

**MEMOIRE  
POUR LE DIPLOME D'ETUDES SPECIALISEES  
D'INNOVATION PHARMACEUTIQUE ET RECHERCHE**

**Soutenu publiquement le 20 octobre 2017  
Par Melle Ariane BONNIN**

**Conformément aux dispositions du Décret du 10 septembre 1990  
tient lieu de**

**THESE EN VUE DU DIPLOME D'ETAT DE DOCTEUR EN PHARMACIE**

---

**Etude des mécanismes d'activation  
protéolytique de la protéine S du coronavirus  
humain 229E**

---

**Membres du jury :**

**Président**    **Dr DUBUISSON Jean**, Directeur de recherche, Centre d'Infection et d'Immunité de Lille

**Assesseur**    **Dr BELOUZARD Sandrine**, Chargée de recherche, Centre d'Infection et d'Immunité de Lille

**Dr GOFFARD Anne**, Maître de conférences des Universités - Praticien Hospitalier, Centre d'Infection et d'Immunité de Lille, CHRU de Lille

**Dr HELLE François**, Maître de conférences, Université de Picardie Jules Verne

**Dr LORIDANT Séverine**, Praticien Hospitalier, CHRU de Lille

**Etude des mécanismes d'activation  
protéolytique de la protéine S du coronavirus  
humain 229E**



## Faculté des Sciences Pharmaceutiques et Biologiques de Lille

3, rue du Professeur Laguesse - B.P. 83 - 59006 LILLE  
CEDEX

☎ 03.20.96.40.40 - 📠 : 03.20.96.43.64



### Université Lille 2 – Droit et Santé

Président :	Professeur Xavier VANDENDRIESSCHE
Vice-présidents :	Professeur Alain DUROCHER Professeur Régis BORDET Professeur Eric BOULANGER Professeur Frédéric LOBEZ Professeur Murielle GARCIN Professeur Annabelle DERAM Professeur Muriel UBEDA SAILLARD Monsieur Ghislain CORNILLON Monsieur Pierre RAVAUX Monsieur Larbi AIT-HENNANI Madame Nathalie ETHUIN Madame Ilona LEMAITRE
Directeur Général des Services :	Monsieur Pierre-Marie ROBERT

### Faculté des Sciences Pharmaceutiques et Biologiques

Doyen :	Professeur Damien CUNY
Vice-Doyen, 1 <sup>er</sup> assesseur :	Professeur Bertrand DECAUDIN
Assesseur en charge de la pédagogie	Dr. Annie STANDAERT
Assesseur en charge de la recherche	Pr. Patricia MELNYK
Assesseur délégué à la scolarité	Dr. Christophe BOCHU
Assesseur délégué en charge des relations internationales	Pr. Philippe CHAVATTE
Assesseur délégué en charge de la vie étudiante	M. Thomas MORGENROTH
Chef des services administratifs :	Monsieur Cyrille PORTA

### Liste des Professeurs des Universités - Praticiens Hospitaliers

Civ.	NOM	Prénom	Laboratoire
Mme	ALLORGE	Delphine	Toxicologie
M.	BROUSSEAU	Thierry	Biochimie
M.	DECAUDIN	Bertrand	Pharmacie Galénique
M.	DEPREUX	Patrick	ICPAL
M.	DINE	Thierry	Pharmacie clinique
Mme	DUPONT-PRADO	Annabelle	Hématologie
M.	GRESSIER	Bernard	Pharmacologie
M.	LUYCKX	Michel	Pharmacie clinique
M.	ODOU	Pascal	Pharmacie Galénique
Mme	RENNEVILLE	Aline	Hématologie
M.	STAELS	Bart	Biologie Cellulaire

### Liste des Professeurs des Universités

Civ.	NOM	Prénom	Laboratoire
M.	ALIOUAT	El Moukhtar	Parasitologie
Mme	AZAROUAL	Nathalie	Physique
M.	BERTHELOT	Pascal	Onco et Neurochimie
M.	CAZIN	Jean-Louis	Pharmacologie – Pharmacie clinique
M.	CHAVATTE	Philippe	ICPAL
M.	COURTECUISSÉ	Régis	Sciences végétales et fongiques
M.	CUNY	Damien	Sciences végétales et fongiques
Mme	DELBAERE	Stéphanie	Physique
M.	DEPREZ	Benoît	Laboratoire de Médicaments et Molécules
Mme	DEPREZ	Rebecca	Laboratoire de Médicaments et Molécules
M.	DUPONT	Frédéric	Sciences végétales et fongiques
M.	DURIEZ	Patrick	Physiologie
M.	FOLIGNE	Benoît	Bactériologie
M.	GARÇON	Guillaume	Toxicologie
Mme	GAYOT	Anne	Pharmacotechnie Industrielle
M.	GOOSSENS	Jean François	Chimie Analytique
M.	HENNEBELLE	Thierry	Pharmacognosie
M.	LEMDANI	Mohamed	Biomathématiques
Mme	LESTAVEL	Sophie	Biologie Cellulaire
M.	LUC	Gerald	Physiologie
Mme	MELNYK	Patricia	Onco et Neurochimie
M.	MILLET	Régis	ICPAL
Mme	MUHR – TAILLEUX	Anne	Biochimie
Mme	PAUMELLE-LESTRELIN	Réjane	Biologie Cellulaire
Mme	PERROY	Anne Catherine	Législation
Mme	ROMOND	Marie Bénédicte	Bactériologie
Mme	SAHPAZ	Sevser	Pharmacognosie
M.	SERGHERAERT	Eric	Législation
Mme	SIEPMANN	Florence	Pharmacotechnie Industrielle
M.	SIEPMANN	Juergen	Pharmacotechnie Industrielle
M	TARTAR	André	Laboratoire de Médicaments et Molécules
M.	WILLAND	Nicolas	Laboratoire de Médicaments et Molécules

### Liste des Maîtres de Conférences - Praticiens Hospitaliers

Civ.	NOM	Prénom	Laboratoire
Mme	BALDUYCK	Malika	Biochimie
Mme	GARAT	Anne	Toxicologie
Mme	GOFFARD	Anne	Bactériologie
M.	LANNOY	Damien	Pharmacie Galénique
Mme	ODOU	Marie Françoise	Bactériologie

M.	SIMON	Nicolas	Pharmacie Galénique
----	-------	---------	---------------------

### Liste des Maîtres de Conférences

Civ.	NOM	Prénom	Laboratoire
Mme	ALIOUAT	Cécile Marie	Parasitologie
M.	ANTHERIEU	Sébastien	Toxicologie
Mme	AUMERCIER	Pierrette	Biochimie
Mme	BANTUBUNGI	Kadiombo	Biologie cellulaire
Mme	BARTHELEMY	Christine	Pharmacie Galénique
Mme	BEHRA	Josette	Bactériologie
M	BELARBI	Karim	Pharmacologie
M.	BERTHET	Jérôme	Physique
M.	BERTIN	Benjamin	Immunologie
M.	BLANCHEMAIN	Nicolas	Pharmacotechnie industrielle
M.	BOCHU	Christophe	Physique
M.	BORDAGE	Simon	Pharmacognosie
M.	BOSC	Damien	Laboratoire de Médicaments et Molécules
M.	BRIAND	Olivier	Biochimie
Mme	CACHERA	Claude	Biochimie
M.	CARNOY	Christophe	Immunologie
Mme	CARON	Sandrine	Biologie cellulaire
Mme	CHABÉ	Magali	Parasitologie
Mme	CHARTON	Julie	Laboratoire de Médicaments et Molécules
M	CHEVALIER	Dany	Toxicologie
M.	COCHELARD	Dominique	Biomathématiques
Mme	DANEL	Cécile	Chimie Analytique
Mme	DEMANCHE	Christine	Parasitologie
Mme	DEMARQUILLY	Catherine	Biomathématiques
Mme	DUMONT	Julie	Biologie cellulaire
Mme	DUTOUT-AGOURIDAS	Laurence	Onco et Neurochimie
M.	EL BAKALI	Jamal	Onco et Neurochimie
M.	FARCE	Amaury	ICPAL
Mme	FLIPO	Marion	Laboratoire de Médicaments et Molécules
Mme	FOULON	Catherine	Chimie Analytique
M.	FURMAN	Christophe	ICPAL
M.	GELEZ	Philippe	Biomathématiques
Mme	GENAY	Stéphanie	Pharmacie Galénique
M.	GERVOIS	Philippe	Biochimie
Mme	GOOSSENS	Laurence	ICPAL
Mme	GRAVE	Béatrice	Toxicologie
Mme	GROSS	Barbara	Biochimie
M.	HAMONIER	Julien	Biomathématiques
Mme	HAMOUDI	Chérifa Mounira	Pharmacotechnie industrielle
Mme	HANNOTHIAUX	Marie-Hélène	Toxicologie
Mme	HELLEBOID	Audrey	Physiologie
M.	HERMANN	Emmanuel	Immunologie
M.	KAMBIA	Kpakpaga Nicolas	Pharmacologie

M.	KARROUT	Youness	Pharmacotechnie Industrielle
Mme	LALLOYER	Fanny	Biochimie
M.	LEBEGUE	Nicolas	Onco et Neurochimie
Mme	LECOEUR	Marie	Chimie Analytique
Mme	LEHMANN	Hélène	Législation
Mme	LELEU-CHAVAIN	Natascha	ICPAL
Mme	LIPKA	Emmanuelle	Chimie Analytique
Mme	MARTIN	Françoise	Physiologie
M.	MOREAU	Pierre Arthur	Sciences végétales et fongiques
Mme	MUSCHERT	Susanne	Pharmacotechnie industrielle
Mme	NIKASINOVIC	Lydia	Toxicologie
Mme	PINÇON	Claire	Biomathématiques
M.	PIVA	Frank	Biochimie
Mme	PLATEL	Anne	Toxicologie
M.	POURCET	Benoît	Biochimie
M.	RAVAUX	Pierre	Biomathématiques
Mme	RAVEZ	Séverine	Onco et Neurochimie
Mme	RIVIERE	Céline	Pharmacognosie
Mme	ROGER	Nadine	Immunologie
M.	ROUMY	Vincent	Pharmacognosie
Mme	SEBTI	Yasmine	Biochimie
Mme	SINGER	Elisabeth	Bactériologie
Mme	STANDAERT	Annie	Parasitologie (80%)
M.	TAGZIRT	Madjid	Hématologie
M.	VILLEMAGNE	Baptiste	Laboratoire de Médicaments et Molécules
M.	WELTI	Stéphane	Sciences végétales et fongiques
M.	YOUS	Saïd	Onco et Neurochimie
M.	ZITOUNI	Djamel	Biomathématiques

### Professeurs Agrégés

Civ.	NOM	Prénom	Laboratoire
Mme	MAYES	Martine	Anglais
M.	MORGENROTH	Thomas	Législation

### Professeurs Certifiés

Civ.	NOM	Prénom	Laboratoire
M.	HUGES	Dominique	Anglais
Mlle	FAUQUANT	Soline	Anglais
M.	OSTYN	Gaël	Anglais

### Professeur Associé - mi-temps

Civ.	NOM	Prénom	Laboratoire
M.	DHANANI	Alban	Droit et Economie Pharmaceutique

### Maîtres de Conférences ASSOCIES - mi-temps

Civ.	NOM	Prénom	Laboratoire
M.	BRICOTEAU	Didier	Biomathématiques
Mme	CUCCHI	Malgorzata	Biomathématiques

M.	FRIMAT	Bruno	Pharmacie Clinique
M.	GILLOT	François	Droit et Economie pharmaceutique
M.	MASCAUT	Daniel	Pharmacie Clinique
M.	ZANETTI	Sébastien	Biomathématiques
M.	BRICOTEAU	Didier	Biomathématiques

### AHU

Civ.	NOM	Prénom	Laboratoire
Mme	DEKYNDT	Bérengère	Pharmacie Galénique
M.	PEREZ	Maxime	Pharmacie Galénique

## ***Faculté des Sciences Pharmaceutiques et Biologiques de Lille***

3, rue du Professeur Laguesse - B.P. 83 - 59006 LILLE CEDEX  
Tel. : 03.20.96.40.40 - Télécopie : 03.20.96.43.64  
<http://pharmacie.univ-lille2.fr>

**L'Université n'entend donner aucune approbation aux opinions émises dans les thèses ; celles-ci sont propres à leurs auteurs.**

## Remerciements

Je remercie le Dr Jean Dubuisson pour m'avoir accueilli dans son équipe. Ses idées (et son impressionnante mémoire) m'ont été et me sont encore très utiles. Je remercie également le Dr François Helle et le Dr Séverine Loridant d'avoir accepté de faire parti de mon jury.

Je remercie le Dr Anne Goffard, ma directrice de thèse, pour m'avoir prise comme étudiante et pour tous ses conseils. Je remercie aussi le Dr Sandrine Belouzard pour tout le temps qu'elle m'a consacré et pour sa patience (et ce n'est pas fini...).

Je remercie également toutes les personnes du labo, qui contribuent à la bonne ambiance, et particulièrement l'équipe corona, dont Anabelle, Adeline et Sandrine avec qui j'ai partagé (entre autres) des bains de boue.

Dans l'ensemble (et chacun se reconnaitra !), je remercie les mange-tôt, les mange-tard, les mamans du labo, ceux qui viennent boire une bière, ceux qui font des gâteaux, ceux qui te font du café, ceux qui aiment discuter, ceux qui sursautent facilement, ceux qui « décorent » les bureaux, et ceux qui se font « décorer » leur bureau, ceux qui dansent en chambre froide, ceux qui font des blagues pourries, ceux des labo d'à côté et ceux qui te parlent anglais.

Je tiens également à remercier ma famille et mes amis et particulièrement mes parents, Sophie, Emmanuel et Claire, qui me soutiennent quoi qu'il arrive et qui m'encouragent.

Et évidemment, je remercie tout ceux qui vont continuer à me supporter jusqu'à ma thèse de science.

# Sommaire

Abréviations .....	10
I. Introduction.....	11
A. Généralités.....	11
1. Découverte des coronavirus humains.....	11
2. Epidémiologie .....	12
3. Clinique .....	16
4. Diagnostic.....	19
5. Traitement et prévention .....	20
6. Structure de la particule virale .....	21
7. Cycle viral .....	23
B. Protéine S.....	27
1. Domaine S1 : Attachement et reconnaissance antigénique.....	28
2. Domaine S2 : Fusion virale.....	31
C. N-glycosylation .....	36
1. Généralités et mécanisme.....	36
2. Rôle de la N-glycosylation.....	38
D. Objectifs .....	40
II. Matériel et méthodes.....	42
A. Cellules.....	42
B. Anticorps .....	42
C. Mutagénèse dirigée.....	42
D. Séquences et alignements.....	43
E. Transfection .....	43
F. Clivage et biotinylation.....	44
G. Analyse par western-blot.....	44
H. Production de pseudo-particules et tests d'infection.....	45
I. Test d'infection en surface.....	46
III. Résultats.....	47
A. Conservation des sites de clivage de la protéine S.....	47
B. Entrée virale par la voie endosomale.....	48
C. Infection en surface médiée par la trypsine.....	50
D. Clivage de la protéine S du HCoV-229E par la trypsine .....	52
IV. Discussion.....	55
V. Bibliographie.....	60
VI. Annexes.....	70

## Abréviations

<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>ERGIC</b>	<i>ER-Golgi intermediate compartment</i>
<b>FP</b>	<i>Fusion peptide</i>
<b>HAT</b>	<i>Human airway trypsin</i>
<b>HCoV-229E</b>	<i>Human Coronavirus 229E</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>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>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>	Réticulum endoplasmique
<b>RTC</b>	Complexe de réplicase - transcription
<b>SP</b>	<i>Signal peptide</i>
<b>TMPRSS2</b>	<i>Transmembrane protease, serine subfamily member 2</i>
<b>TM</b>	<i>Transmembrane domain</i>

# I. Introduction

## A. Généralités

### 1. Découverte des coronavirus humains

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 *alphacoronavirus* et *betacoronavirus* circulent parmi les mammifères, les *gammacoronavirus* parmi les oiseaux et les *deltacoronavirus* infectent les oiseaux et les mammifères (Masters, 2006). Les premiers coronavirus humains 229E (HCoV-229E) et OC43 (HCoV-OC43) furent isolés dans les années 1960 et sont classés parmi les *alphacoronavirus*. En particulier, HCoV-229E a été découvert au Royaume-Uni dans des lavages nasals de personnes ayant un rhume (Hamre and Procknow, 1966). 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 pour *Severe Acute Respiratory Syndrome Coronavirus*. 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 nouveau 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 avec un nombre total de 2040 cas recensés et un taux de mortalité d'environ 35% (WHO, 2017a).

## **2. Epidémiologie**

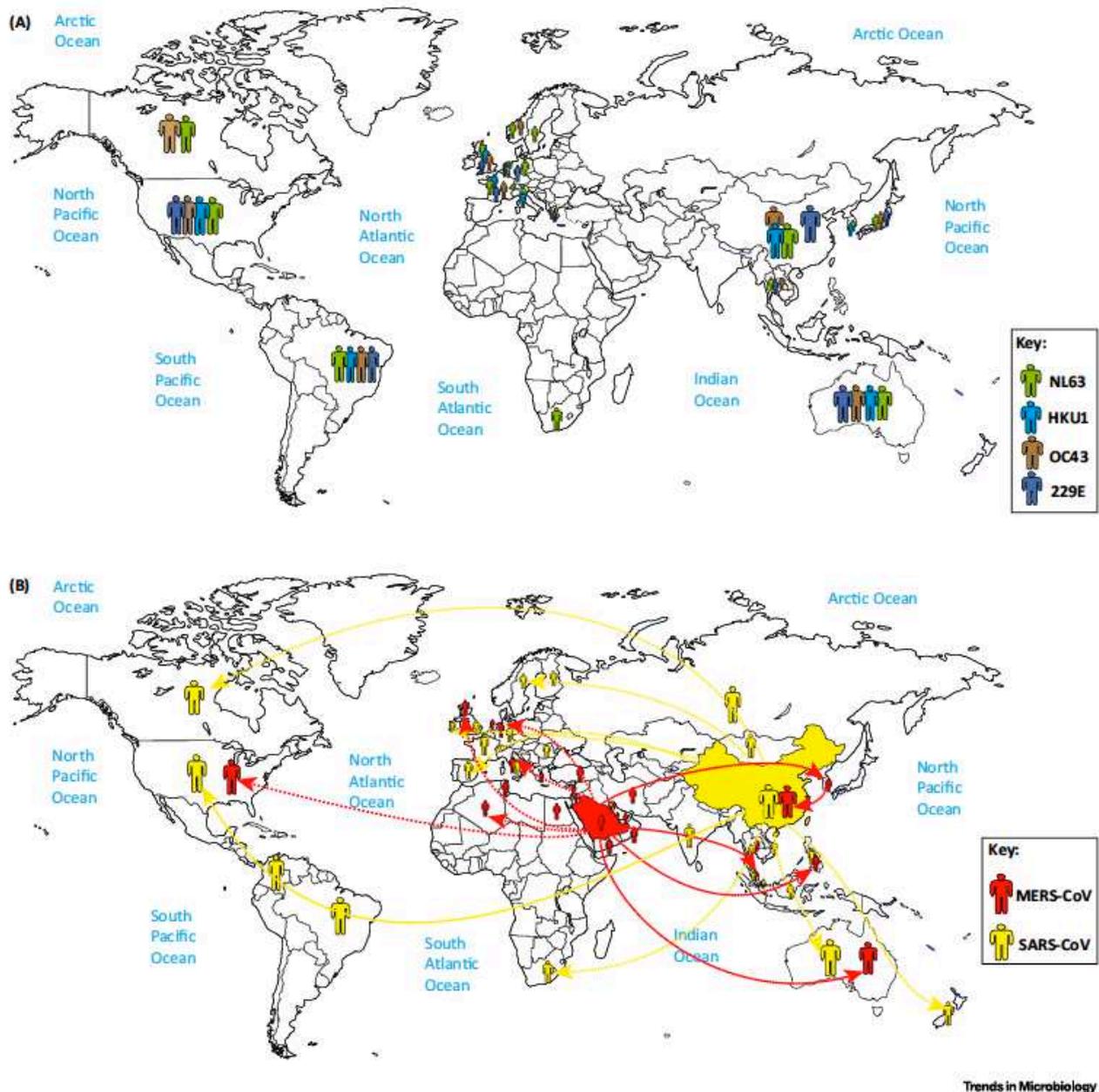
### **Distribution géographique des coronavirus**

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 USA 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 répandu à d'autres pays en Asie, en Amérique du Nord et en Europe (37 pays/régions au total) pendant 9 mois (WHO, 2017b). 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, des cas ont parfois été exportés vers d'autres pays, provoquant des épidémies secondaires (figure 1B) (de Wit et al., 2016).

### **Réservoir et origine des coronavirus**

Les coronavirus infectent une grande diversité d'hôtes, incluant de nombreuses espèces d'oiseaux et de mammifères, dont l'Homme (Monto, 1974). 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. HCoV-229E a cependant été identifié chez le dromadaire (Corman et al., 2016).

Il semble que les 4 genres de coronavirus aient divergé d'un ancêtre commun il y a 300 millions d'années puis ont co-évolué avec leurs hôtes. 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).



**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

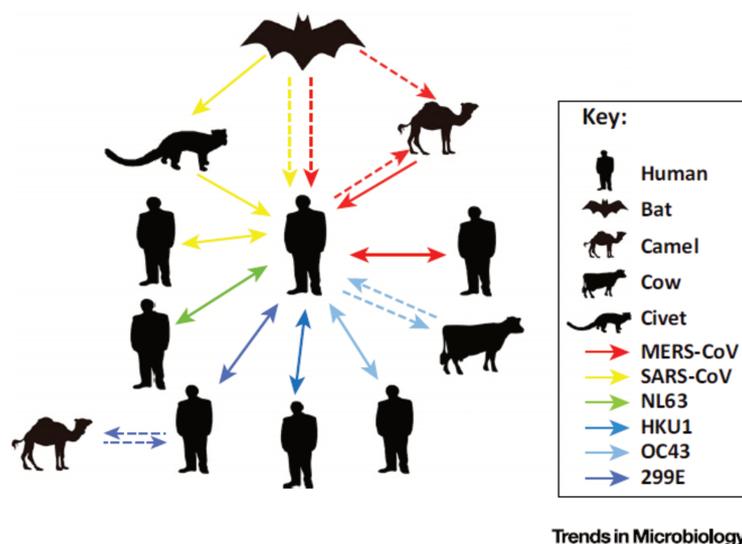
rongeurs (Lau et al., 2015) tandis que les *delta*- et *gammacoronavirus* auraient pour origine les oiseaux (Woo et al., 2012).

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 (Guan et al., 2003 ; Sharif-Yakan and Kanj, 2014). L'ancêtre du SARS-CoV se serait d'abord répandu chez les chauves-souris, puis chez la civette palmée (*Paguma larvata*) et finalement chez les humains. Ce résultat a été confirmé par une étude récente rapportant que les chauve-souris *Horseshoe* chinoises sont le réservoir naturel du SARS-CoV (Ge et al., 2013 ; Poon et al., 2005). Les hôtes intermédiaires ne semblent cependant pas nécessaires pour l'infection des humains (Wu et al., 2016).

Le MERS-CoV vient aussi probablement de coronavirus de chauve-souris (Annan et al., 2013 ; Lu et al., 2013 ; Memish et al., 2013). En effet, après séquençage du génome du virus, 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 Ac neutralisants dirigés contre le MERS-CoV (Reusken et al., 2013). Des Ac 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 du virus des 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. Des études ont montré une transmission interhumaine du MERS-CoV, particulièrement en milieu hospitalier (Guery et al., 2013; Hui et al., 2015; WHO, 2016). Le nombre de cas 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).

De manière générale, il est probable que la plupart des coronavirus humains viennent des chauves-souris (Drexler et al., 2014). 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 (Sabir et al., 2015; Su et al., 2016).



**Figure 2 : Transmission intra- et inter-espèces des coronavirus humains.** Les flèches rouges, jaune, vert, bleu, marron 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 chameaux et 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 issue de Su et al, 2016).

## **Transmission**

Les coronavirus se transmettent par aérosols 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 excrétés à partir du même côté, facilitant la dissémination du virus par la toux et les éternuements (Jonsdottir and Dijkman, 2016). Les coronavirus humains faiblement pathogènes 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).

## **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 (Vabret et al., 2009). 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).

## **3. 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 pneumonie 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. Globalement, il semble que l'infection se produise sans distinction de sexe, d'âge ou de lieu (Dijkman and van der Hoek, 2009). 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; Garbino et al., 2006; Gorse et al., 2009). 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'immunodéficience sont plus vulnérables à l'infection par le MERS-CoV : 75% des patients infectés avaient 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).

## 4. 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 et 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).

## 5. Traitement et 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* chez les primates non humains infectés par le MERS-CoV et dans les essais non randomisés de patients atteint par le SARS-CoV (Zumla et al., 2016).

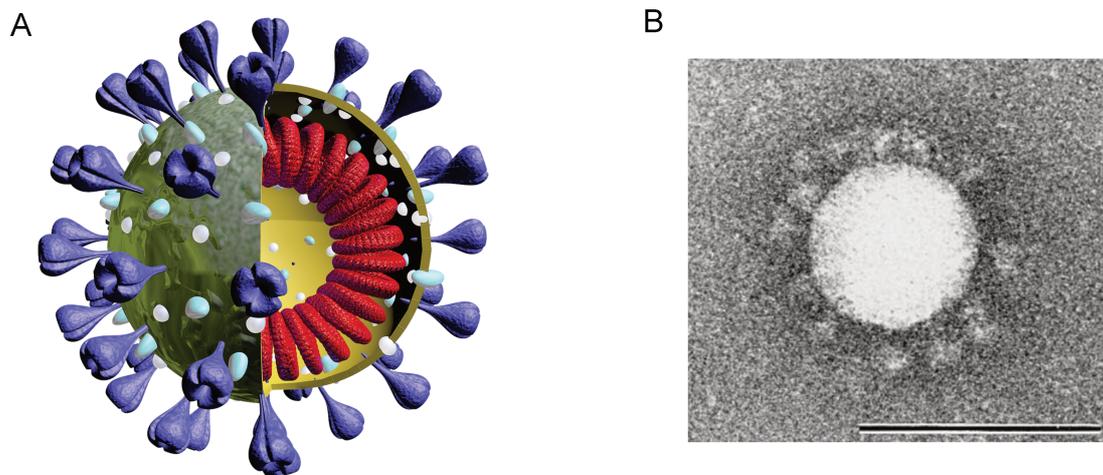
La prévention de l'infection par les coronavirus est importante puisqu'aucun traitement ni vaccin n'est commercialisé. Il a été montré que l'infection par les coronavirus dont HCoV-229E dans l'enfance n'induit pas d'immunité protectrice (Callow et al., 1990). La réinfection a lieu quand le niveau d'anticorps dirigés contre le virus est faible. Il a été montré que les Ac anti-HCoV-229E sont détectés jusqu'à un an après l'infection, indiquant que la protection conférée par ces Ac 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, bien que la dissémination virale aérienne et la transmission indirecte soient peu fréquentes. 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).

## **6. Structure de la particule virale**

Les coronavirus sont des virus sphériques enveloppés d'un diamètre de 80 à 120 nm (Masters, 2006). La capsid virale contenue dans l'enveloppe est de symétrie hélicoïdale. A la surface des particules sont enchâssées entre autres des protéines de surface nommées *spike*, ce qui donne cet aspect de couronne en microscopie électronique et le nom de cette famille virale (Figure 3).

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'. Plusieurs cadres ouverts de lecture sont présents. Deux cadres ouverts de lecture ORF1a et ORF1b (*Open Reading Frame*) chevauchant occupent les deux premiers tiers du génome. 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 (*spike*), M (membrane), N (nucléocapside) et E (enveloppe) et contient le gène codant l'unique protéine accessoire ORF4a (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'(Fehr and Perlman, 2015).



**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 TGEV (*Transmissible gastroenteritis coronavirus*) infectieux purifiés. Coloration négative avec 2% sodium phosphotungstate. Barre noire = 100 nm (Photo de Risco et al, 1995).

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* porcine proche du HCoV-229E) semble être phosphorylée (Calvo et al., 2005), bien que le rôle de cette phosphorylation soit encore mal connu. 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. 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. 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. La protéine de membrane (M) a un rôle essentiel dans l'assemblage des virions. Elle est la protéine structurale la plus abondante dans l'enveloppe virale. Elle a un poids moléculaire de 25 à 30 kDa, et possède 3 domaines transmembranaires. Elle est N-glycosylée à son extrémité N-terminale (Masters, 2006). 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. C'est une protéine transmembranaire dont la topologie est encore débattue. 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).

## 7. Cycle viral

### a) 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 4). La fusion d'un virus enveloppé peut avoir lieu à la surface cellulaire ou dans des endosomes après internalisation de la particule virale. 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.

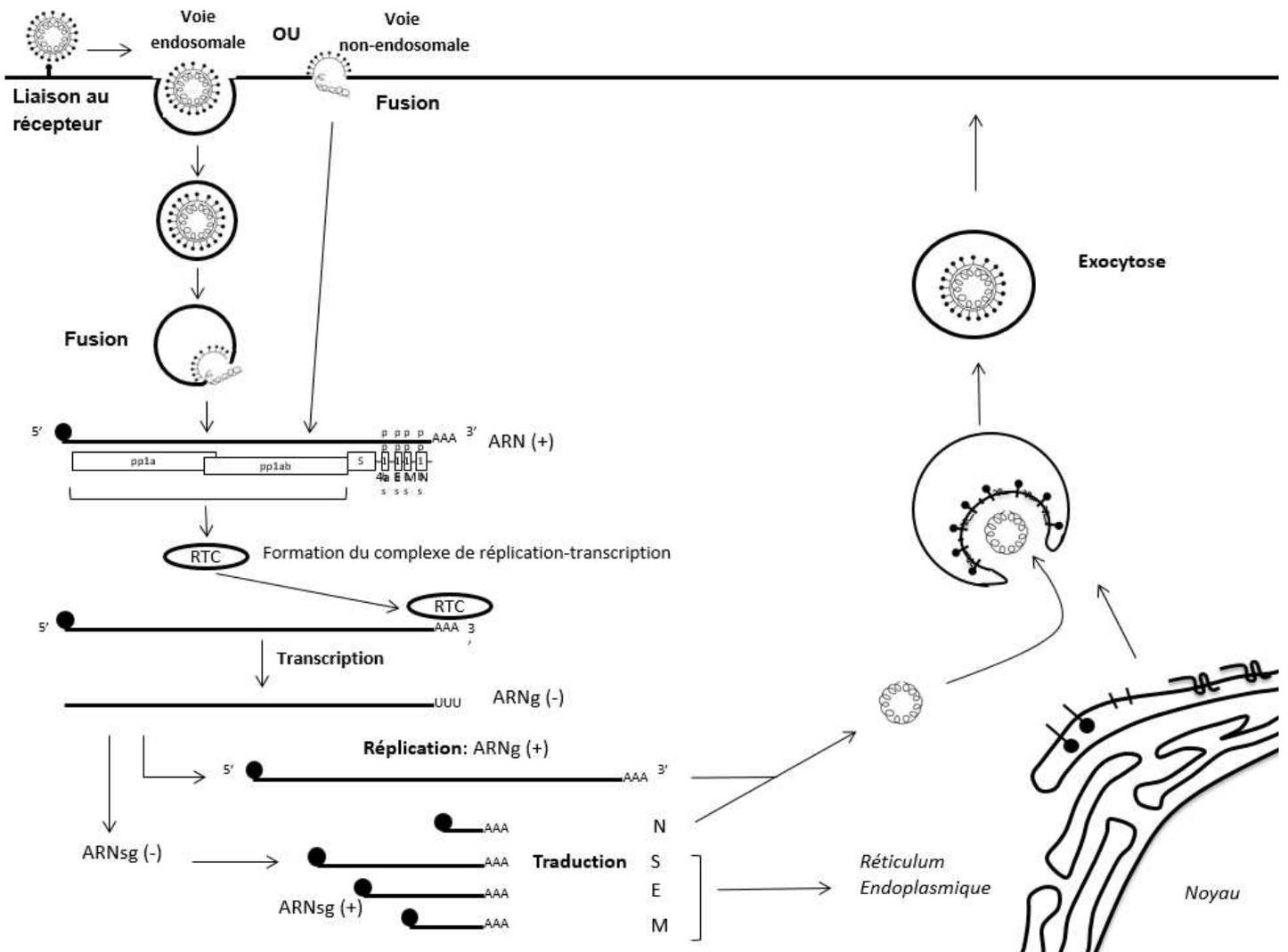
Dans certains cas, la liaison du coronavirus à son récepteur peut suffire à induire le changement de conformation de la protéine S nécessaire à la fusion, comme c'est le cas pour certaines souches du *betacoronavirus Murine Hepatitis Virus* (MHV). Dans d'autres cas, l'activation de la fusion du coronavirus est dépendante du pH acide de certains endosomes de la cellule, ce qui implique une entrée par voie endosomale, comme pour le *gammacoronavirus Infectious Bronchitis Virus* (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 comme TMPRSS2 (*Transmembrane protease, serine subfamily member 2*) ou des protéases endosomales telles que les cathepsines L ou B. Le site de fusion lors de l'entrée virale par voie endosomale dépend probablement en grande partie de la séquence en amont du peptide de fusion et 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 (Kawase et al., 2009) mais 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 surface cellulaire (sérine protéases de type 2 comme TMPRSS2, HAT) (Bertram et al., 2013). Il semble que cette voie soit plus avantageuse pour le virus que la voie endosomale (Shirato et al., 2016).

## **b) 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 chevauchant ORF1a et ORF1b, qui sont traduits en deux polyprotéines pp1a et pp1ab par la machinerie cellulaire (Figure 4). 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. Les deux polyprotéines pp1a et pp1ab subissent ensuite une maturation protéolytique permettant la production de 16 protéines non structurales (nsp). Deux protéases virales sont responsables de cette maturation. La première protéine virale nsp3 contient des domaines protéases homologues à la papaine (*papaine-like protease* PLpro). La deuxième protéine virale nsp5 est une protéase de type sérine qui appartient à la famille de la chymotrypsine (*3C-like proteinase* 3CLpro ou *Main protease* Mpro) (Fehr and Perlman, 2015).



**Figure 4: Schéma général du cycle viral des coronavirus.** Le virus entre dans la cellule cible après reconnaissance du récepteur cellulaire. La fusion entre les membranes virales 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 (nsp) 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 à partir du ERGIC (Endoplasmic Reticulum-Golgi Intermediate Compartment).

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*). D'autres protéines sont importantes pour la réplication. La protéine nsp13, qui a une activité ARN hélicase, est nécessaire pour le déroulement de substrats d'ARN à partir de leur extrémité 5'. La protéine nsp14, qui a une activité d'exoribonuclease (ExoN), 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 (Denison et al., 2011).

Après l'assemblage du RTC, la synthèse d'un ARN viral (-) intermédiaire permet la production d'ARN génomique de pleine longueur, ce mécanisme étant nommé réplication. D'autre part, l'ARN viral (-) intermédiaire permet la production d'ARNs subgénomique (ARNsg) par un mécanisme de synthèse discontinue (Figure 4). Ces différents ARN subgénomiques codent la protéine accessoire ORF4a et les protéines structurales (Fehr and Perlman, 2015; Subissi et al., 2012).

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 au cours de la transcription. La recombinaison joue un rôle majeur dans l'évolution virale (Fehr and Perlman, 2015).

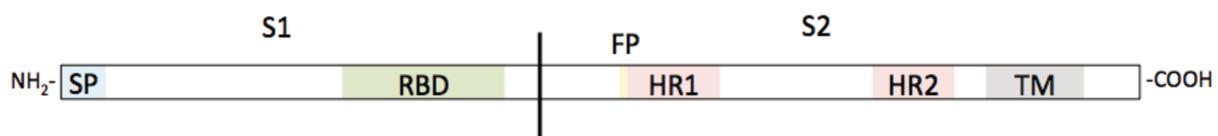
### **c) Assemblage et exocytose**

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 4) 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, formant des vésicules. Les particules virales 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<sub>3</sub>/16M/4N à 1S<sub>3</sub>/25M/4N (où S<sub>3</sub> correspond à des trimères de protéine S) (Neuman et al., 2006).

## B. Protéine S

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 la production du virion 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).



**Figure 5 : Schéma fonctionnel 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.

Sur le plan fonctionnel, la protéine S du HCoV-229E est divisée en deux domaines S1 et S2 (Figure 5). 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'autre 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. Domaine S1 : Attachement et reconnaissance antigénique**

### **a) 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, les structures des protéines S des virus MHV, HKU1 et HCoV-NL63 ont été déterminées en cryo-microscopie électronique (Kirchdoerfer et al., 2016; Ou et al., 2017; Walls et al., 2016a, 2016b). Du point de vue structural, le domaine S1 est constitué de 4 domaines A, B, C et D. Le virus HCoV-NL63 contient en plus un domaine 0 en amont du domaine A à l'extrémité N-terminale de la protéine. Ce domaine est présent chez de nombreux *alphacoronavirus*. Du point de vue phylogénétique, les virus HCoV-229E et HCoV-NL63 sont très proches, ainsi les protéines S partagent 56% d'identité. Cependant le domaine 0 est absent de la protéine S du virus HCoV-229E. Le domaine S1 contient le domaine de liaison du récepteur RBD (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 pour le *betacoronavirus* murin MHV (Murine Hepatitis Virus), ou dans le domaine C-terminal pour HCoV-229E, HCoV-NL63 ou le SARS-CoV. 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. Il semble que le RBD soit situé entre les AA 407 à 547 (Bonavia et al., 2003; Breslin et al., 2003). Il a été montré que les AA 278 à 329 seraient également nécessaires pour l'interaction avec le

récepteur, mais ces résultats doivent encore être confirmés (Hofmann et al., 2006). 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 groupe (Li, 2016). 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  $\beta$ . 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., 2013).

## **b) 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 du tissu. Le SARS-CoV et HCoV-NL63 utilisent l'enzyme de conversion de l'angiotensine 2 (angiotensin-converting enzyme 2 ACE2) pour entrer dans les cellules hôtes (Hofmann et al., 2005; Li et al., 2003), MHV entre à l'aide de CEACAM1 (Carcinoembryonic antigen-related cell adhesion molecule 1) (Williams et al., 1991), et le MERS-CoV se lie à la dipeptidyl peptidase-4 (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 (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). Certains alphacoronavirus sont capables de se lier à la partie d'APN présente à l'extérieur des cellules, sans que les fonctions physiologiques du récepteur ne soient affectées (Chen et al., 2012). 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 spécificité d'espèce de HCoV-229E pour l'APN humaine a été étudiée. Des chimères d'APN de différentes espèces ont permis de déterminer que les acides aminés 288 à 295 de l'APN humain (hAPN) sont nécessaires pour l'entrée virale du HCoV-229E (Kolb et al., 1997). L'introduction dans hAPN d'un séquençon encodant un site potentiel de N-glycosylation au niveau de l'AA 291, comme celui présent sur l'APN porcine, ou entre les AA 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). La raison de l'utilisation de peptidases comme récepteur cellulaire par la plupart des coronavirus n'est pas connue. En effet, le domaine enzymatique de ces peptidases n'est pas impliqué lors de l'entrée virale, comme il a été décrit pour le TGEV (Delmas et al., 1994).

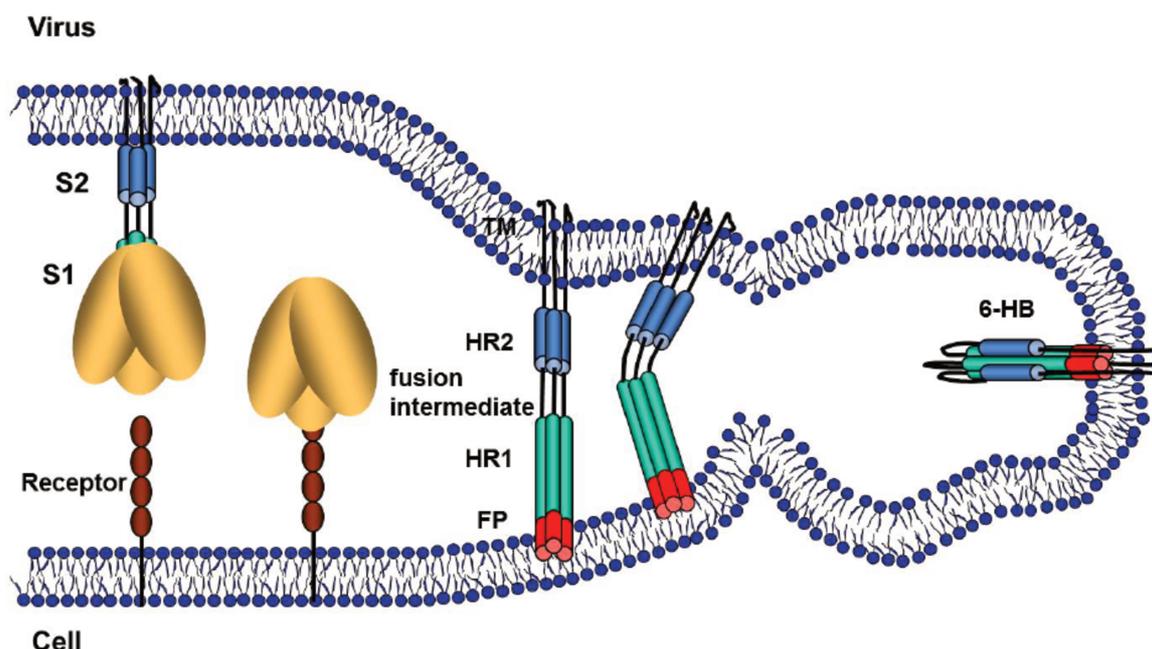
### **c) Reconnaissance antigénique**

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; 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 TGEV et le PEDV que le RBD situé dans le domaine S1 serait la cible la plus importante pour les Ac neutralisants (Du et al., 2009; Reguera et al., 2012). De plus, des variations dans le domaine S1 de la protéine S tel que la N-glycosylation pourraient influencer l'antigénicité (Shih et al., 2006). Pour le moment, le domaine S1 du HCoV-229E a été déterminé comme important lors de la neutralisation par les anticorps, mais les régions précisément impliquées n'ont pas encore été élucidées (Shirato et al., 2012).

## 2. Domaine S2 : Fusion virale

### a) 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).



**Figure 6 : 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. Il y a alors dissociation des domaines S1 et S2 de la protéine S. Le peptide de fusion (FP) exposé va alors s'insérer dans la membrane de la cellule cible, la protéine S est alors dans un état de fusion intermédiaire. Puis un nouveau changement conformationnel rassemblant les *heptad repeat* 1 et 2 (HR1 et HR2) de la protéine S rapproche les membranes virales et cellulaires, la structure finale de la protéine étant nommée *6-helix coiled-coil bundle* (6-HB) (schéma issu de Heald-Sargent et 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 6). Le processus de fusion commence par la liaison au récepteur. D'autres événements peuvent être nécessaires à la fusion, comme un clivage protéolytique par les 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 vers le peptide de fusion, ce qui par conséquent 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 6). 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 formant un trimère interagissent pour former une structure de type *coiled-coil*, tandis que 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).

## **b) 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., 2017; 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. 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, ce 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). Le SARS-CoV peut également être clivé par d'autres protéases comme l'élastase qui permet un clivage quelques acides aminés en amont du site S2' du SARS-CoV, 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 meilleurs candidats appartiennent à la famille des TTSPs (pour *type II transmembrane serine protease*). Parmi ses membres, deux protéines sont capables de cliver la protéine S du SARS-CoV, TMPRSS2 (pour *transmembrane protease/serine subfamily member 2*) et HAT (pour *human airway trypsin* 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'a pas été déterminée. 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 (Belouzard et al., 2012). A ce jour, les processus protéolytiques qui permettent à la cathepsine L d'induire la fusion restent inconnus.

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. 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. 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). Le MERS-CoV est aussi capable de

fusionner à la surface cellulaire en présence de TMPRSS2 (Gierer et al., 2013; Shirato et al., 2013).

Il existe également quelques études sur l'utilisation des protéases cellulaires par HCoV-229E lors de l'entrée virale. 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 capable 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 la cathepsine ne sont pas connus, mais cette étude a montré que les mutations de l'arginine 642 de la souche de laboratoire en méthionine (R642M) et de l'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 au site de clivage (Shirato et al., 2016).

## **C. N-glycosylation**

### **1. Généralités et mécanisme**

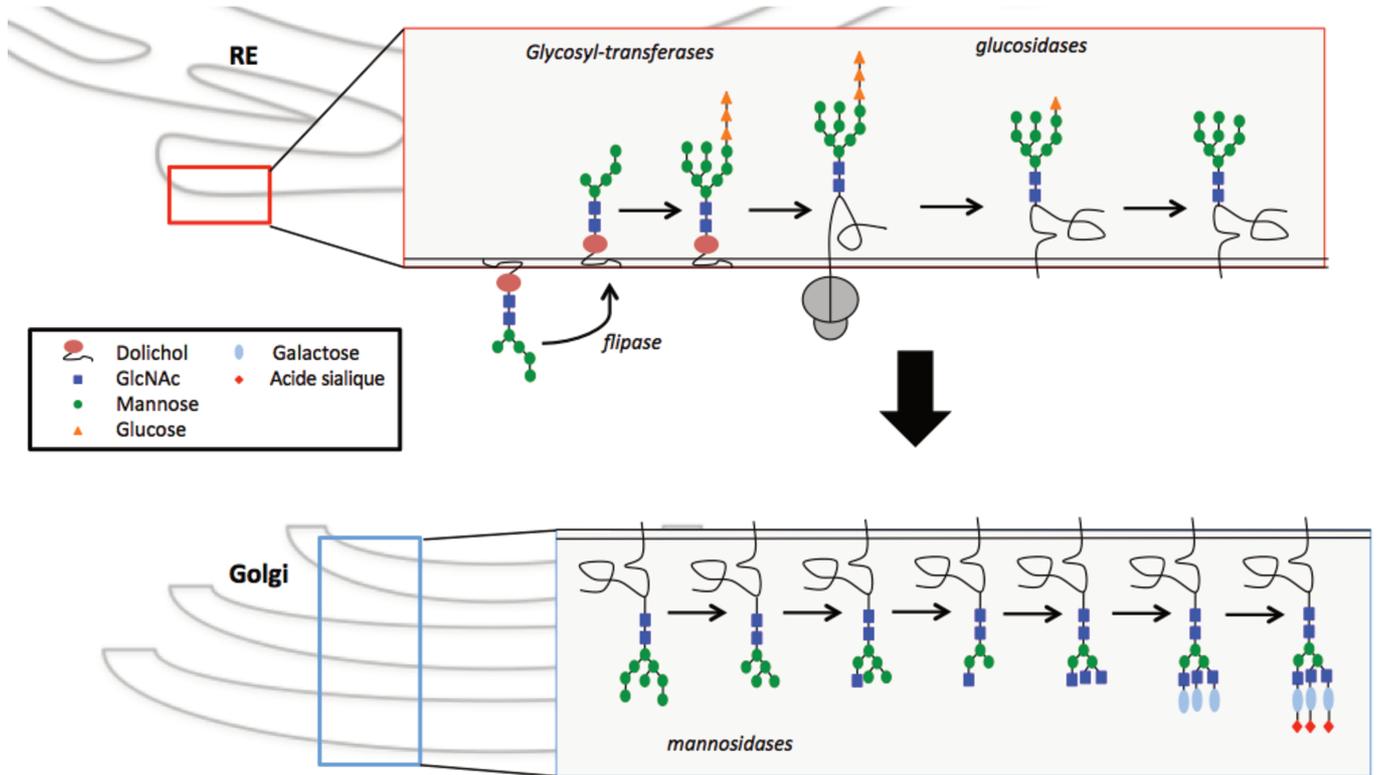
La glycosylation a lieu chez tous les Eucaryotes, ainsi que chez les Archaea et les bactéries (Calo et al., 2010; Nothaft and Szymanski, 2010). Les glycanes, ou chaînes oligosaccharidiques, sont des oses liés de façon covalente à des protéines et qui peuvent avoir des structures variables (Ohtsubo and Marth, 2006). Les protéines peuvent être O- ou N-glycosylées, ou former des glycosaminoglycanes, aussi nommés protéoglycanes, selon le site d'ancrage du 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 7A). L'attachement du glycane sur la protéine nécessite la reconnaissance d'une séquence consensus, dans laquelle 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).

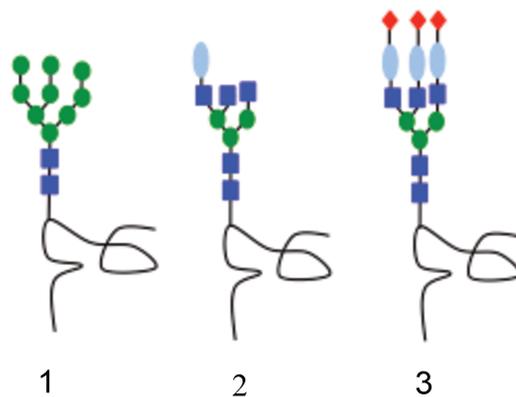
La N-glycosylation débute dans le réticulum endoplasmique par le transfert en bloc d'un précurseur oligosaccharidique préformé (Helenius and Aebi, 2001; Suzuki et al., 1995). Une chaîne oligosaccharidique précurseur est initialement formée dans le cytoplasme. Les 7 premiers oses de cette chaîne sont 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. 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. Elle est d'abord modifiée par des glucosidases qui clivent les 3 glucoses présents sur

la chaîne oligosaccharidique dans le RE puis par une mannosidase qui clive un seul mannose spécifique avant que celle-ci ne transite vers l'appareil de Golgi (Figure 7A).

A



B



**Figure 7** : Biosynthèse des N-glycanes. A. La N-glycosylation est initiée dans le RE. Une chaîne oligosaccharidique précurseur est formée puis transféré sur la glycoprotéine. Le glycan est ensuite mûri dans le RE puis dans l'appareil de Golgi. B. Les différents types de structure du glycan. 1 : Structure riche en mannose, 2 : structure hybride, 3 : structure complexe

Avant l'entrée dans l'appareil de Golgi, toutes les protéines N-glycosylées portent des 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 glycanes.

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 7B). Pour les N-glycanes riche en mannose, seul des mannoses peuvent être clivés mais aucun ose n'est ajouté à la chaîne oligosaccharidique. Les deux derniers types de 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 glycanes par l'endoglycosidase H (endo 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 glycanes après l'action de la mannosidase II, déterminant ainsi la structure du glycanes complexe. La diversité des glycanes formés est très importante, mais rend la détermination de la structure des glycoprotéines difficile.

## **2. Rôle de la N-glycosylation**

### **a) Dans les cellules**

Les 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 glycanes peuvent participer au repliement des glycoprotéines qui les portent, permettant l'acquisition de leur structure tridimensionnelle. Ce processus de contrôle qualité fait intervenir les protéines chaperonnes calnexine et calréticuline qui aident les protéines synthétisées à se replier. Une fois les protéines formées et repliées, les 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 glycanes portés sur la glycoprotéine par 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 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).

## **b) N-glycosylation des protéines virales**

La 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 du virus (Gagneux and Varki, 1999). A l'inverse, les virus possèdent des protéines glycosylées grâce à la machinerie cellulaire, ce qui leur permet de profiter de plusieurs avantages. Le repliement correct des protéines virales et le trafic intracellulaire sont permis par l'utilisation des protéines chaperonnes. Il semble que tous les virus étudiés jusqu'à présent utilisent le cycle calnexine/calréticuline (Larkin and Imperiali, 2011; Vigerust and Shepherd, 2007). Il a été montré pour le virus de l'hépatite C que les N-glycanes présents sur les glycoprotéines d'enveloppe E1 et E2 sont impliqués dans le repliement de la protéine (Goffard et al., 2005).

La N-glycosylation des protéines de surface virales joue un rôle dans les interactions avec le récepteur cellulaire ainsi que dans l'infectivité du virion, comme il a été montré pour les glycoprotéines d'enveloppe du virus de l'immunodéficience humaine (VIH), gp120 et gp160 (Wang et al. 2013; Mathys and Balzarini 2015; Raska et al. 2014), ainsi que pour le virus de l'hépatite C (Ren et al, 2015) et le

SARS-CoV (Han, Lohani, and Cho 2007; Zhou et al. 2010; Shih et al. 2006). De plus, il a également été montré pour le VIH que les N-glycanes des glycoprotéines d'enveloppe sont importantes pour la neutralisation par les anticorps (Ac) au cours de l'infection (Wang et al. 2013; Raska et al. 2014). Le rôle de la glycosylation dans la modulation de l'activation de protéines virales de surface par clivage protéolytique n'a que peu été étudié. Il a cependant été montré la présence d'un N-glycane à proximité d'un site de clivage empêchant l'accès des protéases à la protéine virale HA du virus influenza (Kawaoka et al, 1984 ; Deshpande et al, 1987). Peu de données sont actuellement disponibles sur le rôle des N-glycanes dans les fonctions de la protéine S de HCoV-229E.

## **D. 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. La protéine S du virus n'est pas clivée à sa surface comme celle du SARS-CoV. Pour le 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. Nous avons donc testé si la fusion du virus HCoV-229E était activée par un mécanisme similaire. D'autre part, nous avons aussi analysé le rôle du glycane 19 présent au niveau de la jonction S1/S2 dans ce clivage. L'étude spécifique de l'étape d'entrée du HCoV-229E est permise par la production de pseudo-particules de HCoV-229E (229Epp). Cet outil a été utilisé pour étudier plusieurs virus, dont le virus de l'hépatite C et les coronavirus (Bartosch et al., 2003; Corver et al., 2007). Ces résultats participeront à une meilleure connaissance

du coronavirus respiratoire humain 229E, dans le but de développer des moyens de lutter contre les coronavirus.

## II. Matériel et méthodes

### A. 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).

### B. Anticorps

Aucun anticorps monoclonal dirigé contre la protéine S du HCoV-229E n'étant actuellement disponible, toutes les protéines ont été clonées avec un tag VSVG en position C-terminale permettant leur détection. Les anticorps (Ac) monoclonaux P5D4 (anti-VSVG) et R187 (anti-capside du *murine leukemia virus* (MLV; ATCC CRL1912) ont été produits *in vitro* à l'aide d'un module de production MiniPerm (Heraeus) selon les recommandations du fabricant.

### C. Mutagénèse dirigée

La séquence de la protéine S du HCoV-229E (S-06/2014) a été clonée à partir d'un prélèvement respiratoire provenant d'un patient infecté par HCoV-229E en 2014 et séquencée (Genoscreen). La séquence ADN a été analysée avec le logiciel SerialCloner (version 2.6.1). 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-04/2014 a ensuite été synthétisée par GeneArt afin d'optimiser l'utilisation des codons et insérée dans le vecteur pHCMV entre les enzymes BglII et NheI. La séquence ADN de la protéine S a été utilisée pour construire trois types de mutants. Le mutant de N-glycosylation N19 a été construit en remplaçant le codon AAC du résidu asparagine (Asn) de la séquence consensus Asn-X-Ser/Thr par le

codon CAA du résidu glutamine (Gln). Les mutants dans lesquels un ou plusieurs résidus asparagine sont abolis ont été construits en remplaçant le codon AGA du résidu arginine (Arg) par le codon AAT du résidu asparagine. Un site de reconnaissance par la furine a été introduit en remplaçant les codons par le codon CGC du résidu arginine (Arg – R), par le codon TCG du résidu sérine (Ser - S) ou par le codon AAG du résidu lysine (Lys – K) (Tableau 1, voir en annexe). Les mutations ont été introduites dans la séquence ADN de la protéine S par deux PCR à l'aide du mélange Q5® *High-Fidelity DNA Polymerase* (New England Biolabs). Les fragments ont ensuite été assemblés par une PCR de fusion. Les séquences obtenues codant pour les mutants de glycosylation de la protéine S ont été digérées par BglII et NheI et insérées dans un plasmide pCMV digéré par BglII/NheI. Les plasmides ont été nommés selon la position du site de N-glycosylation aboli. Chaque plasmide obtenu a été vérifié par séquençage (Genoscreen).

#### ***D. Séquences et alignements***

Les alignements de séquences des sites S2' et S1/S2 de la protéine S de chaque coronavirus a été réalisé avec Clustal Omega v1.2.4 (EMBL-EBI). Les séquences utilisées et leur identifiant GenBank sont présentées dans l'annexe, tableau 2.

#### ***E. 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. L'agent de transfection, après avoir été incubé 5 min à température ambiante dans 50 µl de milieu, a été ajouté à l'ADN plasmidique dilué dans du milieu (pour 1 µg d'ADN, 50 µl de milieu) et incubé 20 min à température ambiante. Le mélange a ensuite été ajouté sur les cellules dans du milieu complet. La quantité d'ADN transfectée est d'1 µg par puits pour les plaques 6 puits.

## **F. Clivage et biotinylation**

Des cellules HEK293T ont été réparties en plaques de 6 puits recouvertes de poly-D-lysine, et transfectées avec 1 µg par puits de plasmide d'intérêt. A 48h post-transfection, les cellules ont été lavées 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 Laëmmli (pour une concentration finale de 1 % SDS, 10 % glycérol, 3 mM de Tris-HCl pH 6,8, 5 % 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.

## **G. Analyse par 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 (Ac) 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.

## ***H. Production de pseudo-particules et tests d'infection***

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

En parallèle, les cellules ont été lysées par 250 µl de 1 % de triton X-100 dans du

PBS froid, supplémenté par des inhibiteurs de protéases (Roche) afin de vérifier l'expression des différents mutants en western blot. Pour vérifier la production de pseudoparticules dans les surnageants, elles ont été concentrées par ultracentrifugation sur un coussin à sucrose 30 % de 3h à 175 273 G à 4°C. Chaque culot a été re-suspendu avec du tampon de charge Laemmli. Des western-blot des lysats cellulaire et des pseudo-particules purifiées ont été réalisés comme décrit précédemment.

### ***I. 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 de 25 mM de  $\text{NH}_4\text{Cl}$  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 contenant 0,2% de BSA, 25 mM d'Hepes et 25mM de  $\text{NH}_4\text{Cl}$ . Les cellules ont ensuite été réchauffées par l'ajout de DMEM contenant 2  $\mu\text{g}/\text{ml}$  de trypsine, 0,2% de BSA, 25 mM d'Hepes et 25mM de  $\text{NH}_4\text{Cl}$  puis ont été incubées 5 min dans un bain-marie à 37°C. Les cellules ont été incubées pendant 30 min dans du milieu complet contenant 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 et la luminescence a été mesurée.

### III. Résultats

#### A. Conservation des sites de clivage de la protéine S

A

αCoV	HCov-NL63	734	FGICADGSLIPV <u>RP</u> <u>RN</u> SS	751
	HCov-229E-S-USA/933-40/1993	550	FGVCADGSIIAVQ <u>PR</u> <u>NV</u> S	567
	HCov-229E-S-ATCC	550	FGVCADGSIIAVQ <u>PR</u> <u>NV</u> S	567
	HCov-229E-S-Niigata/01/08	550	FGVCADGSIIAVQ <u>PR</u> <u>NV</u> S	567
	HCov-229E-S-Sendai-H/1948/04	550	FGVCADGSIIAVQ <u>PR</u> <u>NV</u> L	567
	HCov-229E-S-04/2014	550	FGVCADGSIIAVQ <u>PR</u> <u>NV</u> S	567
	MERS-CoV	737	ALPDPSTLT <u>PR</u> <u>SV</u> RSVP	754
	SARS-CoV	658	ASYHTVSL <u>LR</u> STSQKSIV	675
	HCov-HKU1	746	DYN <u>SPSSSS</u> <u>RR</u> <u>KR</u> RSIS	763
	HCov-OC43	749	SGYCVDYSKN <u>RR</u> <u>RR</u> RAIT	766
βCoV				
αCoV	PEDV-CV777	871	FTNVLGASVYDPASG <u>RV</u> <u>VQ</u> KRSVIEDL	897
	FCov-1683	939	GSWLGGLKDILPSHNS <u>KR</u> <u>KY</u> RSAIEDL	965
	CCov-1-71	942	GSWLGGLKDILPSHNS <u>KR</u> <u>KY</u> RSAIEDL	968
	TGEV-Purdue	936	GSWLEGLKYILPSHNS <u>KR</u> <u>KY</u> RSAIEDL	962
	HCov-NL63	850	DYN <u>L</u> SSVLPQ <u>RNI</u> <u>RSS</u> RIAGRSAIEDL	876
	HCov-229E-S-USA/933-40/1993	666	DYN <u>L</u> SSVIPSL <u>PR</u> <u>SGS</u> <u>RV</u> AGRSAIEDI	692
	HCov-229E-S-ATCC	669	DYN <u>L</u> SSVIPSL <u>PT</u> <u>SGS</u> <u>RV</u> AGRSAIEDI	695
	HCov-229E-S-Niigata/01/08	666	DYN <u>L</u> SSVIPSL <u>PR</u> <u>SGS</u> <u>RV</u> AGRSAIEDI	692
	HCov-229E-S-Sendai-H/1948/04	666	DYN <u>L</u> SSVIPSL <u>PR</u> <u>SGS</u> <u>RV</u> AGRSAIEDI	692
	HCov-229E-S-04/2014	666	DYN <u>L</u> SSVIPSL <u>PR</u> <u>SGS</u> <u>RV</u> AGRSAIEDI	692
βCoV				
MERS-CoV	867	GDF <u>N</u> LTLLEPVSI <u>ST</u> GS <u>RS</u> ARSAIEDL	893	
SARS-CoV	777	KYFGGFNFSQILPD <u>LK</u> PT <u>KR</u> SFIEDL	803	
HCov-HKU1	884	DNINFKSLVGLGPHCGSS <u>RS</u> FFEDL	910	
HCov-OC43	888	DINFSPVLGCLGSECS <u>K</u> ASS <u>RS</u> AIEDL	914	

**Figure 8: Alignement de séquences en acides aminés d'alpha- et betacoronavirus.**  
**A.** Alignement de la jonction S1/S2 de coronavirus humains. **B.** Alignement du site S2' des coronavirus humains avec des alphacoronavirus animaux. 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. Inspiré de l'article de Millet et al, 2016.

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 avec un premier clivage S1/S2 suivi d'un deuxième clivage S2'. Nous avons donc analysé les résidus présents au niveau des jonctions S1/S2 et dans la région S2' de la protéine S de HCoV-229E (Figure 8). 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, HCoV-HKU1 et HCoV-OC43 dans cette région. Le virus HCoV-229E quant à lui possède un unique résidu basique (R565) comme le SARS-CoV. L'arginine 565 (Figure 8A) est immédiatement suivie d'un site potentiel de N-glycosylation (N566 – N19) très conservé parmi les différentes souches de HCoV-229E et dans HCoV-NI63. Ce site de N-glycosylation pourrait jouer un rôle dans un éventuel clivage de la jonction S1/S2.

L'alignement de la région S2' montre que le résidu arginine en position 687 sur la séquence d'HCoV-229E (souche 04/2014) est fortement conservé sur toutes les séquences animales et humaines (Figure 8B). Par ailleurs, deux arginines en position 683 et 687 sont hautement conservées sur les séquences d'HCoV-229E. Une quatrième arginine est également présente sur HCoV-NL63.

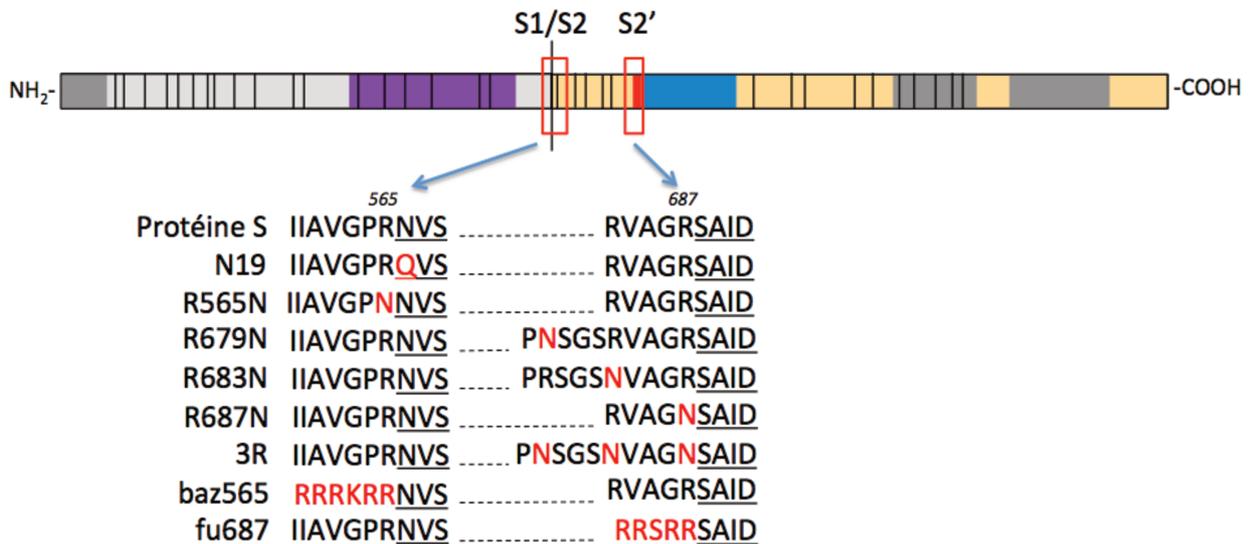
Nous avons étudié si le clivage S1/S2 de la protéine S du virus HCoV-229E était nécessaire à l'activation de la fusion et si la présence du glycane N19 jouait un rôle dans ce clivage. Nous avons aussi déterminé quelle arginine de la région S2' était reconnue par les protéases à sérine pour activer la fusion.

## **B. Entrée virale par la voie endosomale**

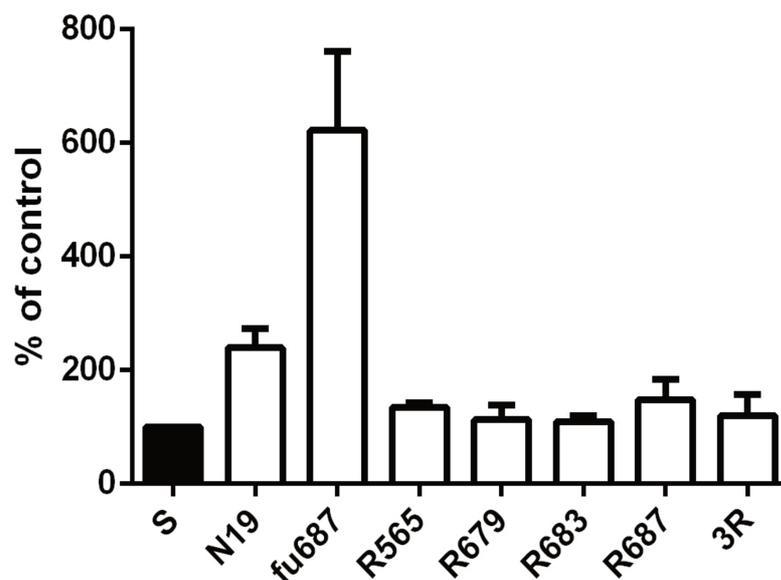
Dans un premier temps, et afin d'étudier le rôle des sites de clivage S1/S2 et S2' *in vitro* dans les étapes précoces de l'entrée du HCoV-229E, différents mutants ont été réalisés dans lesquels les arginines potentiellement utilisées par la trypsine comme site de clivage ont été remplacées par des asparagines (figure 9A). L'arginine présente à la jonction S1/S2 a été mutée en asparagine (R565N) ainsi que les arginines présentes dans la région S2 qui ont aussi été mutées individuellement (R679N, R683N ou R687N) ou les 3 ensembles (mutant 3R) (Figure 9A). De plus, un mutant dans lequel l'asparagine du séquen N-X-S/T a été mutée en glutamine (Q) a été construit (mutant N19). Nous avons d'abord regardé l'effet de ces mutations de la protéine S lorsque le virus entre par la voie endosomale en utilisant les protéases de type cathepsine présentes dans les cellules. Des particules pseudotypées (229Epp) avec la protéine S ou des mutants ont été produites (Figure 9B) en cellules HEK293T. Des cellules HuH7 ont été infectées par les 229Epp puis l'activité

luciférase a été mesurée après 48h. L'activité luciférase de chaque 229Epp pour un mutant donné a été rapportée à celle de la protéine S.

**A**



**B**



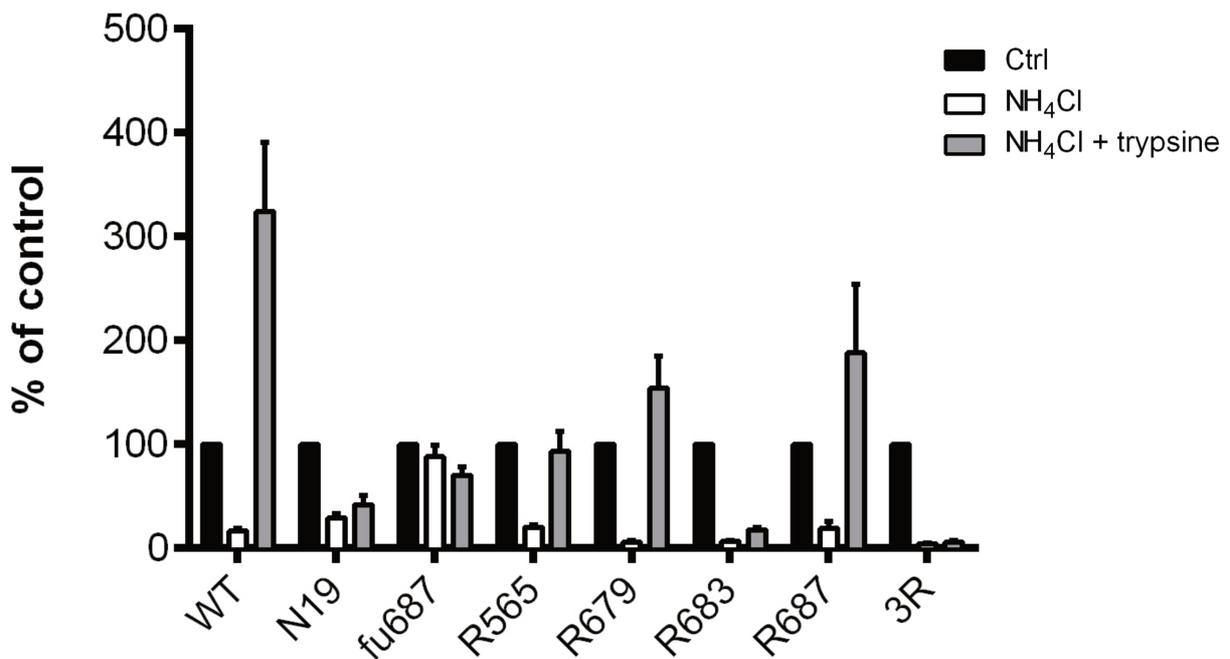
**Figure 9: Infectiosité de pseudoparticules incorporant les mutants.** **A.** Schéma fonctionnel de la protéine S et séquences des régions S1/S2 et S2' de la protéine S et des différents mutants. Les sites potentiels de N-glycosylation sont représentés par un trait noir court, et le site potentiel de clivage S1/S2 est représenté par un grand trait noir. Gris: domaine S1, or: domaine S2, violet: domaine de liaison au récepteur, rouge: peptide de fusion, bleu: HR1. 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. **B.** Des 229Epp ont été produites. Les valeurs luciférase pour chaque mutant ont été rapportées à celles de la protéine S pour obtenir le pourcentage d'infection. L'expérience a été répétée au moins 3 fois.

Les particules pseudotypées incorporant les mutants R565N, R679N, R683N, R687N et 3R ont des niveaux d'infectiosités comparables à celles incorporant la protéine S. Un mutant a aussi été réalisé dans lequel un site de clivage par la furine a été incorporé au niveau de la région S2' (Fu687) (Figure 9B). Ce mutant a un meilleur taux d'infection probablement parce qu'il fusionne plus tôt dans la voie endosomale que la protéine S. Les 229Epp incorporant le mutant N19 sont deux fois plus infectieuses que celles incorporant la protéine S, suggérant que ce glycan modulerait l'infectiosité de la protéine S.

### **C. Infection en surface médiée par la trypsine**

Nous avons induit l'infection par voie directe des 229Epp, c'est à dire à la surface cellulaire en utilisant la trypsine comme protéase (Figure 10). Pour cela, les cellules ont été prétraitées pendant 1h en présence de  $\text{NH}_4\text{Cl}$  afin d'inhiber l'infection par la voie endosomale. Les 229Epp ont été produites comme décrit précédemment et sont ensuite liées à la surface des cellules pendant 2h à 4°C en présence de  $\text{NH}_4\text{Cl}$ . La fusion a été induite par un court traitement par la trypsine à 37°C en présence de  $\text{NH}_4\text{Cl}$ . Les cellules ont ensuite été incubées pendant 30 min en présence de  $\text{NH}_4\text{Cl}$  et l'activité luciférase a été mesurée après 48h. L'activité luciférase de chaque condition pour un mutant donné a été rapportée à celle de son contrôle (pas de traitement par le  $\text{NH}_4\text{Cl}$  ni par la trypsine).

L'abolition de l'entrée endosomale des 229Epp par l'ajout de  $\text{NH}_4\text{Cl}$  est efficace pour les particules pseudotypées avec la protéine S (Figure 10). L'ajout de trypsine restaure l'infection à un niveau jusqu'à trois fois supérieur à celle du contrôle (pas de traitement des 229Epp par  $\text{NH}_4\text{Cl}$  ni par la trypsine).



**Figure 10: Impact de l'infection en surface des 229Epp médiée par la trypsine.** Des cellules HuH7 ont été prétraitées pendant 1h en présence de NH<sub>4</sub>Cl. Les 229Epp ont été liées à la surface des cellules pendant 2h à 4°C en présence de NH<sub>4</sub>Cl, puis la fusion a été induite par un court traitement par la trypsine à 37°C en présence de NH<sub>4</sub>Cl. Les cellules ont ensuite été incubées pendant 30 min en présence de NH<sub>4</sub>Cl et l'activité luciférase a été mesurée après 48h. L'activité luciférase de chaque condition pour un mutant donné a été rapportée à son contrôle (sans NH<sub>4</sub>Cl ni trypsine). L'expérience a été répétée trois fois.

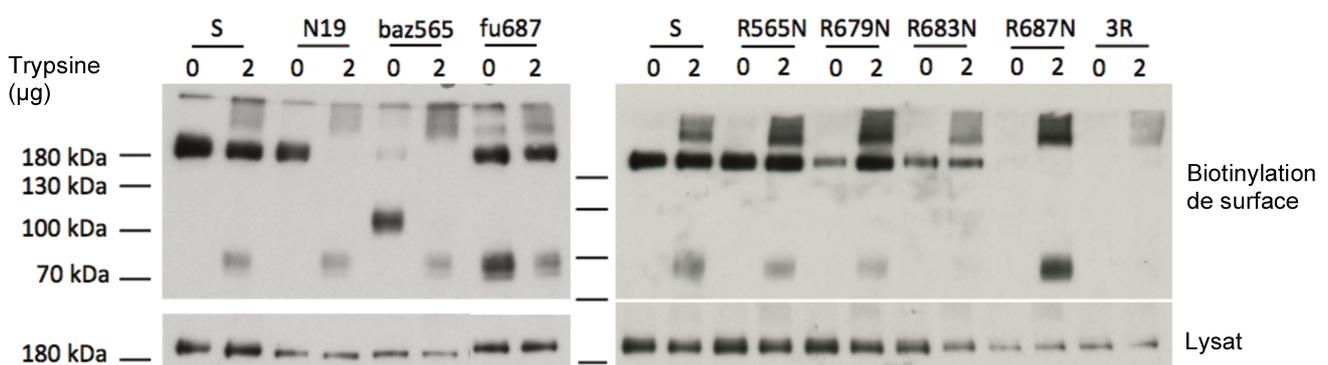
L'infection induite par la trypsine lorsque l'arginine présente à la jonction S1/S2 est mutée (R565N) est diminuée par rapport à celle de la protéine sauvage, suggérant qu'un clivage dans cette région puisse être nécessaire pour l'activation de la fusion. Cependant ce clivage ne semble pas indispensable puisque l'inhibition de l'infection n'est que partielle. De façon surprenante, l'abolition du site de glycosylation N19 inhibe l'infection médiée par la trypsine par rapport à celle de la protéine sauvage. Ce mutant présente pourtant des niveaux d'infection par la voie endosomale supérieure à ceux de la protéine sauvage.

Comme attendu, l'introduction d'un site de reconnaissance par la furine au niveau du site de clivage S2' (fu687) permet une infection similaire quelles que soient les conditions, indépendamment de l'ajout de trypsine ou du blocage de la voie endosomale. D'autre part l'infection médiée par la trypsine pour le mutant 3R dans lequel les 3 arginines présentes dans la région S2' ont été mutées est complètement inhibée. L'abolition individuelle des résidus basiques de la région S2' (R679, R683 ou

R687) inhibe au moins partiellement l'infection médiée par la trypsine par rapport à la protéine sauvage : il n'y a pas de restauration complète de l'entrée virale, quel que soit le résidu muté. Cependant, seule la mutation R683N induit une inhibition complète de l'infection médiée par la trypsine. Ces résultats suggèrent que la trypsine permet d'activer la fusion du virus HCoV-229E en clivant la protéine au niveau de l'arginine 683.

### ***D. Clivage de la protéine S du HCoV-229E par la trypsine***

Nous avons ensuite étudié le clivage de la protéine S par la trypsine (Figure 11). Pour cela, la protéine sauvage et les différents mutants ont été exprimés en cellules HEK293T puis après 48h les cellules ont été traitées par 2 $\mu$ g de trypsine pendant 20 min à 37°C. Après une biotinylation de surface, les cellules ont été lysées. Une partie du lysat a été analysé en western-blot avec un Ac anti-VSVG pour contrôler le niveau d'expression des protéines, l'autre partie du lysat a été incubée en présence de billes de streptavidine pour précipiter les protéines biotinylées. Les protéines biotinylées ainsi collectées ont été analysées en western-blot à l'aide d'un Ac anti-VSVG.



**Figure 11: Clivage de la protéine S du HCoV-229E.** La protéine S et ses mutants 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. Une partie du lysat a été gardé pour servir de contrôle d'expression. Après avoir incubé le reste des 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 Ac anti-tag après western-blot. Cette expérience a été répétée trois fois.

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. En théorie, un clivage en S1/S2 à la position 565 devrait générer deux produits de clivage de 115 et 113 kDa. Un clivage en S2' en position 687 devrait générer deux produits de clivage de 85 et 43 kDa. Une seule protéine est détectable pour chaque clivage (113 kDa pour un clivage en S1/S2 et 85 kDa pour un clivage en S2'), puisque l'Ac reconnaît le tag VSVG positionné à l'extrémité C-terminal de la protéine. Le clivage par la trypsine de la protéine S observé semble correspondre à un clivage au niveau de la région S2', ce qui est confirmé par la taille du fragment observé pour la protéine S possédant un site de clivage par la furine en position 687 (fu687). En effet, pour vérifier la taille des bandes obtenues lorsque la protéine est clivée entre S1 et S2 ou au niveau de S2', des séquences de reconnaissance par la furine ont été introduites par mutation soit au niveau de la jonction S1/S2 (baz565) soit au niveau de la région S2' (fu687) (S2') (Figure 9). L'introduction d'un motif furine en position 565 en absence de trypsine montre une bande unique à environ 115 kDa due au clivage de la jonction S1/S2 par la furine. Le profil obtenu en Western blot après traitement par la trypsine du mutant baz565 montre une bande vers 85 kDa suggérant un clivage S2'. Le mutant baz565 subit donc un clivage S1/S2 important en absence de trypsine, et un clivage S2' en présence de trypsine. L'introduction d'un motif de clivage par la furine en position 687 (fu687) montre une bande vers 200 kDa et entraîne la production d'un produit de clivage de 85 kDa.

Alors que le mutant N19 est légèrement moins exprimé que la protéine S sauvage dans les cellules HEK93T, la suppression du glycane renforce le clivage de la protéine S avec la production d'un produit de clivage d'un poids moléculaire de 85 kDa (Figure 9). Le mutant N19 est totalement clivé tandis que la protéine S ne l'est que partiellement. Ces résultats montrent que l'abolition du glycane en position N19 induit un clivage complet de la protéine S. De plus, les tests utilisés montrent que la protéine S et le mutant N19 sont clivés en S2', mais aucun clivage au niveau S1/S2 n'a été mis en évidence pour le moment.

Nous n'observons pas de produit de clivage par la trypsine pour le mutant R683N, ce qui est en adéquation avec les résultats précédents. En effet, la mutation R683N induit une inhibition complète de l'infection médiée par la trypsine et le clivage

protéolytique de la protéine S est important pour l'entrée virale. Les niveaux d'expression montrent que les mutants R687N et 3R étaient moins exprimés que les autres mutants et que la protéine S. Un produit de clivage dans la région S2' est observé pour le mutant R687N, mais pas le mutant 3R.

Nos résultats montrent qu'il n'y a pas de clivage visible de la protéine S en absence de trypsine et que seul un clivage dans la région S2' est observé. Il semble que l'absence du glycan N19 permette un clivage plus complet de la protéine S mais seul un clivage au niveau de S2' est aussi observé. Enfin, la trypsine semble cliver la protéine au niveau de l'arginine 683.

## IV. 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 (mutant N19), très proche de la jonction S1/S2, est conservé parmi les différentes souches de HCoV-229E et HCoV-NL63 et pourrait moduler le clivage protéolytique de cette jonction. Dans la région S2', un résidu basique (R687 pour HCoV-229E) localisé en N-terminal du peptide de fusion est très conservé parmi les coronavirus. Cependant, la protéine S de HCoV-229E contient deux arginines (R679 et R683) supplémentaires en amont de l'arginine 687. L'arginine 683 est très conservée parmi les souches HCoV-229E tandis que R679 est absente dans certaines souches, comme celle de l'ATCC. La conservation de ces arginines et du séquen N566 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. Le rôle du glycan N19 dans la modulation du clivage a été étudié. De plus, l'arginine impliquée dans le clivage en S2' a été déterminée.

Nous n'avons pas pu observer de clivage de la protéine S à la jonction S1/S2 par la trypsine. Il est possible qu'un intermédiaire transitoire lors du clivage soit présent, mais que nos méthodes d'analyse ne l'ait pas mis en évidence. En revanche, le clivage S1/S2 a été observé pour la protéine S du MERS-CoV et du SARS-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 d'étudier 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 celles incorporant la protéine S sauvage mais l'infection n'est pas complètement

inhibée. Ce résultat suggère que l'arginine 565 n'est pas indispensable à l'activation de la fusion médiée par la trypsine et donc que le clivage S1/S2 n'est pas primordial.

La présence d'un glycane proche de la jonction S1/S2 pourrait influencer l'accès des protéases à la protéine. En 2016, la structure de la protéine S du virus HCoV-NL63 a été déterminée en cryo-microscopie électronique ainsi que sa N-glycosylation (Walls et al., 2016a). La protéine S du virus HCoV-229E a 56% d'identité avec celle de HCoV-NL63. Il a été montré que sur les 39 sites potentiels de N-glycosylation de la protéine S de HCoV-NL63, 34 étaient effectivement glycosylés. Sur ces 34 sites, 26 sont conservés dans la protéine S du virus HCoV-229E dont le site à la jonction S1/S2, ce qui nous suggère que ce séquen est également glycosylé sur la protéine S du HCoV-229E. Le mutant N19 présente des niveaux d'infection par la voie endosomale supérieurs à la protéine sauvage, ce qui suggère que l'abolition du glycane 19 facilite l'entrée par voie endosomale. Il est possible que l'absence de ce glycane facilite l'entrée virale par voie endosomale en permettant un meilleur accès des protéases cellulaires comme la cathepsine au site de clivage S1/S2. A ce jour, les sites de clivage utilisés par les cathepsines pour activer la fusion ne sont pas bien connus mais il a été montré pour le SARS-CoV que la cathepsine L clivait la protéine au niveau de T668 soit une dizaine d'acides aminés en aval de la jonction S1/S2 (Bosch et al., 2008). Il a déjà été montré pour le virus de l'influenza A que la N-glycosylation à proximité du site de clivage HA1/HA2 empêchait l'accès de protéases comme la trypsine ou la furine à la protéine virale HA (Deshpande et al., 1987; Kawaoka et al., 1984; Tse et al., 2014). La mutation du site de glycosylation 19 inhibe l'infection médiée par la trypsine, ce qui indique que le glycane 19 est nécessaire à l'infection médiée par la trypsine. Cependant de façon surprenante, la protéine S est toujours clivée par la trypsine lorsque le site de glycosylation 19 est muté. L'abolition du glycane présent au niveau de la jonction S1/S2 facilite donc la voie endosomale mais inhibe la voie non endosomale. L'ensemble de ces résultats suggère que ce glycane 19 a un rôle dans l'entrée virale mais ne permettent pas de conclure sur le rôle du glycane 19 lors du clivage protéolytique de la protéine S.

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. Ce résultat est concordant 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). 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.

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. D'autre part, nous avons montré que la trypsine clivait la protéine S du virus HCoV-229E au niveau de l'arginine 683. Alors que l'infection à la surface cellulaire médiée par la trypsine est complètement abolie pour les mutants R683N et 3R, ces mutants sont toujours capables d'utiliser la voie d'entrée endosomale. 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 683 est très conservée parmi les différentes souches de HCoV-229E et dans HCoV-NL63. La mutation des deux autres arginines de la région S2' (R679 et R687) inhibe l'infection médiée par la trypsine mais de manière moins importante que lorsque R683 est mutée, et ces mutants sont correctement clivés par la trypsine. Il est possible que les arginines 679 et 687 aient un rôle mineur lors du clivage, cependant, l'arginine 679 ne semble pas être impliquée dans l'entrée virale médiée par TMPRSS2 (Shirato et al., 2016). Pour le mutant R687N, il est possible que la restauration de l'infection ne soit pas complète puisque cette arginine pourrait être incluse dans le peptide de fusion et modifier légèrement sa composition. Il avait précédemment été montré que le produit de clivage de la protéine S de HCoV-229E par la trypsine était de 80 à 85 kDa, ce qui correspond au domaine S2, mais le site exact de clivage n'avait pas été identifié (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 (Bertram et al., 2013). La différence de poids

moléculaire et de nombre de produit de clivage entre cette étude et la nôtre est probablement dû 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-terminal de la protéine S tandis que Bertram et al ont utilisé un sérum dirigé contre HCoV-229E. Nos résultats confirment donc un clivage par la trypsine dans la région S2' et permettent de préciser le résidu impliqué : R683. Il faut noter que l'arginine 683 est décalé de quelques acides aminés par rapport aux arginines du SARS-CoV et du MERS-CoV impliquées dans le clivage (Belouzard et al., 2009). Pour le SARS-CoV, il a été montré que la position du site de clivage au niveau S2' était flexible et qu'en fonction de la protéase utilisée, le clivage pouvait avoir lieu quelques acides aminés en amont du peptide de fusion.

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 permettra 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. Il est probable que TMPRSS2 soit capable d'activer la fusion de HCoV-229E en clivant la protéine de façon similaire à la trypsine au niveau de R683, ce qui sera testé ultérieurement.

Le clivage protéolytique de la protéine S par la trypsine induit un produit de clivage de 80 kDa, ce qui correspond au produit de clivage du mutant fu687 (introduction d'un site de reconnaissance par la furine dans la région S2') et est différent du produit de clivage de baz565 qui est de 120 kDa (introduction d'un site de reconnaissance par la furine dans la région S1/S2). Il a été montré que le clivage protéolytique dans la région S1/S2 des coronavirus hautement pathogènes induit la formation de fortes quantités de produit de clivage, 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' a induit la formation de fortes quantités de produit de clivage. 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.

## V. Bibliographie

- 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.
- 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.
- 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., 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., Bartelink, W., and Rottier, P.J.M. (2008). Cathepsin L Functionally Cleaves the Severe Acute Respiratory Syndrome Coronavirus Class I Fusion Protein Upstream of Rather than Adjacent to the Fusion Peptide. *J. Virol.* *82*, 8887–8890.
- 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.
- 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, 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.
- 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.
- 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.
- 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.
- 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.
- 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., He, Y., Zhou, Y., Liu, S., Zheng, B.-J., and Jiang, S. (2009). The spike protein of SARS-CoV--a target for vaccine and therapeutic development. *Nat. Rev. Microbiol.* 7, 226–236.
- 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., Deijns, 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.
- 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.
- 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.
- 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.
- Gorse, G.J., O'Connor, T.Z., Hall, S.L., Vitale, J.N., and Nichol, K.L. (2009). Human coronavirus and acute respiratory illness in older adults with chronic obstructive pulmonary disease. *J. Infect. Dis.* *199*, 847–857.
- 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.
- 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*.
- Hitzerd, S.M., Verbrugge, S.E., Ossenkoppele, 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., 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.
- 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. *Virology* 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., 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. (2015). 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.
- Li, F. (2016). Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annu. Rev. Virol.* 3, 237–261.
- 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, Y.-Y., Fang, J., and Ao, G.-Z. (2017). 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.
- 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.
- 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.
- Masters, P.S. (2006). The molecular biology of coronaviruses. *Adv. Virus Res.* *66*, 193–292.
- 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., Durosinioun, A., Al Asmari, M., Islam, A., et al. (2013). Middle East Respiratory Syndrome Coronavirus in Bats, Saudi Arabia. *Emerg. Infect. Dis.* *19*, 1819–1823.
- 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.
- 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.
- Nothhaft, 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.
- 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.
- 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.
- 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).
- 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.
- 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.
- 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.
- 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., Khyami, M.A., et al. (2015). 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.
- Sayaka, T.-Y., Tateki, K., Siddell, S.G., and Fumihiko, 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.*
- 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.
- Société Française d'Hygiène Hospitalière (2017). Actualisation des Précautions standard.
- 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.
- Tse, L.V., Hamilton, A.M., Friling, T., and Whittaker, G.R. (2014). A Novel Activation Mechanism of Avian Influenza Virus H9N2 by Furin. *J. Virol.* *88*, 1673–1683.
- 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.
- 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.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*.
- Walls, A.C., Tortorici, Ma., 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.
- 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, N., Shi, X., Jiang, L., Zhang, S., Wang, D., Tong, P., Guo, D., Fu, L., Cui, Y., Liu, X., et al. (2013). 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.
- 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 (2017a). WHO | Middle East respiratory syndrome coronavirus (MERS-CoV).
- WHO (2017b). WHO | SARS (Severe Acute Respiratory Syndrome).
- 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.
- 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.
- 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.

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

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.

Yeager, C.L., Ashmun, R.A., Williams, R.K., Cardellicchio, 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.

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.

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.

## VI. Annexes

Tableau 1 : Séquences des primers. Les nucléotides mutés sont indiqués en gras. F = primer sens, R = primer antisens.

Mutants	Primers
N19	F : CAGCCCAGACA <b>AG</b> TGTCCTACGAC
	R : TAGGACACT <b>TG</b> TCTGGGCTGCACG
R565N	F : TGCAGCCCC <b>AATA</b> ACGTGTCCTACGACAGCGT
	R : GACACGTT <b>ATT</b> GGGCTGCACGGCAATGATGC
Baz565	F : AGCC <b>CGCCGCGCCGCAAGCGCCGCA</b> ACGTGTCCTACGACAGCGT
	R : GTT <b>GCGGCGCTTGCGGCGGCGGCGG</b> GCTGCCGTCGGCGCACACGC
R679N	F : CAGCCTGCCCA <b>AAT</b> TCCGGCTCTAGAGTGGCCGG
	R : TAGAGCCGGA <b>ATT</b> GGGCAGGCTGGGGATCACGC
R683N	F : ATCCGGCTCT <b>AAT</b> GTGGCCGGAAGAAGCGCCAT
	R : TTCCGGCCAC <b>ATT</b> AGAGCCGGATCTGGGCAGGC
R687N	F : TGGCCGGA <b>AAT</b> AGCGCCATCGAAGATATCCT
	R : ATGGCGCT <b>ATT</b> TCCGGCCACTCTAGAGCCGG
3R	F : CCC <b>AAT</b> TCCGGCTCT <b>AAT</b> GTGGCCGGA <b>AAT</b> AGCGCCAT
	R : CAC <b>ATT</b> AGAGCCGGA <b>ATT</b> GGGCAGGCTGGGGATCACGC
Fu687	F : TCTAGAC <b>GCTCGCGCAGA</b> AGCGCCATCGAAGATATCCT
	R : GCTT <b>CTGCGCGAGCGT</b> CTAGAGCCGGATCTGGGCAGGC

Tableau 2 : Numéro d'accèsion d'*alphacoronavirus* et *betacoronavirus*

Numéro d'accèsion	Coronavirus	Souche
AAK38656.1	Virus de la diarrhée épidémique porcine PEDV	CV777
AGT51331.1	Coronavirus humain NL63	
AFH58021.1	Coronavirus félin FCoV	1683
AAV65515.1	Coronavirus canin CCov	1-71
ABG89301.1	Virus de la gastro-entérite transmissible TGEV	Miller 06
CAB91145.1	Virus de la gastro-entérite transmissible TGEV	Purdue
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
AAT98580	Coronavirus humain HKU1	
KF963244.1	Coronavirus humain OC43	FRA_EPI/Caen/2013/15

Université de Lille 2  
FACULTE DES SCIENCES PHARMACEUTIQUES ET BIOLOGIQUES DE LILLE  
**MEMOIRE de DIPLOME D'ETUDES SPECIALISEES**  
**(tenant lieu de Thèse en vue du Diplôme d'Etat de Docteur en Pharmacie)**  
Année Universitaire 2016/2017

**Nom :** Bonnin  
**Prénom :** Ariane

**Etude des mécanismes d'activation protéolytique de la protéine S du coronavirus humain 229E**

**Mots-clés :** HCoV-229E, N-glycosylation, clivage protéolytique, coronavirus, enveloppe

---

**Résumé :**

Le coronavirus humain 229E (HCoV-229E) est à l'origine de rhumes et peut entraîner de sévères complications respiratoires chez les personnes âgées ou ayant une maladie chronique. Les coronavirus sont des virus enveloppés ayant un grand génome à ARN simple brin de polarité positive. Trois protéines virales sont ancrées dans l'enveloppe : la protéine S, la protéine M et la petite protéine E. La protéine S s'assemble en trimères à la surface des virions et est fortement N-glycosylée. Elle joue un rôle majeur dans les étapes précoces de l'infection virale avec le domaine S1 responsable de l'attachement au récepteur et le domaine S2 responsable de la fusion des membranes. La fusion de l'enveloppe virale avec les membranes de la cellule cible est activée par protéolyse de la protéine S. Pour le SARS-CoV (pour *severe acute respiratory syndrome coronavirus*), deux sites de clivages protéolytiques ont été décrits, situé à la jonction S1/S2 et en amont du peptide de fusion (S2'). Comme le SARS-CoV, le virus HCoV-229E a recours aux protéases cellulaires pour initier son infection, cependant les mécanismes de clivage de la protéine sont inconnus. Les coronavirus peuvent utiliser aussi bien les protéases présentes à la surface cellulaire que des protéases endosomales. Nous avons donc étudié le rôle de la région S2' et de la jonction S1/S2 dans l'activation protéolytique de la protéine S de HCoV-229E. Nous avons étudié l'infection par la voie endosomale et l'infection à la surface cellulaire en induisant la fusion avec la trypsine comme protéase modèle. Pour cela, nous avons d'abord cloné la protéine S d'une souche circulante du HCoV-229E. L'analyse de la séquence de la protéine clonée a montré la présence d'un site potentiel de N-glycosylation situé près de la jonction S1/S2 (N19). Nous avons muté individuellement les arginines présentes à la jonction S1/S2 et dans la région S2' ainsi que le site de glycosylation N19. Pour étudier l'entrée virale de ces mutants nous avons utilisé un système de particules rétrovirales pseudotypées. Nos résultats montrent que la protéine S est clivée par la trypsine dans la région S2' pour le virus HCoV-229E au niveau du résidu R683N. La mutation de l'arginine présente au niveau de la jonction S1/S2 a un effet modéré sur l'infection médiée par la trypsine. L'analyse du clivage de la protéine par la trypsine en western-blot n'a pas permis de mettre en évidence de clivage de la protéine au niveau de la jonction S1/S2 mais il est possible que ce clivage soit intermédiaire et très transitoire. L'abolition du glycan 19 semble faciliter l'entrée virale par la voie endosomale mais inhibe la voie non endosomale. Ces résultats suggèrent que ce glycan 19 a un rôle dans l'entrée virale mais aucun lien avec le clivage protéolytique de la protéine S n'a pu être établi. 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 permet d'apporter des indications sur le tropisme et la pathogénicité de ces virus.

---

**Membres du jury :**

**Président**     **Dr DUBUISSON Jean**, Directeur de recherche, Centre d'Infection et d'Immunité de Lille

**Assesseur**    **Dr BELOUZARD Sandrine**, Chargée de recherche, Centre d'Infection et d'Immunité de Lille  
**Dr GOFFARD Anne**, Maître de conférences des Universités - Praticien Hospitalier, Centre d'Infection et d'Immunité de Lille, CHRU de Lille  
**Dr HELLE François**, Maître de conférences, Université de Picardie Jules Verne  
**Dr LORIDANT Séverine**, Praticien Hospitalier, CHRU de Lille