiviral vaccines. *Emerg. Microbes Infect.* *1*, e13.
- Jonsdottir, H.R., and Dijkman, R. (2016). Coronaviruses and the human airway: a universal system for virus-host interaction studies. *Virol. J.* *13*, 24.
- Kawaoka, Y., Naeve, C.W., and Webster, R.G. (1984). Is virulence of H5N2 influenza viruses in chickens associated with loss of carbohydrate from the hemagglutinin? *Virology* *139*, 303–316.
- Kawase, M., Shirato, K., Matsuyama, S., and Taguchi, F. (2009). Protease-mediated entry via the endosome of human coronavirus 229E. *J. Virol.* *83*, 712–721.
- Kirchdoerfer, R.N., Cottrell, C.A., Wang, N., Pallesen, J., Yassine, H.M., Turner, H.L., Corbett, K.S., Graham, B.S., McLellan, J.S., and Ward, A.B. (2016). Pre-fusion

structure of a human coronavirus spike protein. *Nature* 531, 118–121.

Knoops, K., Kikkert, M., van den Worm, S.H.E., Zevenhoven-Dobbe, J.C., van der Meer, Y., Koster, A.J., Mommaas, A.M., and Snijder, E.J. (2008). SARS-Coronavirus Replication Is Supported by a Reticulovesicular Network of Modified Endoplasmic Reticulum. *PLoS Biol.* 6, e226.

Kolb, A.F., Hegyi, A., and Siddell, S.G. (1997). Identification of residues critical for the human coronavirus 229E receptor function of human aminopeptidase N. *J. Gen. Virol.* 78, 2795–2802.

Lamriben, L., Graham, J.B., Adams, B.M., and Hebert, D.N. (2016). N-glycan based ER molecular chaperone and protein quality control system: the calnexin binding cycle. *Traffic Cph. Den.* 17, 308–326.

Larkin, A., and Imperiali, B. (2011). The Expanding Horizons of Asparagine-Linked Glycosylation. *Biochemistry (Mosc.)* 50, 4411–4426.

Lassnig, C., Sanchez, C.M., Egerbacher, M., Walter, I., Majer, S., Kolbe, T., Pallares, P., Enjuanes, L., and Müller, M. (2005). Development of a transgenic mouse model susceptible to human coronavirus 229E. *Proc. Natl. Acad. Sci. U. S. A.* 102, 8275–8280.

Lau, S.K.P., and Chan, J.F.W. (2015). Coronaviruses: emerging and re-emerging pathogens in humans and animals. *Virol. J.* 12, 209.

Lau, S.K.P., Woo, P.C.Y., Li, K.S.M., Huang, Y., Tsoi, H.-W., Wong, B.H.L., Wong, S.S.Y., Leung, S.-Y., Chan, K.-H., and Yuen, K.-Y. (2005). Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl. Acad. Sci.* 102, 14040–14045.

Lau, S.K.P., Woo, P.C.Y., Yip, C.C.Y., Tse, H., Tsoi, H., Cheng, V.C.C., Lee, P., Tang, B.S.F., Cheung, C.H.Y., Lee, R.A., et al. (2006). Coronavirus HKU1 and Other Coronavirus Infections in Hong Kong. *J. Clin. Microbiol.* 44, 2063–2071.

Lau, S.K.P., Woo, P.C.Y., Li, K.S.M., Tsang, A.K.L., Fan, R.Y.Y., Luk, H.K.H., Cai, J.-P., Chan, K.-H., Zheng, B.-J., Wang, M., et al. (2015a). Discovery of a Novel Coronavirus, China Rattus Coronavirus HKU24, from Norway Rats Supports the Murine Origin of Betacoronavirus 1 and Has Implications for the Ancestor of Betacoronavirus Lineage A. *J. Virol.* 89, 3076–3092.

Lau, S.K.P., Feng, Y., Chen, H., Luk, H.K.H., Yang, W.-H., Li, K.S.M., Zhang, Y.-Z., Huang, Y., Song, Z.-Z., Chow, W.-N., et al. (2015b). Severe Acute Respiratory Syndrome (SARS) Coronavirus ORF8 Protein Is Acquired from SARS-Related Coronavirus from Greater Horseshoe Bats through Recombination. *J. Virol.* 89, 10532–10547.

Law, P.T.W., Wong, C.-H., Au, T.C.C., Chuck, C.-P., Kong, S.-K., Chan, P.K.S., To, K.-F., Lo, A.W.I., Chan, J.Y.W., Suen, Y.-K., et al. (2005). The 3a protein of severe acute respiratory syndrome-associated coronavirus induces apoptosis in Vero E6 cells. *J. Gen. Virol.* 86, 1921–1930.

- Li, F. (2016). Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu. Rev. Virol.* 3, 237–261.
- Li, C., Li, W., Lucio de Esesarte, E., Guo, H., van den Elzen, P., Aarts, E., van den Born, E., Rottier, P.J.M., and Bosch, B.-J. (2017a). Cell Attachment Domains of the Porcine Epidemic Diarrhea Virus Spike Protein Are Key Targets of Neutralizing Antibodies. *J. Virol.* 91.
- Li, F., Li, W., Farzan, M., and Harrison, S.C. (2005a). Structure of SARS Coronavirus Spike Receptor-Binding Domain Complexed with Receptor. *Science* 309, 1864–1868.
- Li, W., Moore, M.J., Vasilieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greenough, T.C., et al. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* 426, 450–454.
- Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J.H., Wang, H., Crameri, G., Hu, Z., Zhang, H., et al. (2005b). Bats Are Natural Reservoirs of SARS-Like Coronaviruses. *Science* 310, 676–679.
- Li, Y.-Y., Fang, J., and Ao, G.-Z. (2017b). Cathepsin B and L inhibitors: a patent review (2010 - present). *Expert Opin. Ther. Pat.* 27, 643–656.
- Liu, C., Feng, Y., Gao, F., Zhang, Q., and Wang, M. (2006). Characterization of HCoV-229E fusion core: implications for structure basis of coronavirus membrane fusion. *Biochem. Biophys. Res. Commun.* 345, 1108–1115.
- Lo, Y.-S., Lin, S.-Y., Wang, S.-M., Wang, C.-T., Chiu, Y.-L., Huang, T.-H., and Hou, M.-H. (2013). Oligomerization of the carboxyl terminal domain of the human coronavirus 229E nucleocapsid protein. *FEBS Lett.* 587, 120–127.
- Lu, G., Hu, Y., Wang, Q., Qi, J., Gao, F., Li, Y., Zhang, Y., Zhang, W., Yuan, Y., Bao, J., et al. (2013). Molecular basis of binding between novel human coronavirus MERS-CoV and its receptor CD26. *Nature* 500, 227–231.
- Lu, W., Zheng, B.-J., Xu, K., Schwarz, W., Du, L., Wong, C.K.L., Chen, J., Duan, S., Deubel, V., and Sun, B. (2006). Severe acute respiratory syndrome-associated coronavirus 3a protein forms an ion channel and modulates virus release. *Proc. Natl. Acad. Sci. U. S. A.* 103, 12540–12545.
- Ma, G., Feng, Y., Gao, F., Wang, J., Liu, C., and Li, Y. (2005). Biochemical and biophysical characterization of the transmissible gastroenteritis coronavirus fusion core. *Biochem. Biophys. Res. Commun.* 337, 1301–1307.
- Ma, Y., Wu, L., Shaw, N., Gao, Y., Wang, J., Sun, Y., Lou, Z., Yan, L., Zhang, R., and Rao, Z. (2015). Structural basis and functional analysis of the SARS coronavirus nsp14–nsp10 complex. *Proc. Natl. Acad. Sci. U. S. A.* 112, 9436–9441.
- Madu, I.G., Roth, S.L., Belouzard, S., and Whittaker, G.R. (2009). Characterization of a Highly Conserved Domain within the Severe Acute Respiratory Syndrome

Coronavirus Spike Protein S2 Domain with Characteristics of a Viral Fusion Peptide. *J. Virol.* **83**, 7411–7421.

Mahony, J.B., Petrich, A., and Smieja, M. (2011). Molecular diagnosis of respiratory virus infections. *Crit. Rev. Clin. Lab. Sci.* **48**, 217–249.

Martin J B Raamsman, J.K.L.A. de H.A.A.F. de V.G.G.H.V.P.J.M.R. (2000). Characterization of the Coronavirus Mouse Hepatitis Virus Strain A59 Small Membrane Protein E. *J. Virol.* **74**, 2333.

Mascola, J.R., and Montefiori, D.C. (2003). HIV-1: nature's master of disguise. *Nat. Med.* **9**, 393–394.

Masters, P.S. (2006). The molecular biology of coronaviruses. *Adv. Virus Res.* **66**, 193–292.

Mathys, L., and Balzarini, J. (2015). Several N-Glycans on the HIV Envelope Glycoprotein gp120 Preferentially Locate Near Disulphide Bridges and Are Required for Efficient Infectivity and Virus Transmission. *PLOS ONE* **10**, e0130621.

Matsuyama, S., and Taguchi, F. (2009). Two-Step Conformational Changes in a Coronavirus Envelope Glycoprotein Mediated by Receptor Binding and Proteolysis. *J. Virol.* **83**, 11133–11141.

Matsuyama, S., Ujike, M., Morikawa, S., Tashiro, M., and Taguchi, F. (2005). Protease-mediated enhancement of severe acute respiratory syndrome coronavirus infection. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12543–12547.

Memish, Z.A., Mishra, N., Olival, K.J., Fagbo, S.F., Kapoor, V., Epstein, J.H., AlHakeem, R., Durosinloun, A., Al Asmari, M., Islam, A., et al. (2013). Middle East Respiratory Syndrome Coronavirus in Bats, Saudi Arabia. *Emerg. Infect. Dis.* **19**, 1819–1823.

Menachery, V.D., Graham, R.L., and Baric, R.S. (2017). Jumping species—a mechanism for coronavirus persistence and survival. *Viral Pathog.* **23**, 1–7.

Millet, J.K., and Whittaker, G.R. (2014). Host cell entry of Middle East respiratory syndrome coronavirus after two-step, furin-mediated activation of the spike protein. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 15214–15219.

Millet, J.K., and Whittaker, G.R. (2015). Host cell proteases: critical determinants of coronavirus tropism and pathogenesis. *Virus Res.* **202**, 120–134.

Millet, J.K., Goldstein, M.E., Labitt, R.N., Hsu, H.-L., Daniel, S., and Whittaker, G.R. (2016). A camel-derived MERS-CoV with a variant spike protein cleavage site and distinct fusion activation properties. *Emerg Microbes Infect* **5**, e126.

Mina-Osorio, P. (2008). The moonlighting enzyme CD13: old and new functions to target. *Trends Mol. Med.* **14**, 361–371.

Minskaia, E., Hertzog, T., Gorbalenya, A.E., Campanacci, V., Cambillau, C., Canard, B., and Ziebuhr, J. (2006). Discovery of an RNA virus 3'→5' exoribonuclease that is

- critically involved in coronavirus RNA synthesis. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 5108.
- Monto, A.S. (1974). Medical reviews. Coronaviruses. *Yale J. Biol. Med.* *47*, 234.
- Neuman, B.W., Adair, B.D., Yoshioka, C., Quispe, J.D., Orca, G., Kuhn, P., Milligan, R.A., Yeager, M., and Buchmeier, M.J. (2006). Supramolecular Architecture of Severe Acute Respiratory Syndrome Coronavirus Revealed by Electron Cryomicroscopy. *J. Virol.* *80*, 7918–7928.
- Neuman, B.W., Chamberlain, P., Bowden, F., and Joseph, J. (2014). Atlas of coronavirus replicase structure. *Nidoviruses I* *194*, 49–66.
- Nothaft, H., and Szymanski, C.M. (2010). Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Micro* *8*, 765–778.
- Ohtsubo, K., and Marth, J.D. (2006). Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* *126*, 855–867.
- Oostra, M., de Haan, C.A.M., de Groot, R.J., and Rottier, P.J.M. (2006). Glycosylation of the severe acute respiratory syndrome coronavirus triple-spanning membrane proteins 3a and M. *J. Virol.* *80*, 2326–2336.
- Ou, X., Guan, H., Qin, B., Mu, Z., Wojdyla, J.A., Wang, M., Dominguez, S.R., Qian, Z., and Cui, S. (2017). Crystal structure of the receptor binding domain of the spike glycoprotein of human betacoronavirus HKU1. *Nat. Commun.* *8*, 15216.
- Park, J.-E., Li, K., Barlan, A., Fehr, A.R., Perlman, S., McCray, P.B., and Gallagher, T. (2016). Proteolytic processing of Middle East respiratory syndrome coronavirus spikes expands virus tropism. *Proc. Natl. Acad. Sci.* *113*, 12262–12267.
- Peiris, J.S.M., Guan, Y., and Yuen, K.Y. (2004). Severe acute respiratory syndrome. *Nat. Med.* *10*, S88–S97.
- Peng, G., Sun, D., Rajashankar, K.R., Qian, Z., Holmes, K.V., and Li, F. (2011). Crystal structure of mouse coronavirus receptor-binding domain complexed with its murine receptor. *Proc. Natl. Acad. Sci.* *108*, 10696–10701.
- Pfefferle, S., Oppong, S., Drexler, J.F., Gloza-Rausch, F., Ipsen, A., Seebens, A., Müller, M.A., Annan, A., Vallo, P., Adu-Sarkodie, Y., et al. (2009). Distant Relatives of Severe Acute Respiratory Syndrome Coronavirus and Close Relatives of Human Coronavirus 229E in Bats, Ghana. *Emerg. Infect. Dis.* *15*, 1377–1384.
- Plugge, B., Gazzarrini, S., Nelson, M., Cerana, R., Van, J.L., Etten, Derst, C., DiFrancesco, D., Moroni, A., and Thiel, G. (2000). A Potassium Channel Protein Encoded by Chlorella Virus PBCV-1. *Science* *287*, 1641.
- Poissy, J., Goffard, A., Parmentier-Decrucq, E., Favory, R., Kouv, M., Kipnis, E., Mathieu, D., and Guery, B. (2014). Kinetics and pattern of viral excretion in biological specimens of two MERS-CoV cases. *J. Clin. Virol.* *61*, 275–278.
- Polgár, L. (2013a). Chapter 405 - Catalytic Mechanisms of Cysteine Peptidases. In



Handbook of Proteolytic Enzymes, N.D. Rawlings, and G. Salvesen, eds. (Academic Press), pp. 1773–1784.

Polgár, L. (2013b). Chapter 560 - Catalytic Mechanisms of Serine and Threonine Peptidases. In Handbook of Proteolytic Enzymes, N.D. Rawlings, and G. Salvesen, eds. (Academic Press), pp. 2524–2534.

Ponnusamy, R., Mesters, J.R., Ziebuhr, J., Moll, R., and Hilgenfeld, R. (2006). Non structural proteins 8 and 9 of human coronavirus 229E. *Adv. Exp. Med. Biol.* *581*, 49–54.

Poon, L.L.M., Chu, D.K.W., Chan, K.H., Wong, O.K., Ellis, T.M., Leung, Y.H.C., Lau, S.K.P., Woo, P.C.Y., Suen, K.Y., Yuen, K.Y., et al. (2005). Identification of a Novel Coronavirus in Bats. *J. Virol.* *79*, 2001–2009.

Pozzetto, B., and Groupe Révir (2007). Révir - Référentiel en virologie médicale (Alinéa Plus).

Prill, M.M., Iwane, M.K., Edwards, K.M., Williams, J.V., Weinberg, G.A., Staat, M.A., Willby, M.J., Talbot, H.K., Hall, C.B., Szilagyi, P.G., et al. (2012). Human Coronavirus in Young Children Hospitalized for Acute Respiratory Illness and Asymptomatic Controls. *Pediatr. Infect. Dis. J.* *31*, 235–240.

Raabe, T., Schelle-Prinz, B., and Siddell, S.G. (1990). Nucleotide sequence of the gene encoding the spike glycoprotein of human coronavirus HCV 229E. *J. Gen. Virol.* *71* ( Pt 5), 1065–1073.

Raj, V.S., Mou, H., Smits, S.L., Dekkers, D.H.W., Muller, M.A., Dijkman, R., Muth, D., Demmers, J.A.A., Zaki, A., Fouchier, R.A.M., et al. (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature* *495*, 251–254.

Raska, M., Czernekova, L., Moldoveanu, Z., Zachova, K., Elliott, M.C., Novak, Z., Hall, S., Hoelscher, M., Maboko, L., Brown, R., et al. (2014). Differential glycosylation of envelope gp120 is associated with differential recognition of HIV-1 by virus-specific antibodies and cell infection. *AIDS Res. Ther.* *11*, 23.

Reguera, J., Santiago, C., Mudgal, G., Ordoño, D., Enjuanes, L., and Casasnovas, J.M. (2012). Structural bases of coronavirus attachment to host aminopeptidase N and its inhibition by neutralizing antibodies. *PLoS Pathog.* *8*, e1002859.

Reguera, J., Mudgal, G., Santiago, C., and Casasnovas, J.M. (2014). A structural view of coronavirus–receptor interactions. *Virus Res.* *194*, 3–15.

Reinke, L.M., Spiegel, M., Plegge, T., Hartleib, A., Nehlmeier, I., Gierer, S., Hoffmann, M., Hofmann-Winkler, H., Winkler, M., and Pöhlmann, S. (2017). Different residues in the SARS-CoV spike protein determine cleavage and activation by the host cell protease TMPRSS2. *PLoS ONE* *12*.

Ren, Y., Min, Y.-Q., Liu, M., Chi, L., Zhao, P., and Zhang, X.-L. (2015). N-glycosylation-mutated HCV envelope glycoprotein complex enhances antigen-presenting activity and cellular and neutralizing antibody responses. *Biochim.*

Biophys. Acta.

Reusken, C.B., Haagmans, B.L., Müller, M.A., Gutierrez, C., Godeke, G.-J., Meyer, B., Muth, D., Raj, V.S., Vries, L.S.-D., Corman, V.M., et al. (2013). Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. *Lancet Infect. Dis.* *13*, 859–866.

Risco, C., Antón, I.M., Suñé, C., Pedregosa, A.M., Martín-Alonso, J.M., Parra, F., Carrascosa, J.L., and Enjuanes, L. (1995). Membrane protein molecules of transmissible gastroenteritis coronavirus also expose the carboxy-terminal region on the external surface of the virion. *J. Virol.* *69*, 5269–5277.

Ritchie, G., Harvey, D.J., Feldmann, F., Stroehrer, U., Feldmann, H., Royle, L., Dwek, R.A., and Rudd, P.M. (2010). Identification of N-linked carbohydrates from severe acute respiratory syndrome (SARS) spike glycoprotein. *Virology* *399*, 257–269.

Sabir, J.S.M., Lam, T.T.-Y., Ahmed, M.M.M., Li, L., Shen, Y., E. M. Abo-Aba, S., Qureshi, M.I., Abu-Zeid, M., Zhang, Y., Khiyami, M.A., et al. (2016). Co-circulation of three camel coronavirus species and recombination of MERS-CoVs in Saudi Arabia. *Science* *351*, 81.

Sahar Abdul-Rasool, B.C.F. (2010). Understanding Human Coronavirus HCoV-NL63. *Open Virol. J.* *4*, 76–84.

Samara, E.M., and Abdoun, K.A. (2014). Concerns about Misinterpretation of Recent Scientific Data Implicating Dromedary Camels in Epidemiology of Middle East Respiratory Syndrome (MERS). *MBio* *5*, e01430-14.

Santiago, C., Mudgal, G., Reguera, J., Recacha, R., Albrecht, S., Enjuanes, L., and Casasnovas, J.M. (2017). Allosteric inhibition of aminopeptidase N functions related to tumor growth and virus infection. *Sci. Rep.* *7*, 46045.

Sayaka, T.-Y., Tateki, K., Siddell, S.G., and Fumihito, T. (1991). Localization of major neutralizing epitopes on the S1 polypeptide of the murine coronavirus peplomer glycoprotein. *Virus Res.* *18*, 99–107.

Schachter, H. (2000). The joys of HexNAc. The synthesis and function of N-andO-glycan branches. *Glycoconj. J.* *17*, 465–483.

Sharif-Yakan, A., and Kanj, S.S. (2014). Emergence of MERS-CoV in the Middle East: Origins, Transmission, Treatment, and Perspectives. *PLoS Pathog.* *10*, e1004457.

Shih, Y.-P., Chen, C.-Y., Liu, S.-J., Chen, K.-H., Lee, Y.-M., Chao, Y.-C., and Chen, Y.-M.A. (2006). Identifying epitopes responsible for neutralizing antibody and DC-SIGN binding on the spike glycoprotein of the severe acute respiratory syndrome coronavirus. *J. Virol.* *80*, 10315–10324.

Shirato, K., Kawase, M., Watanabe, O., Hirokawa, C., Matsuyama, S., Nishimura, H., and Taguchi, F. (2012). Differences in neutralizing antigenicity between laboratory and clinical isolates of HCoV-229E isolated in Japan in 2004–2008 depend on the S1

region sequence of the spike protein. *J. Gen. Virol.* **93**, 1908–1917.

Shirato, K., Kawase, M., and Matsuyama, S. (2013). Middle East Respiratory Syndrome Coronavirus Infection Mediated by the Transmembrane Serine Protease TMPRSS2. *J. Virol.* **87**, 12552–12561.

Shirato, K., Kanou, K., Kawase, M., and Matsuyama, S. (2016). Clinical Isolates of Human Coronavirus 229E Bypass the Endosome for Cell Entry. *J. Virol.*

Shirato, K., Kawase, M., and Matsuyama, S. (2017). Wild-type human coronaviruses prefer cell-surface TMPRSS2 to endosomal cathepsins for cell entry. *Virology*.

Shuman, S. (2000). Structure, mechanism, and evolution of the mRNA capping apparatus. In *Progress in Nucleic Acid Research and Molecular Biology*, (Academic Press), pp. 1–40.

Simmons, G., Reeves, J.D., Rennekamp, A.J., Amberg, S.M., Piefer, A.J., and Bates, P. (2004). Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 4240–4245.

Smith, E.C., Blanc, H., Vignuzzi, M., and Denison, M.R. (2013). Coronaviruses Lacking Exoribonuclease Activity Are Susceptible to Lethal Mutagenesis: Evidence for Proofreading and Potential Therapeutics. *PLOS Pathog.* **9**, e1003565.

Snijder, E.J., Decroly, E., and Ziebuhr, J. (2016). Chapter Three - The Nonstructural Proteins Directing Coronavirus RNA Synthesis and Processing. In *Advances in Virus Research*, J. Ziebuhr, ed. (Academic Press), pp. 59–126.

Société Française d'Hygiène Hospitalière (2017). Actualisation des Précautions standard.

Sola, I., Almazán, F., Zúñiga, S., and Enjuanes, L. (2015). Continuous and Discontinuous RNA Synthesis in Coronaviruses. *Annu. Rev. Virol.* **2**, 265–288.

Steiner, D.F. (1998). The proprotein convertases. *Curr. Opin. Chem. Biol.* **2**, 31–39.

Su, S., Wong, G., Shi, W., Liu, J., Lai, A.C.K., Zhou, J., Liu, W., Bi, Y., and Gao, G.F. (2016). Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses. *Trends Microbiol.* **24**, 490–502.

Subissi, L., Decroly, E., Bouvet, M., Gluais, L., Canard, B., and Imbert, I. (2012). Les enzymes de la réplication/transcription chez les coronavirus. *Virologie* **16**, 199–209.

Suzuki, T., Kitajima, K., Inoue, S., and Inoue, Y. (1995). N-Glycosylation/deglycosylation as a mechanism for the post-translational modification/remodification of proteins. *Glycoconj. J.* **12**, 183–193.

Travis R Ruch, C.E.M. (2012). The Coronavirus E Protein: Assembly and Beyond. *Viruses* **4**, 363.

Tusell, S.M., Schittone, S.A., and Holmes, K.V. (2007). Mutational analysis of

- aminopeptidase N, a receptor for several group 1 coronaviruses, identifies key determinants of viral host range. *J. Virol.* *81*, 1261–1273.
- Ujike, M., and Taguchi, F. (2015). Incorporation of Spike and Membrane Glycoproteins into Coronavirus Virions. *Viruses* *7*, 1700–1725.
- Vabret, A., Dina, J., Gouarin, S., Petitjean, J., Tripey, V., Brouard, J., and Freymuth, F. (2008). Human (non-severe acute respiratory syndrome) coronavirus infections in hospitalised children in France. *J. Paediatr. Child Health* *44*, 176–181.
- Vabret, A., Dina, J., Brison, E., Brouard, J., and Freymuth, F. (2009). Coronavirus humains (HCoV). *Pathol. Biol.* *57*, 149–160.
- Vigerust, D.J., and Shepherd, V.L. (2007). Virus glycosylation: role in virulence and immune interactions. *Trends Microbiol.* *15*, 211–218.
- Walls, A.C., Tortorici, M., Bosch, B.-J., Frenz, B., Rottier, P.J., DiMaio, F., Rey, F.A., and Velesler, D. (2016b). Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer. *Nature* *531*, 114–117.
- Walls, A.C., Tortorici, M.A., Frenz, B., Snijder, J., Li, W., Rey, F.A., DiMaio, F., Bosch, B.-J., and Velesler, D. (2016a). Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. *Nat Struct Mol Biol advance online publication*.
- Walsh, E.E., Shin, J.H., and Falsey, A.R. (2013). Clinical impact of human coronaviruses 229E and OC43 infection in diverse adult populations. *J. Infect. Dis.* *208*, 1634–1642.
- Wang, L., Shi, W., Chappell, J.D., Joyce, M.G., Zhang, Y., Kanekiyo, M., Becker, M.M., van Doremalen, N., Fischer, R., Wang, N., et al. (2018). Importance of neutralizing monoclonal antibodies targeting multiple antigenic sites on MERS-CoV Spike to avoid neutralization escape. *J. Virol.*
- Wang, N., Shi, X., Jiang, L., Zhang, S., Wang, D., Tong, P., Guo, D., Fu, L., Cui, Y., Liu, X., et al. (2013a). Structure of MERS-CoV spike receptor-binding domain complexed with human receptor DPP4. *Cell Res.* *23*, 986–993.
- Wang, Q., Qi, J., Yuan, Y., Xuan, Y., Han, P., Wan, Y., Ji, W., Li, Y., Wu, Y., Wang, J., et al. (2014). Bat origins of MERS-CoV supported by bat coronavirus HKU4 usage of human receptor CD26. *Cell Host Microbe* *16*, 328–337.
- Wang, W., Nie, J., Prochnow, C., Truong, C., Jia, Z., Wang, S., Chen, X.S., and Wang, Y. (2013b). A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. *Retrovirology* *10*, 14.
- Wei, X., Decker, J.M., Wang, S., Hui, H., Kappes, J.C., Wu, X., Salazar-Gonzalez, J.F., Salazar, M.G., Kilby, J.M., Saag, M.S., et al. (2003). Antibody neutralization and escape by HIV-1. *Nature* *422*, 307.
- Weiss, S.R., and Leibowitz, J.L. (2011). Coronavirus pathogenesis. *Adv. Virus Res.* *81*, 85–164.

- Wentworth, D.E., and Holmes, K.V. (2001). Molecular determinants of species specificity in the coronavirus receptor aminopeptidase N (CD13): influence of N-linked glycosylation. *J. Virol.* *75*, 9741–9752.
- White, J.M., and Whittaker, G.R. (2016). Fusion of Enveloped Viruses in Endosomes. *Traffic Cph. Den.* *17*, 593–614.
- WHO (2004). WHO | Severe acute respiratory syndrome (SARS).
- WHO (2015a). Summary and risk assessment of current situation in Republic of Korea and China.
- WHO (2015b). WHO | Investigation of cases of human infection with MERS-CoV.
- WHO (2015c). OMS | Prévention et lutte contre les infections lors de la prise en charge de cas probables ou confirmés d'infection par le coronavirus du syndrome respiratoire du Moyen-Orient (MERS-CoV).
- WHO (2016). Global summary and risk assessment.
- WHO (2017). WHO | SARS (Severe Acute Respiratory Syndrome).
- WHO (2018). WHO | Middle East respiratory syndrome coronavirus (MERS-CoV).
- Williams, R.K., Jiang, G.S., and Holmes, K.V. (1991). Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. *Proc. Natl. Acad. Sci. U. S. A.* *88*, 5533–5536.
- Wilson, L., Gage, P., and Ewart, G. (2006). Hexamethylene amiloride blocks E protein ion channels and inhibits coronavirus replication. *Virology* *353*, 294–306.
- de Wit, E., van Doremalen, N., Falzarano, D., and Munster, V.J. (2016). SARS and MERS: recent insights into emerging coronaviruses. *Nat Rev Micro* *14*, 523–534.
- Wong, A.H.M., Zhou, D., and Rini, J.M. (2012). The X-ray crystal structure of human aminopeptidase N reveals a novel dimer and the basis for peptide processing. *J. Biol. Chem.* *287*, 36804–36813.
- Wong, A.H.M., Tomlinson, A.C.A., Zhou, D., Satkunarajah, M., Chen, K., Sharon, C., Desforges, M., Talbot, P.J., and Rini, J.M. (2017). Receptor-binding loops in alphacoronavirus adaptation and evolution. *Nat. Commun.* *8*, 1735.
- Woo, P.C.Y., Lau, S.K.P., Chu, C., Chan, K., Tsoi, H., Huang, Y., Wong, B.H.L., Poon, R.W.S., Cai, J.J., Luk, W., et al. (2005). Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J. Virol.* *79*, 884–895.
- Woo, P.C.Y., Lau, S.K.P., Yip, C.C.Y., Huang, Y., and Yuen, K.-Y. (2009). More and More Coronaviruses: Human Coronavirus HKU1. *Viruses* *1*, 57–71.
- Woo, P.C.Y., Lau, S.K.P., Lam, C.S.F., Lau, C.C.Y., Tsang, A.K.L., Lau, J.H.N., Bai, R., Teng, J.L.L., Tsang, C.C.C., Wang, M., et al. (2012). Discovery of Seven Novel Mammalian and Avian Coronaviruses in the Genus Deltacoronavirus Supports Bat

Coronaviruses as the Gene Source of Alphacoronavirus and Betacoronavirus and Avian Coronaviruses as the Gene Source of Gammacoronavirus and Deltacoronavirus. *J. Virol.* 86, 3995–4008.

Wrensch, F., Winkler, M., and Pöhlmann, S. (2014). IFITM Proteins Inhibit Entry Driven by the MERS-Coronavirus Spike Protein: Evidence for Cholesterol-Independent Mechanisms. *Viruses* 6, 3683–3698.

Wu, K., Li, W., Peng, G., and Li, F. (2009). Crystal structure of NL63 respiratory coronavirus receptor-binding domain complexed with its human receptor. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19970–19974.

Wu, Z., Yang, L., Ren, X., Zhang, J., Yang, F., Zhang, S., and Jin, Q. (2016). ORF8-Related Genetic Evidence for Chinese Horseshoe Bats as the Source of Human Severe Acute Respiratory Syndrome Coronavirus. *J. Infect. Dis.* 213, 579–583.

Xiong, X., Tortorici, M.A., Snijder, J., Yoshioka, C., Walls, A.C., Li, W., McGuire, A.T., Rey, F.A., Bosch, B.-J., and Velesler, D. (2017). Glycan shield and fusion activation of a deltacoronavirus spike glycoprotein fine-tuned for enteric infections. *J. Virol.* JVI.01628-17.

Yamada, Y., and Liu, D.X. (2009). Proteolytic Activation of the Spike Protein at a Novel RRRR/S Motif Is Implicated in Furin-Dependent Entry, Syncytium Formation, and Infectivity of Coronavirus Infectious Bronchitis Virus in Cultured Cells. *J. Virol.* 83, 8744–8758.

Yan, A., and Lennarz, W.J. (2005). Unraveling the Mechanism of Protein N-Glycosylation. *J. Biol. Chem.* 280, 3121–3124.

Yáñez-Mó, M., Gutiérrez-López, M.D., and Cabañas, C. (2011). Functional interplay between tetraspanins and proteases. *Cell. Mol. Life Sci. CMLS* 68, 3323–3335.

Yang, Y.-M., Hsu, C.-Y., Lai, C.-C., Yen, M.-F., Wikramaratna, P.S., Chen, H.-H., and Wang, T.-H. (2017). Impact of Comorbidity on Fatality Rate of Patients with Middle East Respiratory Syndrome. *Sci. Rep.* 7, 11307.

YangYang, DengYao, WenBo, WangHuijuan, MengXin, LanJiaming, F, G., and TanWenjie (2014). The Amino Acids 736–761 of the MERS-CoV Spike Protein Induce Neutralizing Antibodies: Implications for the Development of Vaccines and Antiviral Agents. *Viral Immunol.* 27, 543–550.

Yap, S.S.L., Nguyen-Khuong, T., Rudd, P.M., and Alonso, S. (2017). Dengue Virus Glycosylation: What Do We Know? *Front. Microbiol.* 8, 1415.

Yeager, C.L., Ashmun, R.A., Williams, R.K., Cardellichio, C.B., Shapiro, L.H., Look, A.T., and Holmes, K.V. (1992). Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* 357, 420–422.

Yin, Y., Zhang, X., Qiao, Y., Wang, X., Su, Y., Chen, S., Qin, T., Peng, D., and Liu, X. (2017). Glycosylation at 11Asn on hemagglutinin of H5N1 influenza virus contributes to its biological characteristics. *Vet. Res.* 48, 81.

- Yuan, Y., Cao, D., Zhang, Y., Ma, J., Qi, J., Wang, Q., Lu, G., Wu, Y., Yan, J., Shi, Y., et al. (2017). Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat. Commun.* 8, 15092.
- Zaki, A.M., van Boheemen, S., Bestebroer, T.M., Osterhaus, A.D.M.E., and Fouchier, R.A.M. (2012). Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *N. Engl. J. Med.* 367, 1814–1820.
- Zeng, R., Yang, R.-F., Shi, M.-D., Jiang, M.-R., Xie, Y.-H., Ruan, H.-Q., Jiang, X.-S., Shi, L., Zhou, H., Zhang, L., et al. (2004). Characterization of the 3a Protein of SARS-associated Coronavirus in Infected Vero E6 Cells and SARS Patients. *J. Mol. Biol.* 341, 271–279.
- Zhang, R., Wang, K., Lv, W., Yu, W., Xie, S., Xu, K., Schwarz, W., Xiong, S., and Sun, B. (2014). The ORF4a protein of human coronavirus 229E functions as a viroporin that regulates viral production. *Viral Membr. Proteins - Channels Cell. Netw.* 1838, 1088–1095.
- Zhang, S., Tuo, J., Huang, X., Zhu, X., Zhang, D., Zhou, K., Yuan, L., Luo, H., Zheng, B., Yuen, K., et al. (2018). Epidemiology characteristics of human coronaviruses in patients with respiratory infection symptoms and phylogenetic analysis of HCoV-OC43 during 2010-2015 in Guangzhou. *PLOS ONE* 13, e0191789.
- Zhang, X., Chen, S., Yang, D., Wang, X., Zhu, J., Peng, D., and Liu, X. (2015). Role of stem glycans attached to haemagglutinin in the biological characteristics of H5N1 avian influenza virus. *J. Gen. Virol.* 96, 1248–1257.
- Zheng, J., Yamada, Y., Fung, T.S., Huang, M., Chia, R., and Liu, D.X. (2018). Identification of N-linked glycosylation sites in the spike protein and their functional impact on the replication and infectivity of coronavirus infectious bronchitis virus in cell culture. *Virology* 513, 65–74.
- Zhou, Y., Lu, K., Pfefferle, S., Bertram, S., Glowacka, I., Drosten, C., Pöhlmann, S., and Simmons, G. (2010). A single asparagine-linked glycosylation site of the severe acute respiratory syndrome coronavirus spike glycoprotein facilitates inhibition by mannose-binding lectin through multiple mechanisms. *J. Virol.* 84, 8753–8764.
- Zöller, M. (2008). Tetraspanins: push and pull in suppressing and promoting metastasis. *Nat. Rev. Cancer* 9, 40.
- Zumla, A., Hui, D.S., and Perlman, S. (2015). Middle East Respiratory Syndrome. *Lancet Lond. Engl.* 386, 995–1007.
- Zumla, A., Chan, J.F.W., Azhar, E.I., Hui, D.S.C., and Yuen, K.-Y. (2016). Coronaviruses - drug discovery and therapeutic options. *Nat Rev Drug Discov* 15, 327–347.





## 7 ANNEXES

**Tableau 3:** Séquences des primers. Les nucléotides mutés sont indiqués en gras. F = primer sens, R = primer antisens.

Mutants	Primers
R565N	F : TGCAGCCCA <b>A</b> ATAACGTGTCCTACGACAGCGT R : GACACGTT <b>A</b> TTGGGCTGCACGGCAATGATGC
Fur565	F : AGCC <b>G</b> CC <b>G</b> CC <b>G</b> CC <b>G</b> CAAG <b>G</b> CC <b>G</b> CAACGTGTCCTACGACAGCGT R : GTT <b>G</b> CC <b>G</b> CC <b>G</b> CTT <b>G</b> CC <b>G</b> CC <b>G</b> CC <b>G</b> GGCTGCCGTCGGCCGACACGC
R679N	F : CAGCCTGCCCA <b>A</b> TTCCGGCTCTAGAGTGGCCGG R : TAGAGCCGGA <b>A</b> TTGGGCAGGCTGGGGATCACGC
R683N	F : ATCCGGCTCT <b>A</b> ATGTGGCCGGAAGAAGCGCCAT R : TTCCGGCCAC <b>A</b> TTAGAGCCGGATCTGGGCAGGC
R687N	F : TGGCCGGA <b>A</b> ATAGCGCCATCGAAGATATCCT R : ATGGCGCT <b>A</b> TTCCGGCCACTCTAGAGCCGG
3R	F : CCC <b>A</b> ATTCCGGCTCT <b>A</b> ATGTGGCCGGA <b>A</b> ATAGCGCCAT R : CAC <b>A</b> TTAGAGCCGGA <b>A</b> TTGGGCAGGCTGGGGATCACGC
Fu687	F : TCTAGAC <b>G</b> CT <b>C</b> G <b>C</b> G <b>A</b> AGCGCCATCGAAGATATCCT R : GCT <b>T</b> CT <b>G</b> CC <b>G</b> AG <b>C</b> GTCTAGAGCCGGATCTGGGCAGGC
N2	F : CAACGGCACCC <b>A</b> AACCAGCCACA R : TGTGGCTGGTT <b>T</b> GGGTGCCGTTG
N3	F : CTGCTGACCC <b>A</b> AACCTCCAGCGTG R : GCTGGAGGTT <b>T</b> GGGTCAGCAGGAA
N4	F : GTGTACTTCC <b>A</b> AGGCACCGGCAGA R : GCCGGTGCCT <b>T</b> GGAAGTACACGAA
N5	F : TTCTACAGCC <b>A</b> AGCCAGCAGCGAC R : GCTGCTGGCT <b>T</b> GGCTGTAGAAGCC
N6	F : TACTGCACCC <b>A</b> AAACACCCTGGTG R : CAGGGTGTT <b>T</b> GGGTGCAGTAGAA
N7	F : CTGCTTCGTG <b>C</b> AAACCACCATCG R : CGATGGTGGTT <b>T</b> GCACGAAGCAG
N8	F : CACCATCGGCC <b>A</b> AGAGACAACCA

	R : TGGTTGTCTCT <b>TTGG</b> CCGATGGTG
N9	F : GTGAACTT <b>CCA</b> AGTGACCAACGCC R : GTTGGTCACT <b>TTG</b> GAAGTTCACGGC
N10	F : GTGCTGGT <b>CCA</b> AGTGTCCCAGACC R : CTGGGACACT <b>TTG</b> GACCAGCACGTC
N11	F : CTCCGTGAT <b>CCA</b> AATCACCCCTGGCCAACTTCAA R : CCAGGGTGAT <b>TTG</b> GATCACGGAGGGCCGGCAGT
N12	F : GCCAACTT <b>CCA</b> AGAGACAAAGGGCCCTCTGTGC R : CCTTTGTCTCT <b>TTG</b> GAAGTTGGCCAGGGTGATAT
N13	F : AGCTTCATG <b>CA</b> AGTGACCCTGAAC R : CAGGGTCACT <b>TTG</b> CATGAAGCTGGA
N14	F : AGAATCAGC <b>CA</b> AGACACCTTCCTG R : GAAGGTGTCT <b>TTG</b> GCTGATTCTGAT
N15	F : GACGTGACAC <b>CA</b> AGGCACCATCTAC R : ATGGTGCCT <b>TTG</b> TGTCACGTCCTTG
N16	F : CTGAGCGAG <b>CA</b> ATTCACCAGCTAC R : CTGGTGAAT <b>TTG</b> CTCGCTCAGCATG
N17	F : CTACGCCAGC <b>CA</b> AGGCACATACTGCACCGA R : TGTATGTGCCT <b>TTG</b> GCTGGCGTAGAAGAACTTGG
N18	F : CGGCACATAC <b>CA</b> ATGCACCCGACGCCGTGCTGAC R : CGTCGGTGCAT <b>TTG</b> GTATGTGCCGTTGCTGGCGT

**Tableau 4** : Numéro d'accèsion d'alphacoronavirus et betacoronavirus

<b>Numéro d'accèsion</b>	<b>Coronavirus</b>	<b>Souche</b>
AGT51331.1	Coronavirus humain NL63	
KF514433.1	Coronavirus humain 229E	USA/93-40/1993
AB691763.1	Coronavirus humain 229E	ATCC
BAL45641.1	Coronavirus humain 229E	Niigata/01/08
AB691766.1	Coronavirus humain 229E	Sendai-H/1948/04
	Coronavirus humain 229E	04/2014
KF745068.1	Coronavirus du syndrome respiratoire du Moyen-Orient MERS-CoV	FRA/UAE
AAP13441.1	Coronavirus du syndrome respiratoire aigu sévère SARS-CoV	Urbani

**Tableau 5:** Analyses statistique des pentes

Comparaison	Différence des pentes	p-value	IC95% de la différence	
N2 vs S / ctrl	14.4074	0.5376	-31.4408	60.2555
N2 vs S / E	-15.8718	0.4477	-56.8797	25.1360
N2 vs S / F	-52.8836	0.0036	-88.3975	-17.3698
N3 vs S / ctrl	6.4246	0.7586	-34.5832	47.4325
N3 vs S / E	-28.0660	0.1796	-69.0738	12.9419
N3 vs S / F	-49.1698	0.0067	-84.6836	-13.6560
N4 vs S / ctrl	13.8135	0.5104	-27.3569	54.9839
N4 vs S / E	-78.4201	0.0002	-119.43	-37.4123
N4 vs S / F	-71.4600	<.0001	-106.97	-35.9462
N5 vs S / ctrl	25.0987	0.2300	-15.9091	66.1066
N5 vs S / E	-45.9785	0.0280	-86.9863	-4.9706
N5 vs S / F	-59.4204	0.0012	-95.3687	-23.4721
N6 vs S / ctrl	-5.5206	0.7917	-46.5285	35.4872
N6 vs S / E	13.7784	0.5098	-27.2295	54.7862
N6 vs S / F	-3.0677	0.8753	-41.4270	35.2917
N7 vs S / ctrl	0.2100	0.9920	-40.7978	41.2179
N7 vs S / E	-49.2596	0.0186	-90.2674	-8.2517
N7 vs S / F	-58.7054	0.0014	-94.6537	-22.7572
N8 vs S / ctrl	9.4534	0.6859	-36.3947	55.3016
N8 vs S / E	-17.6827	0.3977	-58.6906	23.3251
N8 vs S / F	-45.2388	0.0126	-80.7526	-9.7249
N9 vs S / ctrl	-13.9291	0.5512	-59.7773	31.9190
N9 vs S / E	30.7475	0.1415	-10.2604	71.7553
N9 vs S / F	28.7101	0.1130	-6.8038	64.2239
N10 vs S / ctrl	16.4111	0.4325	-24.5967	57.4190
N10 vs S / E	-57.9528	0.0057	-98.9607	-16.9450
N10 vs S / F	-75.2159	<.0001	-110.73	-39.7020
N11 vs S / ctrl	-7.5956	0.7452	-53.4437	38.2526
N11 vs S / E	-110.80	<.0001	-151.80	-69.7882
N11 vs S / F	-114.68	<.0001	-150.19	-79.1614
N12 vs S / ctrl	4.8878	0.8151	-36.1200	45.8956

N12 vs S / E	-66.3936	0.0015	-107.40	-25.3857
N12 vs S / F	-101.11	<.0001	-136.63	-65.5991
N13 vs S / ctrl	0.1924	0.9927	-40.8154	41.2003
N13 vs S / E	-34.0523	0.1035	-75.0602	6.9555
N13 vs S / F	-24.2048	0.1814	-59.7187	11.3090
N14 vs S / ctrl	8.5065	0.6841	-32.5014	49.5143
N14 vs S / E	18.7516	0.4224	-27.0965	64.5998
N14 vs S / F	-89.4945	<.0001	-127.00	-51.9854
N15 vs S / ctrl	-6.4652	0.7571	-47.4731	34.5426
N15 vs S / E	-130.26	<.0001	-171.27	-89.2515
N15 vs S / F	-151.50	<.0001	-187.01	-115.98
N16 vs S / ctrl	2.9836	0.8865	-38.0242	43.9915
N16 vs S / E	-69.1368	0.0010	-110.14	-28.1290
N16 vs S / F	-93.9581	<.0001	-129.47	-58.4443
N17 vs S / ctrl	-16.9416	0.4182	-57.9897	24.1064
N17 vs S / E	-121.66	<.0001	-162.67	-80.6496
N17 vs S / F	-115.25	<.0001	-150.76	-79.7345
N18 vs S / ctrl	-22.3326	0.2855	-63.3405	18.6752
N18 vs S / E	-90.8074	<.0001	-131.82	-49.7995
N18 vs S / F	-95.8919	<.0001	-131.41	-60.3780
N19 vs S / ctrl	-6.7829	0.7456	-47.7908	34.2249

**Article** : HCoV-229E spike protein fusion activation by trypsin-like serine proteases is mediated by proteolytic processing in the S2' region.

1     **HCoV-229E spike protein fusion activation by trypsin-like serine proteases is mediated**  
2                                     **by proteolytic processing in the S2' region.**  
3  
4  
5     Ariane Bonnin, Adeline Danneels, Jean Dubuisson, Anne Goffard<sup>#</sup> and Sandrine Belouzard<sup>#</sup>  
6     \*.  
7  
8     Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 8204 - CIIL-  
9     Centre d'Infection et d'Immunité de Lille, F-59000 Lille, France.  
10  
11    \*Corresponding authors:  
12    *E-mail: [sandrine.belouzard@ibl.cnrs.fr](mailto:sandrine.belouzard@ibl.cnrs.fr)*  
13    <sup>#</sup>*Equally contributed to the work.*  
14  
15    Keywords  
16    HCoV-229E, fusion, entry, spike protein  
17  
18    Subject category  
19        Animal  
20        RNA Viruses  
21  
22    Word count : 2082  
23  
24    Depositories:  
25    The GenBank accession number for the France/Lille/2014 spike protein sequence is :  
26    MH048989  
27  
28    Abbreviations :  
29    HCoV-229E, Human coronavirus 229E; SARS-CoV, severe acute respiratory syndrome  
30    coronavirus; MERS-CoV, Middle East respiratory syndrome coronavirus; MHV, mouse  
31    hepatitis virus; TTSP, transmembrane serine protease; TMPRSS2, transmembrane protease  
32    serine 2

**33 Abstract**

34

35 Human coronavirus 229E (HCoV-229E) is responsible for common colds. As other  
36 coronaviruses, HCoV-229E exploits cellular proteases to activate its fusion mediated by the  
37 spike protein. We analysed the proteolytic processing of the HCoV-229E spike protein by  
38 trypsin-like serine proteases leading to activation of the fusion process. Unlike other  
39 coronaviruses, HCoV-229E fusion activation appears to be a one step process. Indeed,  
40 cleavage of the S1/S2 interface does not seem to be a pre-requisite, and the fusion activation  
41 strongly relies on the S2' region, with arginine residue 683 acting as the recognition site.

42

**43 Main text**

44

45 *Coronavirinae* are enveloped virus with a long positive sense RNA genome. As for other  
46 enveloped viruses, they need to fuse their viral envelope with a host cell membrane to deliver  
47 their genome to the target cell. Coronavirus spike proteins are major mediators of virus entry  
48 and fulfil a dual function by mediating receptor binding and fusion. Spike proteins consist of  
49 two distinct domains carrying these functions. The amino-terminal subunit, called S1 domain,  
50 is responsible for receptor binding, whereas the C-terminal part, called S2 domain, contains  
51 the fusion machinery. Proteolytic priming of coronavirus spike protein is a key feature of the  
52 fusion process, and coronaviruses exploit a wide variety of host proteases including  
53 cathepsins, furin and members of the transmembrane serine protease (TTSP) family [1] to  
54 achieve this requirement. It is likely that the proteolytic processing differs according to the  
55 protease used. Depending on the cellular localization of the protease used to activate the  
56 fusion, entry can occur either at the cell surface or after internalization of the virus by  
57 endocytosis. For SARS-CoV spike protein, two cleavage events mediated by trypsin have  
58 been identified. The SARS-CoV virion harbors an uncleaved spike protein at its surface. First,  
59 the spike protein is primed by cleavage of the S1 and S2 domain, then to fully activate fusion,  
60 a second cleavage occurs at a position called S2' (R797) exposing the fusion peptide [2, 3].  
61 Activation of the SARS-CoV spike protein fusion by elastase also involves recognition of a  
62 cleavage site in the S2' region [4]. The importance of the S2' cleavage has also been  
63 demonstrated for MERS-CoV and MHV [5, 6].

64 HCoV-229E was identified in 1966 in students suffering from common cold [7], but children,  
65 elderly and immune-compromised persons can develop a more severe disease [8]. It has been  
66 shown that HCoV-229E requires endosomal cathepsin L activity to infect Hela cells [9].  
67 However, Shirato et al. have recently shown that clinical isolates of HCoV-229E

68 preferentially uses the transmembrane protease serine 2 (TMPRSS2), a trypsin-like protease,  
69 rather than endosomal cathepsin for entry [10]. TMPRSS2 is known to be present at the  
70 surface of the human airway epithelia and is likely the main activator of HCoV-229E fusion  
71 *in vivo*. In addition, TMPRSS2-mediated entry at the cell surface protects the virus from the  
72 inhibitory effects of interferon-induced transmembrane proteins (IFITMs) known to inhibit  
73 entry of several enveloped virus [11]. However, the detailed proteolytic processing of HCoV-  
74 229E spike protein to activate its fusion remains to elucidate.

75 To investigate this proteolytic maturation, we first cloned the spike protein sequence from a  
76 patient infected by HCoV-229E hospitalized in Lille, France in 2014 (Spike  
77 France/Lille/2014, accession number MH048989). The codon-optimized sequence of the  
78 France/Lille/2014 spike protein was inserted into the pCDNA3.1(+) vector and fused with a  
79 sequence encoding a C-terminal VSVG tag to facilitate the detection of the protein in  
80 immunoblot. Then, we analysed the proteolytic processing of the spike protein required for  
81 entry mediated by trypsin/trypsin like protease, focusing on the role of the S1/S2 junction and  
82 the S2' region. The sequence of the French isolate spike protein shows 99% identity with the  
83 spike of the Niigata, 0349 and USA/932-72/1993 isolates and only 94% of identity with the  
84 VR-740 laboratory strain. HCoV-229E spike protein contains 3 arginine residues in the S2'  
85 region (R679, R683 and R687) which are conserved in different strains of HCoV-229E,  
86 except for R679 that is absent in the VR-740 laboratory strain (Fig. 1(a)). First, to confirm the  
87 role of the HCoV-229E spike protein S2' region in entry, we inserted a furin cleavage site  
88 (Fur687) in the spike protein (Fig.1(a)), at the N-terminal position of the potential fusion  
89 peptide. To analyse the effect of the mutation in the context of viral particles, we used a  
90 murine leukemia virus (MLV)-based pseudotyping system that produces particles inducing  
91 luciferase expression after transduction of permissive cells. Pseudoparticles fully mimic the  
92 functionality of the wild-type virus in terms of cell entry and were used because of their  
93 technical advantage. Huh-7 cells were inoculated with pseudoparticles harbouring the wild  
94 type protein or the Fur687 mutant (Fig 1(b)). Neutralization of endosomal pH by ammonium  
95 chloride treatment inhibited transduction by HCoV-229E pseudoparticles showing that  
96 HCoV-229E infects Huh-7 cells via the endosomal route and uses low-pH active endosomal  
97 protease to mediate its fusion. As shown in Figure 1(b), pseudotyped virions with HCoV-  
98 229E spike protein containing a furin cleavage site at the S2' position were insensitive to  
99 endosomal pH neutralization, but were affected by furin convertase inhibitor. This result  
100 shows that the spike protein with the furin cleavage site does not rely on endosomal cathepsin  
101 for entry but relies on furin and this confirms the involvement of the S2' region for fusion



102 activation of HCoV-229E, as observed for other coronaviruses [2, 5, 6, 12]. It is interesting to  
103 note that the spike proteins incorporated on the pseudotyped virions were not cleaved by furin  
104 in producing cells, indicating that the protein is not processed during exit of the producer cell  
105 but during entry into the target cell (Fig 1(d)). This suggests that during exit and secretion, the  
106 cleavage site is inaccessible to furin and that during entry, conformational change induced by  
107 receptor binding may expose the cleavage site. To investigate the role of the S1/S2 junction  
108 and the specific role of the 3 arginine residues located in the S2' region (R679, R683 and  
109 R687) in the proteolytic activation of HCoV-229E spike protein by trypsin-like serine-  
110 proteases, the arginine residues present at S1/S2 and at S2' were mutated individually  
111 (R565N, R679N, R683N or R687N). We chose to mutate the arginine residues into  
112 asparagine to minimize the effect of the mutations on the conformation/folding of the protein.  
113 As seen in Fig 1(c), these arginine residues were not implicated in entry when the virions use  
114 the endocytic pathway to invade the cells. The endosomal route of entry of the different  
115 mutants was confirmed by inhibition of their entry by ammonium chloride treatment (Fig  
116 1(c)). Indeed, when HCoV-229E is internalized, the virus is dependent on the endosomal  
117 cathepsin L to mediate its fusion. However, unlike other proteases, cathepsin L has a limited  
118 prime-site specificity in terms of recognition sites and the precise sites of proteolysis by  
119 cathepsin L remains unclear.

120 In 2005, Simmons et al. showed that SARS-CoV fusion can be induced at the cell surface  
121 when virions are bound to the plasma membrane at 4°C and briefly treated by exogenous  
122 protease such as trypsin [13]. Particles pseudotyped with the different S1/S2 and S2' arginine  
123 spike mutants were bound at the surface of Huh-7 cells treated with ammonium chloride to  
124 neutralize the endosomal pH and block the endosomal entry pathway. Then, fusion was  
125 induced by a brief trypsin treatment (Fig 2(a)). For the mutant R683N, trypsin treatment was  
126 not able to bypass the requirement for cathepsin L activity during the entry process,  
127 suggesting that proteolysis by trypsin occurs at R683 residue. Mutations of R679 or R687 had  
128 no effect on entry mediated by trypsin. Spike proteins of field strains of HCoV-229E contain  
129 an arginine at the position 679 whereas the laboratory strain VR-740 contains a threonine.  
130 Field strains were reported to use preferentially the TPPS pathway [10], this result suggest  
131 that acquisition of the arginine residue at the position 679 is not involved in this adaptation. It  
132 is surprising that mutation of R687 had no effect on entry mediated by trypsin. Indeed, R687  
133 is believed to be located at the N-terminus of the fusion peptide and the arginine residue  
134 located at the same position in the SARS-CoV spike protein is the major recognition site of  
135 trypsin. Unlike SARS-CoV, mutation of the arginine residue located at the S1/S2 junction

136 (R565N) has little effect on entry mediated at the cell surface by trypsin. *In vivo*, the major  
137 route of entry of respiratory coronavirus is believed to be mediated by the TTPS pathway. To  
138 confirm our results observed with trypsin, we expressed the human TMPRSS2 in Huh-7 cells  
139 with a lentiviral vector. Transduction of cells with a lentivirus expressing the human  
140 TMPRSS2 resulted in an increase in transduction by particles pseudotyped with the wild type  
141 S protein which was insensitive to ammonium chloride treatment (Fig 2(b)). As previously  
142 observed for the SARS-CoV [13, 14], HCoV-229E pseudotype entry mediated at the cell  
143 surface was more efficient than entry by the endocytic pathway. As seen with infection  
144 mediated by trypsin, mutation of the R683 abolished the capacity of the virus to use  
145 TMPRSS2 to infect the cells since its infectivity was sensitive to ammonium chloride  
146 treatment of the cells (Fig 2(c)). To further confirm that cleavage mediated by trypsin-like  
147 serine protease occurs at R683N, HEK293T cells transiently expressing the wild type or the  
148 different mutant proteins were treated by trypsin for 1h and then proteins at the cell surface  
149 were biotinylated at 4°C. After precipitation of biotinylated protein with streptavidin-  
150 conjugated beads, cleavage products were analysed in western blotting with an antibody  
151 directed against a C-terminal tag of the protein. Spike proteins with a furin cleavage site  
152 introduced at the S1/S2 boundary (Fur565) or at the S2' position (Fur687) were used as  
153 controls (Fig. 3). Cleavage by furin between the S1 and S2 domain gave rise to a band with a  
154 size below 135 kDa corresponding to the S2 domain. Cleavage at the S2' position (Fur687)  
155 induces a cleavage product with a smaller size below 100 kDa. Contrary to other  
156 coronaviruses for which trypsin induces mainly a processing at the S1/S2 junction, cleavage  
157 of the wild-type HCoV-229E spike protein by trypsin resulted in the appearance of a cleavage  
158 product in western blot with a size similar to the one observed for the mutant Fur687,  
159 corresponding to the cleavage of the protein at the S2' position. Mutation of R683 abolished  
160 the cleavage induced by trypsin confirming the proteolytic processing of the spike protein at  
161 this position by trypsin. The absence of any S2 fragment detection suggests that either the S2  
162 fragment is very short-lived, being rapidly processed at the S2' position or that the S1/S2  
163 cleavage is not required for the S2' cleavage. If the S2 fragment was short-lived, then  
164 mutation of R683 should lead to the accumulation of the intermediate cleavage product.  
165 However, we did not observe any accumulation of cleavage product of bigger size when R683  
166 was mutated, arguing again against the requirement of the S1/S2 cleavage to prime the  
167 cleavage at the S2' position as shown for the SARS-CoV and MERS-CoV spike proteins [2,  
168 6].

169 Taken together, our results show that activation of the fusion process by trypsin and trypsin-  
170 like serine protease, particularly TMPRSS2 expressed in human airway epithelia, strongly  
171 rely on the R683 residue as a cleavage site. It has been shown that the position of the cleavage  
172 site in the S2' region of the SARS-CoV spike protein can be shifted from the position relative  
173 to the fusion peptide but with some effects on fusion activation and infection. Indeed, for the  
174 SARS-CoV spike protein, cleavage at the N-terminus of the fusion peptide results in  
175 improved fusion and infection capacities [4]. For HCoV-229E spike protein, the cleavage site  
176 by trypsin is not located at the N-terminus of the fusion peptide but is present 3 residues  
177 upstream. For the SARS-CoV spike protein, even if the location of the cleavage site presents  
178 some flexibility, certain position in the vicinity of the fusion peptide were also not accessible  
179 to protease. Indeed, no cleavage product was observed when R683 was mutated (Fig 3),  
180 suggesting that both arginine residues 679 and 687 were not accessible to the protease,  
181 otherwise, if these residues could be cleaved but without being functional to activate the  
182 fusion, the same band would have been observed with the R683N mutant.

183

#### 184 **Author statements**

185 The authors declare no conflict of interest.

186

#### 187 **Funding**

188 A part of this study was supported by a grant “ Soutien aux thématiques émergentes 2014”  
189 from the University Hospital of Lille, France.

190

#### 191 **Acknowledgements**

192 We thank all the members of the molecular and cellular virology laboratory for helpful  
193 discussions and input during the course of this work.

194

#### 195 **References**

- 196 1. **Simmons G, Zmora P, Gierer S, Heurich A, Pöhlmann S.** Proteolytic activation of the  
197 SARS-coronavirus spike protein: Cutting enzymes at the cutting edge of antiviral  
198 research. *Antiviral Res* 2013;100:605–614.
- 199 2. **Belouzard S, Chu VC, Whittaker GR.** Activation of the SARS coronavirus spike  
200 protein via sequential proteolytic cleavage at two distinct sites. *PNAS* 2009;106:5871–  
201 5876.
- 202 3. **Madu IG, Roth SL, Belouzard S, Whittaker GR.** Characterization of a highly  
203 conserved domain within the severe acute respiratory syndrome coronavirus spike protein  
204 S2 domain with characteristics of a viral fusion peptide. *J Virol* 2009;83:7411–7421.

- 205 4. **Belouzard S, Madu I, Whittaker GR.** Elastase-mediated activation of the severe acute  
206 respiratory syndrome coronavirus spike protein at discrete sites within the S2 domain. *J*  
207 *Biol Chem* 2010;285:22758–22763.
- 208 5. **Burkard C, Verheije MH, Wicht O, van Kasteren SI, van Kuppeveld FJ, et al.**  
209 Coronavirus Cell Entry Occurs through the Endo-/Lysosomal Pathway in a Proteolysis-  
210 Dependent Manner. *PLoS Pathog* 2014;10:e1004502.
- 211 6. **Millet JK, Whittaker GR.** Host cell entry of Middle East respiratory syndrome  
212 coronavirus after two-step, furin-mediated activation of the spike protein. *PNAS*  
213 2014;111:15214–9.
- 214 7. **Hamre D, Procknow JJ.** A new virus isolated from the human respiratory tract. *Proc*  
215 *Soc Exp Biol Med Soc Exp Biol Med N Y N* 1966;121:190–193.
- 216 8. **van der Hoek L.** Human coronaviruses: what do they cause? *Antivir Ther* 2007;12:651–  
217 658.
- 218 9. **Kawase M, Shirato K, Matsuyama S, Taguchi F.** Protease-Mediated Entry via the  
219 Endosome of Human Coronavirus 229E. *J Virol* 2008;83:712–721.
- 220 10. **Shirato K, Kanou K, Kawase M, Matsuyama S.** Clinical Isolates of Human  
221 Coronavirus 229E Bypass the Endosome for Cell Entry. *J Virol* 2016;JVI.01387–16.
- 222 11. **Bertram S, Dijkman R, Habjan M, Heurich A, Gierer S, et al.** TMPRSS2 Activates  
223 the Human Coronavirus 229E for Cathepsin-Independent Host Cell Entry and Is  
224 Expressed in Viral Target Cells in the Respiratory Epithelium. *J Virol* 2013;87:6150–  
225 6160.
- 226 12. **Yamada Y, Liu DX.** Proteolytic activation of the spike protein at a novel RRRR/S motif  
227 is implicated in furin-dependent entry, syncytium formation, and infectivity of  
228 coronavirus infectious bronchitis virus in cultured cells. *J Virol* 2009;83:8744–8758.
- 229 13. **Simmons G, Gosalia DN, Rennekamp AJ, Reeves JD, Diamond SL, et al.** Inhibitors of  
230 cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proc Natl Acad*  
231 *Sci U S A* 2005;102:11876–11881.
- 232 14. **Matsuyama S, Ujike M, Morikawa S, Tashiro M, Taguchi F.** Protease-mediated  
233 enhancement of severe acute respiratory syndrome coronavirus infection. *Proc Natl Acad*  
234 *Sci U S A* 2005;102:12543–12547.

235

236 **FIGURE LEGEND**237 **Figure 1**

238 (A) Alignment of the S1/S2 and S2' region of different coronavirus spike protein (HCoV-  
239 229E isolated in France, 2014, HCoV-229E Niigata/01/08, HCoV-229E 0349, HCoV-229E  
240 USA/932-72/1993, HCoV-229E laboratory strain VR-740, MERS-CoV and SARS-CoV).  
241 R565 at the S1/S2 boundary is in red and R679, R683 and R687 are presented in green. For

7

242 SARS-CoV and MERS-CoV the arginine at the S2' position is also presented in green. The  
243 amino sequence between the S1/S2 boundary and the S2' region is not totally presented and  
244 the omitted sequence is depicted by a dotted line. (B) MLV-based particles pseudotyped with  
245 the wild type spike protein (WT) or with a spike protein in which a furin cleavage site was  
246 introduced in the S2' region (Fur687) were produced and used to inoculate Huh-7 cells in  
247 presence or absence of 25 mM NH<sub>4</sub>Cl or 5 μM of furin inhibitor. Results are presented as the  
248 percentage of transduction in absence of drug (DMSO) and are expressed as the mean of three  
249 independent experiments. Error bars represent the standard error of the means (SEM). Data  
250 were analysed by using an ANOVA test (\*\*\*) P<0.001). (C) Endosomal entry of particles  
251 pseudotyped with the mutated spike proteins. Pseudoparticles were produced and used to  
252 inoculate Huh-7 cells at 37°C in presence or absence of 25 mM NH<sub>4</sub>Cl. 48h later, the cells  
253 were lysed and luciferase activity was measured. Results are presented as relative light units  
254 (RLU) and are expressed as the mean of five independent experiments. Error bars represent  
255 the standard error of the means (SEM). Data were analysed by using an ANOVA test (\*\*\*)  
256 P<0.001). (D) Analysis of pseudoparticles formation. Pseudoparticles were produced with the  
257 different mutated spike proteins and concentrated on sucrose cushion, then pseudoparticles  
258 were lysed and the production was analysed by western blotting with anti-VSVG antibody to  
259 detect the spike protein and anti-gag antibody.

260

261 **Figure 2**

262 (A). Trypsin-mediated entry at the cell surface. Particles pseudotyped with the different  
263 mutated spike proteins were bound at the cell surface at 4°C in presence of 25 mM NH<sub>4</sub>Cl  
264 and were treated with a 5 min incubation with 2 μg/ml of trypsin to induce fusion. 48h later,  
265 the cells were lysed and luciferase activity was measured. Results are presented as the  
266 percentage of transduction observed with particles pseudotyped with the wild type protein and  
267 are expressed as the mean of three independent experiments. Error bars represent the standard  
268 error of the means (SEM). Data were analysed by using an ANOVA test (\*\* P<0.01) (B)  
269 Huh-7 cells were transduced or not with a lentivirus expressing the human TMPRSS2. Then  
270 cells were infected with particles pseudotyped with the wild type spike protein in presence or  
271 absence of 25 mM NH<sub>4</sub>Cl. Results are presented as relative light units (RLU) and are  
272 expressed as the mean of three independent experiments. Data were analysed by using an  
273 ANOVA test (\*\*\*) P<0.001). (C) Huh-7 cells transduced or not with a lentivirus expressing  
274 TMPRSS2 were inoculated with particles pseudotyped with the different mutated spike  
275 proteins. Results are presented as the percentage of infection observed with particles

8

276 pseudotyped with the wild type protein and are expressed as the mean of three independent  
277 experiments. Error bars represent the standard error of the means (SEM). Data were analysed  
278 by using an ANOVA test (\*\*\*) P<0.001).

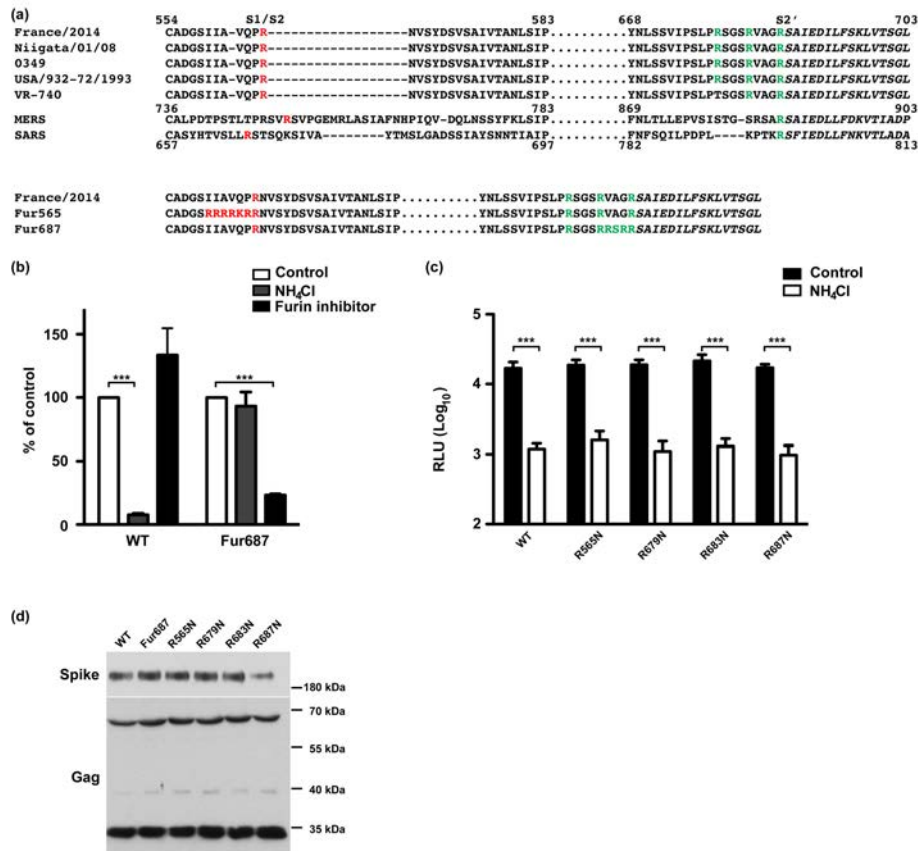
279

280 **Figure 3**

281 Spike protein cleavage with trypsin. HEK293T cells were transfected with plasmid encoding  
282 the different mutated spike protein. 48h later, cells were treated for 1h with 2µg/ml of trypsin  
283 at 37°C. Then cell surface protein were biotinylated at 4°C and the cells were lysed.  
284 Biotinylated protein were precipitated with streptavidin-conjugated beads and analysed by  
285 western blotting with the C-terminal part of the protein detected with an anti-VSVG antibody.

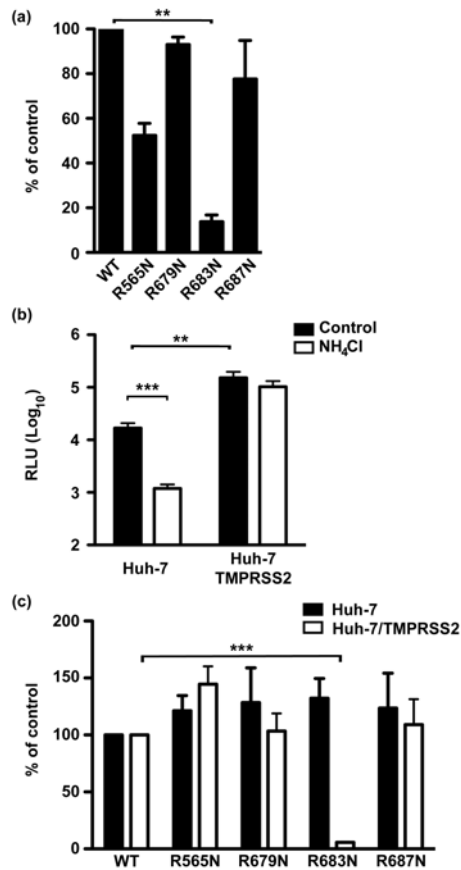
286

287



**Figure 1**

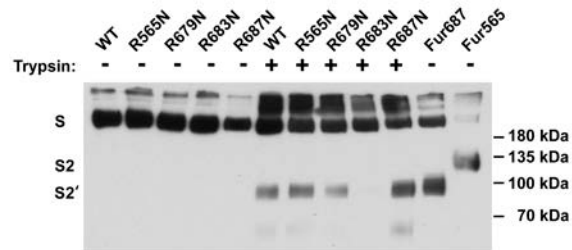
(A) Alignment of the S1/S2 and S2' region of different coronavirus spike protein (HCoV-229E isolated in France, 2014, HCoV-229E Niigata/01/08, HCoV-229E 0349, HCoV-229E USA/932-72/1993, HCoV-229E laboratory strain VR-740, MERS-CoV and SARS-CoV). R565 at the S1/S2 boundary is in red and R679, R683 and R687 are presented in green. For SARS-CoV and MERS-CoV the arginine at the S2' position is also presented in green. The amino sequence between the S1/S2 boundary and the S2' region is not totally presented and the omitted sequence is depicted by a dotted line. (B) MLV-based particles pseudotyped with the wild type spike protein (WT) or with a spike protein in which a furin cleavage site was introduced in the S2' region (Fur687) were produced and used to inoculate Huh-7 cells in presence or absence of 25 mM NH<sub>4</sub>Cl or 5  $\mu$ M of furin inhibitor. Results are presented as the percentage of transduction in absence of drug (DMSO) and are expressed as the mean of three independent experiments. Error bars represent the standard error of the means (SEM). Data were analysed by using an ANOVA test (\*\*\*) P<0.001). (C) Endosomal entry of particles pseudotyped with the mutated spike proteins. Pseudoparticles were produced and used to inoculate Huh-7 cells at 37°C in presence or absence of 25 mM NH<sub>4</sub>Cl. 48h later, the cells were lysed and luciferase activity was measured. Results are presented as relative light units (RLU) and are expressed as the mean of five independent experiments. Error bars represent the standard error of the means (SEM). Data were analysed by using an ANOVA test (\*\*\*) P<0.001). (D) Analysis of pseudoparticles formation. Pseudoparticles were produced with the different mutated spike proteins and concentrated on sucrose cushion, then pseudoparticles were lysed and the production was analysed by western blotting with anti-VSVG antibody to detect the spike protein and anti-gag antibody.



**Figure 2**

(A). Trypsin-mediated entry at the cell surface. Particles pseudotyped with the different mutated spike proteins were bound at the cell surface at 4°C in presence of 25 mM NH<sub>4</sub>Cl and were treated with a 5 min incubation with 2 µg/ml of trypsin to induce fusion. 48h later, the cells were lysed and luciferase activity was measured. Results are presented as the percentage of transduction observed with particles pseudotyped with the wild type protein and are expressed as the mean of three independent experiments. Error bars represent the standard error of the means (SEM). Data were analysed by using an ANOVA test (\*\* P<0.01) (B) Huh-7 cells were transduced or not with a lentivirus expressing the human TMPRSS2. Then cells were infected with particles pseudotyped with the wild type spike protein in presence or absence of 25 mM NH<sub>4</sub>Cl. Results are presented as relative light units (RLU) and are expressed as the mean of three independent experiments. Data were analysed by using an ANOVA test (\*\*\*) P<0.001). (C) Huh-7 cells transduced or not with a lentivirus expressing TMPRSS2 were inoculated with particles pseudotyped with the different mutated spike proteins. Results are presented as the percentage of infection observed with particles pseudotyped with the wild type protein and are expressed as the mean of three independent experiments. Error bars represent the standard error of the means (SEM). Data were analysed by using an ANOVA test (\*\*\*) P<0.001).





**Figure 3**

Spike protein cleavage with trypsin. HEK293T cells were transfected with plasmid encoding the different mutated spike protein. 48h later, cells were treated for 1h with 2 $\mu$ g/ml of trypsin at 37°C. Then cell surface protein were biotinylated at 4°C and the cells were lysed. Biotinylated protein were precipitated with streptavidin-conjugated beads and analysed by western blotting with the C-terminal part of the protein detected with an anti-VSVG antibody.