Intracellular trafficking of the M protein of Middle East respiratory syndrome coronavirus (MERS-CoV)

Anabelle Perrier

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LILLE

École Doctorale Biologie–Santé

Pour l'obtention du grade de :

Docteur de L'UNIVERSITÉ DE LILLE

Thèse soutenue le 20 septembre 2019

Trafic intracellulaire de la protéine M du coronavirus du syndrome respiratoire du Moyen-Orient (MERS-CoV)

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

Directrice de Thèse : **Dr. Sandrine Belouzard**

Présentée et soutenue publiquement par :

Anabelle Perrier

Membres du Jury :

Président	Pr. Didier Hober
Rapporteur	Pr. Sophie Le Poder
Rapporteur	Pr. Hans Nauwynck
Examinateur	Dr. François Helle
Directrice de thèse	Dr. Sandrine Belouzard

Remerciements

J'ai souvent pensé au moment où je devrais rédiger ce texte, et cela me paraissait si lointain que j'ai du mal à réaliser que je suis réellement en train de le faire. Sauf qu'évidemment, j'allais bien devoir finir ma thèse un jour ! (ouf).
Cela dit, j'ai du mal à imaginer « l'après » labo. Les matins où je ne vous rejoindrais plus à moitié réveillée pour boire le café, où l'on ne passera pas la journée à travailler en discutant, et tout ça me semble encore un peu abstrait.

Mais il est malgré tout temps de vous dire au revoir, même si je compte bien revenir vous embêter régulièrement ! Considérez donc ce maigre mot de remerciements comme un numéro d'adieu de ma petite publication improbable et vide de sens du MCV Lab (que vous aviez la gentillesse de lire malgré tout), et bouclons la boucle.

Evidemment, privilège de Grand-chef, je vais commencer par Jean, alias John Dubosquet. Merci avant tout de m'avoir accueillie dans le laboratoire, merci d'être toujours aussi patient, calme et réceptif, et merci de répondre aux questions de virologie bien plus rapidement et précisément que Google et PubMed réunis. Bientôt Grand-chef du CIIL, ça y est le labo prend le pouvoir !

Je vais faire le plus important en deuxième, ce qui correspond finalement assez bien à mes talents d'organisation. La plus importante sans aucun doute, c'est Sandrine, alias Séverine Balouzerd. Tu me suis depuis ma plus tendre enfance de chercheuse (autrement dit, le M1), et tu as bizarrement accepté de me garder tout ce temps, durant le M2 et la thèse. Tu m'as appris absolument tout ce que je sais, de comment pipeter à comment cultiver des cellules, en passant par les rapports que tu as dû relire (je me souviendrais toujours du VersionFinale.turêves.doc), sans parler des présentations que tu as subies des dizaines de fois.

Bref, si je sais faire quoi que ce soit dans un laboratoire, c'est grâce à toi. Mais au-delà des connaissances que tu m'as transmises, tu as surtout toujours été présente lorsque j'avais besoin de toi, que ce soit pour une question, pour une manip qui ne marche pas, pour gérer le stress, ou pour courir dans la boue.

Je garde précieusement sur mon ordinateur un fichier « Sandrine me préfère », comme preuve, au cas où un quelconque futur étudiant essaierai un jour de me voler ma place de préférée. Ce qui est évidemment hors de question.
Merci pour TOUT.

Je n'aurais probablement pas passé d'aussi bonnes années ici si je n'avais pas profité de la (délicieuse) compagnie des andouillettes. Cœur de l'équipe Corona, que dis-je, piliers, fers de lance, atout charme… bref, indispensables à ce qu'ont été ces quatre années : Ariane et Adeline, alias Marianne Ninbo et Aline Daniel. Les pauses café ou thé, les soirées à discuter, le week-end andouillette, la ch'ti délire, le dépannage de manips, les chansons des compétentes, l'inventaire des produits chimiques… tant de souvenirs ! Merci mille fois pour votre aide précieuse et votre soutien pendant ces années (même si

Ariane est partie avant la fin !), je n'aurais pas réussi sans vous. Et merci d'avance pour les nombreuses années à venir. Vivement le prochain week-end andouillettes !

Je ne peux pas poursuivre sans remercier mon Grand-tuteur Yves, alias Yvan Chromé, alias GT. Certes, au départ, ce sobriquet faisait de toi en quelque sorte mon « grand-père de la bio », ce qui n'était pas hyper flatteur (surtout que tu as plutôt l'âge de mon père, et que tu es vintage, pas vieux !). Puis finalement, tout le monde a finit par s'habituer à ce que je t'appelle GT, et tu as tenu ton rôle à merveille.
Merci pour tes protocoles hyper rapides, tes nombreuses explications, et tes idées (on s'en souviendra du mutant K199…), mais surtout merci pour ta bonne humeur et ton humour douteux (oui, j'ai écrit douteux). Nous avons partagé beaucoup de bon moments, entre les nombreuses pauses café et les bières du vendredi (ou de tous les autres jours d'ailleurs), ou les sorties un peu plus culturelles.
Plus que 7 ans avant la retraite, profite bien de la pince sacrée des immunofluorescences tant qu'il est encore temps !

Merci aussi aux mange-tôt avec qui j'ai partagé tant de moments.
Nathalie, alias Natacha Kalinsov (je tairais ton autre surnom, au cas où des enfants liraient ceci), merci pour tes histoires toujours si drôles, ta bonne humeur (parce que même quand tu râles, c'est marrant), et surtout pour ton chocolat à l'extrait naturel de Margarita, dont la saveur inédite restera à jamais gravée dans ma mémoire. Karin (respo team building) et Muriel, alias Corine Poincaré et Ariel Lamort, merci pour toutes ces discussions sur le cinéma, les livres, les séries, les anecdotes sur vos enfants et toutes les conversations plus improbables que l'on a eu au fil du temps (les rêves incluant des kalachnikov au P2, par exemple).

Un grand, très grand merci plus général à tous les membres du labo : à Claire, Laurence et Virginie pour vos rires si communicatifs, à Cécile-Marie et Cyrine pour votre gentillesse, à Soph pour tes dépannages à la McGyver, à Lowiese qui chante et souris toujours en travaillant, à Anne pour ton humour, à Thibault (fayot de légende) pour les discussions très drôles et les longues histoires sur les américains, à Kévin pour ta gentillesse et ton humour (et un peu ta maladresse aussi, parce qu'on aime tant se moquer), Karoline pour tes histoires de voyage et tous les mots rigolos que tu m'as appris, Thomas qui plane mais qui est toujours de bonne humeur (et merci parce qu'il me fallait bien un lyonnais dont je pouvais me moquer !).

Merci aussi à tout ceux avec qui j'ai partagé un bout de thèse : Rayan pour ta bienveillance et les soirées dans ton bar, Najet la catsitter officielle de Caliméro, Laura et Laure pour votre soutien et mon surnom de Bébé Corona, Jeanne pour tes références improbables et tes (nombreux) gâteaux, Maliki et Juliano pour votre bonne humeur et pour avoir porté tous les trucs lourds à ma place, Lydia madame joie de vivre et ses malheurs en cuisine, Czes et les longs débriefs de Game of Thrones, mais aussi à Rehab, à Juliette, merci à tous.

Bonne chance aux petits nouveaux Martin (aka Flash) et Esther (aka Esclave), sachez que vous avez bien de la chance d'avoir atterrit ici, alors profitez en !

Je vous laisse à tous la garde de Chouchouvirus, puisque c'est la mascotte du labo. Mais n'oubliez pas de l'emmener en vacances et de m'envoyer les photos !

Un immense merci à nos voisins de palier avec qui j'ai aussi partagé les hauts et les bas de ma thèse : d'abord merci aux Françoisettes ! Alex, on a commencé ensemble mais je te double sur le sprint final ! Depuis notre rencontre un peu par hasard à une journée des doctorants, les conversations furent nombreuses et plus ou moins intelligentes, mais toujours drôles. Merci pour ton soutien pendant ces années, et je ne manquerais pas de revenir pour ton jour J ! Gauthier, même si on a mis un peu de temps à se rencontrer, je suis très heureuse d'avoir pu partager autant de conversations et de bons moments au labo et en dehors (d'ailleurs tu te souviens de la fois où tu avais dit « Sainté » au lieu de « Santé » ?! Ah, on avait ri). Malgré tout, je dois bien avouer que tes jeux de mots pourris me font bien rire. Jonathan merci à toi aussi pour ta gentillesse et tes longues (très longues) histoires ponctuées d'anecdotes improbables, et tes passions étranges (je te ramènerais des cartes si j'en trouve encore).

Merci aussi à Anaïs, Bernardo, Kamel, Sophie et tous les autres, pour vos sourires et les bons moments passés ensembles !

Merci également à ma famille qui a subit les longues explications du sujet de ma thèse avec patience à chaque Noël ou anniversaire, et qui s'est déplacée jusque dans le Grand Nord pour ma soutenance ! Merci pour votre soutien sans faille.

Et évidemment par dessus tout je remercie Valentin, pour sa patience et sa gentillesse chaque jour, quand il a dû subir la mauvaise humeur, les récits de manips ratées, la fatigue, les allers retours au labo le week-end, et plus généralement, pour m'avoir supportée et continuer à le faire tous les jours ! Je n'y serais pas arrivée sans toi, merci infiniment.

MERCI !

« Le commencement de toutes les sciences, c'est l'étonnement de ce que les choses sont ce qu'elles sont. »

Aristote

Table of Contents

Résumé

Les coronavirus sont une famille de virus émergents, comme l'ont montré les émergences récentes des coronavirus SARS-CoV (Severe acute respiratory syndrome coronavirus) et MERS-CoV (Middle-East respiratory syndrome coronavirus), pathogènes pour l'homme. Nous ne disposons ni d'antiviraux spécifiques ni de vaccins pour lutter contre les coronavirus.

Des quatre protéines structurales du virus, la protéine M est le moteur de l'assemblage viral. Exprimée seule en cellules, elle peut dépasser le site d'assemblage dans la voie de sécrétion. Nous avons confirmé la localisation de MERS-M dans le TGN (trans-golgi network) et identifié deux signaux dans son domaine C-terminal impliqués dans son trafic : un signal DxE d'export du réticulum, et un signal KxGxYR de rétention dans le TGN. Le signal DxE avait déjà été identifié sur une autre protéine virale, tandis que le signal KxGxYR est un nouveau motif. Pour confirmer son rôle nous avons construit des chimères entre MERS-M et la protéine M du coronavirus IBV (Infectious bronchitis virus), localisée dans le compartiment intermédiaire entre le RE et le Golgi (ERGIC), et démontré que pour les deux protéines le domaine C-terminal est déterminant pour leur localisation spécifique.

Un second projet de caractérisation de l'activité antivirale de la digoxigénine contre le HCoV-229E a été initié. Nous avons démontré que cette drogue inhibe le virus à une étape post-entrée, avec une concentration inhibitrice médiane (IC50) de 250nM, et qu'elle n'est pas toxique pour les cellules à cette concentration. Elle inhibe aussi l'infection par le virus de l'hépatite C, et elle semble cibler une étape précoce de la réplication des virus RNA (+).

Mots clés : Coronavirus – trafic intracellulaire –protéine M – MERS – antiviraux

Résumé en français

Les coronavirus font partie de l'ordre des *Nidovirales*, la famille *Coronaviridae* et la sous-famille des *Orthocoronavirinae*, qui regroupe quatre genres : *Alphacoronavirus, Betacoronavirus, Gammacoronavirus* et *Deltacoronavirus*. Ils peuvent infecter une grande variété d'oiseaux et de mammifères, dont l'homme. Ils provoquent en général des symptômes respiratoires et entériques, mais peuvent aussi dans certains cas causer des problèmes rénaux ou hépatiques.

Jusqu'à récemment, la recherche sur les coronavirus était principalement d'intérêt vétérinaire, particulièrement à cause des coronavirus de porc et de poulet. Le coronavirus de la bronchite aviaire (IBV) par exemple, connu depuis les années 30, est une maladie du poulet extrêmement contagieuse qui cause de nombreux problèmes à l'industrie avicole (Cook et al., 2012).

Jusqu'en 2002, seulement quatre coronavirus infectant l'homme étaient connus : HCoV-229E (Human Coronavirus 229E), HCoV-OC43 (Human Coronavirus OC43), HCoV-HKU1 (Human Coronavirus HKU1) and HCoV-NL63 (Human Coronavirus NL63). Tous causent des symptômes respiratoires bénins de type rhume, mais peuvent être plus graves chez des patients immunodéprimés (Walsh et al., 2013).

Les deux dernières décennies ont cependant vu l'émergence de deux coronavirus hautement pathogènes infectant l'homme : le coronavirus du syndrôme respiratoire aigu sévère (SARS-CoV) en 2002, et le coronavirus du syndrôme respiratoire du Moyen-Orient (MERS-CoV) en 2012. L'épidémie de SARS-CoV, qui a débuté dans la province du Guangdong au sud de la Chine en novembre 2002, a causé plus de 8000 cas, avec un taux de mortalité de 10%. L'épidémie a été stoppée grâce à de strictes mesures d'hygiène et de quarantaine. L'épidémie de MERS-CoV a débuté en 2012 en Arabie Saoudite, et s'est rapidement répandue dans tout le Moyen-Orient. Encore une fois, de strictes mesures de confinement ont permis de limiter l'étendue de l'épidémie bien qu'il y ait eu des cas importés (après un voyage au Moyen-Orient) portant le nombre de pays touchés à 27. D'ailleurs, en 2015, un patient revenant du Moyen-Orient a déclenché une épidémie de plus de 180 cas en Corée du Sud, démontrant le potentiel épidémique du MERS-CoV. A l'heure actuelle, le MERS-CoV a causé plus de 2400 cas et 800 morts, ce qui correspond à un taux de mortalité de 35%. L'épidémie est toujours en cours, avec une majorité de nouveaux cas en Arabie Saoudite.

L'émergence récente du SARS-CoV et du MERS-CoV démontre que les coronavirus sont une famille de pathogènes émergents, et deviennent un problème de santé publique. Le MERS-CoV est d'ailleurs sur la liste de surveillance de l'OMS, et a été classé avec d'autres pathogènes tels qu'Ebola ou Zika parmi les pathogènes susceptibles de créer une urgence sanitaire.

En effet, nous ne disposons ni de traitement spécifique ni de vaccins pour lutter contre les coronavirus. Il est donc nécessaire de mieux comprendre la biologie de ces virus émergents afin de pouvoir développer des stratégies antivirales et vaccinales permettant de stopper leur propagation.

Nous avons mené deux projets en parallèle : un projet visant à identifier des signaux impliqués dans le trafic intracellulaire de la protéine M du MERS-CoV, et un second projet ayant pour objectif de caractériser le mécanisme d'action de la digoxigénine, un composé ayant une activité antivirale contre le coronavirus humain HCoV-229E.

Projet I : Etude du trafic intracellulaire de la protéine M du MERS-CoV

Les coronavirus ont quatre protéines structurales : la protéine de spike (S), la protéine de membrane (M), la protéine d'enveloppe (E) et la protéine de nucléocapside (N). Les protéines S, M et E sont enchâssées dans l'enveloppe lipidique du virus, et la protéine N est associée à l'ARN viral. Les protéines membranaires sont produites dans le réticulum endoplasmique (RE) et doivent être transportées dans la voie de sécrétion jusqu'au site d'assemblage, le compartiment intermédiaire entre le réticulum et le Golgi (ERGIC). Ensuite, un jeu d'interaction entre les différentes protéines est nécessaire à l'assemblage des nouveaux virions. La protéine M est essentielle à cette étape, puisqu'elle est capable d'interagir avec toutes les autres protéines structurales, S, N et E (Corse and Machamer, 2003; De Haan et al., 2000; de Haan et al., 1999b; Tseng et al., 2010a). Elle est donc considérée comme le moteur de l'étape d'assemblage.

D'autre part, il a été observé pour plusieurs coronavirus que la protéine M était capable de dépasser le site d'assemblage dans la voie de sécrétion et d'atteindre le Golgi ou le TGN (Klumperman et al., 1994a). De ce fait, la protéine M doit porter des signaux reconnus par les protéines cellulaires régulant le transport dans la voie de sécrétion, permettant son adressage spécifique dans ces compartiments.

Comprendre comment les protéines virales sont adressées près du site d'assemblage du virus et identifier les signaux impliqués dans leur transport contribuerait à la meilleure compréhension du cycle viral des coronavirus, et en particulier à celle de l'étape d'assemblage. Cela est particulièrement vrai pour la protéine M, puisqu'elle joue un rôle central dans les mécanismes d'assemblage. Cependant, peu de protéines M ont été étudiées dans cette optique, et la grande majorité des signaux impliqués dans leur transport sont encore inconnus.

En particulier, le trafic intracellulaire de la protéine M du MERS-CoV et plus généralement l'étape d'assemblage de ce virus n'ont pas été étudiés attentivement pour l'instant, puisque l'émergence du virus est très récente.

L'objectif de notre étude était donc d'étudier le trafic intracellulaire de la protéine M du MERS-CoV et d'identifier les signaux d'adressage impliqués dans sa localisation.

Pour initier notre étude du trafic intracellulaire de MERS-M, nous avons cloné la séquence codant pour la protéine M dans des vecteurs pCDNA3.1 avec différents tags en C-terminal ou N-terminal de la protéine. L'ajout d'un tag nous permettait d'une part d'avoir plus de flexibilité pour l'utilisation d'anticorps pour les co-marquages, et d'autre part d'utiliser des propriétés spécifiques à certains tags.

Nous avons donc construit pCDNA3.1-MERS-M, pCDNA3.1-HA-MERS-M, pCDNA3.1-V5-MERS-M, pCDNA3.1-MERS-M-V5, et pCDNA3.1-MERS-M-VSVG. Après transfection en cellules HeLa, nous avons d'abord caractérisé la localisation subcellulaire de la protéine HA-MERS-M en utilisant des marqueurs de différents compartiments : TGN46 (TGN), GFP-CI-MPR (TGN), ERGIC-53 (ERGIC), calréticuline CRT (RE) et CD4 (surface). La localisation a été analysée par immunofluorescence et le niveau de colocalisation a été quantifié avec ImageJ en utilisant le *Pearson Correlation Coefficient* (PCC). La protéine HA-MERS-M colocalise fortement avec les marqueurs TGN46 et GFP-CI-MPR, avec des PCC respectivement de 0.878 (+/-0.014) et 0.852 (+/-0.012), indiquant que la protéine est localisée dans le TGN.

Pour s'assurer que l'ajout d'un tag n'avait pas d'effet sur la localisation de la protéine M, nous avons comparé les localisations subcellulaires de toutes les protéines

taguées à celle de la protéine M non taguée. Toutes les protéines, taguées ou non, colocalisent avec les marqueurs du TGN avec un PCC >0,8, indiquant qu'elles sont toutes localisées dans le TGN peu importe la présence ou la position du tag. L'ajout d'un tag en N-terminal ou C-terminal n'a donc pas d'effet sur la localisation de la protéine.

Pour identifier des signaux impliqués dans le trafic de la protéine, nous avons construit des mutants de troncations du domaine C-terminal de MERS-M, et observé leur localisation subcellulaire.

Nous avons observé que la protéine HA-MΔ20 (avec les 20 derniers acides aminés en C-terminal tronqués) n'était plus localisée dans le TGN mais dans le RE, colocalisant avec le marqueur CRT. Cette localisation dans le RE pouvait être due soit à un défaut d'export du réticulum, soit à un problème de repliement de la protéine. Pour vérifier cela, nous avons construit la protéine MΔ20-VSVG, le tag VSVG ayant la particularité de contenir un signal DxE d'export du réticulum, et avons étudié la localisation subcellulaire de ce mutant. Cette protéine MΔ20-VSVG n'est plus bloquée dans le réticulum, mais exportée à la surface cellulaire, où elle colocalise avec le marqueur CD4, démontrant que la localisation de HA-MΔ20 dans le RE était due à un défaut d'export et non à un mauvais repliement.

Ces résultats indiquaient la présence d'un probable signal d'export du RE dans les 20 derniers acides aminés de la protéine. En étudiant cette partie de la séquence, nous avons identifié un signal di-acidique DxE semblable à celui présent sur le tag VSVG. Nous avons donc vérifié si ce signal DxE pouvait être un signal d'export fonctionnel pour MERS-M. Pour se faire, nous avons muté les deux résidus impliqués en alanine et étudié la localisation subcellulaire du mutant HA-MERS-M-D211A ;E213A. Ce mutant colocalise avec le marqueur CRT, indiquant que la protéine est bloquée dans le réticulum. L'ajout d'un tag VSVG (construction MERS-M-D211A ;E213A-VSVG) permet de restaurer la localisation de la protéine sauvage dans le TGN, confirmant que DxE est un signal d'export fonctionnel pour MERS-M.

La localisation de la protéine MERS-MΔ20-VSVG à la surface cellulaire et non dans le TGN suggérait que les 20 derniers acides aminés de la protéine contiennent également un motif déterminant la localisation de la protéine dans le TGN. Nous avons donc construits de nouveaux mutants avec de plus petites troncations de ce domaine :

MERS-MΔ5, MERS-MΔ10 et MERS-MΔ15. Comme toutes ces protéines étaient localisées dans le TGN comme la protéine M sauvage, nous avons effectué des mutations individuelles en alanine des 5 acides aminés AGNYR localisés entre les délétions Δ20 et Δ15 dans le contexte de la protéine M-Δ15. Ainsi les mutations des résidus G201, Y203 et R204 induisent une relocalisation de la protéine vers la surface cellulaire.

Nous avons ensuite introduit ces mutations dans la protéine de pleine longueur mais nous avons également muté certains résidus en amont des 20 derniers acides aminés. Ces résidus Y195, R197 et K199 ont été sélectionnés car ils sont très conservés chez les *Betacoronavirus*. Les constructions MERS-M-Y195A, MERS-M-R197A, MERS-M-K199A, MERS-M-G201A, MERS-M-Y203A, et MERS-M-R204A ont été transfectées en cellules HeLa et la localisation subcellulaire des protéines mutantes a été analysée par immunofluorescence et microscopie confocale.

Les protéines MERS-M-Y195A et MERS-M-R197A colocalisent avec le marqueur TGN46, comme la protéine M sauvage, indiquant que les résidus Y195 et R197 ne sont pas impliqués dans la localisation de la protéine dans le TGN. En revanche, pour les protéines MERS-M-K199A, MERS-M-G201A, MERS-M-Y203A, et MERS-M-R204A on observe un export des protéines à la surface cellulaire, confirmé par la colocalisation avec le marqueur CD4.

Un quadruple mutant MERS-M-K199A;G201A;Y203A;R204A-VSVG (M-KGYR) a donc été construit, et sa localisation a été analysée par immunofluorescence. Comme les mutants individuels, le quadruple mutant est exporté à la surface cellulaire et colocalise avec le marqueur CD4. La quantification par PCC montre que le niveau de colocalisation avec CD4 est similaire pour les mutants individuels et le quadruple mutant.

La quantité de protéine en surface pour les différents mutants a été analysée dans des expériences de biotinylation des protéines de surface. Les constructions MERS-M-D211A ;E213A, MERS-MΔ20, MERS-M-K199A, MERS-M-G201A, MERS-M-Y203A, MERS-M-R204A, et MERS-M-KGYR, toutes avec un tag V5 N-terminal et une mutation du site de N-glycosylation de la protéine afin de faciliter la quantification en western blot, ont été transfectées dans des cellules HeLa. Les protéines localisées au niveau de la surface cellulaire ont été marquées avec de la biotine à 4°C. Ensuite, les protéines

biotinylées ont été précipitées avec des billes d'agarose couplées à de la streptavidine et analysées en immunoblot.

Seulement 1% de la protéine M sauvage est exprimée à la surface cellulaire. En revanche, la mutation des résidus K199, G201, Y203 and R204 seuls ou en combinaison provoque l'export de la protéine à la surface cellulaire: de 5 à 10% du total pour les mutants individuels, et environ 13% pour le quadruple mutant. Le mutant DxE n'est presque pas détecté à la surface cellulaire, ce qui est en accord avec son blocage dans le RE observé en immunofluorescence. En revanche, comme observé avec le PCC en immunofluorescence, le mutant V5-MΔ20 est légèrement plus exporté à la surface que la protéine M sauvage. Ces résultats confirment que les résidus K199, G201, Y203 et R204 sont impliqués dans la localisation spécifique de la protéine M au niveau du TGN.

Puisque les mécanismes de trafic intracellulaire sont des processus hautement dynamiques, cette rétention de la protéine M dans le TGN peut être le résultat d'un équilibre entre des mouvements antérogrades et rétrogrades dans la voie de sécrétion. Si une protéine est exportée à la surface cellulaire, puis internalisée par endocytose et recyclée vers le TGN, à l'équilibre, elle sera localisée dans le TGN. L'internalisation des protéines est médiée par des signaux dont la mutation peut entrainer l'accumulation de ces protéines à la membrane plasmique. Nous avons donc testé si le motif KxGxYR était un signal d'endocytose.

Pour vérifier cette hypothèse nous avons réalisé un test d'internalisation, pour comparer les quantités de protéine M sauvage et protéine M-KGYR internalisées au cours du temps. Pour cela, les protéines en surface de cellules exprimant les protéines MERS-M ou MERS-M-KGYR ont été biotinylées à 4°C, puis les cellules ont été incubées à 37°C pour permettre l'endocytose. La biotine des protéines non-internalisées a été clivée, et les protéines internalisées ont été précipitées et détectées par immunoblot, puis le pourcentage de protéine totale internalisé a été quantifié. Ainsi, nous avons montré que la mutation du motif KxGxYR n'avait aucun effet sur l'internalisation de la protéine M. D'autre part, pour la protéine M du virus de l'hépatite murine (MHV) il a été montré que des mutants de la protéine M ne formant pas de multimères étaient adressés vers la surface cellulaire (Locker et al., 1995). Nous avons donc analysé l'oligomérisation des protéines MERS-M et MERS-M-KGYR et n'avons pas détecté de différence entre le mutant et la protéine sauvage.

Pour confirmer le rôle du signal KxGxYR en tant que motif de rétention dans le TGN, nous avons voulu transférer le signal sur une autre protéine. Dans ce but, nous avons construit des chimères entre la protéine MERS-M et la protéine M du coronavirus IBV. De cette manière, la structure à trois domaines transmembranaires de la protéine est conservée. De plus, comme la protéine MERS-M est localisée dans le TGN et la protéine IBV-M est localisée dans le *cis*-Golgi et le ERGIC, la localisation des chimères peut être différenciée comme étant de type MERS-M ou de type IBV-M. D'autre part, il a été montré que le premier segment transmembranaire de IBV-M, et plus particulièrement 4 résidus polaires, étaient responsables de la rétention intracellulaire de la protéine (Machamer et al., 1993a).

Nous avons donc construit des chimères dans lesquelles les domaines C-terminaux des deux protéines ont été échangés : MERS-M/IBV-M et IBV-M/MERS-M. La même construction a été faite mais avec le signal KxGxYR muté : IBV-M/MERS-M-KGYR. De même, nous avons construit des chimères dans lesquelles les premiers segments transmembranaires ont été échangés : TM1-MERS/IBV-M et TM1-IBV/MERS-M-KGYR. La localisation subcellulaire de ces protéines chimériques a été analysée par immunofluorescence. Comme attendu, la protéine IBV-M/MERS-M colocalise avec TGN46, indiquant que la protéine chimérique est localisée dans le TGN comme MERS-M sauvage. En revanche, la protéine chimérique portant le signal KxGxYR muté ne localise ni avec TGN46 ni avec ERGIC-53, mais est exportée à la surface cellulaire. Ces résultats confirment le rôle du signal KxGxYR comme motif de rétention dans le TGN, même dans le contexte de la protéine chimérique.

Plus surprenant, la protéine MERS-M/IBV-M colocalise avec le marqueur ERGIC53, indiquant que cette protéine est localisée dans le ERGIC, comme la protéine IBV-M sauvage. Ce résultat suggère que le domaine C-terminal de la protéine IBV-M est déterminant pour sa localisation spécifique dans le ERGIC. De plus, le remplacement du premier segment transmembranaire de la protéine MERS-M-KGYR (TM1-IBV/MERS-M-KGYR) par celui de IBV n'induit pas de rétention intracellulaire de la protéine. Ces deux résultats indiquent que le premier segment transmembranaire d'IBV-M n'est probablement pas impliqué dans la rétention intracellulaire de la protéine.

L'ensemble de nos résultats indique que pour MERS-M et IBV-M, le domaine C-terminal est essentiel à la localisation spécifique des protéines. Pour MERS-M, nous avons identifié deux signaux localisés dans les 20 derniers acides aminés de la protéine : un signal DxE d'export du réticulum endoplasmique, et un signal KxGxYR de rétention dans le TGN. Pour IBV-M, nous avons déterminé que le domaine C-terminal est essentiel pour la localisation de la protéine dans le ERGIC, mais nous n'avons pas identifié les signaux impliqués.

Pour aller plus loin, il serait intéressant d'étudier le rôle du signal KxGxYR dans les interactions M-S, M-E et M-N, particulièrement parce que le domaine C-terminal de la protéine M est suspecté d'être impliqué dans ces interactions. D'autre part, utiliser la génétique inverse afin d'insérer les mutations KxGxYR et DxE dans le génome de pleine longueur permettrait d'étudier l'impact de ces mutations dans le contexte de l'infection.

Enfin, poursuivre la recherche de signaux impliqués dans le trafic de MERS-M et IBV-M dans leur compartiment spécifique serait intéressant, puisqu'il est probable que leur localisation soit le résultat d'une coopération entre plusieurs signaux localisés dans différents domaines de la protéine.

Projet II : Caractérisation d'un inhibiteur de l'infection au HCoV-229E

Comme mentionné précédemment, les coronavirus sont une famille de virus émergents, et nous n'avons à l'heure actuelle ni traitement spécifique ni vaccin disponible. C'est pourquoi la recherche sur les coronavirus en général et en particulier la recherche d'inhibiteurs spécifiques de l'infection par les coronavirus sont d'une grande importance.

C'est dans ce contexte que notre laboratoire a initié un projet visant à identifier des molécules ayant un effet inhibiteur sur les coronavirus, en utilisant le coronavirus HCoV-229E comme modèle. Nous avons choisi ce virus car c'est un pathogène humain qui peut être manipulé dans un laboratoire de biosécurité de niveau 2 (BSL-2) rendant l'expérimentation plus rapide et moins contraignante par rapport aux coronavirus plus pathogènes SARS-CoV et MERS-CoV qui doivent être manipulés dans un laboratoire de biosécurité de niveau 3 (BSL-3).

Afin d'identifier des molécules à visée thérapeutique, un crible d'une banque de molécules contre l'infection par le virus HCoV-229E a été réalisé au laboratoire. Un crible de la banque PRESTWICK chemical library (PRESTWICK Chemical) contenant 1120 composés a été réalisé à deux concentrations différentes lors de l'infection. Le second projet de ma thèse a pour but de caractériser l'effet antiviral de certaines des molécules identifiées lors de ces cribles. Ainsi, quatre molécules ont été retenues afin de valider leur effet antiviral en test de dose-réponse : la trifluoperazine dihydrochloride (TH), le perhexiline maleate (PM), l'astemizole et la digoxigénine.

La trifluopérazine est un antagoniste de la dopamine avec des propriétés antipsychotiques, utilisé dans le traitement de la schizophrénie (Howland, 2016).

Le perhexiline maleate est utilisé dans le traitement des angines de poitrine. Il agit en se fixant à l'enzyme mitochondriale carnitine palmitoyltransferase (CPT)-1, et provoque une augmentation de la production d'ATP pour la même consommation d'oxygène, et donc une augmentation de l'efficacité de contraction du myocarde (Horowitz and Mashford, 1979).

L'astémizole est un inhibiteur réversible des récepteurs de l'histamine H1. Il a des propriétés antiallergiques et est donc utilisé dans le traitement des rhinites et conjonctivites allergiques, ainsi que de l'asthme (Janssens, 1993).

La digoxigénine est un stéroïde issu des feuilles et fleurs de *Digitalis lanata*. C'est un dérivé de la digoxine, un glycoside cardiaque utilisé dans le traitement des insuffisances cardiaques. La famille des glycosides cardiaques comprend de nombreuses molécules, toutes avec une activité similaire au niveau du cœur, et extraites de plusieurs plantes dont : *Strophanthus gratus, Acokanthera oblongifolia* and *Acokanthera schimperi* (ouabain), *Digitalis lanata* and *Digitalis purpurea* (digoxine, digitoxine), *Scilla maritima* (proscillaridine A), *Nerium oleander* (oléandrine, oléandrigenine); mais aussi retrouvées dans les poisons de certaines grenouilles venimeuses (bufaline, marinobufagénine) (Winnicka et al.).

La digoxine a été isolée dans les années 30, mais les propriétés des plantes de la famille des digitales *(Digitalis*) sont connues depuis des siècles. Elles ont été utilisées comme traitement pour le cœur, mais aussi comme poisons puisqu'elles sont toxiques à haute dose (Bessen, 1986). En effet, même si la digoxine est encore largement utilisée dans le monde pour le traitement des insuffisances cardiaques et des arythmies (elle est d'ailleurs sur la liste de l'OMS des médicaments essentiels), elle a un index

thérapeutique très étroit, et peut être toxique et rapidement causer des effets secondaires chez les patients si elle est administrée à trop haute dose.

La cible cellulaire des glycosides cardiaques est la pompe Na^+/K^+ ATPase : ils sont capables de se fixer sur sa sous-unité α et de bloquer son activité. Cela cause une augmentation du Na^+ intracellulaire et en conséquence une augmentation de la concentration de Ca^{2+} en raison d'une activité inversée de l'échangeur sodium/calcium. Cette forte concentration intracellulaire en calcium dans les myocytes augmente la contractilité cardiaque.

Mais la pompe Na^+/K^+ ATPase fonctionne également comme un récepteur classique et la fixation des glycosides cardiaques peut aussi induire le déclenchement de nombreuses cascades de signalisation. De plus, leur potentiel en tant qu'anti-cancéreux ou antiviraux a déjà été démontré dans la littérature (Amarelle and Lecuona, 2018; Laird et al., 2014; Winnicka et al.).

Dans notre étude, nous avons évalué le potentiel inhibiteur de quatre drogues, la digoxigénine, la trifluoperazine dihydrochloride (TH), le perhexiline maléate (PM), et l' astémizole, contre l'infection par des coronavirus *in vitro*.

Les premiers tests de dose réponse et de toxicité ont montré que toutes les drogues avaient un bon effet inhibiteur sur HCoV-229E, à des concentrations non toxiques. La digoxigénine a montré l'effet le plus important, avec une concentration inhibitrice médiane (IC50) de 250nM.

Nous avons ensuite voulu identifier l'étape du cycle à laquelle les drogues exercent leur effet inhibiteur. La TH et le PM semblent agir sur plusieurs étapes du cycle viral, ce qui aurait rendu la caractérisation de leur mécanisme d'action complexe. L'astémizole semble en revanche avoir un effet sur l'entrée du virus, puisqu'elle inhibe l'infection lorsqu'elle est ajoutée pour une pré-incubation des cellules, ou au moment de l'inoculation. D'autre part, la digoxigénine a un effet clair sur une étape post-entrée, puisqu'elle est efficace uniquement lorsqu'elle est ajoutée en post-inoculation.

Aux vues de ces premiers résultats, nous avons sélectionné la digoxigénine pour caractériser plus avant son mécanisme d'inhibition de l'infection par HCoV-229E.

Dans un premier temps nous avons évalué plus précisément son IC50, sa toxicité vis-à-vis des cellules, ainsi que l'étape du cycle viral ciblée. Les résultats démontrent que la digoxigénine a un effet inhibiteur sur HCoV-229E, avec un IC50 d'environ 250 nM, indépendamment du type cellulaire.

En ce qui concerne la toxicité, le test MTS montre une diminution progressive de la viabilité cellulaire à partir d'une dose de 1 µM de digoxigénine pour les temps d'incubation 24 et 48h, et dès 0,3125 µM pour le temps d'incubation 72h. Cette baisse est cependant très modérée à 24h, et plus marquée à 72h. Il est aussi important de noter que ces doses et ces temps d'incubations sont bien supérieurs à ceux utilisés pour les tests d'infection (maximum 8h d'incubation, IC50 de 250nM).

Pour déterminer quelle étape du cycle viral est inhibée par la digoxigénine, nous avons infecté des cellules Huh-7 avec le virus HCoV-229E en ajoutant la drogue durant les différentes étapes : en pré-incubation avec les cellules (1h), pendant l'infection (1h), en post-infection (6h) ou pendant toute l'expérience (8h). La digoxigénine montre un effet antiviral post-infection très clair sur l'infection par HCoV-229E, et ne semble pas avoir d'effet lorsqu'elle est ajoutée lors d'autres étapes du cycle.

Afin de confirmer que la digoxigénine n'inhibe que l'étape post-entrée, nous avons testé son effet sur l'infection par des pseudoparticles. L'utilisation des pseudoparticules permet d'évaluer l'effet de la drogue sur l'étape d'entrée du virus. Des particules rétrovirales pseudotypées soit avec la glycoprotéine d'enveloppe du VSV (VSVGpp, contrôle) soit avec la protéine S du HCoV-229E (229Epp) ont été produites. Des cellules Huh-7 ont été ensemencées en plaques 96 puits et infectées pendant 3h avec des VSVGpp ou des 229Epp, en présence de différentes concentrations de digoxigénine. Etonnamment, la digoxigénine inhibe l'infection avec les VSVGpp et les 229Epp à un niveau comparable. Cependant, l'IC50 est un peu plus élevé que pour l'inhibition de l'infection par le virus HCoV-229E. Nous pensons que l'inhibition observée n'est pas due à un effet spécifique sur la glycoprotéine de surface, mais qu'il est plus probable que la drogue ait un effet sur une étape de transcription/rétrotranscription qui a lieu après l'entrée des pseudoparticules dans les cellules.

Nous avons également testé l'effet inhibiteur de la digoxigénine sur un autre virus à ARN (+), le virus de l'hépatite C (HCV).

La digoxigénine inhibe également l'infection avec HCV de façon dose-dépendante, avec un IC50 qui semble un peu supérieur à celui pour HCoV-229E. De même, la digoxigénine inhibe pour HCV l'étape post-infection. Ce résultat suggère que la digoxigénine pourrait être un inhibiteur des virus à ARN (+).

Afin de préciser l'étape post-entrée inhibée par la drogue A, nous avons utilisé des cellules Huh-7 exprimant de façon stable un réplicon HCV. Un réplicon subgénomique est une forme modifiée du génome du virus de l'hépatite C qui se réplique de façon autonome dans des cellules, mais ne produit pas de particules infectieuses. Il est défectif pour l'assemblage, et permet l'étude de la réplication uniquement, en contournant les autres étapes de cycle viral. Ce réplicon contient un gène rapporteur (GFP) et est déficient pour l'assemblage. La mesure de l'expression de la GFP reflète directement la réplication du réplicon dans les cellules. Ces cellules ont donc été incubées en présence de 0,25 ou 1,25 µM de digoxigénine, du DMSO ou du Daclatasvir (inhibiteur de la réplication du HCV, contrôle positif). Le niveau d'expression de la GFP a été mesuré par comptage des cellules positives au microscope et par western blot.

Les deux méthodes de quantification montrent que l'incubation des cellules avec la digoxigénine, indépendamment de la dose, n'a pas d'effet sur l'expression de la GFP, donc sur la réplication du réplicon. Ce résultat indique que la digoxigénine ne peut pas inhiber la réplication dans un contexte où elle est déjà « stable », et suggère donc qu'elle pourrait agir sur une étape précoce de la réplication virale.

L'ensemble des résultats démontre jusqu'ici que la digoxigénine est un inhibiteur de l'infection par HCoV-229E et par HCV, et qu'elle n'est pas toxique pour les cellules aux doses actives. Son effet inhibiteur agit sur une étape post-entrée, probablement une étape précoce de la réplication d'après nos expériences sur les cellules exprimant le réplicon HCV. De plus, il est probable que cette drogue puisse inhiber plus largement d'autres virus à ARN (+).

Afin d'identifier le mécanisme d'inhibition de la digoxigénine sur l'infection par HCoV-229E, nous avons tenté de générer des mutants de résistance en incubant des

cellules infectées avec des doses croissantes de digoxigénine. Nous n'avons pour l'instant pas réussi à obtenir des virus résistants à la digoxigénine, mais si l'on en obtient l'intérêt serait de séquencer leur génome, et particulièrement les gènes des protéines non-structurales, afin d'identifier des mutations. Cela nous donnerait des informations concernant l'étape de la réplication du virus qui est inhibée par la drogue.

Pour aller plus loin, l'utilisation de siRNA dirigés contre la sous-unité α de la pompe Na^+/K^+ ATPase serait intéressante pour confirmer que l'effet inhibiteur de la digoxigénine passe bien par sa cible cellulaire connue. D'autre part, l'électroporation d'ARN du HCoV-229E ou l'utilisation d'un réplicon subgénomique permettrait de confirmer que la digoxigénine agit sur une étape post-entrée de la réplication du HCoV-229E. Enfin, tester la digoxigénine sur d'autres coronavirus, particulièrement sur le MERS-CoV et/ou le SARS-CoV, serait intéressant pour évaluer si cette drogue est un inhibiteur potentiel des coronavirus hautement pathogènes.

L'intérêt de la médecine humaine pour les coronavirus s'est développé dans les deux dernières décennies, suite à l'émergence du SARS-CoV et du MERS-CoV, qui ont causé des épidémies de pneumonies sévères. Avant cela, la recherche sur les coronavirus était plutôt d'intérêt vétérinaire, principalement à cause de leur impact sur les industries du porc et du poulet. Cependant, même si le premier coronavirus a été identifié dans les années 1930, nous ne comprenons pas encore complètement leur cycle de réplication, leur transmission ou leurs mécanismes de passage de la barrière d'espèce. Il est donc nécessaire de poursuivre la recherche sur la biologie des coronavirus afin de mieux comprendre cette famille de virus émergents.

De plus, nous ne disposons actuellement ni de composés antiviraux spécifiques, ni de vaccins pour lutter contre les coronavirus humains. Les traitements sont le plus souvent symptomatiques, et les épidémies de SARS-CoV et de MERS-CoV ont été relativement contenues uniquement grâce à de strictes mesures d'hygiène et de quarantaine. En ce qui concerne les vaccins, quelques vaccins sont disponibles pour des coronavirus animaux, par exemple IBV, mais leur efficacité sur le terrain est extrêmement limitée entre autres par l'émergence continuelle de nouveaux variants du virus. Pour les coronavirus humains, bien que de nombreux candidats vaccins soient en développement, aucun n'est disponible à l'heure actuelle.

Abbreviations

ACE2	*Angiotensin converting enzyme 2*
APN	*Aminopeptidase N*
Bat-SL-CoVs	*Bat SARS-like coronaviruses*
BCoV	*Bovine coronavirus*
BdCoV	*Bottlenose dolphin coronavirus*
BWCoV	*Beluga whale coronavirus*
CCoV	*Canine coronavirus*
CEACAM-1	*Carcinoembryonic antigen-cell adhesion molecule*
CHIKV	*Chikungunya virus*
CIE	*Clathrin-independent endocytosis*
CME	*Clathrin-mediated endocytosis*
CMs	*Convoluted membranes*
CPT1	*Carnitine palmitoyltransferase 1*
CRCoV	*Canine respiratory coronavirus*
DC SIGN	*Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin*
DMV	*Double membrane vesicle*
DPP4	*Dipeptidyl peptidase 4*
ECoV	*Equine coronavirus*
EEs	*Early endosomes*
ER	*Endoplasmic reticulum*
ERC	*Endocytic recycling compartment*
ERES	*ER-exit sites*
ERGIC	*ER-Golgi intermediate compartment*

FCoV	*Feline coronavirus*
FECV	*Feline enteric coronavirus*
FIPV	*Feline infectious peritonitis virus*
GEF	*Guanine nucleotide exchange factor*
gRNA	*Genomic RNAs*
HCV	*Hepatitis C virus*
IBV	*Infectious bronchitis virus*
IC50	*Median inhibition concentration*
IP3R	*Inositol 1,4,5-triphosphate receptor*
LCVCs	*Large virion-containing vesicles*
LEs	*Late endosomes*
MERS-CoV	*Middle-East respiratory syndrome coronavirus*
MHV	*Murine hepatitis virus*
MLV	*Murine leukemia virus*
NCX	*Na^{+}/Ca^{2+} exchanger*
Neu 5,9 Ac2	*N-acetyl-9-O-acetylneuraminic acid*
NLS	*Nuclear localization signal*
PDCoV	*Porcine deltacoronavirus*
PEDV	*Porcine epidemic diarrhea virus*
PHEV	P*orcine hemagglutinating encephalomyelitis virus*
PI3K	*Phosphatidylinositol 3-kinase*
PLC	*Phospholipase C*
PM	*Perhexiline maleate*
PRCV	*Porcine respiratory coronavirus*
RBD	*Receptor binding domain*

RdRp	*RNA-dependent RNA polymerase*
REs	*Recycling endosomes*
RTC	*Replicase-transcriptase complex*
SADS-CoV	*Swine acute diarrhea syndrome coronavirus*
SARS-CoV	*Severe acute respiratory syndrome coronavirus*
SARSr-CiCoV	*SARS-related civet coronavirus*
SeCoV	*Swine enteric coronaviruses*
sgRNA	*Subgenomic RNAs*
SNX	*Sorting nexins*
TGEV	*Transmissible gastroenteritis virus*
TGN	*Trans-Golgi network*
TH	*Trifluoperazine dihydrochloride*
TM	*Transmembrane segment*
TMPRSS2	*Transmembrane protease, serine 2*
TRS	*Transcriptional regulatory sequence*
UTR	*Untranslated region*
VLPs	*Virus-like particles*
Vps	*Vacuolar protein sorting*
VSV-G	*Vesicular stomatitis virus glycoprotein*

Table of Figures

1. Introduction

1.1 Taxonomy and history

Coronaviruses are members of the *Nidovirales* order, the family *Coronaviridae* and the subfamily *Orthocoronavirinae*, and are composed of four genera: *alphacoronavirus, betacoronavirus, gammacoronavirus* and *deltacoronavirus.* Each genus is then divided into several subgenera (ICTV, 2018). The name coronavirus comes from the Latin *"corona*", meaning crown, and refers to the appearance of the viruses when observed in electron microscopy (Cavanagh, 2005). This appearance is due to the protrusions formed by proteins of the envelope at the surface of the virion.

Coronaviruses are widespread pathogens that can infect a wide variety of species among mammals and birds, including humans (Figure 1). *Alpha*- and *betacoronaviruses* mostly infect mammals whereas *gamma*- and *deltacoronaviruses* mostly infect birds (with the exception of porcine deltacoronavirus (PDCoV), beluga whale coronavirus BWCoV SW1, bottlenose dolphin-CoV HKU22 (Woo et al., 2014) and the Asian leopard cat-CoV (Dong et al., 2007)).

Six coronaviruses are now known to infect humans. The first human coronaviruses discovered were HCoV-229E and HCoV-OC43. They were successively isolated in 1966 and 1967 from the respiratory tract of patients with a common cold (Hamre and Procknow, 1966; McIntosh et al., 1967) and respectively classified as *alphacoronavirus* and *betacoronavirus*. Since these coronaviruses only caused mild symptoms in humans, initially the research on this family of viruses was mostly of veterinary interest. This was until the emergence, in 2002, of the first highly pathogenic coronavirus in humans: the *betacoronavirus* SARS-CoV (*severe acute respiratory syndrome coronavirus)* (Peiris et al., 2003). This emergence boosted the interest in coronaviruses, and led to the identification of two other human coronaviruses causing common cold, the *alphacoronavirus* HCoV-NL63 in 2004 (van der Hoek et al., 2004) and the *betacoronavirus* HCoV-HKU1 in 2005 (Woo et al., 2005a). Less than a decade later, another highly pathogenic human coronavirus emerged in the Middle East: MERS-CoV (*Middle East Respiratory Syndrome Coronavirus*), a *betacoronavirus* as SARS-CoV (Zaki et al., 2012). Coronaviruses are now a viral family of importance both to veterinary and human medicine, with a wide range of hosts and clinical presentations. More importantly, these viruses have a huge potential for cross-species transmission.

Figure 1 - ***Classification, host and receptor of mentioned coronaviruses***

GENUS	SUBGENUS	SPECIES	STRAINS	HOST	ENTRY RECEPTOR
Alphacoronavirus	Duvinacovirus	Human coronavirus 229E		Human	APN
	Pedacovirus	Porcine epidemic diarrhea virus (PEDV)		Pig	APN
	Rhinacovirus	Bat coronavirus HKU2		Bat	unknown
	Setracovirus	Human coronavirus NL63		Human	ACE2
		NL63-related bat coronavirus		Bat	unknown
	Tegacovirus	Alphacoronavirus 1	Transmissible gastroenteritis virus (TGEV)	Pig	APN
			Feline coronavirus (FCoV)	Cat	APN
			Canine coronavirus (CCoV)	Dog	APN
			Porcine respiratory coronavirus (PRCV)	Pig	APN
Betacoronavirus	Embecovirus	Betacoronavirus 1	Bovine coronavirus (BCoV)	Cow	Neu 5,9 Ac2
			Human coronavirus OC43	Human	Neu 5,9 Ac2
			Canine respiratory coronavirus (CRCoV)	Dog	unknown
			Porcine hemagglutinating encephalomyelitis virus (PHEV)	Pig	unknown
			Equine coronavirus (ECoV)	Horse	unknown
		China rattus coronavirus HKU24		Rat	unknown
		Human coronavirus HKU1		Human	unknown
		Murine coronavirus	Murine hepatitis virus (MHV)	Mouse	CEACAM1
	Merbecovirus	Hedgehog coronavirus 1		Hedgehog	unknown
		Middle East respiratory syndrome-related coronavirus	Middle East respiratory syndrome coronavirus (MERS-CoV)	Human	DPP4
		Bat coronavirus HKU5		Bat	unknown
		Bat coronavirus HKU4		Bat	DPP4
	Sarbecovirus	Severe acute respiratory syndrome-related coronavirus	Severe acute respiratory syndrome coronavirus (SARS-CoV)	Human	ACE2
			SARS-related civet coronavirus (SARSr-CiCov)	Civet	ACE2
			Bat SARS-like coronaviruses (WIV1, WIV16)	Bat	ACE2
Deltacoronavirus	Buldecovirus	Coronavirus HKU15	Porcine deltacoronavirus (PDCoV)	Pig	unknown
Gammacoronavirus	Cegacovirus	Beluga whale coronavirus SW1	Beluga whale coronavirus SW1	Beluga Whale	unknown
			Bottlenose dolphin coronavirus HKU22 (BdCoV)	Dolphin	unknown
	Igacovirus	Avian coronavirus	Infectious bronchitis virus (IBV)	Bird	unknown

1.2 Animal coronaviruses

Coronaviruses have a broad host range among both domestic and wild animals. Non-exhaustively, they can infect bats (Bat-CoVs), swine (TGEV, PEDV), cats (FCoV), mice (MHV), birds (IBV), dogs (CCoV), horses (ECoV), cattle (BCoV), dolphins (BdCoV HKU22)… Most animal coronaviruses typically cause enteric symptoms, such as *porcine epidemic diarrhea virus (*PEDV) and *transmissible gastroenteritis virus* (TGEV) in swine (Wang et al., 2019b) or *feline coronavirus* (FCoV) in cats. But there are a few exceptions, such as *infectious bronchitis virus* (IBV) in chicken, *porcine respiratory coronavirus* (PRCV) in pigs, and *canine respiratory coronavirus* (CRCoV) in dogs that may cause respiratory symptoms, some strains of *Murine hepatitis virus* (MHV) that can cause hepatitis or encephalitis in mice, and *feline infectious peritonitis virus* that can cause fibrinous and granulomatous serositis.

A few animal coronaviruses are extensively studied for different reasons: MHV because it is used as a model for multiple sclerosis, IBV in chicken and several coronaviruses in swine because they may cause a huge economic burden respectively to poultry and pork industries, and FCoV in cats because it presents a rare and complex pathogenesis, that is still not fully understood.

1.2.1 Swine coronaviruses

Seven coronaviruses are now known to infect pigs. Among them five are *alphacoronaviruses*: the *transmissible gastroenteritis virus* (TGEV*), the porcine respiratory coronavirus* (PRCV), the *porcine epidemic diarrhea virus* (PEDV), the *swine enteric coronaviruses* (SeCoV) and the *swine acute diarrhea syndrome coronavirus* (SADS-CoV). In addition, there is one *betacoronavirus,* the *porcine hemagglutinating encephalomyelitis virus* (PHEV), and one *deltacoronavirus,* the *porcine deltacoronavirus* (PDCoV) (Wang et al., 2019b).

Many swine coronaviruses cause enteric symptoms and provoked epidemics as soon as the 1950s. Indeed, the first that was identified is TGEV, after an epidemic of transmissible gastroenteritis (Doyle and Hutchings, 1946). Subsequently in 1972, a very

similar disease rapidly spread in Europe and Asia and the causative agent was identified as another coronavirus, PEDV. More recently, three porcine coronaviruses causing enteric symptoms emerged or re-emerged, demonstrating once again that coronaviruses are a family of emergent pathogens. *Porcine deltacoronavirus* (PDCoV) was detected in pig fecal samples collected in 2009 (Woo et al., 2012), but the first known epidemic caused by that virus occurred in 2014 in the USA (Wang, 2014). A more virulent variant of PEDV also re-emerged in 2010, and caused major outbreaks (Sun et al., 2012). Even more recently, in 2016, *severe acute diarrhea syndrome coronavirus* (SADS-CoV) emerged in China in the Guangdong province (the same location as SARS-CoV) causing high mortality among piglets (Pan et al., 2017). Diseases caused by TGEV, PEDV, SADS-CoV, and PDCoV cannot be differentiated without laboratory diagnosis

A few other swine coronaviruses cause other symptoms besides gastroenteritis. PHEV causes vomiting and wasting disease progressing to encephalomyelitis, and was first described in piglets in Canada in 1958 (Greig et al., 1962). In 1984, PRCV was identified in pigs, this time causing respiratory symptoms (Pensaert et al., 1986). PRCV is believed to have derived from TGEV.

All emerging or re-emerging coronaviruses in pigs were detected in China, which is not surprising considering that China concentrates more than 50% of the world's pig population (FAO 2019).

Transmissible gastroenteritis virus (TGEV) and *porcine respiratory coronavirus (PRCV)*

TGEV was the first described coronavirus to cause severe diarrhea, vomiting and dehydration among pigs. Transmission mostly occurs through fecal-oral route. In adults, diarrhea is transient and mild, whereas in seronegative piglets under 2 weeks, the mortality rate is close to 100%. Following infection, there is destruction of the enterocytes of the villi, leading to malabsorption and maldigestion and subsequently to dehydration and loss of electrolytes (Xia et al., 2018).

During the 1980s, a closely related coronavirus, believed to have derived from TGEV, was discovered in Belgium: PRCV. TGEV causes a strictly enteric pathology but replicates both in epithelial villus cells of the small intestine and in lungs cells. On the other hand, PRCV replicates almost exclusively in the respiratory tract and to a lesser extent in the small intestine (Sánchez et al., 1992). The infection with PRCV is mostly

unapparent, but seroconversion against TGEV is always detected. The extremely rapid spread of PRCV in Europe and USA suggests a strong airborne transmission, especially because unlike TGEV the virus was equally present in countries with high standards in terms of hygiene (Laude et al., 1993).

TGEV still causes sporadic outbreaks in China (Hou et al., 2012; Weiwei et al., 2014) but is generally less present than it used to be before the 1980s (Chen et al., 2019). This is probably linked to the cross protection that PRCV confers against TGEV, inducing the production of antibodies that neutralize both PRCV and TGEV at the same titers (Cox et al., 1993). Comparison of TGEV and PRCV on the molecular level showed that only a few genetic modifications of TGEV led to PRCV, although the latter is presenting a completely different pathogenesis and host-pathogen interaction. Indeed, comparison of TGEV and PRCV genomes showed a similar general organization, with differences mainly composed of deletions and point mutations in the PRCV genome. Actually, the sequence divergence between TGEV and PRCV is around 3%, which is only slightly higher than the divergence between two TGEV strains. The deletions vary from one strain of PRCV to another, but are always located in the 5' end of the S gene and in the ORF3 (Chen et al., 2019). The spike protein of PRCV has lost sialic acid binding, possibly explaining the difference in tropism between PRCV and TGEV (Schultze et al., 1996).

***Porcine epidemic diarrhea virus* (PEDV)**

Several PED epidemics were observed during the 1970s in Europe, with the first starting in 1972 in England and rapidly spreading to other European countries (Choudhury et al., 2016). Nevertheless, since the 1990s PEDV is rare in Europe except in a few pockets, for example in Italy (Martelli et al., 2008). On the other hand, PEDV was first reported in Asia in the 1980s and still has an important impact on the Chinese pork industry, causing high mortality among piglets. This is especially true since China saw the emergence of a highly pathogenic variant of PEDV in the 2010s (Sun et al., 2012). This strain caused serious outbreaks in other Asian countries including Japan and South Korea (Lee and Lee, 2014; Yamane et al., 2016), and then appeared in the US in 2013 where it spreads rapidly across the country, causing the death of more than 8 million newborn piglets (Stevenson et al., 2013).

PEDV can infect pigs of all ages, causing watery diarrhea, vomiting, anorexia (which affects growth performance for growing pigs), and depression. In piglets up to

one week, diarrhea and vomiting rapidly leads to severe dehydration and death (Pensaert and Martelli, 2016). The younger the piglet, the higher is the mortality rate, with almost 100% for piglets of 1-3 days. Transmission of the virus occurs mostly by fecal-oral route. PEDV tropism is restricted to the enterocytes and epithelial cells located in the intestinal villi. It is not clear why PEDV causes more severe pathology in nursing piglets compared to weaned piglets, but it is postulated that a slower enterocyte regeneration in neonatal pigs play a role in the process.

Importantly, recently emerged highly virulent strains are highly infectious (very low minimal infectious dose) and the disease they cause progresses more rapidly than for previously known PEDV strains. Additionally, the virus replicates throughout the intestine and causes severe intestinal lesions (Lin et al., 2016).

***Porcine hemagglutinating encephalomyelitis virus* (PHEV)**

PHEV was identified as the causative agent of encephalitis in piglets in 1962 (Greig et al., 1962). As other swine coronaviruses, PHEV causes enteric symptoms such as vomiting and diarrhea, but this virus is also capable of spreading to the central nervous system, where it can subsequently cause encephalomyelitis (Hirano et al., 2001). The neurological symptoms include ataxia, hyperesthesia and paddling (Mora-Díaz et al., 2019), and lead to death after 2 to 3 days. Serological surveys showed that PHEV is widespread, and mostly subclinical, mothers likely conferring protection to newborn piglets through maternal antibodies. Thus, the number of clinical outbreaks is low, with only notable outbreaks in Canada, Argentina, and China (Sasseville et al., 2001). The virus primarily replicates in the respiratory tract, thereby causing non-specific respiratory symptoms such as cough and sneezing at the onset of infection. It can then spread from the primary sites of infection through the peripheral nervous system to the central nervous system. PHEV is the only known neurotropic coronavirus in swine, and also exhibits neurotropism in experimentally infected mice and rats (Hirano et al., 2001).

***Swine acute diarrhea syndrome coronavirus* (SADS-CoV)**

Swine acute diarrhea syndrome coronavirus was first isolated in China in the Guangdong province (Gong et al., 2017) and caused epidemics among piglets, that remained confined to China. Similarly to other enteric swine coronaviruses, SADS-CoV causes acute diarrhea and vomiting, leading to weight loss, dehydration and death in piglets. Mortality for piglets under five days was 90%, whereas for piglets older than

eight days the mortality dropped to 5% and the virus causes mild symptoms for adult pigs (Zhou et al., 2018). The virus likely originates from bats, since it shares 95% nucleotide identity with the Bat-CoV HKU2, and other related bat-coronaviruses were found to share 96-98% identity with SADS-CoV (Cui et al., 2019). One of the few studies on SADS-CoV showed that the virus is mainly found in the intestine, but can also be found in a lot of other tissues, including liver, spleen, heart, kidneys, stomach and lungs (Xu et al., 2019).

***Porcine deltacoronavirus* (PDCoV)**

Porcine deltacoronavirus was detected for the first time in fecal samples of pigs in Hong-Kong (Woo et al., 2012) and the first known epidemic occurred in the USA in 2014 (Wang, 2014) where the virus spread rapidly across the country. The virus was associated with diarrhea and vomiting, with high mortality among piglets, and is therefore clinically indistinguishable from TGEV or PEDV. Nevertheless, despite co-circulation of PDCoV with PEDV at the same period in the US, there was no cross-protection between the viruses. The transmission and pathogenesis of PDCoV is highly similar to TGEV or PEDV, with infection of enterocytes in the small intestine leading to villous atrophy and subsequent enteric symptoms. In addition to the intestinal lesions, PDCoV also causes lesions in the stomach, and milder lesions in the lungs. Nevertheless, PDCoV seems to be less pathogenic that PEDV, since the mortality rate among suckling piglets is around 40% (Jung et al., 2016).

***Swine enteric coronaviruses* (SeCoV)**

Swine enteric coronaviruses were retrospectively identified in Italy and in Germany (Akimkin et al., 2016; Boniotti et al., 2016). The Italian strain apparently circulated from 2009 to 2012, and then disappeared. The German strain was identified in fecal samples from 2012, sharing 99,5% identity with the Italian strain. Although slightly different, both strains are chimeric with most of the genome originating from TGEV, and the S and 3A genes being derived from PEDV. A closely related strain (>98%) was subsequently identified in Central Eastern Europe (Belsham et al., 2016).

1.2.2 Feline coronavirus (FCoV)

Peritonitis in cats was first observed in the 1950s (Holzworth, 1963), and the responsible virus was identified in 1970 (Ward, 1970). It is currently among the most

important pathogens in cats, and is distributed worldwide. *Feline coronavirus (*FCoV) is an *alphacoronavirus* that can infect both wild (Heeney et al., 1990; Hofmann-Lehmann et al., 1996; Kennedy et al., 2003; Leutenegger et al., 1999; Munson et al., 2004; Paul-Murphy et al., 1994) and domestic *felidae* (Pedersen, 2009; Taharaguchi et al., 2012), although it is more prevalent in domestic cats. The virus belongs to the same species as TGEV and CCoV, which are all regrouped under *Alphacoronavirus 1*.

Seroprevalence of FCoV is highly variable depending on the location and the number of cats in the household. In animal shelters or households with multiple cats, almost 90% of cats are seropositive (reviewed in Drechsler et al., 2011). The transmission occurs through fecal-oral route, which explains the close relationship between the number of cats and the seroprevalence of the disease. Housing and hygiene practices therefore also have a huge impact on the number of contaminated cats in the setting.

There are two existing serotypes of FCoV based on neutralizing antibodies. Serotype I is closely related to TGEV, has a specific feline spike protein, and is responsible for most naturally occurring infections (Benetka et al., 2004), whereas serotype II has a recombinant spike protein containing Canine-CoV sequences, and is mostly used as a model because of its rapid propagation in cell culture (Fiscus and Teramoto, 1987). Interestingly, feline coronaviruses can also be divided into two biotypes displaying very different pathologies: the enteric biotype (FECV), that is widespread and causes asymptomatic or mild disease when infecting the gut, and the infectious peritonitis biotype (FIPV) that causes a systemic infection that is usually fatal (Drechsler et al., 2011; Pedersen et al., 1984). Importantly, both biotypes exist in both serotypes.

Even though FECV or FIPV strains differ from one location to another, they are always extremely similar (97.3–99.5% identity) in the same setting. This, and the identification of deletions in the FIPV genome when comparing it to the FECV genome, led to the commonly admitted fact that an FIPV strain arises from an FECV strain within the same animal (Vennema et al., 1998). Importantly, FECV causes persistent infections in cats, which explains that mutations can gradually appear, and an FIPV strain can arise. The risk of developing peritonitis in cats infected by FECV ranges from 5 to 12% (Addie et al., 1995; Foley et al., 1997) but can dramatically decrease to less than 1% depending

on environment factors such as population density, husbandry practices, or length of stay in the shelter. The risk is higher for kittens and cats under two years of age, or immunosuppressed cats.

FECV is widespread in cats and transmitted horizontally quite easily. But interestingly, even though FIPV is highly infectious when inoculated into naïve cats, horizontal transmission doesn't often occur. In that light, FIPV seems to be highly tissue bound, and shedding in feces could only occur if specific lesions affect the intestinal tract.

FIPV causes fibrinous and granulomatous serositis, and can be found under two forms: an effusive or wet form, and a non-effusive of dry form, both causing a severe systemic disease. The more common in natural infections is the wet form, so called because of the presence of abundant protein-rich effusions in body cavities (Tekes and Thiel, 2016). This leads to a distended abdomen, accompanied by perivascular inflammation observable because of white/yellow plaques of slightly granular exudate on kidneys, splenic capsule and virtually any other organ. In the dry form of the disease, there is little to no effusions, the inflammatory reaction being restricted to specific sites such as kidneys, eyes or brain. The development of one form of the disease rather than the other seems to be dependent on the robustness of the host immune system. It has been reported that the wet form of FIP often develops as a last stage of the dry form, probably because of a collapse of the immune system (Pedersen, 2014). Symptoms of the dry form include fever, loss of appetite, weight loss, stunted growth, depression, and distended abdomen. Adding to this, other symptoms depend on which organs are affected by the inflammatory response (kidney, intestine, eyes, brain…etc).

The disease is invariably fatal after the apparition of certain clinical signs, and available treatments are only symptomatic and palliative. Nevertheless, spontaneous recovery has been reported in a percentage of cats (Tekes and Thiel, 2016)**.** The vast majority of cats that develop FIP die days, weeks or months after infection. There are only a few spontaneous remissions. Treatment is mainly focused on reducing inflammation, therefore corticosteroids are sometimes used although no beneficial effect was demonstrated. Immunosuppressants such as prednisone are sometimes used, but can only slow the progression of the disease (Pedersen, 2009). Ribavirin and other antivirals showed effectiveness *in vitro*, but are toxic in cats. Feline interferon was one of the treatment options, but was shown to have no beneficial effect (Ritz et al., 2007),

whereas human interferon is ineffective or contraindicated depending on the administration route. New promising antivirals such as 3C-like protease inhibitors are under development (Kim et al., 2016) but are not available for use to date.

Mutations/deletions in the 3c, 7b and spike genes have been associated with the shift from FECV to FIPV biotype (Pedersen et al., 2009; Tekes and Thiel, 2016; Vennema et al., 1998), although this is still debated (Hora et al., 2016). Generally, mutations and deletions in the 3c gene lead to a premature stop codon or a frame shift, leading to a truncated protein. Other mutations in the spike gene or the 7b accessory gene also were observed in the FIPV genome (Borschensky and Reinacher, 2014; Desmarets et al., 2016; Oguma et al., 2018).

The 3c gene encodes a small protein of yet unknown function, but it is postulated to play a role in virulence (Hsieh et al., 2013) and in viral replication in the intestines (Pedersen et al., 2012), which is consistent with both the shift to a more virulent pathogenesis and the shift of tropism from intestinal epithelia for FECV to macrophages for FIPV. It is worth noting that FECV has also been found in monocytic cells but at extremely low levels (Kipar et al., 2010). Interestingly, it was shown in experimentally infected cats, that although in most cats viral shedding persists in feces long after infection, in other cats there was a delay of shedding in feces, and detection of non-enterotropic FECV mutants in blood (likely from monocytic cells), therefore supporting the hypothesis of a gradual adaptation to replication in immune cells (Desmarets et al., 2016). The increased tropism for immune cells is accompanied by a loss of enteric tropism, probably due to observed mutations in the spike protein (Chang et al., 2012; Oguma et al., 2018).

Generally, the 3c and 7b mutations could play a role in the virulence of the FIPV biotype but could also be consequences of the systemic infections with FIPV, rather than causes of the biotype switch (Tekes and Thiel, 2016).

On the other hand, the change of tropism from enterocytes for FECV to enterocytes and macrophages for FIPV is likely due to mutations in the spike protein, considering its major role in entry and tropism. Interestingly, two point mutations located in the putative fusion peptide were identified in the S gene that can distinguish the vast majority of FIPVs from FECVs (Chang et al., 2012). Additionally, it was demonstrated that all FECVs had a conserved furin cleavage site located at the S1/S2

interface, and that mutation at this junction arises during development of FIP (Licitra et al., 2013).

There seems to be an important selection pressure upon FECV infection *in vivo*, thereby leading to the emergence of a lot of quasispecies of the virus in each cat. The virus evolves rapidly, accumulating FIPV-specific mutations that arise gradually, and ultimately lead to the emergence of the FIPV biotype.

One modified-live non-adjuvanted intranasal serotype II vaccine against FIPV has been licensed in 1991 (Gerber et al., 1990). This vaccine was obtained from serial passages in cell culture and UV irradiation of a virulent DF2-FIPV strain. In that study, it gave protection upon homologous challenge in cats that are naïve for FCoV, with 85% of vaccinated cats surviving FIPV infection when only 17% of unvaccinated cats did (although the study was done on a small population). The vaccine is licensed for kittens older than 16 weeks of age. This might explain in part the failure of the vaccine in the field, since kittens born in settings where FCoV is endemic become infected before 16 weeks of age. Since the vaccine is ineffective in cats that were already exposed to FCoV, antibody testing should be performed before vaccination, which makes its use more difficult. That is why the vaccination for FIPV is currently not recommended.

1.2.3 Infectious bronchitis virus (IBV)

Infectious bronchitis virus (IBV) is a *gammacoronavirus* that can infect wild birds (Domanska-Blicharz et al., 2014; Hepojoki et al., 2017; Muradrasoli et al., 2010) and domestic birds, both galliform and non-galliform. Domestic chicken (*Gallus gallus*) and pheasants (*Phasianus* spp) are considered to be the naturals hosts of IBV. However, IBV-like viruses have been identified in a number of other species including geese, peafowl, teal, pigeon, turkey, penguins, quail, duck, and Amazon parrot (Circella et al., 2007; Dea and Tijssen, 1989; Jonassen et al., 2005), infection being asymptomatic in some of these species. All of these viruses were regrouped with IBV strains under the subgenus *Igacovirus* and the species *Avian coronavirus* in *gammacoronaviruses* (ICTV 2018). Considering its omnipresence in both domestic and wild birds (the latter likely playing a role in its propagation), it is not surprising that IBV is distributed globally, and is especially prevalent in domestic fowl in countries with a developed poultry industry.

IBV can infect chickens of all breeds and ages, as well as both sexes, but is more severe in baby chicks in which the mortality rate ranges from 25 to 30%. It can increase to up to 80% depending on the host status (age, immunity), the virulence and tropism of the serotype (nephropathogenic strains being generally more deadly), and the environment (temperature, stress) (Bande et al., 2016). Moreover, co-infections with other viruses or bacterial secondary infections can also cause fatal outcomes (Sid et al., 2015).

The virus is extremely contagious (morbidity up to 100%), and likely spreads mostly by aerosol, since it can be detected in the respiratory tract up to 28 days after infection. Nevertheless, because of a high viral shedding in feces (virus could be detected up to 20 weeks after infection (Alexander and Gough, 1977)), the importance of the fecal-oral transmission should not be overlooked (Cook and Mockett, 1995). A few studies also looked at vertical transmission of the virus, and researchers could recover the virus from eggs laid by IBV-infected hens, and from semen of infected cockerels (Cook, 1971; Cook and Garside, 1967). Although interesting, this type of transmission is likely negligible in regard of the lateral transmission of the virus.

The incubation period is rather short, but variable depending on the strains and route of infection (18-36h). The virus primarily infects the upper respiratory tract, and then spreads to the kidney, reproductive tract, and other organs (Ignjatovic and Sapats, 2000). The clinical signs include sneezing, coughing, weight loss, and nasal discharge (Bande et al., 2016). The infected birds look lethargic, with ruffled feathers. Some other clinical signs can appear, depending on the strain tropism: nephropathogenic strains will provoke wet droppings and excessive water intake, whereas infection of the reproductive tract will cause lesions of the oviduct and therefore misshaped or rough-shelled eggs. IBV can be recovered from many organ tissues, not always with associated clinical signs. The ability of the virus to replicate at many different sites may be due to the attachment of the virus to host cells being dependent on sialic acids (neuraminic acid) (Schultze et al., 1992).

IBV is one of the most studied coronaviruses since its discovery in the 1930s in the USA (Schalk and Hawn, 1931), because of the economic losses that it causes to the poultry industry worldwide. Indeed, in addition to a high mortality rate in broiler chickens, infected chickens display a stunted growth and loss of weight. The infection

also diminishes the quantity and quality of the egg production, because of its impact on the reproductive tract. Antibodies to IBV were detected in humans working in close contact with poultry, but the virus did not cause any disease (Miller and Yates, 1968).

In some locations, the incidence of infection approaches 100% in chickens, despite a number of available vaccines. The variety of antigenic strains (mostly due to variability in the spike protein) and the low cross-protection between vaccines explains this low effectiveness (Cavanagh, 2007). Moreover, there is a continuous emergence of new IBV serotypes: these new strains emerge after selection pressure, which means that they possess new features such as immune response evasion, increased virulence or diversified tropism. Indeed, the conditions for IBV strains in the field are favorable to recombination events: very high number of chickens, often kept in high-density settings, co-circulation of different strains… etc. Thus, developing effective vaccines against IBV strains has been, and still is, a big challenge. Live attenuated vaccines induce a strong immune response, but raise concerns because of the risk of mutations and reversion to virulence. On the other hand, inactivated vaccines seem to be safer, but induce a weaker immune response that needs to be compensated by multiple dosing and/or formulation (Bande et al., 2015).

1.2.4 Murine hepatitis virus (MHV)

The *betacoronavirus murine hepatitis virus* (MHV) was first isolated in 1949 (Cheever et al., 1949) and after electron microscopy became available it was observed that the virus was morphologically similar to IBV (David-Ferreira and Manaker, 1965). This first isolated strain was JHM, named after J. Howard Mueller, a neurotropic strain causing encephalomyelitis and hepatitis in mice. During the following years, other MHV strains with different tropisms were identified (Gledhill and Andrewes, 1951; Manaker et al., 1961).

We now know a variety of MHV strains with a tropism falling into one of two biotypes: strictly enterotropic or polytropic. Polytropic strains commonly cause hepatitis in mice, although they usually infect primarily cells of the upper respiratory tract, and then spread to tissues like the liver, intestine, vascular endothelium or the central nervous system (Barthold and Smith, 1984). MHV enterotropic strains are highly

contagious. These viruses can infect enterocytes and are mostly asymptomatic among adult mice, but lethal in neonates (in 24-48h) with a mortality rate close to 100%. However, when the virus became endemic in a population, there is no apparent clinical disease as neonates are then protected via maternal immunity (Barthold and Smith, 1984). Polytropic strains are less contagious, and require close contact in order to spread from mouse to mouse. The virus usually enters and primarily replicates in epithelial cells of the nose, and neurotropic strains can sometimes directly spread to the olfactory tract of the brain. In most cases, the virus disseminates through blood to the lungs and often to liver, hemopoietic and lymphoid tissues. The clinical signs depend on numerous factors such as age, mouse and virus strain (Taguchi et al., 1976). The infection is mostly unapparent in immunocompetent adult mice, but can lead to neurologic, hepatic or enteric clinical signs in immunodeficient mice (Sebesteny and Hill, 1974).

MHV is one of the most extensively studied coronaviruses, and is widely used as a model for molecular biology of coronaviruses, but also for diseases such as hepatitis and multiple sclerosis. The virus is still present in laboratory mice around the world but with a low prevalence (Pritchett-Corning et al., 2009). Although the disease is mostly asymptomatic in immunocompetent adult mice, it used to cause problems for research purposes because of the induced modulation of the immune system, especially the T cell response (Bergmann et al., 2001). Nevertheless nowadays, rigorous hygiene and containment procedure in laboratory settings largely prevent infection with MHV.

1.3 Human coronaviruses

1.3.1 Endemic human coronaviruses

Four coronaviruses are known to be endemic in humans, and to cause mild common cold-like respiratory symptoms. Among these, HCoV-229E and HCoV-OC43 are known since the 1960s, whereas HCoV-NL63 and HCoV-HKU1 were discovered respectively in 2004 and 2005, in the aftermath of the SARS-CoV outbreak. HCoV-NL63 and HCoV-229E are classified as *alphacoronaviruses* and HCoV-OC43 and HCoV-HKU1 as *betacoronaviruses.* These four viruses are distributed globally, and are associated with mild respiratory symptoms in immunocompetent adults, but can be more severe in children, elderly, and immunocompromised patients (Glezen et al., 2000; Jevšnik et al., 2012; Walsh et al., 2013).

Symptoms of HCoV-OC43 and HCoV-229E in adults are mostly common cold-like and include general malaise, headache, nasal discharge, sneezing, a sore throat and sometimes cough and fever. It was shown in experimentally infected individuals that HCoV-229E causes common cold in 50% of cases (Bradburne et al., 1967). Concerning HCoV-NL63, common symptoms in adults are cough, rhinorrhea, fever, and hypoxia. Obstructive laryngitis can also be observed, particularly in children. Even though the virus was detected in patients of all ages, it seems to be more associated with children, since most infected patients are under 5 years-old (Hoek et al., 2006). Interestingly, it was first isolated in nasopharyngeal aspirates of a 7-month-old child presenting bronchiolitis and fever (van der Hoek et al., 2004). The virus was also detected in samples dating back to 1988, from a child with pneumonia (Fouchier et al., 2004), suggesting that it has been circulating in the human population for a while. HCoV-HKU1 was discovered in Hong-Kong shortly after HCoV-NL63, and has roughly the same clinical presentation. The virus usually causes upper respiratory infection, but pneumonia or acute bronchiolitis were also observed (Lau et al., 2006). Considering the symptoms they provoke, these human coronaviruses are clinically indistinguishable from rhinoviruses or influenza viruses.

HCoV-OC43, HCoV-229E, HCoV-NL63 and HCoV-HKU1 all circulate particularly during winter in temperate climate countries, (Hendley et al., 1972) although one

epidemic of HCoV-NL63 was also detected in spring and summer in Hong-Kong (Chiu et al., 2005). These human coronaviruses are well adapted to humans and circulate worldwide, and are estimated to be responsible for 10-30% of common colds. Therefore, they have an important economic impact because of the burden they represent to public health, and generally the loss of productivity that they cause. Moreover, these viruses are still a threat to children, elderly and immunocompromised people.

1.3.2 Emergent highly pathogenic coronaviruses

In recent years, two highly pathogenic human coronaviruses emerged in the human population, causing epidemics of severe pneumonia mostly in China and in the Middle East.

1.3.2.1 *Severe acute respiratory syndrome coronavirus* (SARS-CoV)

First, in November 2002, an epidemic of severe acute respiratory syndrome occurred in the Guangdong province in the south of China, caused by a new *betacoronavirus* named SARS-CoV (*Severe acute respiratory syndrome coronavirus*) (Drosten et al., 2003; Peiris et al., 2003). The virus rapidly spread to 26 countries causing more than 8000 cases with a mortality rate of 10%. All cases outside of China were imported, and since most of them were rapidly quarantined, the virus did not spread further in these countries. There was human-to-human transmission only in a few other countries, including Canada, Vietnam, Hong-Kong and Singapore (WHO).

The epidemic was particularly alarming because the virus readily spread from human-to-human, mostly in healthcare settings. Rigorous quarantine of infected patients and strict hygiene measures for healthcare workers allowed containing of the epidemic, which ended in July 2003. Since then, only four primary cases of SARS were recorded: three were due to laboratory accidents (Senio, 2003), and one was probably due to zoonotic transmission regarding the circumstances, but this was not confirmed (WHO).

Since SARS-CoV demonstrated strong human-to-human transmission, it is considered as a virus with an extremely high epidemic potential. Human-to-human transmission probably mostly occurs via respiratory droplets. Airborne transmission

was also suggested, but is not the main transmission route. The virus has been detected in respiratory secretions, feces, urine and tears of infected patients.

The incubation period of SARS ranges from 2 to 14 days (mean 6 days), and the patients are more infectious at five days or more after the onset of disease, which corresponds to the peak of viral loads in secretions. The estimated number of secondary cases from an infected patient is 2.2 to 3.6, which corresponds to an epidemic capable of infecting the entire population if left unchecked (R_0>1) (Lipsitch et al., 2003).

Though these estimations underestimate the impact of the superspreading events, which can have dramatic consequences on an epidemic. Superspreaders are patients that infect a disproportionate number of contacts when compared with other patients, and this phenomenon was observed for several infectious diseases over history (Stein, 2011). Nevertheless, the underlying mechanisms that allow superspreaders to infect so many people are still unknown. Several superspreading events were reported for SARS-CoV. First, in Singapore, only five primary cases were responsible for the infection of 103 secondary cases (Centers for Disease Control and Prevention (CDC), 2003) whereas 81% of patients had no evidence of infecting others. In Hong Kong, 138 people were infected in a teaching hospital, all cases tracing back to only one index patient (Wong and Hui, 2004). Two other superspreading events occurred in Hong-Kong, one causing 13 cases in a hotel, and another more than 180 cases in a housing complex. Lastly, one superspreading event took place on a flight between Hong-Kong and Beijing, causing 22 cases. Interestingly, one case from the hotel was subsequently responsible for the spread of the virus to Canada, and this resulted in 128 cases in a Toronto hospital (Wong et al., 2015). Superspreading events therefore have a huge impact on the spread of a pathogen. Importantly, this was observed in the recent years for both SARS-CoV and MERS-CoV, and to a lesser extent for Ebola virus, but also less recently for tuberculosis or measles. For SARS and MERS, superspreading events contributed massively to the onset of the epidemics, and to their worldwide spreading.

Clinical presentation of SARS corresponds to viral pneumonia: fever, myalgia, general malaise, headache and cough (Cheng et al., 2007; Tsang et al., 2003). Rhinorrhea and sore throat can also be observed. There is often a clinical deterioration one week after the onset of symptoms, marked by diarrhea (Cheng et al., 2004). More severe cases developed severe pulmonary lesions and acute respiratory distress syndrome. An average 10% of infected patients died of the disease, but this is highly dependent on age

and conditions of the patients. The estimated fatality rate for patients under 60 years is around 7-13%, but for patients over 60 years it goes up to 43-55% (Donnelly et al., 2003). Generally, the number of cases was much higher in elderly people than in young people. This is probably linked to the fact that the most important determinant of the infection outcome is the presence of comorbidities, such as heart or respiratory conditions.

1.3.2.2 *Middle East respiratory syndrome coronavirus* (MERS-CoV)

In April 2012, another highly pathogenic coronavirus caused an epidemic of severe pneumonia in the Middle East. The virus, thus named *Middle East respiratory syndrome coronavirus* (MERS-CoV), was isolated in Saudi Arabia in a 60 years old patient presenting a severe acute respiratory syndrome and renal failure, who died of the infection (Zaki et al., 2012). The clinical presentation was highly similar to the one of SARS-CoV, but the agent was identified as a new *betacoronavirus* (at first named HCoV-EMC).

Typical symptoms include fever, cough and sometimes pneumonia. Gastrointestinal symptoms, particularly diarrhea, have been reported (WHO). It is believed that a number of MERS-CoV cases are asymptomatic, however it has not been confirmed yet. More severe clinical presentation is seen in the elderly, immunocompromised patients and patients with comorbidities (heart, kidney or pulmonary condition, diabetes, hypertension… etc.). A study in Saudi Arabia strongly linked severe cases of MERS-CoV with these comorbidities (Assiri et al., 2013). In these severe cases, respiratory distress, kidney failure, septic shock and ultimately multi-organ failure leading to death have been reported. The incubation period ranges from 2 to 13 days, with a mean of five days.

The virus spread from Saudi Arabia to other countries of the Middle East including Jordan, Iran, United Arab Emirates, Oman… etc. A few cases were reported in Europe and in the US, but all were imported cases with a history of travelling to the Middle East (Figure 2). In most cases, the patient was strictly quarantined and the virus didn't spread further. The MERS-CoV epidemic is still ongoing, counting 2 428 cases and 838 deaths as of July 8th 2019, bringing the mortality rate to 35%. Even though 27

countries reported MERS-CoV cases, Saudi Arabia still concentrates 80% of total cases. MERS-CoV does not seem to spread readily from human-to-human with a $R_0<1$, unlike SARS-CoV.

The MERS-CoV epidemic containment relies essentially on cases identification and quarantine. Indeed, a large outbreak in South Korea in 2015 showed that the virus still has epidemic potential if left unchecked. In May 2015, a 68 years-old patient returning from the Middle East was diagnosed with MERS-CoV in South Korea. From that one index patient, twenty-nine secondary infections were reported, and two of these were subsequently responsible for 106 other infections. In total, more than 180 cases were declared in the country, and 75% of these cases could be traced back to the three superspreaders. During this outbreak, most infections occurred in the hospital, a phenomenon probably enhanced by the multiple changes of healthcare institution of the patients (Lee and Wong, 2015).

For human-to-human transmission, close contact seems to be necessary. Logically, secondary cases are usually, as for SARS-CoV, family members, healthcare workers and/or other patients in the same healthcare settings. Surprisingly, during the South Korea outbreak, close contact with an infected patient occurred in only 10% of cases. On the other hand, sharing the same healthcare environment was a major determinant for transmission. Therefore, it has been postulated that fomite transmission was a possible route of MERS-CoV spread, and the virus was subsequently shown to survive on inanimate surfaces up to 48h (van Doremalen et al., 2013), thereby explaining contamination without close contacts in healthcare settings (or on aircrafts as observed for SARS-CoV (Olsen et al., 2003)). The main transmission routes of MERS-CoV are therefore supposed to be respiratory droplets upon close contact (through sneezing and coughing) and fomites.

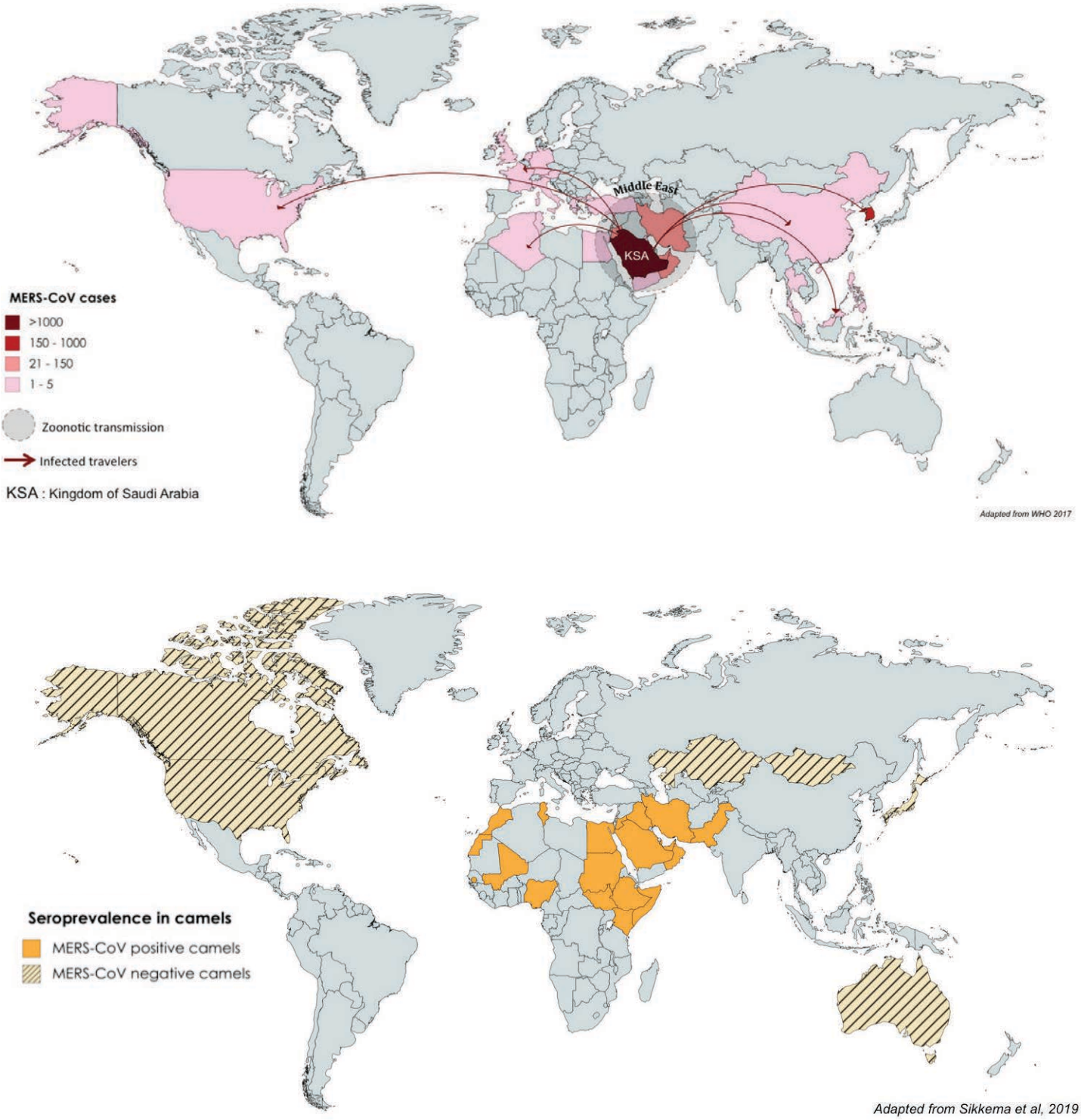


Figure 2 – *Geographic repartition of MERS-CoV cases and seropositive camels*

Upper panel - Most MERS-CoV cases are restricted to the Middle East and South Korea, with only sporadic imported cases in the rest of the world. *Lower panel* - Continuous zoonotic transmission can still occur in the Middle East as the virus circulates among camels in that area. Camels throughout Africa were also shown to be seropositive for MERS-CoV.

1.3.3 Treatments and vaccines

1.3.3.1 Treatments

There is neither vaccine nor specific antiviral agent currently available to treat coronaviruses. The treatments for SARS-CoV, MERS-CoV and severe cases of other human coronaviruses are therefore mostly symptomatic.

In the urgency of the SARS-CoV epidemic, a number of treatments were administered without previous proof of efficacy against the virus. Notably, ribavirin was widely used because of its wide spectrum of antiviral activity, but caused serious adverse effects (anemia or bradycardia for example) without really improving the clinical presentation. It was often administrated along with corticosteroids, which were used to mitigate the damaging effect of inflammation and to treat the respiratory distress syndrome. Used alone, corticosteroids seemed to worsen the clinical presentation rather than improve it, with higher SARS-CoV RNA levels and increased 30-day mortality rate in patients treated with methylprednisolone (Zumla et al., 2016). On the other hand, the use of IFN-α in combination with corticosteroids was shown to be more efficient than corticosteroids alone, notably improving oxygen saturation and recovering from lung opacities, but this was in an uncontrolled study (Loutfy et al., 2003). The combination lopinavir/ritonavir, classically used in the treatment of HIV, proved somewhat efficient in the treatment of SARS-CoV (Lai, 2005) if used besides the ribavirin/corticosteroid combination. In a few cases, treatment with convalescent plasma was also performed and showed a reduction in mortality among treated patients, but this was dependent on the availability and the time of administration (Zumla et al., 2016). Interestingly, traditional Chinese medicine was also used in some areas, and the combination of western medicine and traditional Chinese medicine showed better results than western medicine alone (Lai, 2005; Lin et al., 2003), although it is difficult to assess its impact because of varying conditions.

Similarly, ribavirin was used during the MERS-CoV outbreak and showed no improvement of clinical outcome, even if combined with IFN α2a, IFN α2b, or IFN β1a. A combination of ribavirin, lopinavir/ritonavir and IFN α2a proved effective in one patient with severe MERS, who had resolved viremia two days after the onset of treatment

(Spanakis et al., 2014). The use of corticosteroids showed no improvement of the clinical outcome regardless of the combination with antivirals or IFNs.

Importantly, as these treatments were administered in the urgency of the epidemics, their efficacy was not assessed in controlled randomized studies.

New treatment approaches are in testing since the aftermath of the SARS epidemic and the emergence of MERS. Nevertheless, all of these antivirals were only tested *in vitro* or *in vivo* in animals, and their efficacy in humans is left to prove.

Despite their high diversity, coronaviruses share essential elements that are therefore main targets for antiviral therapy. Among these, the replication complex, and notably the 3C-like protease is an interesting target, and 3CL protease inhibitors are currently in testing (Konno et al., 2016; Kumar et al., 2017), with among them lopinavir. Lopinavir is a protease inhibitor used in the treatment of HIV, which previously showed beneficial effect in SARS patients. Lopinavir is used in combination with ritonavir that acts as an enhancer, and they inhibit HIV replication by forming inhibitor/enzyme complexes with the aspartyl protease of HIV. Interestingly, it was also shown that lopinavir and ritonavir can both bind to the 3CLpro of SARS-CoV and inhibit its activity (Nukoolkarn et al., 2008; Zhang and Yap, 2004). The efficacy of the combination lopinavir/ritonavir and IFN-β1b is currently under investigation in a phase 2 clinical trial (ClinicalTrials.gov Identifier: NCT02845843). Inhibitors of the RNA-dependent RNA polymerase (RdRp) (Peters et al., 2015) or the helicase (Adedeji et al., 2012) are also in testing. Targeting RNA synthesis, and formation of double-membrane vesicles (DMVs) is also interesting, and a promising compound named K22 was shown to inhibit six coronaviruses (HCoV-229E, SARS-CoV, MERS-CoV, MHV-A59, IBV-Beaudette and FCoV) and seems to have a broad range of antiviral activity against Nidoviruses (Lundin et al., 2014; Rappe et al., 2018).

Another key target on coronaviruses is the surface spike protein, because of its major role in entry and immune response. Numerous approaches are in development to inhibit either S protein binding to its receptor (antibodies against RBD, Griffithsin) or fusion of viral and cellular membranes mediated by the domain S2 (peptides). A limitation of this approach is the narrow spectrum resulting of the high variability of the spike protein, and especially the S1 domain. Interestingly, Griffithsin is a lectin, which displays a broad range of activity because of its binding to N-glycans on the heavily

glycosylated S protein, rather than to specific sequences in the RBD or other (O'Keefe et al., 2010).

Small interfering RNAs (siRNAs) are also an option, aiming to interfere with the expression of viral proteins, but the optimal delivery method in humans is still uncertain, and the spectrum of activity is relatively narrow (Zumla et al., 2016).

Using repurposing of existing drugs, numerous host-based antivirals were discovered to have an activity against coronaviruses *in vitro* or in animal models. Among these, it was shown that some coronaviruses are inhibited *in vitro* by nitazoxanide, which potentiates the innate immune response and particularly IFN production (Rossignol, 2016), chlorpromazine that inhibits clathrin-mediated endocytosis (de Wilde et al., 2014), host protease inhibitors that block the fusion step or spike protein cleavage (cathepsins, TMPRSS2 or furin inhibitors (Zumla et al., 2016)) or cyclosporine A that inhibits cyclophilins, which are necessary to the replication of various RNA viruses (Li et al., 2018). Targeting host-cell receptors using specific peptides or antibodies in order to block S-mediated binding or fusion is also in testing.

Efforts to develop antivirals against coronaviruses are focused on SARS-CoV and MERS-CoV, because of the high pathogenicity of these viruses for humans, and their high epidemic potential. Nevertheless, antivirals against less pathogenic human coronaviruses are also needed, especially for treatment of severe infections in susceptible patients. More generally, because of the emerging potential of coronaviruses and the currently circulating viruses, there is an urgent need for antivirals against coronaviruses.

1.3.3.2 Vaccines

Logically, research efforts are also mostly focused on the development of vaccines against SARS-CoV and MERS-CoV. In order to design efficient vaccines against coronaviruses, it is important to know which type of immune response is required against the virus. For coronaviruses, a combination of virus-specific humoral and cellular immune response seems to be necessary for protection (Cho et al., 2018). It was shown for SARS-CoV that T-cell immunity is a crucial determinant of disease outcome (Liu et al., 2017). Early work on MERS-CoV mice models also showed that the T cell response is necessary for viral clearance, and that the disease was more severe in absence of a type I IFN response (Zhao et al., 2014).

All data point to a crucial role of neutralizing antibodies in disease outcome. For example, administration of convalescent camel plasma to MERS-infected mice (harboring hDPP4) decreased lung pathology and enhanced virus clearance, with an efficacy directly linked to the quantity of neutralizing antibodies (Zhao et al., 2015). Moreover, it was found that serum antibody responses in patients were higher in survivors than in fatal cases of MERS-CoV (Min et al., 2016), and convalescent plasma was used to treat a few patients during the SARS epidemic (Cheng et al., 2005; Yeh et al., 2005). Nevertheless, benefits of administration of convalescent plasma to MERS-CoV patients is left to be proven in controlled clinical trials, adding to the fact that obtaining efficient neutralizing sera is difficult (Ko et al., 2018; Okba et al., 2017). Still, this type of passive immunization remains an option, until neutralizing antibodies or vaccines become commercially available. Importantly, one anti-MERS-S antibody (SAB-301) recently completed its phase I clinical trial, and seems to be safe in healthy adults (Beigel et al., 2018).

Many vaccines against SARS-CoV or MERS-CoV are currently under development.

Live attenuated or inactivated vaccines are interesting because they preserve the structure of the virus and the immunogenic components, thereby triggering a strong innate and adaptive immune response. Indeed, antibodies against at least eight different proteins were detected in SARS patients' sera (Guo et al., 2004). Nevertheless, safety issues limit the use of live attenuated vaccines, namely the risk of viruses reverting to virulence through mutation and/or recombination. Moreover, the recombination events are especially common in coronaviruses evolution, because of their unique replication strategy based on subgenomic RNA production. Interestingly, a team engineered two SARS-CoV viruses with a "rewired" transcription regulatory network (TRN), thereby modifying the TRS-based replication step. The first modified virus had 3 of the 6 nucleotides of the TRS mutated (3nt-TRN), and was shown to be attenuated and relatively incompetent for recombination with wild-type virus. Nevertheless, this virus reverted via mutation after serial passages in mice. Therefore, a second virus was engineered with the six nucleotides of the TRS mutated, and one nucleotide added, thereby creating a new 7nt TRS sequence. In addition, the reverting mutations yielded on the 3nt-TRN virus after passage in mice were also inserted in this virus. This virus was also attenuated, recombination incompetent with the wild type, and more stable

upon serial passages (Graham et al., 2018; Yount et al., 2006). If proven stable and safe, this modified virus could make an interesting vaccine candidate.

Other attempts at engineering live vaccines targeted the E protein, generating SARS-CoV-ΔE and MERS-CoV-ΔE. Both were attenuated, and SARS-CoV-ΔE was shown to be protective in hamsters (Lamirande et al., 2008). MERS-CoV-ΔE was propagation defective, and replicated at lower titers than the wild type. The development of these vaccine candidates was not pursued further. In the same fashion, SARS-CoV and MERS-CoV lacking nsp14 (exonuclease), nsp16 (involved in RNA capping), or accessory proteins were produced, showed protection in humanized mice, but were not developed further (Schindewolf and Menachery, 2019).

Inactivated vaccines are interesting because they are in theory easily prepared, and should induce a strong immune response. Many inactivated vaccines were produced with different techniques, but proofs of efficacy against live SARS-CoV challenge are rare: in mice, such a vaccine protected against pulmonary disease and induced antibody production (Stadler et al., 2005), but since mouse models usually harbor a mild disease, it is questionable whether this vaccine would be efficient for hosts with a more severe clinical presentation. Another inactivated vaccine was also tested in ferrets, with only mild protection. An inactivated vaccine was tested on 36 patients, and proved to be safe and immunogenic (Lin et al., 2007), but in absence of a subsequent challenge with live SARS-CoV, there is no data on efficacy.

Moreover, inactivated SARS-CoV vaccines raised concerns because it was observed that eosinophil-related lung pathology appeared in vaccinated animals (Bolles et al., 2011). The same observation was made after vaccination with an inactivated MERS-CoV (Agrawal et al., 2016). This reduced the interest in that type of vaccine candidates, even though other vaccine formulations are developed to avoid eosinophil-related lung pathology (Schindewolf and Menachery, 2019).

Apart from whole-virus vaccines, subunit, recombinant or DNA vaccines are interesting tools. Generally, the S protein of coronaviruses is considered as a good vaccine candidate because it is highly immunogenic and is therefore the main target of neutralizing antibodies. For these reasons, many types of vaccines based on the S protein are currently in development. Interestingly, the N protein is also used, as it is

more conserved than S across CoV genera, and was shown to induce cell-mediated immunity (Schindewolf and Menachery, 2019).

Subunit vaccines are mostly focusing on the receptor-binding domain (RBD) of the S protein. Intranasal administration of MERS-CoV RBD protein in mice induced neutralizing antibodies and an important cell-mediated response. This RBD was then fused to a part of human IgG1 to increase its half-life, and was ultimately engineered to fold into trimers. This vaccine was shown to give long lasting protection to hDPP4-mice upon challenge (Tai et al., 2016). Using the whole S protein is also interesting because of the presence of other epitopes outside of the RBD. However, despite protection against challenge after vaccination in hamsters, antibody-dependent enhancement (ADE) of infection was observed with SARS-CoV i*n vitro* (Kam et al., 2007). Although this phenomenon has not been reported for MERS-CoV yet, it could be an issue.

DNA vaccines are immunogenic and safe, with only mild adverse effects after injection with electroporation. They are capable of inducing cellular and humoral immunity in mice (some in non-human primates), and they allow for the conservation of native post-translational modifications.

A SARS-CoV DNA vaccine based on a vector expressing the full-length S protein (VRC SARS DNA) was tested in ten healthy adults and was shown to be safe and immunogenic in a phase I clinical trial. Similarly, a DNA vaccine based on an optimized plasmid coding for the S protein of MERS-CoV completed its phase I clinical trial (Muthumani et al., 2015). This vaccine induced both cellular immunity and neutralizing antibodies in mice, macaques, and camels, and showed protection against MERS-CoV challenge in vaccinated rhesus macaques. Other similar DNA vaccines with the full length S protein expressed via a different vector are also in development (Schindewolf and Menachery, 2019). Interestingly, a DNA vaccine based on the S1 domain is also under investigation. It induced higher levels of neutralizing antibodies when compared with the full-length S expressed in the same vector (Al-Amri et al., 2017; Chi et al., 2017).

Viral vector vaccines, using the backbone of an attenuated virus expressing one or more immunogenic proteins of the targeted virus, are also in development. This approach induces a robust immune response, enhanced by both cellular entry and the viral backbone components.

Vaccines based on human adenoviral vectors were developed for MERS-CoV, and were proven to be efficient in mice. However, their efficacy in humans may be reduced by pre-existing immunity against adenoviruses (Schindewolf and Menachery, 2019). In order to bypass this problem, chimpanzee adenoviruses were also developed as vaccine platforms. Hence, a vaccine based on the replication-deficient simian adenovirus vector ChAdOx1, containing the MERS S protein (ChAdOx1 MERS) proved its efficacy in mice and is currently undergoing a phase I clinical trial (Munster et al., 2017).

A recombinant MERS vaccine based on a modified vaccinia virus Ankara (MVA) vector expressing the MERS-CoV spike (MVA-MERS-S) was demonstrated to be immunogenic in mice, and has been safely evaluated in camels and humans along with ChAdOx1 MERS (Alharbi et al., 2017). MVA-MERS-S is also currently undergoing a phase I clinical trial.

Another recombinant vaccine was engineered using a live attenuated measles platform harboring the SARS-CoV S protein, and showed full protection against SARS-CoV challenge in mice (Escriou et al., 2014). Similar vaccines expressing the S or N protein of MERS-CoV were later developed and proved efficient in mice (Bodmer et al., 2018).

Finally, virus-like particles (VLPs) are also an interesting vaccine candidate, as they can harbor different viral proteins while conserving the virion structure. Most tests were done using the baculovirus expression system for production of the VLPs. VLPs were produced via infection of insect cells with a recombinant baculovirus co-expressing E, M and S of MERS-CoV, and were injected into rhesus macaques. This vaccination induced the production of neutralizing antibodies, and a Th1-mediated immune response (Wang et al., 2016).

In conclusions, many vaccines produced by different approaches have been tested for both SARS-CoV and MERS-CoV, and the most advanced vaccine candidates showed promising results in the first clinical trials. Nevertheless, problems such as protection of the elderly and immunocompromised patients, along with the risk of immunopathology induced by vaccination still need to be addressed.

1.4 Cross-species transmission

70% of newly emerged human pathogens come from animals (Jones et al., 2008), making the host jump a major mechanism leading to new human diseases. Host jump is essentially based on equilibrium between conserving key components and making sufficient changes to overcome the species barrier.

RNA viruses in general have an extremely high mutation rate, and are therefore evolving very fast. Nevertheless, coronaviruses possess a complex replication machinery, which performs proofreading, allowing them to replicate their large 30kb genome without losing essential features to mutation. On the other hand, mechanisms such as gene duplication, alternative open reading frames and mostly recombination events are common (Menachery et al., 2017a).

As a result of both their unique replication strategy and their large genome, coronaviruses have a high frequency of recombination, allowing them to switch and adapt even faster to new hosts (Graham and Baric, 2010). Indeed, the major reason for the high rate of recombination observed in coronaviruses may be the production of subgenomics RNAs during replication: template switching is most likely to happen for closely related genes, therefore increasing the chances for homologous recombination with other viruses.

These characteristics explain the broad host range of coronaviruses and the recurrent cross-species transmission events. This was illustrated particularly in the last decades, with regular spillovers between animals, and from animals to humans. An overview of coronaviruses hosts and supposed/demonstrated cross-species transmissions is presented in Figure 4.

1.4.1 Zoonotic transmission

Concerning human pathogens, we recently saw the emergence of the two highly pathogenic human coronaviruses SARS-CoV and MERS-CoV, respectively in 2002 and 2012. Both of theses viruses likely originated from bats, but it was shown that the viruses did not jump directly from bats to humans, but went through an intermediate host. Hosts and transmission routes of SARS-CoV and MERS-CoV are summarized in Figure 3.

For SARS-CoV, it was observed that animal food handlers who had close contact with wild animals were over-represented in early SARS-CoV cases, which led to the idea that the virus was of zoonotic origin. Masked palm civets were rapidly identified as a possible source of the virus, as a SARS-CoV-like virus was isolated from palm civets (SARSr-CiCoV), with 99.8% nucleotide identity to the human strain (Guan et al., 2003). This close proximity with the human strain, and also between palm civet strains, indicates that the virus has not been circulating for a long time in the palm civets population. Although 80% of palm civets present on live markets during the epidemic tested positive for SARS-CoV, most palm civets on farms were negative, as were their wild counterparts (Shi and Hu, 2008). Other carnivorous wild animals from markets tested positive for SARS-CoV, such as raccoon dogs or Chinese ferret badgers (Guan et al., 2003; Wang et al., 2005; Xu et al., 2004). Since these animals are sold in markets alongside palm civets, it confirmed that transmission of the virus to humans occurred by exposition to infected animals sold on Chinese markets. Therefore, Guangdong province is still considered as a potential location of re-emergence for SARS-CoV.

Nevertheless, the circulation of the virus among palm civets being recent, they did not constitute the reservoir of the virus. The attention therefore shifted to bats, as they are known to be the reservoir of a number of zoonotic viruses. SARS-CoV related viruses were then identified in horseshoe bats, with nucleotide identity of 87–92% (Li et al., 2005a) and these bats were considered to be the natural reservoir of the virus.

In 2013, Ge *et al.* identified two SARS-like-CoV (SL-CoVs) in bats of the Yunnan province in China. One of these strains (Rs3367) had 95% amino acid identity with SARS-CoV, which was higher than any previously identified strain, particularly concerning the RBD. Also, for the first time, they isolated a live virus from bat fecal samples: this isolate had 99,9% identity with Rs3367 and was named SL-CoV-WIV1. *In vitro* infectivity tests showed that SL-CoV-WIV1 was able to use human, civet and bat ACE2 for cell entry (Ge et al., 2013). In 2016, the same team isolated another SL-CoV from the same location as WIV1, named WIV16. This virus is even more closely related to human and civet SARS-CoV (96%) especially concerning the S gene, and could use bat, civet and human ACE2, as WIV1 (Yang et al., 2015).

Importantly, although very close to SARS-CoV, none of these SL-CoVs are considered to be the direct ancestor of the human virus, as they still have too many

differences with SARS-CoV especially in the S, ORF3 and ORF8 genes. Interestingly, by accumulating data in the same location in Yunnan, China, the team eventually identified many SL-CoVs strains, each harboring high identity with different regions of the SARS-CoV genome. These strains are thought to be the building blocks that allowed the emergence of SARS-CoV by recombination (Hu et al., 2017).

The results of these studies support the bat-origin of SARS-CoV, but also shed light on the cross-species transmission mechanism. It was believed at first that the inability of SL-CoVs to use human ACE2 was the main barrier blocking their direct spillover to humans, hence explaining the need of an intermediate host. However, the recent discovery of SL-CoVs that are able to use the human ACE2 receptor suggest that a direct bat-to-human infection is possible. Furthermore, the risk of new emergence of SL-CoVs in the human population must be taken into account for the control of public health, especially considering the large pool of bat SL-CoVs and their wide distribution.

For MERS-CoV, the first identification of the virus showed that its closest relatives were coronaviruses HKU4 and HKU5, previously isolated from two species of bats (Zaki et al., 2012). Therefore, it was soon postulated that the virus was of zoonotic origin, with bats being a possible reservoir. It was later shown that HKU4, but not HKU5, could bind human DPP4/CD26 (Wang et al., 2014), but could not enter into humans cells because human cellular proteases could not activate HKU4 spike protein (Yang et al., 2014). Another bat coronavirus even more closely related to MERS-CoV was identified in *Neoromicia capensis* bat in Africa, and named NeoCoV (Corman et al., 2014). Interestingly, this virus was identified in a region where camels are traded, and were found to be positive for MERS-CoV.

Indeed, dromedary camels were soon suspected to be an intermediate host since many primary cases were associated with contact with camels (especially many camel workers). MERS-CoV viruses were then identified in camels, and camels in both Africa and the Middle East tested largely positive for MERS-CoV serology, supporting their role in zoonotic transmission (Figure 2, lower panel (Chu et al., 2014; Müller et al., 2014; Reusken et al., 2013). This high prevalence among dromedary camels suggests that the virus has been circulating in their population for a while. Interestingly, one study analyzed dromedary camels' sera dating back to 1992-2010, in Saudi Arabia. The analyses revealed that 72-100% of dromedary camels already had anti-MERS-CoV antibodies during this period (Alagaili et al., 2014).

More strikingly, identical strains of MERS-CoV were isolated from both camels and humans who had close contact with each other (Azhar et al., 2014; Ferguson and Kerkhove, 2014), further supporting camel-to-human transmission. Transmission from camels to humans likely occurred by close contact ((Mackay and Arden, 2015), Figure 3). Additionally, the WHO recommends not consuming raw camel milk, meat or urine in case this could also be a transmission route. The disease caused by MERS-CoV appears mild in camels, with only infection of the upper respiratory tract and common-cold like symptoms (Adney et al., 2014), thereby producing droplets upon sneezing, that can play a role in transmission.

The high prevalence of MERS-CoV among dromedary camels and the close contact of the species with humans explain the recurrent zoonotic transmission, and therefore the fact that new MERS-CoV cases continue to emerge despite quarantine of human cases. In the case of SARS-CoV, without a sustained transmission from the intermediate host and since human/bat interactions are limited, there were no more human infections after the outbreak was contained by stopping human-to-human transmission (de Wit et al., 2016).

Both SARS-CoV and MERS-CoV likely originated from bats, and more generally bats are considered as an important source of zoonotic viruses. They make up 20% of mammalian diversity and have a broad geographic range, living either as isolated individuals or large colonies. Importantly, being the only flying mammals, they can span thousands of miles, thereby enhancing their relationships to other mammals and birds. The fact that bats are amongst the oldest mammals is also of importance, considering their long co-evolution with pathogens (Menachery et al., 2017a). Surprisingly, although bats are a carrier of numerous viruses, they rarely exhibit any sign of disease. Diverse reasons are hypothesized to explain this phenomenon, such as altered innate immunity, enhanced response to oxidative stress or high levels of type I interferon. It is also believed that these mechanisms are responsible for enhancing the chances for host jump and pathogenic response in other hosts.

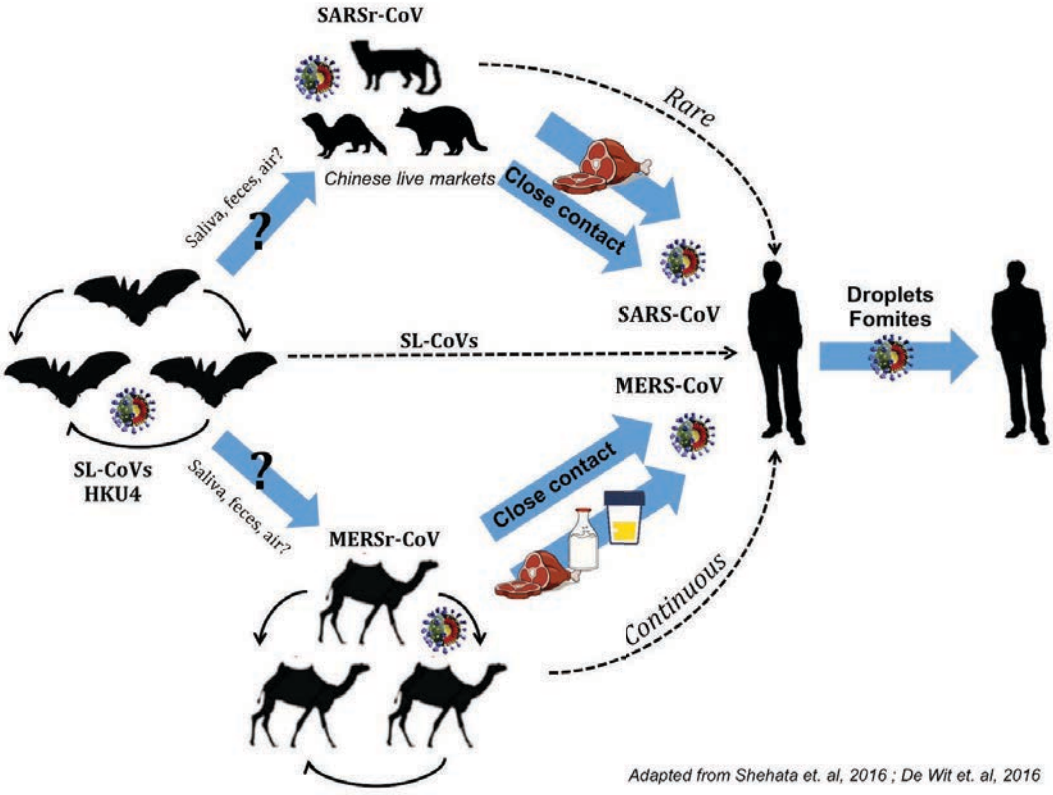


Figure 3 - *Transmission routes of MERS-CoV and SARS-CoV*

Both MERS-CoV and SARS-CoV originated from bats and were transmitted to humans through an intermediate host. For SARS-CoV, bats seem to be the actual reservoir of the virus, and transmission to intermediate host such as palm civets (and subsequently to humans) is rare. For MERS-CoV the virus circulates among the dromedary camels, which are a continuous source of infections for humans. Zoonotic transmission occurs through close contact, and possibly meat, milk or urine consumption. Human to human transmission likely occurs via droplets and fomites, but is more efficient for SARS-CoV than for MERS-CoV.

Contrary to the newly emerged SARS-CoV and MERS-CoV that do not spread readily from human to human, the four less pathogenic human coronaviruses HCoV-229E, HCoV-NL63, HCoV-HKU1 and HCoV-OC43 are considered to be endemic in the human population. In light of what was demonstrated for SARS-CoV and MERS-CoV, it is also postulated that these human coronaviruses are of zoonotic origin.

It is postulated that HCoV-NL63 originated from bats (Huynh et al., 2012), since the virus can readily replicate *in vitro* in bat cell lines. Furthermore, bat coronaviruses closely related to HCoV-NL63 were identified more recently (Tao et al., 2017) and interestingly, sequence analysis showed that HCoV-NL63 is a recombinant between NL63-like bat coronaviruses and 229E-like bat coronaviruses, with recombination mostly in the S gene. No intermediary host has been identified yet (Figure 4).

A coronavirus closely related to HCoV-229E was identified in 2007 in alpacas (Crossley et al., 2012, 2010) and other related coronaviruses were identified in bats at the same period (Corman et al., 2015; Pfefferle et al., 2009a). Although the identification of 229E-like bat coronaviruses with absence of disease suggested that these viruses were endemic in bats and thus made them a good reservoir candidate, the presence of a close alpaca-CoV was difficult to explain since they share no habitat with bats. The most probable explanation was an infection acquired from humans (Corman et al., 2018a). After the MERS-CoV outbreak, HCoV-229E related coronaviruses were also identified in dromedary camels (Corman et al., 2016), as for MERS-CoV. Both camelid and human viruses could have originated from bats, or from one another in either direction. Nevertheless, detailed analyses suggest that 229E-like viruses in camels have evolved towards the human genotype, suggesting that dromedary camels are indeed the zoonotic source of human infection.

The *betacoronavirus* HCoV-OC43 is believed to have switched from bovine to humans rather recently (Vijgen et al., 2005). Indeed, HCoV-OC43 is remarkably closely related to bovine coronavirus (BCoV), both genetically and antigenically (Brandão et al., 2006; Vijgen et al., 2006). The main hypothesis is that HCoV-OC43 originates from BCoV (Bidokhti et al., 2013), however recombination events with the PHEV pig coronavirus or CRCoV canine coronavirus likely occurred.

For HKU1, no viruses were identified in other animals that are close enough to be considered as ancestors or spillovers. Nevertheless, HKU1 is still closely related to MHV and other rodent coronaviruses, suggesting that it may originate from them (Wang et al., 2015). Furthermore, it is likely that HCoV-HKU1 originated from several recombination events with animal coronaviruses, including MHV (Su et al., 2016; Woo et al., 2005b) therefore suggesting a rodent origin for this virus (Figure 4).

More generally, lineage A *betacoronaviruses* (including HCoV-HKU1, HCoV-OC43 and BCoV) are thought to have originated from rodents, as supported by the occurrence of recombination events between HKU1 and MHV, and by the recent discovery of *China Rattus coronavirus* HKU24, which is closely related to both BCoV and HCoV-OC43 (Figure 4).

1.4.2 Animal to animal transmission

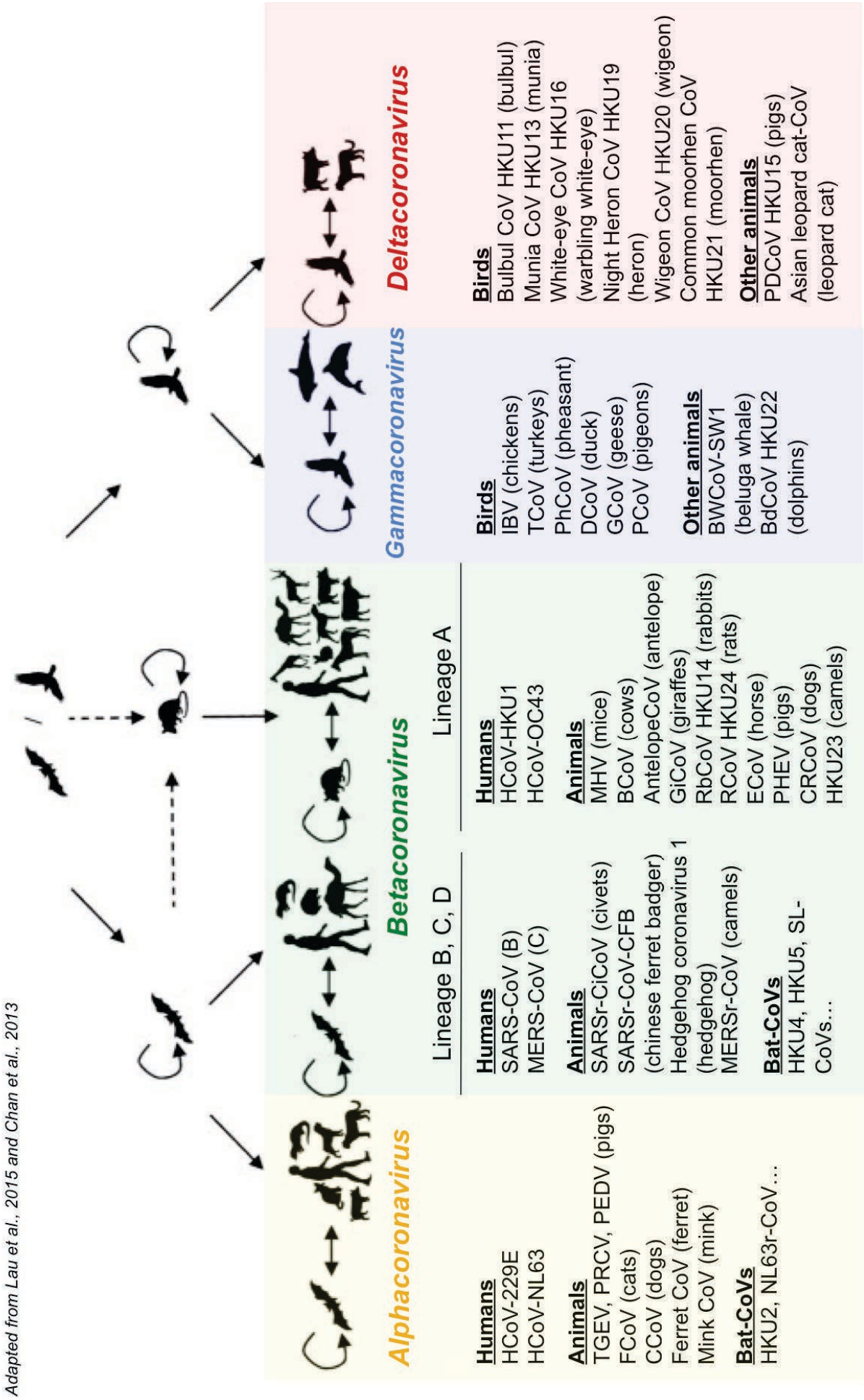
The recombination events between coronaviruses referred to earlier require a co-infection in the same host in order to take place, thus relying on close contact between animals. These events seem to have happened regularly during coronaviruses evolution, thereby allowing multiple host jumps, and broadening the coronaviruses host spectrum.

It has been shown that a recombination event occurred between FCoV and CCoV, giving rise to FCoV type II, and it has also been demonstrated that cats can be infected with CCoV under experimental conditions (Barlough et al., 1984) but it gave them no protection against a subsequent challenge to FCoV. This is not surprising considering the high identity between theses viruses, and more generally the close relationship between *alphacoronaviruses* TGEV, HCoV-229E, FCoV and CCoV. However, when cats were experimentally infected with HCoV-229E, no clinical signs nor cross protection to FCoV were detected (Barlough et al., 1985). Studies also demonstrated that TGEV, CCoV and FCoV can all infect pigs, dogs, and cats, causing variable clinical outcomes and cross protection (Saif, 2004). TGEV displayed cross reactivity with PRCV, FIPV, and CCoV in different immunoassays (Horzinek et al., 1982). Even for PEDV, which showed no cross-reactivity with TGEV or any other coronavirus, some studies reported low cross reactivity with FIPV (Zhou et al., 1988).

Another interesting example is the closely related viruses that form the *Betacoronavirus 1* species. BCoV was identified in and mostly infects cattle. But coronaviruses antigenically indistinguishable from BCoV were also isolated from wild ruminants such as sambar deer or waterbuck, and are able to infect domestic calves (Tsunemitsu et al., 1995). Interestingly other coronaviruses with high similarity to BCoV were identified in other species. One was identified in dogs with more than 98% identity to BCoV and HCoV-OC43 (Erles et al., 2003). Even more surprisingly, BCoV is able to infect and cause diarrhea in baby turkeys, but not in baby chicks in experimental settings (Ismail et al., 2001).

Once more, these cross-infections demonstrate the high potential of coronaviruses for cross-species transmission, and the long co-evolution of coronaviruses with their hosts.

Figure 4- ***Overview of Coronaviruses hosts and animal transmission***



Adapted from Lau et al., 2015 and Chan et al., 2013

1.4.3 Molecular determinants of host jump

For efficient cross-species transmission to take place, there are three pre-requisites: recognition of the receptor on the cells of the new host, efficient replication using the host cell factors, and evasion of the immune system.

The mechanism of host jump relies mainly on the spike protein (S). Indeed, the spike protein is the main determinant of host and cell tropism because of its essential role in the entry step of the viral cycle. Moreover, it is also the main target of neutralizing antibodies: conservation of antigenic sites between S proteins of different coronaviruses thus explains the cross-reactivity sometimes observed.

The S protein can be divided into S1 and S2 domains: S1 being responsible for binding to the receptor, and S2 involved in membrane fusion. Therefore, mutations in the receptor-binding domain (RBD) located in the S1 part of the protein are critical for a tropism change. Logically, S1 is highly variable across and within coronaviruses genera. On the other hand, the S2 domain is much more conserved among all coronaviruses, even if mechanisms of fusion activation can vary largely from one coronavirus to another.

In that light, cross-infection and/or cross-antigenicity between HCoV-229E, TGEV, FCoV and CCoV is due to a high identity of their S proteins, illustrated by the fact that they all use their host *aminopeptidase N* (APN) receptor upon entry (Graham and Baric, 2010). Although each of their spike proteins recognizes a different species-specific region of their APN receptors, all of these viruses can also use the feline APN for entry, and it is therefore postulated that they originate from a common ancestor. Use of conserved epitopes on receptors could therefore contribute to the cross-species infections in those cases. Similarly, HCoV-OC43 and BCoV both interact with sialic acids to allow entry into host cells (Schultze et al., 1991a; Vlasak et al., 1988).

Nevertheless, there are also examples of closely related coronaviruses that do not use the same receptor, such as HCoV-229E using APN and HCoV-NL63 using ACE2, and examples of coronaviruses from different genera that use the same receptor, such as ACE2 used by both HCoV-NL63 (α-CoV) and SARS-CoV (β-CoV).

More generally, closely related coronaviruses probably derived from a common ancestor through mutations and/or recombination mostly of the S protein, gradually

allowing the emergence of host or tropism variants. Examples of such an event are the chimeric virus SeCoV, composed of a TGEV backbone and the S and 3a proteins of PEDV, or the FCoV serotype II that has a spike protein containing canine sequences (Belsham et al., 2016; Le Poder, 2011).

It was shown that MERS-CoV is capable of using goat, camelid, cow, sheep, horse, pig, monkey and human DPP4 as entry receptors (Barlan et al., 2014; Eckerle et al., 2014). Importantly, binding to the host-cell receptor is not sufficient to allow entry into host cells. The subsequent fusion step sometimes requires activation of the spike protein by cellular proteases, which are also a determinant of host jump. For example, the bat coronavirus HKU4 is able to bind to the human DPP4 receptor, but cannot enter into human cells, because it is not adapted to human cellular proteases. Likewise, entry mediated by DPP4 receptors of species for which MERS-CoV has a low affinity is enhanced by the addition of human TMPRSS2 or trypsin proteases.

Mechanisms of host jumping involved in the emergence of SARS-CoV were extensively studied and well characterized. Analyses of receptor usage were performed on both human SARS-CoV and palm civet SARSr-Ci-CoV. It was demonstrated that human SARS-CoV could use both human and palm civet ACE2 for entry, whereas palm civet SARSr-Ci-CoV can only bind to palm civet ACE2, and not to human ACE2 (Belouzard et al., 2012a). These observations further supported the hypothesis of a host jump from palm civets to humans. Subsequently, it was shown that only two point mutations in the S1-RBD region were responsible for the adaptation of the virus to human ACE2, allowing the host jump (Li et al., 2005b). Other adaptive mutations strengthening the binding to the either human or palm civet receptor were also identified.

On the other hand, the host jump from bats to palm civets and other carnivorous animals is not fully understood yet, as the receptor of Bat-SL-CoVs is not known. Regardless, it was shown that modifying bat-SL-CoVs with the human SARS-CoV-RBD allowed entry of the virus through interaction with human ACE2 (Becker et al., 2008)

The spike protein is a good example of the critical equilibrium that is needed between conservation of essential domains and changes in others to adapt to a new host. Though the spike protein is critical to the host jump, full adaptation to the new host is mediated by changes in other proteins involved in pathogenesis.

1.5 Biology of coronaviruses

1.5.1 Virion structure

Coronavirus particles are pleomorphic but roughly spherical, and have a diameter ranging from 50 to 150 nm, with a mean around 80-90 nm (Neuman et al., 2006). Three proteins are anchored into the lipid envelope of the virion: the spike protein (S), the membrane protein (M) and the envelope protein (E) (Figure 5). Some coronaviruses such as MHV also have a fourth protein anchored in the envelope, the hemagglutinin-esterase protein (HE).

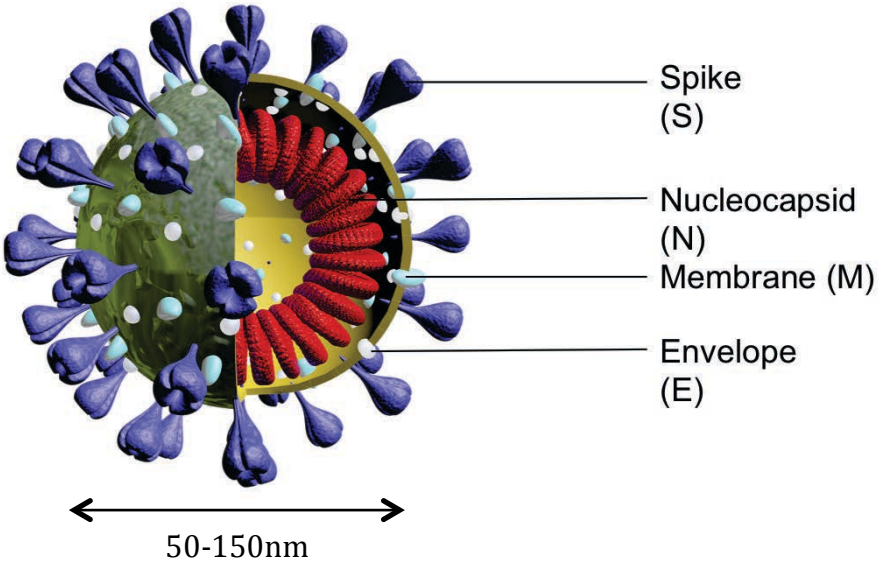


Figure 5 - ***Schematic structure of a coronavirus particle***

M, E and S proteins are anchored in a lipidic envelope, containing the viral capsid composed of the nucleocapsid protein and the viral RNA. The average size of the virions ranges from 50 to 150nm, excluding the S protrusions of 17-20nm.

Coronavirus virions have a distinctive appearance because of the club-like protrusions of spike protein trimers, which project 17-20 nm from the viral envelope (Masters, 2006a) (Figure 6A). Additionally, some coronaviruses display another set of smaller projections, which correspond to the hemagglutinin-esterase (HE) protein (Sugiyama and Amano, 1981) (Figure 6B). Inside the viral envelope, the nucleocapsid protein (N) is associated with viral RNA, and together they form a helical capsid, with a diameter of 9-16 nm. The helical capsid is a common feature of RNA (-) viruses, but is more surprising for an RNA (+) virus.

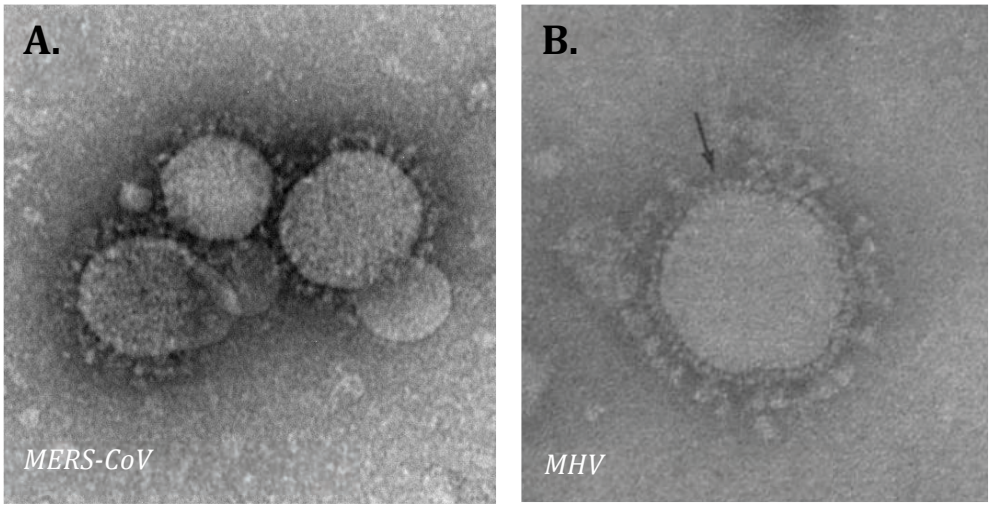


Figure 6 - ***Coronavirus particles by electron microscopy.***
(A) *MERS-CoV particles after negative staining, (CDC).* Projections formed by S protein trimers are visible. (B) *MHV virus after negative staining* (*Sugiyama and Amano, 1981)*. The arrow indicates smaller projections corresponding to the HE protein.

1.5.2 Genome

Coronaviruses have a single stranded non-segmented positive RNA genome. Their genome is extremely large for an RNA molecule, with a length of 26,2-31,7 kb (Figure 7). As it is a single stranded RNA (+) genome that has a 5' cap and a 3' poly(A) tail, it resembles cellular mRNAs (Masters, 2006a). The coronavirus genome is therefore infectious when transfected into cells, as for other RNA (+) viruses. Both 5' and the 3' ends of the genome contain an untranslated region (UTR) containing secondary structures that are essential to the replication and transcription of the RNA genome. Additionally, the 5' end contains a leader sequence, essential to the discontinuous replication mechanism. The replicase gene occupies two-thirds of the genome at the 5' end, with two overlapping open reading frames ORF1a and ORF1b, coding for polyproteins that are further cleaved into 16 non-structural proteins (nsps). The genes encoding the structural and accessory proteins occupy the last third of the genome at the 3' end (Fehr and Perlman, 2015) (Figure 7).

The gene order is conserved in all coronaviruses species: 5'-leader-UTR-replicase-(HE)-S-E-M-N-3'-UTR-poly(A) tail, with the accessory genes dispersed at the 3' end in between the structural proteins genes. At the beginning of each gene coding for a structural or accessory protein, there is a transcriptional regulatory sequence (TRS), which is essential for their expression.

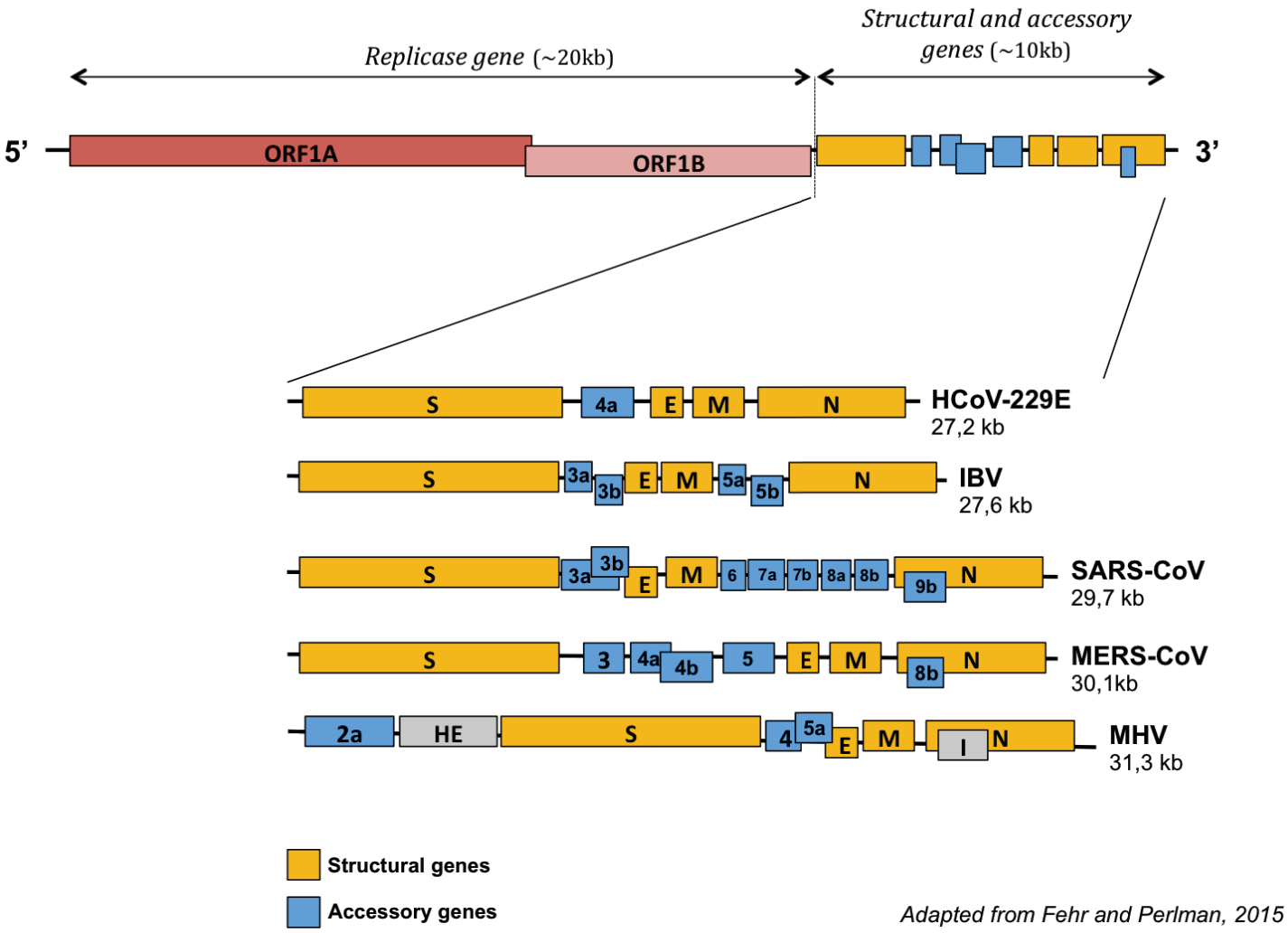


Figure 7- ***Coronavirus genome organization***

The upper panel shows the MERS-CoV genome as an example, with the replicase gene (in red) in the 5' two/thirds of the genome, and the structural (yellow) and accessory (blue) genes in the last 3' third. All coronaviruses genomes have a size around 30kb, with a variable number of accessory genes, as shown here with HCoV-229E (α-CoV), SARS-CoV, MERS-CoV and MHV (β-CoV) and IBV (γ-CoV).

1.5.3 Structural proteins

1.5.3.1 S protein

The spike (S) protein is a type I transmembrane glycoprotein, ranging in size from 180 to 200 kDa. It contains a large ectodomain, one transmembrane domain, and a short intracellular tail (Figure 8). It is highly glycosylated, with 21 to 35 N-glycosylation sites, and assembles in trimers at the surface of the virion (Belouzard et al., 2012a). The S protein plays an essential role in the entry step of the viral cycle, being involved in both attachment of the virus to the cellular receptor and fusion of the viral envelope with the cellular membrane. The structures of spike proteins of many coronaviruses

have been recently resolved through cryo-EM studies, including HKU1 (Kirchdoerfer et al., 2016; Ou et al., 2017), NL63 (Walls et al., 2016a), IBV (Shang et al., 2018a), PDCoV (Shang et al., 2018b), MHV (Walls et al., 2016b), SARS-CoV and MERS-CoV (Yuan et al., 2017). The protein ectodomain is divided into two functional domains, S1 and S2. S1 is in the N-terminal part of the protein, and is involved in the attachment to the cellular receptor, whereas S2 is in the C-terminal part of the protein and is involved mainly in the fusion step. In some coronaviruses, the spike protein is cleaved at the S1/S2 interface through a furin-cleavage site.

The S1 domain is highly variable among and across coronaviruses species: it is the main determinant of host and cellular tropism, since it contains the *receptor-binding domain* (RBD). The S1 domain can be further divided into four domains: A, B, C and D, with an additional domain 0 on many alphacoronaviruses (Hulswit et al., 2016). In spike protein trimers, the B domains are clustered on top of the S2 domains close to the axis, whereas the A domains are further from the center. Despite its high variability in sequence, the overall structure of the S1 domain is conserved among coronaviruses of the same genus. For *betacoronaviruses*, the A domain is made of a β-sandwich fold, and the B domain of a core of antiparallel β-sheets (Kirchdoerfer et al., 2016; Walls et al., 2016b; Wang et al., 2013). Domains C and D are made of discontinuous parts of the S sequence and form β-sheet structures adjacent to the S2 trimer.

The S protein is the main target of neutralizing antibodies, especially the S1 domain and the RBD, which contain most targeted epitopes (Du et al., 2017; Qiu et al., 2005). As this domain needs to remain accessible for receptor binding, it is exposed and targeted by the immune system, contrary to the S2 domain located at the core of the spike protein trimers. Interestingly, the many glycosylation sites present on the spike protein and especially the S1 domain could act as a glycan shield masking epitopes from the immune system (Walls et al., 2016a; Xiong et al., 2018).

The S2 domain involved in fusion is much more conserved, and contains two-heptad repeats, which is a characteristic of class I fusion proteins (Bosch et al., 2003).

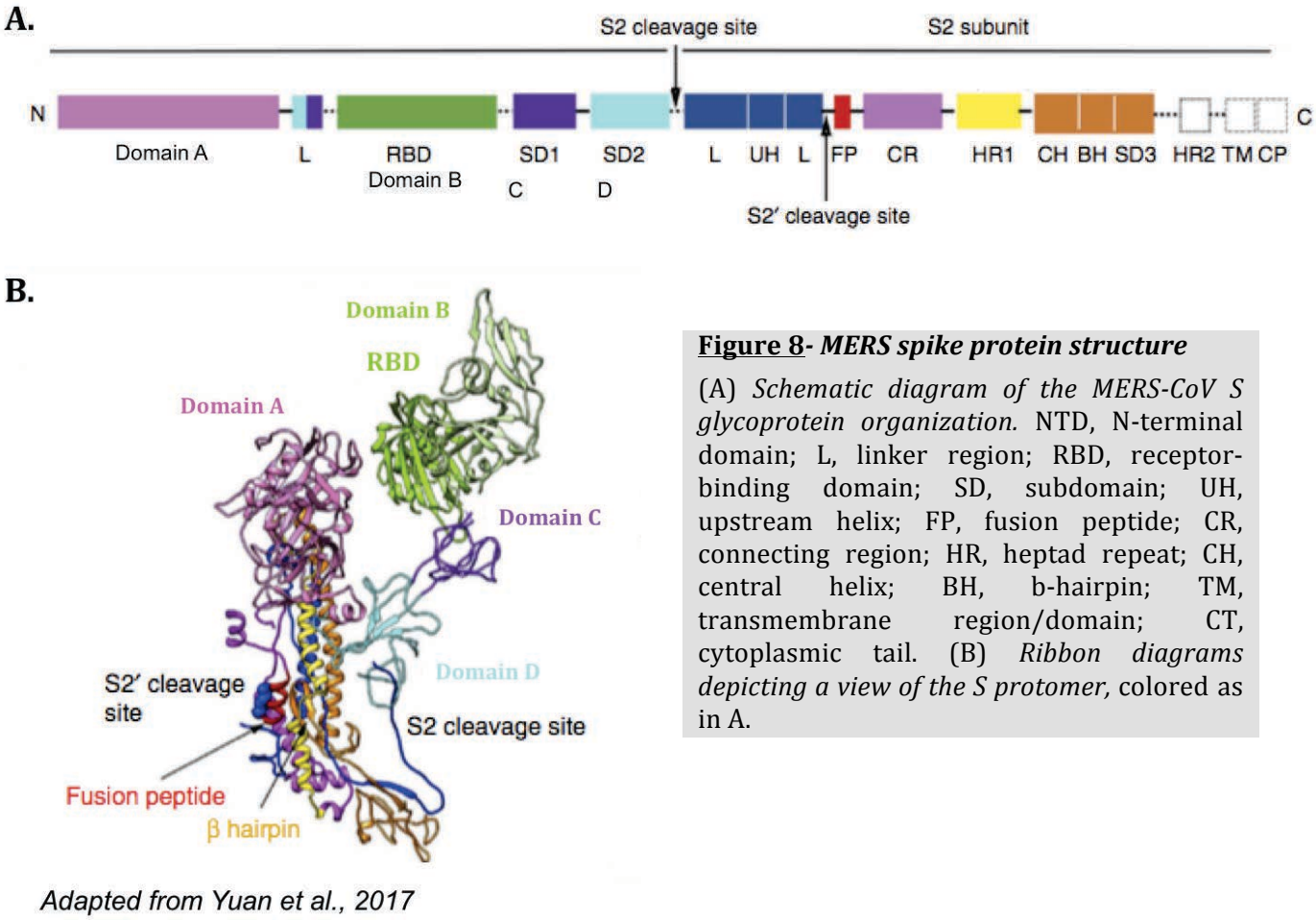


Figure 8- ***MERS spike protein structure***

(A) *Schematic diagram of the MERS-CoV S glycoprotein organization.* NTD, N-terminal domain; L, linker region; RBD, receptor-binding domain; SD, subdomain; UH, upstream helix; FP, fusion peptide; CR, connecting region; HR, heptad repeat; CH, central helix; BH, b-hairpin; TM, transmembrane region/domain; CT, cytoplasmic tail. (B) *Ribbon diagrams depicting a view of the S protomer,* colored as in A.

Adapted from Yuan et al., 2017

1.5.3.2 HE protein

Many *betacoronaviruses* of lineage A possess an additional fourth protein anchored in the viral envelope, the *hemagglutinin-esterase* (HE) protein. This protein was first isolated on PHEV and BCoV in the 1980s (Callebaut and Pensaert, 1980; King and Brian, 1982). The protein can be visualized in electron microscopy, and it forms a second set of projections on the viral surface, but smaller than the spike trimers (Figure 6B). The HE protein is composed of a large ectodomain, a transmembrane segment and a very short C-terminal domain. Monomers of HE are 65kDa after N-glycosylation (48kDa before), and they assemble into homodimers (Masters, 2006a). Because of its hemagglutinating properties, corresponding to binding to sialic acids on glycoproteins, it was believed that the HE protein could duplicate or replace the role of S protein in viral entry. Nevertheless, it was shown that HE was unessential (and only slightly beneficial) to virus growth, and that in absence of the S protein it is not sufficient to allow infection in cell cultures (Popova and Zhang, 2002).

HE proteins also have an acetylesterase activity: HE of BCoV, HCoV-OC43, ECoV and MHV-DVIM are sialate-9-O- acetylesterases, whereas HE proteins of RCoV, MHV-

JHM and MHV-S are sialate-4-O-acetylesterases. HE protein was also linked to a receptor-destroying activity in PHEV, BCoV and HKU1 (Huang et al., 2015; Schultze et al., 1991b), similar to the activity of HEF protein of influenza C viruses. Actually, the coronavirus HE protein is related to hemagglutinin-esterases of toroviruses and influenza C viruses (HA), suggesting a recent gene transfer, either from virus to virus or from a cellular source. HE protein of MHV seems to enhance the neurovirulence of the virus in mice, but to be a burden to the virus *in vitro,* since mutants with an inactive HE had a growth advantage when compared to wild-type virus (Kazi et al., 2005).

1.5.3.3 E protein

The envelope (E) protein is a small protein ranging in size from 8.4 to 12 kDa, which is not abundant in the viral envelope. It is not well conserved across coronavirus genera, but all E proteins have the same general structure: a short hydrophilic N-terminal domain, a transmembrane domain, and a hydrophilic C-terminal tail making up most of the protein (although the presence of only one transmembrane domain is still debated) (Liao et al., 2006; Ruch and Machamer, 2012a). A few cysteine residues in the C-terminal tail are targets for palmitoylation, and it was shown that this post-translational modification had no effect on trafficking of E, but was important for VLP formation (Boscarino et al., 2008).

E protein seems to have an important role in virion assembly and secretion (Fischer et al., 1998). It was shown for several coronaviruses that E and M proteins alone were sufficient to induce the secretion of virus-like particles (VLPs), indicating a major role of these proteins during the assembly step. Nevertheless, the role of E in the assembly process is still not fully understood. Beyond assembly, the E proteins of several coronaviruses such as SARS-CoV and MERS-CoV were shown to have an ion-channel activity (Torres et al., 2006), which could be involved in their role in viral egress and notably in the trafficking of newly formed virions through the secretory pathway, via bending and scission of the intracellular membranes. Furthermore, ion-channel activity of E was linked to virulence for several coronaviruses including SARS-CoV (Nieto-Torres et al., 2014). This ion channel activity and the role of E in assembly were both linked to high oligomer forms of E protein.

On the other hand, monomeric forms of IBV-E seem to be involved in modification of the secretory pathway. Indeed, using a mutant virus in which E transmembrane domain was replaced by the one of VSV-G, it was shown that this

domain is capable of altering the delivery of cargo to the cell surface, via blockade of cargo movement and disassembly of the Golgi. Also, IBV-E seems to be involved in retention of IBV-S into the Golgi (Ruch and Machamer, 2011). Furthermore, recent work reported that IBV-E is capable of neutralizing Golgi pH, thereby preventing premature cleavage of the S protein (Westerbeck and Machamer, 2019a).

Additionally, the E protein seems to have a role in modulating the host response. For example, it was demonstrated that a SARS-CoVΔE induced more stress and apoptosis than the wild type virus (DeDiego et al., 2011).

1.5.3.4 N protein

The N protein encapsulates the viral RNA, forming a helical nucleocapsid located inside the virion. Its size ranges from 43 to 50 kDa. Sequences analysis showed that N protein is composed of three highly conserved domains: two structural regions, the N-terminal domain (NTD) and the C-terminal domain (CTD), separated by a disordered central domain (linker region, LKR) (McBride et al., 2014). All three domains were shown to interact with RNA in different coronaviruses, and the CTD domain was shown to be involved in the oligomerization of the protein, which stabilizes it. Both RNA binding and oligomerization are necessary for the formation of the nucleocapsid.

N is a phosphorylated protein, (four sites identified on IBV-N and TGEV-N) but the precise role of these phosphorylations is still unclear. It is postulated that they could be involved in the association with membranes, conformational changes, RNA binding and/or virion assembly (Chen et al., 2005; Hogue, 1995).

The N protein is likely to be involved in replication of viral RNA, since it is capable of binding the TRS-L and the presence of N protein increases RNA synthesis. However, it was shown that the replicase gene products alone could still produce sgRNAs, suggesting that N is non-essential to transcription, its role being probably modulatory (Thiel et al., 2001). It is hypothesized that this role could be in selecting preferentially the viral mRNAs for translation, and/or facilitating template switching. Early studies on coronavirus assembly mostly showed that N protein was dispensable for VLP formation, but more recent studies showed for SARS-CoV that co-expression of N protein with E and M greatly increases VLP yields (Kuo et al., 2016a).

The N protein also has a role in pathogenesis and an effect on the host cell. For instance, SARS-CoV N is capable of modulating the cell-cycle and blocking the

progression to the S phase, and is also an interferon antagonist (Kopecky-Bromberg et al., 2007; Surjit et al., 2005). Additionally, the N protein of SARS-CoV was reported to localize to the nuclei, where it could have another function (You et al., 2007).

1.5.3.5 M protein

The M protein is the most abundant protein of the viral envelope. It ranges from 25 to 30 kDa when unglycosylated. Multiple forms of higher molecular weights are observed upon SDS-PAGE migration, forming a smear, and corresponding to either multimeric or glycosylated forms. M protein can be divided into three domains: a short N-terminal domain, a transmembrane domain with three transmembrane segments, and a long C-terminal domain that makes up more than half of the protein (Figure 9).

The C-terminal domain is located inside of the virion and on the cytoplasmic side when inserted into host cell membranes. The C-terminal domain contains a highly conserved amphipathic region just after the third transmembrane domain. This region is associated with membranes, and for one-third of TGEV M proteins the C-terminal domain was detected at the virion surface, suggesting that this region could constitute a fourth transmembrane segment conferring a new C-exo conformation (Risco et al., 1995) (Figure 9). Nevertheless, this has not been further demonstrated.

There is diversity in nucleotide sequences of M proteins across but not among coronavirus genera, and the hydrophobicity profiles are all remarkably similar, supporting a common general structure. Interestingly, cryo-electron microscopy analysis showed that M proteins can be found into an elongated (M_{LONG}) or a compact ($M_{COMPACT}$) form (Neuman et al., 2011). It seems that the difference between the two forms is a conformation change affecting the C-terminal domain, by either stretching it (M_{LONG}) or collapsing its structure ($M_{COMPACT}$). M_{LONG} is the main form found on virions and VLPs, and is responsible for interaction with N (and possibly S). Additionally, M_{LONG} dimers are able to induce membrane curvature. $M_{COMPACT}$ is the minor form, and has an indistinct shape suggesting aggregates. It does not seem to associate with N or to induce membrane curvature, and its role is yet unknown. Neuman *et al*. could partially convert M_{LONG} to $M_{COMPACT}$ by transient pH acidification, thereby weakening or disrupting M-N interactions, and therefore propose a model in which M-N interactions drive the

budding and are then disrupted before membrane fusion (Neuman and Buchmeier, 2016).

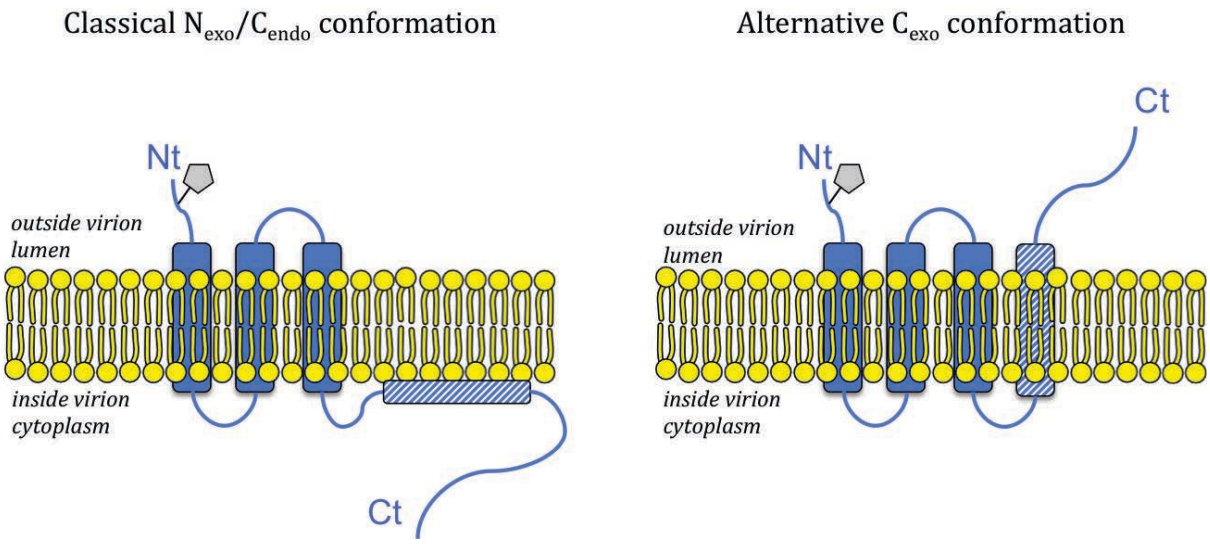


Figure 9- ***M protein topology***

Left panel - The most abundant and classical topology of MERS-M protein, in N_{exo}/C_{endo} with 3 transmembrane segments. *Right panel* - The alternative topology proposed for a subset of TGEV-M proteins, in C_{exo} with four transmembrane segments.
The amphipathic domain is showed in striped blue, and the glycosylation site (N-linked or O-linked) is represented in grey.

M proteins are generally glycosylated on one site located in their N-terminal domain (Figure 9). *Alphacoronaviruses* and *gammacoronaviruses* are generally N-glycosylated. *Betacoronaviruses* MHV and HCoV-OC43 are O-glycosylated (Locker et al., 1992a; Mounir and Talbot, 1992). On the other hand, *betacoronaviruses* SARS-CoV (Oostra et al., 2006) and MERS-CoV have a potential N-glycosylation site on their M protein (Yang et al., 2013a). Interestingly, IBV-M is the only described coronavirus with two N-glycosylation sites in its N-terminal domain (Liang et al., 2019; Stern and Sefton, 1982). M protein glycosylation is dispensable for both assembly interactions and intracellular trafficking (de Haan et al., 1998b).

The N-terminal domain also contains a cleavable signal peptide in TGEV, FIPV and CCoV-M proteins, although it may not be essential for membrane insertion (Laude et al., 1987). Most M proteins do not have a signal peptide, and it was demonstrated for IBV that the first or third transmembrane segments alone can induce anchoring of the protein in membranes (Rottier et al., 1984) and for MHV that each of the transmembrane segments could induce anchoring (Rottier et al., 1990).

The M protein is the motor of the assembly of viral particles. This step of the viral cycle requires interactions between structural proteins, which are orchestrated by the M

protein. Indeed, M has been shown to interact with all other structural proteins, and also to form homodimers (Corse and Machamer, 2003; De Haan et al., 2000; de Haan et al., 1999b; Tseng et al., 2010a). Moreover, it was shown that only M and E were essential for VLP production (Ho et al., 2004; Vennema et al., 1996a) and that S was dispensable but incorporated when present (see Assembly).

Another demonstrated function of the M protein is an interferon antagonist activity, with inhibition of IFN-β synthesis via inhibition of IRF-3 (Yang et al., 2013a).

1.5.4 Non-structural proteins

All non-structural proteins (nsps) are encoded in the 5' part of the coronavirus genome, within ORF1a and ORF1b. ORF1b is expressed after a ribosomal frameshift, allowing the production of pp1ab (see Replication). Nsps are released after cleavage of the polyproteins pp1a and pp1ab via an autoprocessing mechanism of the viral proteases. The papain-like protease (PLpro) cleaves at the sites nsp 1/2, nsp 2/3 and nsp 3/4, and the main protease (Mpro) or 3C-like protease (3CLpro) cleaves the remaining 11 sites. PLpro is encoded within the nsp3 gene and Mpro within the nsp5. Most coronaviruses encode two PLpro on nsp3, except SARS-CoV and MERS-CoV, which encode only one (Fehr and Perlman, 2015). Notably, nsp3 also plays a role in capsid formation through its interaction with N protein, and was shown to have a deubiquitinase activity.

Non-structural proteins 12 to 16 regroup the basic functions necessary to RNA synthesis: nsp12 encodes the viral RNA-dependent RNA-polymerase (RdRp) and nsp13 the RNA helicase, respectively necessary for polymerization of new RNA and for unwinding of RNA templates. Nsp14 encodes an exoribonuclease that is involved in the proofreading and therefore the fidelity of RNA replication, an essential feature for coronaviruses because of their large genome. Nsp14 also has an N7-methyl-transferase (N7MTase) activity, which combined with nsp16 that encodes the 2'O-methyl-transferase (2'O-MT) are involved in the formation of the 5' cap. The capping of viral RNAs is crucial for protection against cellular sensors, notably MD5A (Züst et al., 2011) that would destroy any RNA identified as non-self. Nsp15 encodes an endoribonuclease that was shown to colocalize and interact with proteins of the replication complexes, suggesting a role in replication that is not yet understood (Athmer et al., 2017; Zhang et al., 2018).

Non-structural proteins 7 to 10 have enhancing abilities on RNA replication, and are mostly co-factors for other proteins of the replication complex. For instance, nsp10 is a cofactor working in heterodimers with nsp14 and nsp16, and stimulates their activities. Nsp7 and 8 form hexadecameric complexes that may function as an RNA processivity clamp for the RdRp, as the RdRp alone exhibits poor processive synthesis, but is much faster in presence of nsp7 and nsp8 (Subissi et al., 2014). Other reports suggest that the nsp7/nsp8 complex could have an RNA polymerase activity of its own (te Velthuis et al., 2012). Nsp9 has been described as a dimer that binds aspecifically to nucleic acids, with a preference for single stranded RNA. It was also shown to be interacting with nsp12 and nsp8, but its function in coronavirus replication is still unclear. Nsp11 is the region of the viral RNA where the ribosomal frameshift occurs. In SARS-CoV it is theoretically a 13-residues peptide, but it has never been detected in infected cells (Neuman et al., 2014).

Nsp1 mostly has a role in inhibiting the innate immune response of the host, by promoting mRNA degradation (Huang et al., 2011; Narayanan et al., 2015). Surprisingly, this protein is absent in *gammacoronaviruses*. Nsp2 has no proven functional role, and it was demonstrated that nsp2 is dispensable for replication of both MHV and SARS-CoV, but that virus growth was slower in its absence (Graham et al., 2006).

Nsp4 and nsp6 are potential transmembrane scaffold proteins likely involved in membrane rearrangements occurring upon coronavirus replication and notably in the formation of DMVs, along with nsp3 (Angelini et al., 2013; Beachboard et al., 2015; Doyle et al., 2018; Hagemeijer et al., 2014). These transmembrane proteins are also involved in the anchoring of the replicase complex into cellular membranes. Importantly, nsp3 and nsp4 interact together, and their absence abolishes replication of SARS-CoV (Sakai et al., 2017). Nsp6 was also described as stimulating the formation of autophagosomes smaller than normal, that could promote replication complex assembly (Cottam et al., 2011, 2014).

1.5.5 Accessory proteins

Accessory proteins are usually dispensable to viral replication, as their name indicates. Interestingly, a few accessory proteins are incorporated into virions, thereby

making them “structural accessory proteins”, as in dispensable for viral replication but incorporated into virus particles. Among them are MHV I and SARS-CoV 3a (Ito et al., 2005; Shen et al., 2005), SARS-CoV 7a (Huang et al., 2006; Schaecher et al., 2007) and possibly SARS-CoV 9b (Xu et al., 2009).

Although accessory proteins are dispensable to viral replication, mutational analysis and reverse genetics studies repeatedly showed that they are critical to viral pathogenesis, and often regulate host response. Depending on the strains and genera, the number and type of accessory proteins in coronaviruses varies greatly, but similar effects can be observed from one coronavirus to another.

1.5.5.1 Membrane rearrangements

Rearrangement of intracellular membranes is a critical feature of RNA (+) viruses’ replication, and is usually mediated by a few non-structural (as described above for nsp3, nsp4 and nsp6) and/or accessory proteins.

For instance, it was observed in electron microscopy that ORF3a of SARS-CoV could induce vesicle formation, and that it was necessary and sufficient to induce Golgi fragmentation. Furthermore, ORF3a localizes to vesicles exhibiting late endosome markers. Therefore, OFR3a seems to have an important role in Golgi fragmentation and vesicle formation (Freundt et al., 2010). Interestingly, this protein is incorporated into virions (Ito et al., 2005; Shen et al., 2005) and is structurally very similar to M proteins. ORF6 of SARS-CoV was also shown to be involved in formation of DMVs (Zhou et al., 2010).

Membrane rearrangements are crucial to coronaviruses replication, and numerous different structures were observed in electron microscopy. A few proteins have been identified as playing a role in these modifications, but a lot is yet to discover concerning these mechanisms.

1.5.5.2 Interferon antagonism

Interfering with the host’s innate immune response is a crucial mechanism of survival for viruses. It is therefore not surprising that many viruses have evolved to develop many ways of inhibiting antiviral responses, the main strategy being inhibition of interferon, in various ways. This is especially striking for highly pathogenic viruses

such as SARS-CoV and MERS-CoV, but it's not limited to coronaviruses (Nipah virus, Ebola virus… etc. (Basler, 2015; Shaw, 2009)).

For instance, SARS-CoV has many proteins antagonizing interferon signaling, including structural proteins (M, N), non-structural proteins (nsp1, nsp7, nsp15) and accessory proteins. Among the accessory proteins, protein 3b inhibits type I interferon. ORF6 was shown to enhance viral replication both *in vitro* and *in vivo*, and to interfere with interferon antiviral effects through inhibition of STAT 1 nuclear import (Kopecky-Bromberg et al., 2007). ORF8a and ORF8b also inhibit IFN-β response, via interaction with IRF3 (Wong et al., 2018a).

For MERS-CoV, it was shown that accessory proteins ORF4a, ORF4b and ORF5 are interferon antagonists (Yang et al., 2013a). The ORF4a seems to have the stronger inhibitory effect via inhibition of IFN β, IRF3/7, NF-κB and the ISRE signaling pathways (Niemeyer et al., 2013). ORF4a was also shown to bind to RNA and to protect dsRNA from being detected by cellular sensors such as RIG-I and MDA5 (Siu et al., 2014). Interestingly, ORF4b localizes mostly to the nucleus, as ORF3b of SARS-CoV, and when ORF4b is expressed, NF-κB is located in the cytoplasm and its effect is inhibited. Without ORF4b or with an NLS-mutant of ORF4b, NF- κB was located to the nucleus and led to the production of inflammatory cytokines (Canton et al., 2018; Comar et al., 2019). Consistently, a mutant lacking all accessory proteins was shown to be attenuated *in vivo*, and to elicit robust IFN response (Menachery et al., 2017b).

MHV ORF5a has a type I interferon antagonist activity: indeed, an MHV-A59 strain deficient for ORF5a became partly sensitive to IFN I (Koetzner et al., 2010). This also implied involvement of another protein in the overall interferon antagonism, which could be protein 2a (NS2, specific of lineage A *betacoronaviruses*), that is homologous to MERS-CoV ORF4b (NS4b) (Goldstein et al., 2017). This protein has a phosphodiesterase activity specific for 2'5'–oligoadenylate and could inhibit the IFN-inducible OAS-RNase-L pathway, thereby inhibiting degradation of viral RNA (Liu et al., 2014).

For IBV, it was demonstrated that viruses lacking the ORF5b induced a more robust production of IFN I *in vitro*, indicating a role of this protein in inhibition of the host's immune response. Similarly, ORF3a was shown to inhibit interferon production (Kint et al., 2015). In this regard, ORF5b and ORF3a might compensate the absence of an nsp1 protein in *gammacoronaviruses* (Kint et al., 2016).

Generally, most coronaviruses exhibit several proteins with an interferon antagonist effect, demonstrating the importance of these mechanisms to coronaviruses replication.

1.5.5.3 Tropism and virulence

Many accessory proteins of coronaviruses have an effect on virulence, typically on cell apoptosis and stress. Maybe more surprisingly, a few accessory proteins also seem to be involved in cell tropism, and therefore viral pathogenesis.

It was suggested that ORF3a could have an ion channel activity, involved in SARS-CoV mediated apoptosis (Chen et al., 2009). Similarly, ORF3b, ORF7a, ORF8a, and ORF9b induce apoptosis in a caspase-dependent way, probably by inhibiting pro-survival factors. Overexpression of ORF7a also led to an arrest of the cell cycle in G0/G1 phase, blocking the passage in S phase. ORF7b of SARS-CoV could be an attenuating factor *in vivo*, as its deletion allows viruses to grow at higher titers *in vitro* and *in vivo* in hamsters, when compared to the wild type virus (Pfefferle et al., 2009b). Mutants of IBV with a truncated 3b gene were shown to be more virulent in chicken embryos, suggesting that IBV 3b could have a similar attenuating effect (Liu et al., 2014).

Protein 2a of lineage A *betacoronaviruses* is encoded directly in 3' of the replicase gene. Its expression is non-essential to MHV replication *in vitro*. Nevertheless, absence of protein 2a doesn't modify replication in the brain, but mutations of 2a in its catalytic domain impair virus growth in the liver (Roth-Cross et al., 2009). This is probably due to its phosphodiesterase activity, which could have different effects depending on cell types. This protein is therefore considered as a virulence factor altering tropism, although the precise mechanism is still unknown.

1.6 Viral cycle

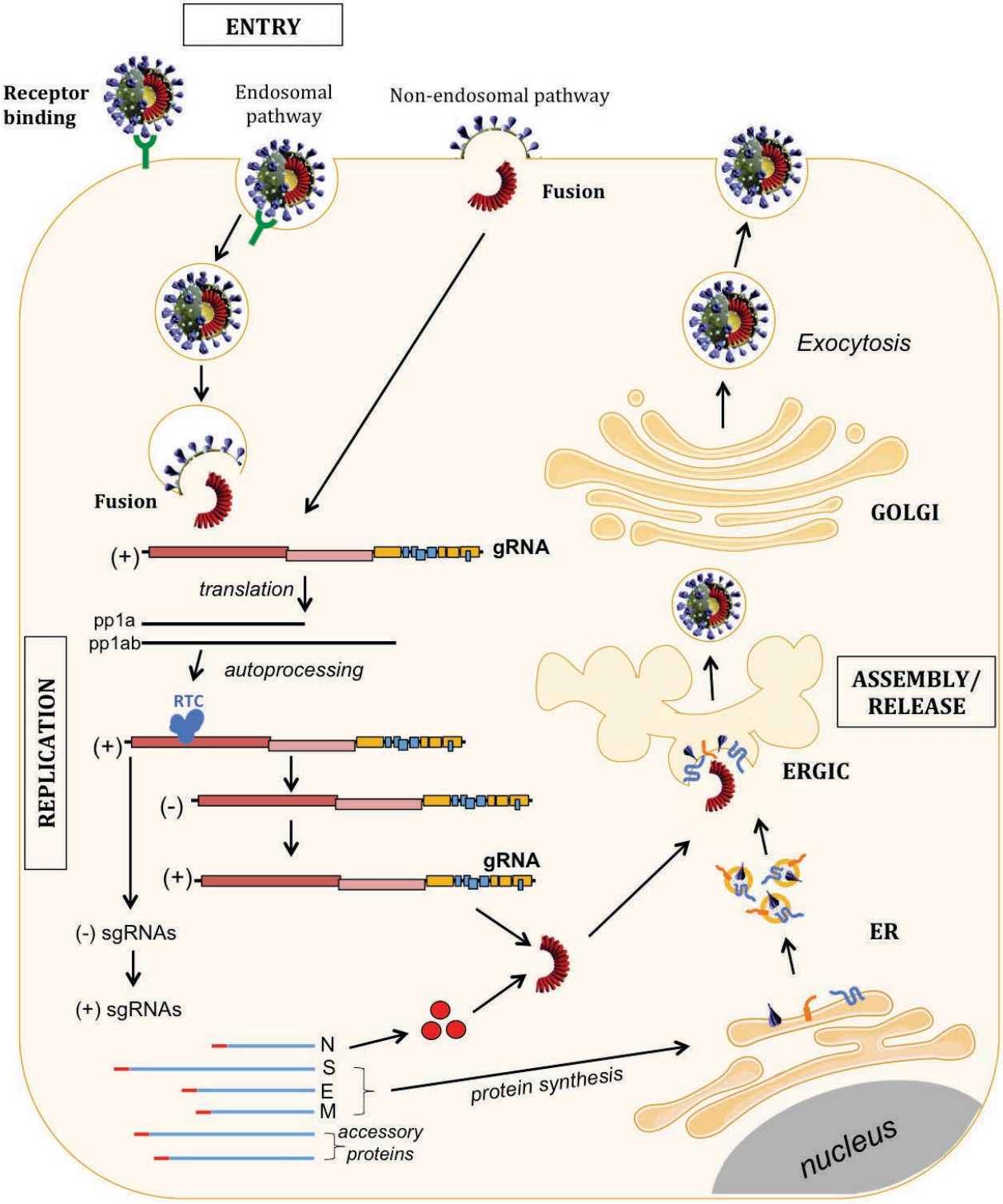


Figure 10 - ***Viral cycle of Coronaviruses***

The cycle is initiated by binding of the S protein to the cellular receptor. Fusion is triggered by various mechanisms, occurring either directly at the cell surface or in endosomes. After release, the viral genome is translated and the replicase complex is formed after autoprocessing of the polyproteins. The replication step then produces gRNAs and sgRNAs, both in (-) and (+) form. Envelope proteins are synthesized in the ER while N protein is produced and assembles with RNA in the cytoplasm. Assembly then occurs through interactions between structural proteins, and newly formed virions bud from the ERGIC and are released through exocytosis.

The main steps of the coronaviruses life cycle are depicted in Figure 10.

1.6.1 Entry

Viral entry is mediated by the interaction between viral and cellular proteins, and for enveloped viruses, fusion between their envelope and the cellular membrane is necessary in order to release their genome into the cytoplasm. The mechanisms of viral entry into host cells can therefore be divided into two main steps: binding to the cell receptor, and fusion of viral and cellular membranes. The spike protein of coronaviruses is essential to both steps (Belouzard et al., 2012a).

1.6.1.1 Receptor binding

Thus far, four proteins were identified as coronaviruses receptors: *Carcinoembryonic antigen-cell adhesion molecule* 1 (CEACAM1) for MHV, *aminopeptidase N* (APN) for HCoV-229E, FCoV, PEDV and other *alphacoronaviruses*, *angiotensin-converting enzyme 2* (ACE2) for HCoV-NL63 and SARS-CoV, and *dipeptidyl peptidase 4* (DPP4) also known as CD26, for MERS-CoV (Figure 1). Interestingly, three of these receptors are ectopeptidases (APN, ACE2 and DPP4) but their enzymatic activity was shown to be dispensable to infection (Bosch et al., 2014; Delmas et al., 1994).

In addition, some coronaviruses are able to bind to sialic acids: they are the only known receptor for BCoV and HCoV-OC43. TGEV, IBV and MERS-CoV also exhibit sialic acid binding, but the exact role of this activity for entry is unknown, although it is postulated that it could be involved in pre-attachment or early attachment phases (Li et al., 2017). Additionally, both IBV and FCoV were shown to bind to heparan sulfates (de Haan et al., 2008; Madu et al., 2007).

Other molecules were shown to be alternative receptors, co-receptors or facilitators of coronaviruses entry. The C-type lectin *Dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin* (DC-SIGN) has been identified as an alternate receptor/enhancer for FcoV, SARS-CoV and HCoV-229E (Han et al., 2007; Regan and Whittaker, 2008; Zhang et al., 2012). This lectin is able to recognize high mannose glycosylation patterns, and since the spike protein of coronaviruses is highly glycosylated, it is able to bind to DC/L-SIGN. Infected dendritic cells or macrophages can then transfer virions to other cells, primarily susceptible or not (Bosch et al., 2014).

For MERS-CoV, *Carcinoembryonic antigen-cell adhesion molecule* 5 (CEACAM5) was identified as an attachment factor. Indeed, it co-precipitated with the spike protein, and the disruption of this interaction reduced MERS-CoV entry. Nevertheless, CEACAM5 cannot support MERS-CoV entry in the absence of DPP4, but facilitates entry when present (Chan et al., 2016). Interestingly, tetraspanin CD9 was also shown to facilitate MERS-CoV entry by keeping the spike protein in close proximity to the cellular protease TMPRSS2 (Earnest et al., 2017), thus allowing rapid fusion at the cell surface. In absence of CD9, the spike protein was not cleaved by TTSPs, and the virus entered via the endosomal pathway. More generally, tetraspanins enriched microdomains in the membrane seem to be involved in coronavirus entry, notably in proteolytic priming (Earnest et al., 2015).

As discussed before (see Protein S), the spike protein can be divided into two domains, S1 being involved in receptor binding and S2 involved in fusion. The S1 domain can be further divided into four domains: A, B, C and D, with an additional domain 0 on many alphacoronaviruses (including HCoV-NL63). Position of the RBD varies across coronaviruses species, but it is located either in domain A (MHV) or domain B (SARS-CoV, HCoV-229E). Entry into host cell is determined by interaction between the cellular receptor and the RBD of the spike protein. This binding is the main determinant of host and cell tropism. Additionally, domain 0 is capable of binding to sialic acids, thereby also having a role in entry.

1.6.1.2 Fusion priming/triggering

After recognition and binding to the cellular receptor, fusion of the viral envelope with the cell membrane is necessary. Fusion can occur either directly at the cell surface (non-endosomal pathway) or in the endosomal compartment after endocytosis (endosomal pathway)(Figure 10). The choice of pathway can be dependent on the cell type and the requirement to activate the fusion process. Indeed, the fusion step requires important conformational changes of the S protein. Depending on the strains and species, receptor binding can be sufficient to induce theses changes, but for others, triggering of fusion by endosomal pH acidification or proteolytic cleavage is necessary.

In that last case, host-cell proteases are critical to the entry step. Activation by proteolytic cleavage can occur at two different sites of the S protein: either at the

interface between S1 and S2 domain (S1/S2 site) or at another site within S2, adjacent to the fusion peptide (S2') (Belouzard et al., 2009). The cleavage can be performed by a variety of proteases, such as cathepsins, furin-like pro-protein convertases (PCs), Type II-transmembrane serine proteases (TTSP). For activation of the spike protein by serine proteases, one or two cleavages are necessary. The first cleavage in S1/S2 is necessary for some coronaviruses to be further cleaved in S2', and dispensable for other S proteins (HCoV-229E S), some being already cleaved on virions (IBV-S). Importantly on the other hand, the second cleavage in S2' is crucial for fusion as it exposes the fusion peptide (Belouzard et al., 2009). The mechanism of cleavage in the endosomal compartment by cathepsins has not been characterized yet.

An example of entry solely triggered by receptor binding is the MHV strain JHM. MHV-JHM is able to use both endosomal and non-endosomal pathways, therefore its entry is pH-independent. Furthermore, it was shown that receptor-binding was sufficient to trigger fusion, since a soluble form of CEACAM1 could induce conformational changes in S2 (Matsuyama and Taguchi, 2002). On the other hand, the strain MHV-2 requires proteolytic cleavage of its spike protein in order to induce fusion, and this cleavage is performed by cathepsins in the endosomal pathway.

IBV-S protein is cleaved by furin at the S1/S2 site during assembly and exocytosis, the S protein on virions is therefore already cleaved (Cavanagh et al., 1986). After binding to the receptor, the entry goes through the endosomal pathway and it was shown for IBV strains Beaudette and M41 that entry is driven by pH acidification (Chu et al., 2006). Recent work on the Beaudette strain confirmed that IBV entry is dependent on low pH and on clathrin-mediated endocytosis, and that the virus follows the endosomal pathway (Wang et al., 2019a). Notably, it was also shown that the Beaudette strain relies on a cleavage by furin at the S2' for entry (Yamada and Liu, 2009), whereas this furin site is absent in all other IBV strains (Belouzard et al., 2009).

It is likely that coronaviruses can use different entry pathways. It is the case of SARS-CoV, for which fusion can occur both directly at the cell surface and through the endosomal pathway. Indeed, inhibitors of endosomal acidification or inhibitors of cathepsins can block SARS-CoV entry (Matsuyama et al., 2005). But after treatment of cells with trypsin-like proteases or the transmembrane protease serine 2 (TMPRSS2), fusion and entry could occur independently of the endosomal pathway (Bertram et al.,

2011; Matsuyama et al., 2010). Both S1/S2 and S2' sites were shown to be cleaved on SARS-CoV spike protein, and it is suggested that fusion activation requires a sequential cleavage on both sites (Belouzard et al., 2009). The proposed model suggests that S protein is cleaved at the S1/S2 site after binding to the receptor, then the spike protein is cleaved at the S2' site, therefore exposing the fusion peptide. Overall, two entry pathways have been identified: the direct fusion at the cell surface after cleavage by TMPRSS2 or other proteases is called the "early pathway", whereas the fusion occurring in the endosomal compartment after cleavage by cathepsins is called the "late pathway" because of the extensive endosomal maturation that is needed (Millet and Whittaker, 2018).

The overall mechanism is the same for MERS-CoV as for SARS-CoV, with sequential cleavages at the S1/S2 boundary and then at the S2' position, and the "early" and "late" pathways. There are only a few discrepancies. For example, furin can cleave MERS-CoV spike at the S1/S2 site during protein biosynthesis, and then at the S2' site during entry into a new host cell (Millet and Whittaker, 2014). More generally, it seems that MERS-CoV S can be cleaved at the S2' site by a variety of proteases at the cell surface (furin, trypsin, elastase, HAT, TMPRSS2…etc.), and that if none is available the virus will undergo endocytosis and cathepsins L will cleave the spike protein at the S2' site, inducing fusion in endosomes. It seems that the fact that the MERS-CoV spike protein is already cleaved in S1/S2 on virions accelerates the mechanism, and that fusion subsequently occurs in early rather than late endosomes. Moreover, protease usage likely depends on the cell type and the availability of the different proteases that the virus encounters (Park et al., 2016). Coronaviruses probably evolved to be able to use a wide variety of proteases to enter cells.

In humans, TMPRSS2 is expressed in most respiratory tissues (including upper airways, bronchi, and lung). Studies in TMPRSS2 KO mice infected with SARS-CoV or MERS-CoV showed that they had reduced weight-loss and viral kinetics when compared to control mice, showing that absence of TMPRSS2 reduces but does not abolish viral replication (Iwata-Yoshikawa et al., 2019). This suggests that the redundancy of the entry pathways is important for efficient infection, moreover as proteases availability varies across cell types (Park et al., 2016).

Also, studies on human coronaviruses HCoV-229E, HCoV-HKU1 and HCoV-OC43 indicate that the "early pathway", with fusion occurring at the cell surface, is more advantageous *in vivo*. This was inferred from experiments comparing entry of clinical isolates and of cell-culture adapted strains, showing that clinical isolates were mostly bypassing the endosomal pathway (Shirato et al., 2017, 2018). Several mutations were identified in the cell-culture adapted viruses when compared with the clinical isolates, which probably play a role in the accessibility of the cathepsins L cleavage sites. Taken together these results suggest that circulating HCoVs preferentially use cell surface protease to induce fusion and entry in the airway epithelial cells, but the redundancy of the pathways is likely of importance for efficient infection, as discussed above.

1.6.1.3 Conformational changes and fusion

After triggering of the fusion process, the spike protein undergoes important conformational changes. The first conformational change will allow exposition of the fusion peptide, which will be inserted into the host-cell membrane. At this stage, the spike protein displays a pre-hairpin conformation, with the HR1 and HR2 domains aligned. These domains will then fold back together and form a six-helix-bundle, thereby pulling the viral and cellular membranes into close proximity, ultimately resulting in the formation of a fusion pore (White and Whittaker, 2016). After this fusion of the two membranes, the viral nucleocapsid is released in the cell cytoplasm.

1.6.2 Replication

The replication step occurs in the cell cytoplasm, and in close association to membranes (Ulasli et al., 2010). After its release, the viral genome acts as an mRNA and ORF1a and ORF1b are translated into polyproteins pp1a and pp1ab (Figure 10). In order to translate both overlapping ORFs, the genome contains a slippery sequence (5'-UUUAAAC-3'), followed by a "pseudoknot" structure (two interleaved stemloops) at the ORF1a/ORF1b interface. In most cases, the ribosome begins translation of ORF1a, unties the pseudoknot and continues translation until it encounters the ORF1a stop codon, thereby producing the pp1a. But it was shown *in vitro* that in 25-30% of cases, the pseudoknot blocks the ribosome, causing it to pause on the slippery sequence. In that case there is a -1 ribosomal frameshift occurring before the ribosome unwinds the

pseudoknot. The ribosome therefore “misses” the ORF1a stop codon, and continues translation onto the ORF1b, thereby producing the pp1ab (Brian and Baric, 2005).

The polyproteins pp1a and pp1ab are then cleaved by viral proteases into non-structural proteins respectively 1 to 11 and 1 to 16. Interestingly, these viral proteases are within the polyproteins, and are cleaved and released by their own proteolytic activity, a mechanism referred to as “autoprocessing” (Muramatsu et al., 2013). These proteases are the papain-like protease (PLpro) and the main protease (Mpro) also called the 3C-like protease (3CL-pro). Most of the proteins encoded in the polyproteins are part of the replicase-transcriptase complex (RTC) and are involved in the replication of the genome and the synthesis of subgenomic RNAs (RNA-dependent RNA polymerase (RdRp), helicase, exoribonuclease…etc.). Other nsps are involved in the membrane rearrangements that are necessary to coronaviruses replication.

After translation and assembly of the replication complex, RNA-dependent-RNA-synthesis will begin. Two processes will occur in parallel: genome replication, corresponding to the production of genomic RNA (gRNA), and the synthesis of subgenomic RNAs (sgRNA) (Figure 10). From the viral genome, a full-length negative stranded form of gRNA is produced, serving as a template for production of the (+) gRNA. The full-length (+) gRNAs will later be packaged into newly formed virions. In parallel, and from the viral genome also, there is synthesis of negative sgRNAs that will serve as templates for production of high amounts of the positive sense sgRNAs. These (+) sgRNAs act as mRNAs encoding the structural and accessory proteins located in 3’ of the viral genome. SgRNAs are all 3’ co-terminal with the full-length genome, and all have at their 5’ end a common leader sequence (TRS-L). Indeed, these sgRNAs are produced through a discontinuous mechanism, and the presence of a TRS sequence at the beginning of each gene (called “body”, TRS-B), together with the presence of a TRS leader sequence (TRS-L) in the 5’ region of the genome are thought to be mediating this process. Indeed, it is hypothesized that the RdRp pauses at any TRS-B sequence, and then either goes back to the TRS-L or continues elongation to the next TRS-B. These sgRNAs would therefore be synthesized by the fusion of noncontiguous sequences, in that case the TRS-L with the TRS-B of each gene, this being possible because of an identical core sequence present in all TRS-B and in the TRS-L (Sola et al., 2015). Nevertheless, it is still unclear how this mechanism would be regulated, and for instance

how the RdRp would sometimes go back to the leader sequence and sometimes bypass the TRS-B. This mechanism of template switching from TRS-B to the leader is specific to *Nidoviruses* (Ulferts and Ziebuhr, 2011) and is one of the features that make recombination a common event in this family of viruses.

As for many RNA (+) viruses, coronaviruses replication relies on numerous rearrangements of the intracellular membranes, with which the replicase complex is associated. Double membrane vesicles (DMVs) were observed into host cells for MHV, but more recently it was shown for SARS-CoV that an important network of modified reticular membranes, named convoluted membranes (CMs), are also connected to the DMVs and continuous with the ER. Other rearrangements of the ER were observed in IBV-infected cells, such as spherules associated to zippered ER (Maier et al., 2014). It is believed that these membrane rearrangements form "replicative factories" that protect the viral RNA from antiviral mechanisms, and put all elements necessary to replication into close proximity. Notably, dsRNA, which is considered as a replicative intermediate, was detected inside DMVs (Knoops et al., 2008). DMVs and CMs appear quite early during infection, but other structures such as large virion-containing vesicles (LVCVs) appear later. Non-structural proteins such as nsp3, nsp4 and nsp6 are likely to play a role in membrane rearrangements upon coronavirus infection, and in association of the replication complex with these structures. It is hypothesized that these proteins are able to induce membrane curvatures using a multimeric scaffold mechanism, similar to the one of cellular complexes COPI and COPII, or by inserting an amphipathic domain into the lipid bilayer (Hagemeijer et al., 2012).

1.6.3 Assembly and Release

From their respective sgRNAs, structural and accessory proteins are synthesized. S, M and E proteins are synthesized and directly inserted into the ER membrane, as they are membrane-bound proteins, whereas N protein is synthesized in the cytoplasm, where it subsequently binds viral RNA. The assembly step then occurs in the *ER-Golgi intermediate compartment* (ERGIC) (Klumperman et al., 1994a; Stertz et al., 2007; Tooze et al., 1984) (Figure 10) and has two pre-requirements: (1) trafficking of all structural proteins to the assembly site and (2) interactions between those proteins.

1.6.3.1 Intracellular trafficking to the assembly site

After synthesis in the ER, membrane proteins are packaged into COPII vesicles, budding from specific ER-exit-sites (ERES). Classically, it is accepted that membrane proteins are transported to the plasma membrane through the secretory pathway in the absence of specific signals. Therefore, in order to assemble in the ERGIC, structural proteins must contain signals that direct their retention in or near the assembly compartment (Cohen et al., 2011; Lontok et al., 2004; Swift and Machamer, 1991). Importantly, these retention mechanisms are believed to be highly dynamic, likely involving different retention/export signals inducing retrograde and anterograde transport through the secretory pathway (Ujike and Taguchi, 2015).

When co-expressed in cells, S, M and E proteins do co-localize in the same compartment (Hsieh et al., 2008). But interestingly, when proteins are expressed alone in cells, they do not necessarily localize to the assembly site.

1.6.3.1.1 M proteins

The M proteins of several coronaviruses such as MERS-CoV and MHV were shown to go beyond the budding site in the secretory pathway, localizing to the *trans-Golgi network* (TGN) (Klumperman et al., 1994a; Mayer et al., 1988; Rottier and Rose, 1987; Yang et al., 2013a). The IBV-M protein on the other hand, was shown to localize to the ERGIC and *cis*-Golgi (Machamer et al., 1993a). The SARS-CoV M protein localizes mainly to the Golgi complex, but a fraction is also transported to the cell surface (Voss et al., 2006). The M protein was also reported to localize, at least partly, at the cell surface in cells infected with TGEV (Laviada et al., 1990) or FCoV (Jacobse-Geels and Horzinek, 1983).

Signals involved in the specific localization of the M protein were identified on IBV-M, MHV-M and SARS-M.

For IBV-M, four polar residues located in the first transmembrane segment (TM1) were identified as necessary for the specific localization of M to the *cis*-Golgi, and the TM1 was shown to be sufficient to retain intracellularly a chimeric VSV-G/M protein (Machamer and Rose, 1987a; Machamer et al., 1993a; Swift and Machamer, 1991). Nevertheless, the effect of the mutation of these polar residues on the retention of the

full-length protein could not be assessed, as the mutant proteins were all retained in the ER probably because of improper folding (Swift and Machamer, 1991).

For MHV-M, both the first transmembrane segment and a part of the C-terminal domain seem to be involved in the specific retention of the protein in the *trans*-Golgi network. A C-terminal deletion of 18 amino acids resulted in the export of the mutant protein at the cell surface (Armstrong and Patel, 1991a), indicating that this part of the domain is important for *trans*-Golgi addressing. More precisely, one amino acid in the C-terminal region, Y211, was mutated in other studies and the resulting mutant Y211G was partly exported at the cell surface (de Haan et al., 1998b, 1999b) but this was not investigated further. In other reports, mutants lacking the TM1 were only partly retained in the Golgi region, indicating that this domain is also involved in retention (Armstrong et al., 1990; Locker et al., 1994a). Nevertheless these proteins were not exported to the cell surface, but rather localized to the endosomes or ER. Moreover, when a chimeric protein composed of the VSV-G protein ectodomain and MHV-M TM1 was constructed in the same manner as for IBV-M, it was exported at the cell surface, meaning that MHV-M TM1 was not sufficient to retain the chimeric protein intracellularly (Swift and Machamer, 1991). This is surprising, as MHV-M TM1 possesses three of the four polar residues described as critical for IBV-M *cis*-Golgi retention.

Taken together, these results suggest that the MHV-M requires a part of its C-terminal domain and possibly its TM1 to be efficiently retained in the trans-Golgi, and this mechanism is thought to be similar to the one of a TGN resident protein rather than a protein being recycled from the plasma membrane (Locker et al., 1994a) as the wild-type protein was not detected at the cell surface in internalization assays.

SARS–CoV M localizes mostly in the Golgi compartment, but is also partly exported to the cell surface, although this was not observed in all studies (Nal et al., 2005; Voss et al., 2006). Tseng *et al.* constructed mutants containing only the first 50, 75 or 100 N-terminal amino acids, which localized to the Golgi region, suggesting a role of the transmembrane domain in Golgi retention. Importantly, a mutant lacking the three TMs and containing only the C-terminal domain (MΔ1-100) also located to the Golgi region, with a bit of export to the plasma membrane, suggesting that the C-terminal domain also plays a role in Golgi retention. Interestingly, another mutant lacking the 3

TMs and the amphipathic domain (MΔ1-159) was not retained in the Golgi, and less than 10% of the mutant proteins were still associated to membranes (Tseng et al., 2010a).

MERS-M protein localization was not as extensively studied as for MHV, IBV and SARS-CoV, but the protein was shown to localize to the *trans*-Golgi network (TGN) (Yang et al., 2013a).

Interestingly, it was shown for IBV-M and MHV-M that oligomerization of the protein correlates with its retention to the Golgi, with proteins properly located being detected in multimers while proteins exported at the cell surface are not (Locker et al., 1995; Weisz et al., 1993). This indicates that M protein retention could be mediated in part by its oligomerization, the formation of large complexes possibly preventing incorporation into transport vesicles (Ujike and Taguchi, 2015), as it was shown for Golgi glycosyltransferases (Banfield, 2011). Nevertheless, it is possible that the oligomers also bind to Golgi components or resident proteins.

1.6.3.1.2 E protein

The E protein expressed alone in cells has been reported to accumulate in the ER and ERGIC (SARS-E, MHV-E) or in Golgi compartments (IBV-E, MERS-E). Moreover, specific targeting signals were identified in its cytoplasmic tail (Cohen et al., 2011; Corse and Machamer, 2000; Lim and Liu, 2001; Nal et al., 2005).

IBV-E was reported to localize in the *medial*-Golgi, with partial overlap with the *cis*- and *trans*-Golgi (Corse and Machamer, 2000). IBV-E was also detected in the ER and ERGIC in other studies, at an early time point of 7h post-transfection. A lysine at the C-terminal extremity of the protein was shown to be responsible for this ER/ERGIC localization, behaving as a dilysine-like motif for ER targeting (Lim and Liu, 2001). Nevertheless, even in presence of this putative signal, the protein reaches the Golgi complex after 7h. It was later postulated that the ER localization observed by *Lim and Liu* could also be due to the high expression levels induce by the vaccinia/T7 expression system (Liao et al., 2006).

MHV E was shown to localize mostly to the ERGIC and *cis*- and *medial*-Golgi, and to partially localize to *trans*-Golgi (Venkatagopalan et al., 2015). Time-course

experiments showed that the E protein remains at the site of assembly throughout infection.

SARS-CoV E was shown to localize mostly to the *cis*-Golgi (Cohen et al., 2011), and to the ERGIC (Nieto-Torres et al., 2011) and the C-terminal domain of the protein was shown to be involved in this retention.

Interestingly, a secondary structure of the C-terminal domain is highly conserved among all coronaviruses E proteins. It is a β-hairpin composed of two β-strands, with a conserved proline in between. It was demonstrated for several coronaviruses including SARS-CoV and IBV that this structure is necessary and sufficient to induce retention of E in the Golgi region (Cohen et al., 2011; Li et al., 2014).

1.6.3.1.3 S protein

When the S protein is expressed alone in cells, it localizes at the cell surface and/or is present throughout the secretory pathway, especially in the ER and ERGIC. Interestingly, ER retrieval signals were identified in the cytoplasmic tails of IBV-S and SARS-CoV S (Lontok et al., 2004). For IBV-S, this signal is a dilysine motif (-KKxx-COOH) that can retain a reporter protein into the ERGIC, and therefore prevent export to the cell surface. SARS-CoV S cytoplasmic tail contains a dibasic motif (-KxHxx-COOH) that also promotes retention of the protein in the ERGIC, as the dilysine motif of IBV-S (McBride et al., 2007). A SARS-CoV S protein with the retrieval signal mutated was shown to be less colocalized with M protein and thus less incorporated into VLPs (Ujike et al., 2016). Interestingly this KxHxx-COOH motif is conserved in MERS-CoV spike protein.

The mechanism underlying the function of this ER retrieval motif is well described: the signal is directly recognized by COPI complex and the S proteins are incorporated into COPI vesicles and retrieved from the Golgi to the ER through retrograde transport. The protein therefore cycles between Golgi and ER, leading to the localization of the S protein near the ERGIC budding compartment at steady state (Cosson and Letourneur, 1994). Nevertheless, further studies showed that mutation of this signal had little or no effect in the context of the whole virus. S proteins with a mutated signal were incorporated into virions at a similar level when compared with wild type proteins, but exhibited higher fusion activity (Shirato et al., 2011).

A tyrosine based signal YxxF was also identified in IBV-S, and thought to promote recycling of the protein from the plasma membrane, and to function as an endocytosis signal (Lontok et al., 2004). Alternatively, other reports showed that YxxF functions as a retention signal rather than an endocytosis signal (Winter et al., 2008). A similar intracellular retention motif YxxI was identified in the cytoplasmic tail of TGEV-S, although the precise intracellular localization was not assessed. Contrary to SARS-CoV S and to a lesser extent IBV-S, no TGEV-S was detected at the cell surface (Schwegmann-Wessels et al., 2004). It is believed that the absence of a similar motif could explain the extended expression of SARS-CoV S at the cell surface when compared with spike proteins of other coronaviruses.

Yee-Joo Tan hypothesized that the spike protein of SARS-CoV could be internalized through an interaction with the accessory protein 3a (Tan, 2005). This protein contains a tyrosine-based endocytosis signal and was shown to be expressed at the cell surface and then internalized. Furthermore, co-immunoprecipitation experiments showed that 3a interacts with the spike protein (Tan et al., 2004). This hypothesis was not tested yet.

Interestingly, very recent work on PEDV using viruses mutated for both ER retrieval and tyrosine-based signals showed that mutant S proteins were less incorporated into virions, resulting in a reduced pathogenicity of the virus in inoculated pigs (Hou et al., 2019).

1.6.3.2 Protein-protein interactions

Experiments using co-expression of coronavirus structural proteins in cells are widely used to observe the production of virus-like particles (VLPs), and this is the main technique that is used to study viral assembly requirements. Generally, co-expression of E and M proteins is sufficient to induce the formation of coronavirus VLPs (Baudoux et al., 1998a; Corse and Machamer, 2000; Ho et al., 2004; Vennema et al., 1996a) and the S and N proteins are incorporated when present, but dispensable. This was shown for several coronaviruses including MHV, IBV, BCoV, TGEV and SARS-CoV. However, it is more controversial for SARS-CoV as others reported that the N protein could be necessary to the formation of VLPs along with E and M (Siu et al., 2008a), or that M and N were sufficient to the production of VLPs (Huang et al., 2004). For IBV, a study also reports formation of VLPs after co-expression of only M and S proteins with a baculovirus system (Liu et al., 2013).

Interestingly, it was demonstrated that the M protein of MHV and SARS-CoV could self-assemble and be released into vesicles smaller than VLPs (De Haan et al., 2000; Tseng et al., 2010a). For MHV and IBV, E protein alone is also secreted in similar small vesicles (Corse and Machamer, 2000; Maeda et al., 1999).

Still, it is worth noting that many of these original VLP experiments were performed using vaccinia-virus systems, hence in conditions of high overexpression. Therefore, interpretation of these experiments is complicated by the fact that E and M expressed alone can also be secreted into small vesicles (likely exosomes). In that light, it is not surprising that more recent experiments based on co-transfection of plasmids showed that N greatly enhances VLP yields (Ruch and Machamer, 2012a).

Since E and M proteins are sufficient to produce VLPs, their precise role in the assembly process was investigated further by generating mutant proteins and recombinant viruses.

For the E protein of MHV, Fischer *et al*. performed clustered charged-to-alanine mutations on the E gene (in the C-terminal part of the protein), and incorporated these mutations in the whole virus genome. Of the four resulting mutants, one was lethal and one was similar to wild type. The other two were partially temperature-sensitive at 37°C and 39°C when compared to the wild type MHV-59. In order to test the thermolability of these mutants, both were grown at 33°C or 39°C, heat-treated at 40°C, and then the plaque titers at 37°C were measured. Interestingly, viruses grown at 33°C were much more thermolabile than the wild type, whereas viruses grown at 39°C were similar to wild type. However, the infectious titers of the stocks grown at 39°C were lower than those of stocks grown at 33°C. The hypothesis proposed by the authors is that the mutants had a flawed assembly: grown at 33°C viruses assembled in a flawed manner that was stable at 33°C but not 39°C; whereas viruses grown at 39°C assembled correctly but at low levels and were stable at 39°C. Interestingly one mutant displayed pinched and elongated shapes when observed in electron microscopy (Fischer et al., 1998).

To go further, viruses lacking the E gene were constructed for TGEV, MHV, SARS-CoV and MERS-CoV (Almazán et al., 2013; DeDiego et al., 2007; Kuo and Masters, 2003; Ortego et al., 2007).

Interestingly, the deletion was lethal for TGEV (Ortego et al., 2007) and MERS-CoV (Almazán et al., 2013) but not for MHV (Kuo and Masters, 2003). Although severely impaired, all MHV-ΔE viruses replicated at low titers and were infectious (Kuo and Masters, 2003). For SARS-CoV, the ΔE mutant replicated at lower titers because of an assembly defect, but the decrease was much smaller than for MHV (DeDiego et al., 2007). The MERS-ΔE mutant was replication competent but propagation defective (Almazán et al., 2013).

Therefore it was concluded that E greatly enhances viral assembly but is not essential to this mechanism. Interestingly, E proteins from other coronaviruses (SARS-CoV, IBV, BCoV) were shown to efficiently replace MHV E (Kuo et al., 2007), suggesting that they share fundamental functions. More surprisingly, in stocks of the MHV-ΔE mutants described above, faster growing mutants were isolated: all of them displayed a duplication of the M gene resulting in the expression of a truncated M protein named M* (Kuo and Masters, 2010). This M* protein was incorporated into virions and greatly increased viral replication when compared with MHV-ΔE, suggesting that it could replace the E protein to some extent.

Concerning the M protein, mutants with deletions in different domains were generated in order to investigate its role in assembly. Deletion of the last amino acid of the C-terminal domain of MHV-M (MΔ1) dramatically reduced VLP formation, but surprisingly, the recombinant virus exhibited a wild type phenotype, suggesting that in the context of the whole virus, other factors could compensate for this mutation (de Haan et al., 1998b). Interestingly, a deletion of the two last amino acids of the C-terminal domain (MΔ2) resulted in a complete abolishment of VLP production, and a recombinant virus could not be generated (de Haan et al., 1998b). Consistently, a conserved sequence located in the C-terminal domain of M was shown to be important for VLP formation with E protein, and some deficient M mutants were stabilized when the N protein was also present (Arndt et al., 2010a).

The M protein is believed to be the motor of viral particle assembly. Accordingly, in most experiments involving VLPs, the M protein was necessary. This central role of the protein in the assembly process is the result of its capacity to interact with all other structural proteins.

The existence of interactions between M and E can be inferred from their capacity to induce VLPs, but proper interaction was only demonstrated for several coronaviruses. Among them, it was shown for IBV that E and M cytoplasmic tails are involved in their interaction (Corse and Machamer, 2003; Lim and Liu, 2001). For SARS-CoV, the transmembrane domain of E is necessary for interaction with M (Chen et al., 2009). Importantly, in this paper the E protein was considered to have two membrane spanning domains and it was concluded that both were involved in interaction with M.

The disproportionate quantities of M in virions and VLPs when compared to E suggest that homotypic M-M interactions are the driving force of assembly (Figure 11). Nevertheless, since E protein is necessary to VLP formation, it was suggested that E could have a role either in membrane curvature or in particle scission, at precise sites in a lattice composed of M proteins (Ruch and Machamer, 2012a; Vennema et al., 1996a) (Figure 11).

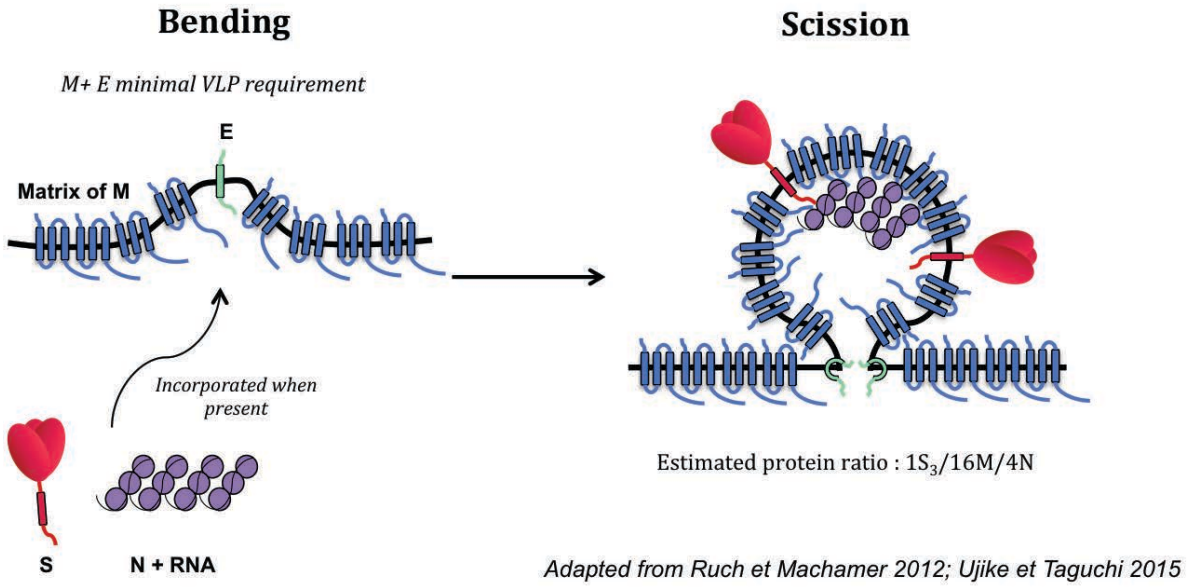


Figure 11 – ***Membrane bending and scission during assembly***

Membrane bending is likely induced by E protein inducing curvature at precise locations in a matrix composed of M protein, which constitutes the driving force of the process. Through interactions with M, N and S can be incorporated into forming virions. The E proteins could then induce scission in order to release the newly formed virions. The E protein is present in very low amounts, with an approximate E:S:M:N of 1/20/300/60.

The S protein is incorporated into VLPs when present, and this incorporation is dependent on an interaction with the M protein (Figure 11). For MHV, it has been shown that the transmembrane domain and the amphipathic region of the C-terminal domain of M are involved in this interaction (de Haan et al., 1999b). It was subsequently shown that the C-terminal domain of both MHV and FCoV spike was also involved in the interaction with M (Bosch et al., 2005; Godeke et al., 2000).

The spike proteins contain a highly conserved cysteine-rich region in their cytoplasmic domain, adjacent to the transmembrane domain. This motif was shown to be palmitoylated for MHV, TGEV and SARS-CoV spike proteins. Inhibition of palmitoylation of MHV-S disrupted M-S interaction and S incorporation into VLPs (Thorp et al., 2006), whereas for SARS-CoV S and TGEV-S, depalmitoylation impaired incorporation of S into VLPs but not its co-localization/interaction with the M protein (Gelhaus et al., 2014; Ujike et al., 2012).

Incorporation of genomic RNA into newly formed virions or VLPs relies on the recognition of a specific RNA packaging signal. Unsurprisingly, this recognition was shown to be mediated by the N protein for MHV (Kuo et al., 2014, 2016b), MERS-CoV (Hsin et al., 2018) and SARS-CoV (Hsieh et al., 2005). Other reports on MHV point to a role of the M protein which could induce packaging of RNA even in the absence of N (Narayanan et al., 2003). Although N protein binds the viral RNA on the packaging signal, it is postulated that the M protein truly directs encapsidation in the viral particle through a specific interaction with the N protein via their C-terminal domains (Escors et al., 2001; Kuo and Masters, 2002).

Kuo *et al*. further mapped the interactions between M and all other structural proteins by constructing chimeric MHV viruses in which total or part of the M or N proteins was replaced by their distantly related (38% amino acid identity) SARS-CoV counterparts (Kuo et al., 2016a).

Concerning M-N interactions, mutants containing the whole MHV-M and the N3 domain of SARS-CoV N, or SARS-CoV M and the N3 domain of MHV N were lethal. On the other hand, an MHV mutant containing SARS-CoV M and SARS-CoV N3 domain (MN3) was viable, although it grew at lower titers when compared to the wild type. Both of these results confirm that the C-terminal domain N3 is crucial for M-N interaction.

Further mapping of M C-terminal regions involved in its interaction with N was attempted by constructing MHV mutants containing the N3 domain of SARS-CoV N, and only parts of the SARS-CoV M C-terminal domain. All of these mutants were lethal, suggesting that the region of M interacting with N is more complex than a linear sequence of amino acids. On the other hand, substitution of the N-terminal domain of the M protein had no effect on M-N interactions, despite high divergence and different glycosylation of this domain in each virus.

The MN3 mutant also gave insight into M-E interactions, as analysis of revertants (MN3rev) showed that most of them had single mutations in the E protein. These grew in larger plaques when compared to other revertants, suggesting that MHV-E does not properly interact with SARS-CoV M, but that few mutations (notably F20S) can partially rescue this interaction. This was surprising considering earlier work in which MHV-E could be successfully replaced by SARS-CoV E (Kuo et al., 2007). Even more surprisingly, an MHV mutant containing both E and M of SARS-CoV was also severely impaired in growth.

Concerning M-S interactions, it was observed that the MN3 mutant incorporated less S protein when compared to the wild type, thereby explaining that it grows to lower titers than the wild type. This was already surprising as SARS-CoV M and MHV-S were previously shown to efficiently interact (Yao et al., 2013). Furthermore, another chimera containing the SARS-CoV S protein endodomain and transmembrane segment in addition to M and N3, did not rescue M-S interaction as would be expected from earlier work (Bosch et al., 2005), and was even lethal.

1.6.3.3 Viral egress and release

After assembly of the envelope proteins and the ribonucleocapsid in the ERGIC, viral particles bud from this compartment and traffic to the plasma membrane, where they are released from the cell.

As described for assembly, the E protein is likely involved in membrane curvature, and in scission of membrane at the base of newly formed virions, promoting their release (Figure 10). But additional data suggest that E could also be involved in viral particles trafficking and release.

E was reported to alter the host secretory pathway, by inducing both fragmentation of the Golgi and blockade of cargo trafficking through it. It has been demonstrated that although it is the C-terminal domain of E that plays a role in assembly, it is the transmembrane domain that is involved in these mechanisms (Ruch and Machamer, 2011). Additionally, in the case of a mutant virus containing an E protein with its transmembrane domain replaced or mutated, the spike protein showed higher expression at the cell surface and/or aberrant cleavage (Westerbeck and Machamer, 2019a). This retention of S reduces the amount of antigen at the cell surface, and also reduces syncytia formation. The effect of IBV-E on the Golgi was shown to be mediated by neutralization of pH, and surprisingly this was mediated by a monomeric form of E, suggesting that its ion-channel activity is not involved in this mechanism (Westerbeck and Machamer, 2019a).

The benefit of the modifications of the secretory pathway for the virus is not fully understood. Interestingly, the fragmentation of the Golgi was shown to correlate with the formation of large vacuoles containing viruses (Ulasli et al., 2010), indicating that this promotes viral secretion. Also, the large size of virions when compared to other cargo could require a specific mechanism of trafficking.

Recombinant TGEV-ΔE and MERS-CoV-ΔE were shown to grow only in cells expressing the E protein (Almazán et al., 2013; Ortego et al., 2007). Analysis of TGEV-ΔE infected cells showed that viral-like particles assembled but accumulated intracellularly in the ERGIC. These particles resembled immature virus, and contained RNA and all structural proteins (except E) but were not released from the cell (Ortego et al., 2007). In other words, viruses lacking the E protein are replication competent but propagation defective.

1.7 Intracellular trafficking in the secretory pathway

In eukaryotic cells, secretory and membrane proteins undergo synthesis, proper folding and delivery through the secretory pathway. Since the experiments of Palade (Palade, 1975), it is admitted that this pathway includes the ER, the Golgi complex (the intermediary ERGIC was discovered more recently), and the plasma membrane, and that transport from one compartment to another is mediated by transport vesicles (Mellman and Warren, 2000). Interestingly, there is both anterograde transport, from ER to Golgi to plasma membrane, and retrograde transports from the surface to the Golgi or from the Golgi back to the ER.

The secretory pathway has a main role of transport, but is also necessary for protein folding and post-translational modifications. The variety and quantity of cargo proteins going through the secretory pathway is huge, and their transport and addressing to specific compartments is therefore highly regulated.

An overview of compartments and routes of intracellular trafficking in the secretory pathway is presented in Figure 13.

Endoplasmic reticulum (ER)

Many types of proteins can be synthesized in the ER and go through the secretory pathway: soluble proteins, transmembrane proteins (with various number of TMs), GPI-anchored proteins… etc. They represent one third of the total cell proteome.

Membrane-bound proteins are synthesized in the ER where they are directly inserted into the membrane. The first requirement in order to exit the ER is that proteins must be properly folded. Misfolded proteins are translocated back into the cytoplasm and degraded by the ubiquitin-proteasome system. After this quality control for proper folding, the proteins left will be able to exit the ER at specific sites devoided of ribosomes: ER-exit-sites (ERES). Anterograde transport from the ER to the Golgi complex is mediated by proteins of the COPII coat, which concentrate at these exit sites.

The COPII complex assembly is initiated by activation of Sar1 (a small GTPase) mediated by Sec12 that acts as a guanine nucleotide exchange factor (GEF) (Figure 12). Sar1-GTP then initiates vesicle formation by causing membrane curvature, and recruiting proteins of the inner layer of the coat, Sec23/24. These proteins provide the

cargo-binding function. Proteins of the outer layer, Sec13/31, are subsequently recruited in heterotetramers and polymerize into a lattice structure that drives membrane bending. Sar1 then mediates fission of the membranes, thereby releasing the COPII-coated vesicle from the ER membrane (Venditti et al., 2014a).

But in order to be incorporated into COPII vesicles exiting the ER, cargo proteins must contain ER-export signals that can interact directly with COPII proteins (usually Sec24 or Sar1) or indirectly via binding to cargo adaptors or transmembrane receptors (Sato and Nakano, 2007). Among these signals, we can cite LxxLE, di-acidic DxE, YNNSNP, di-hydrophobic motifs (FF, YY, LL, or FY) and 3R motifs (Barlowe, 2003). These signals are located in the cytoplasmic domain of the proteins that is exposed. The presence of these signals conditions the selective sorting of cargo proteins into the transport vesicles. Hence, the coat proteins have a double function: they are both responsible for formation of the transport vesicles (through membrane curvature and scission), and for the specific recruitment of cargo proteins in it.

ER-Golgi intermediate compartment (ERGIC)

After budding from the ER, these vesicles are addressed to the Golgi, but first transit through the ERGIC. This compartment was discovered more recently than the ER and Golgi, and its function is still a matter of debate. Indeed, two models are proposed for transport of vesicles from the ER to the Golgi through the ERGIC. The first is the transport complex model, in which the ERGIC is considered as a transient compartment that arises from the fusion of COPII transport vesicles. These bigger cargo containers then migrate to the Golgi using microtubules, or give rise to cis-Golgi by fusing together. Nevertheless, this model is largely based on experiments with overexpression of viral proteins, and does not correlate with many other experiments, especially those on the trafficking of ERGIC-53 (Appenzeller-Herzog, 2006).

On the other hand, new experiments are more in favor of the older stable compartment model (Klumperman et al., 1998). In this model, the ERGIC is considered as a real stable compartment, in close proximity but distinct from the ER. It operates as a first sorting station after ER-export, mediating both anterograde and retrograde transport. In that case, ER-to-Golgi transport occurs in two-steps: a first short-range step of transport of COPII vesicles from ERES to the ERGIC, and a second long-range transport of large COPI-coated anterograde carriers from the ERGIC to the *cis*-Golgi.

Regardless of the model, the ERGIC is the first sorting station after the ER, and mediates anterograde transport to the *cis*-Golgi and retrograde transport to the ER. Both of these mechanisms likely occur through COPI coated vesicles (Appenzeller-Herzog, 2006).

COPI coat assembly is initiated by activation of Arf1, mediated by the GEFs Gea1 or Gea2. Arf1-GTP then recruits proteins of the inner coat Sec21/Sec26/Ret2/Ret3, mediating cargo-binding activity. Subsequently, Sec27/Ret1/Sec28 proteins of the outer coat are recruited, and they assemble in a triskelion-like structure (Figure 12).

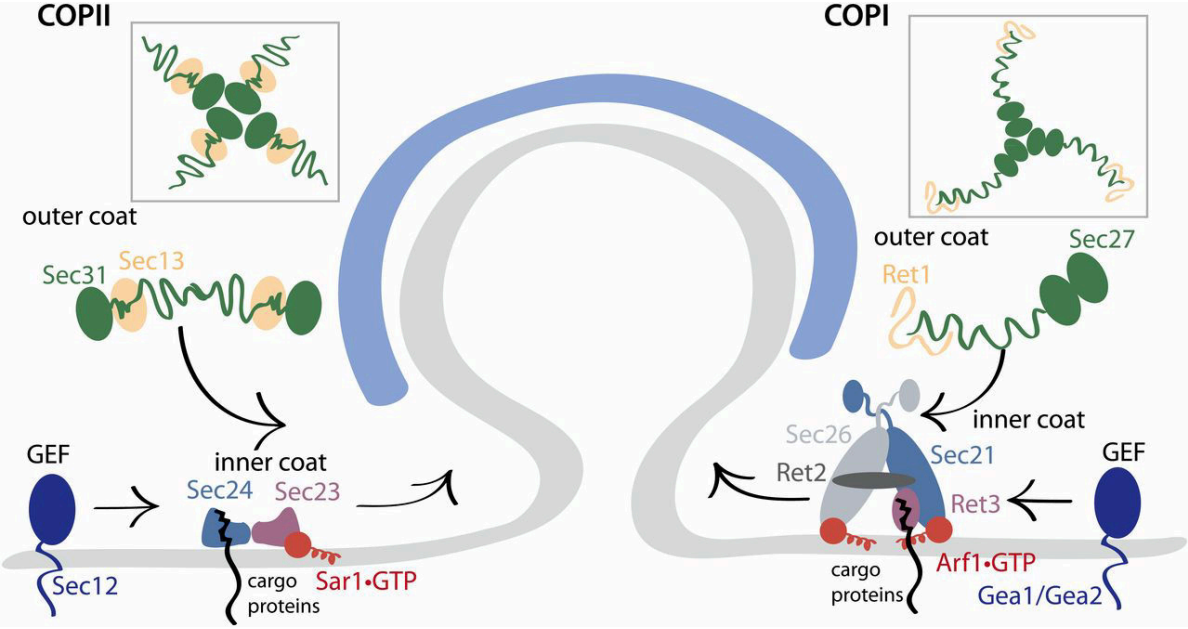


Figure 12- ***Assembly of the coat complexes COPI and COPII***

Activation of both COPI and COPII complexes are mediated by a small GTPase (SAR1/Arf) after its activation by a GEF. Then proteins of the inner coat are recruited, and mediate cargo-binding activity. Subsequently proteins of the outer layer are recruited and form lattice structures driving membrane binding. Finally, after concentrations of cargo proteins into the forming vesicle, the small GTPase mediates membrane fission. From *Barlowe and Miller*, 2013.

The Golgi complex

The Golgi is composed of a set of cisternae, the *cis*-Golgi, the *medial*-Golgi, and the *trans*-Golgi. When the vesicles arrive from the ERGIC, they fuse on the *cis*-Golgi face, and then progress to the *medial*- and *trans*-Golgi. But then again, two models are proposed for the anterograde movement of cargo proteins from *cis*- to *trans*-Golgi: the stable compartment model, postulating that proteins are transported from one cisternae to the next by vesicular transport; and the cisternal maturation model which suggests that the *cis*-Golgi cisternae maturates into *medial*- and then *trans*-Golgi, by acquiring and losing

Golgi-resident proteins specific of each cisternae (Boncompain and Perez, 2013). Although many studies supported each model, two studies showed direct evidence of a maturation of Golgi cisternae (Losev et al., 2006; Matsuura-Tokita et al., 2006). A proposed addition to this maturation model is the presence of export sites on every cisternae of the Golgi, thereby explaining different transport kinetics of cargo proteins.

The main role of the Golgi is the sorting of these cargo proteins, but it also has an important role for post-translational modification, especially glycosylation. Simultaneously to their transport, proteins are sequentially modified in each cisternae. Proteins that are destined to the ER are transported from the Golgi back to the ER through COPI vesicles (retrograde transport). Some retention signals directly interact with COPI coat proteins, such as the C-terminal KKXX motif which mediates retrieval of ER resident proteins.

For the other proteins, after passage through or maturation into *cis*- *medial*- and *trans*-Golgi cisternae, they are inserted into newly formed vesicles in the *trans*-Golgi network (TGN). The TGN will then address each protein to its final destination, and Golgi resident proteins will be retained in their specific Golgi cisternae. A known Golgi retention signal identified in plants is the KxD/E motif, that interacts directly with COPI coat proteins and mediates Golgi retention of endomembrane proteins (Gao et al., 2014). Interestingly this signal is found in many proteins from plants, yeast and animals.

In non-polarized cells, it is admitted that in absence of specific signals (return to the ER, resident proteins, or endosomes/lysosomes), a protein will be exported to the cell surface: it is the constitutive secretory pathway.

The endosomal network

Transport vesicles emerge from the *trans*-Golgi network and are then addressed to their next destinations: the plasma membrane (or more precise domains of the membrane in polarized cells), early endosomes, or the endosomal-lysosomal pathway (Gadila and Kim, 2016). The endosomal network is considered as another crucial sorting station, receiving cargo proteins both from the TGN via anterograde transport and the plasma membrane via endocytosis. This network includes early endosomes (EEs), late endosomes (LEs), lysosomes and recycling endosomes, each having special functions mediated by different resident proteins and pH acidification properties. Importantly, endosomes display homotypic fusion abilities (SNARE-mediated), and are therefore highly dynamic structures (Naslavsky and Caplan, 2018).

The first endosomal sorting station is the early endosome, to which all internalized proteins and some TGN-exported proteins converge. This compartment is actually formed by constant fusion of incoming endocytic vesicles containing internalized proteins. Proteins arriving to early endosomes will either be retained in it, which ultimately leads to lysosome degradation, be recycled to the cell membrane, or sent back to the TGN through retrograde transport. In order to mediate these different types of transport, the early endosomes are composed of specialized microdomains with specific proteins and functions (Scott et al., 2014), into which cargo proteins are sorted before being further transported.

Endo-lysosomal pathway

Proteins remaining in the endosomes will follow the endosomal-lysosomal pathway and be degraded. There are two hypothesis concerning this pathway: either the proteins are transported from the early endosomes to the late endosomes through vesicular transport, or early endosomes undergo maturation and gradually acquire late endosome markers (mechanism known as Rab conversion), before ultimately fusing with lysosomes. The second hypothesis is more robust, since structures harboring both Rab5 (EEs marker) and Rab7 (LEs marker) were identified, along with proteins involved in the switch (Poteryaev et al., 2010). In that scenario, multivesicular bodies (MVBs) are maturation intermediates rather than transport carriers.

Lysosomal hydrolases, which are essential to the degradation function of the lysosome, are addressed to the lysosome by a specific sorting in the TGN. Indeed, these proteins undergo glycosylation in the Golgi and ultimately harbor mannose-6-phosphate (M6P). In the TGN, M6P-receptors will bind the M6P on the enzymes, and mediate the addressing to the endo-lysosomal pathway. Binding to other receptors such as sortilin or LIMP-2 can also mediate lysosome addressing (Braulke and Bonifacino, 2009). Targeting to the lysosome can be mediated by signals located in the C-terminal domain of proteins, including di-leucine and tyrosine-based signals, recognized by adaptor proteins of the clathrin-coat.

Recycling to the plasma membrane

The internalization and recycling pathways constitute one of the major mechanisms of membrane traffic. Both proteins and lipids are internalized in huge amounts by endocytosis, and therefore need to be efficiently recycled back to the plasma

membrane. In the early endosomes, internalized receptors will be detached from their ligands due to acidic pH: ligands will be addressed to lysosomes for degradation, and receptors will be recycled. This recycling can occur directly through rapid traffic to the plasma membrane in a Rab4-dependent manner (fast recycling), or indirectly through the recycling endosomes (slow recycling). Recycling endosomes cluster in a perinuclear region close to the MTOC called the endocytic recycling compartment (ERC) (Naslavsky and Caplan, 2018).

The general assumption is that most membrane proteins are returned to the plasma membrane in absence of a specific sequence that signals for recycling, because targeting of a membrane protein to late endosomes or other intracellular destinations on the other hand requires a specific targeting motif (Maxfield and McGraw, 2004). Nevertheless, a few recent studies now hint that in some cases specific signals might be required for recycling (reviewed in (Naslavsky and Caplan, 2018)

Retrograde transport to the TGN

Endosomes-to-Golgi retrieval is mediated by the retromer-protein complex located in the endosomes. Originally identified in yeast, this complex is conserved among all eukaryotes (Gallon and Cullen, 2015). It is composed of a vacuolar protein sorting (Vps) trimer containing Vps26, Vps29, Vps35, and a heterodimeric subcomplex of sorting nexins (SNX). The Vps trimer, and more specifically Vps35, is able to bind a variety of cargo, and the SNX complex is necessary for recruitment of the retromer to the endosomal membrane. The retromer is involved in fast recycling to the plasma membrane (along with SNX27 and other proteins), and in retrograde transport from endosomes to the TGN. One well-described example of the latter is the retrieval of M6PR to the TGN after addressing of lysosomal enzymes to early endosomes. Indeed, these proteins contain M6P and need to be associated to their M6P-receptor (M6PR) in the TGN, for their specific sorting and addressing to lysosomes. The retromer-complex then mediates M6PR retrieval from the endosomes to the TGN. Similarly, cargoes destined to anterograde transport from early to late endosomes are associated with sortilin in the TGN, and after transport sortilin is also retrieved to the TGN through the retromer-complex.

Endocytosis

Endocytosis is a crucial mechanism for maintaining the composition and distribution of lipids and proteins in the membrane, and therefore has an effect on many cellular functions, including communication with the cell exterior. The tight regulation of both endocytosis and exocytosis controls the interactions of the cell with the exterior (Doherty and McMahon, 2009). Additionally, endocytosis is involved in the signaling and regulation of many other cellular processes, and is also an entry route for many pathogens. Endocytosis occurs in all cell types, but the mechanisms and proteins involved vary, as do the proportions of each type of endocytosis (clathrin-dependent or independent).

Clathrin-mediated endocytosis (CME) accounts for a large proportion of endocytic events, and more than 50 proteins were shown to be involved in this process (Kaksonen and Roux, 2018). First, proteins of the coat, including clathrin, cluster on the inner side of the plasma membrane. Nucleation of clathrin at the site of internalization is mediated by the AP-2 adaptor. Cargo proteins are concentrated to the coated region in order to be incorporated in the forming vesicle. Simultaneously, clathrin polymerizes, forming a curved lattice and stabilizing the deformation of the membrane during pit formation. The clathrin-coated pit undergoes progressive invagination leading up to scission: clathrin polymerization and accessory proteins aid to the formation of the vesicle neck, and the scission is then performed by dynamin, a GTPase that forms a helical polymer around the neck and mediates fusion of the membranes put in apposition. The clathrin-coated vesicles is thereby released from the membrane and internalized.

Clathrin-independent endocytosis (CIE) regroups various mechanisms, which are less characterized than CME. The most common clathrin-independent mechanism of endocytosis is the caveolar-type endocytosis, involving invaginations of 60-80nm at the cell surface, which are particularly abundant in some cells types (smooth muscle, fibroblasts, adipocytes). Multicaveolar assemblies are often observed, with many caveolae connected together at the plasma membrane. Other clathrin independent endocytic pathways such as flotillin-dependent, Arf6-dependent, or IL2Rβ pathway have been described. Additionally, phagocytosis and macropinocytosis are other types of endocytosis that can occur in specialized cells, in which materials or particles can be

internalized after that membrane invaginations form around the cargo (phagocytosis) or extracellular material is aspecifically internalized (macropinocytosis) (Doherty and McMahon, 2009).

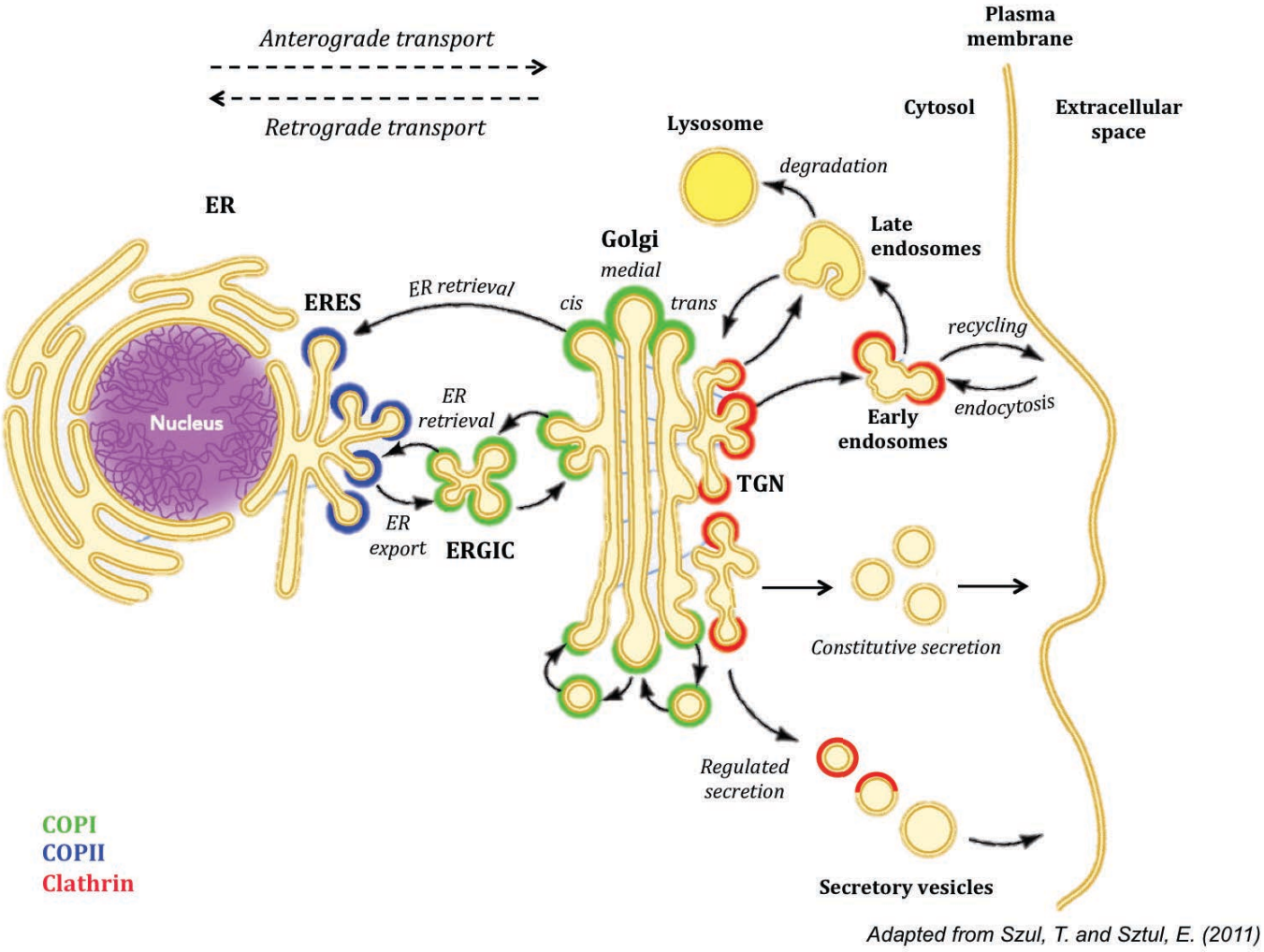


Figure 13- ***Trafficking through the secretory pathway***

Proteins synthesized in the ER are sent to the ERGIC after control of proper folding. A first sorting occurs, with retrograde transport of some ER resident proteins and anterograde transport of proteins to the *cis*-Golgi. Proteins are modified in each cisternae and sorted in the TGN. Again, proteins are returned to the ER, or to previous Golgi cisternae. Other proteins are sent to lysosomes, regulated secretion pathway, endosomes, or secreted to the cell surface.

2. Aims of the PhD project

Research on coronaviruses is of growing importance to human medicine, in light of the recent emergence of highly pathogenic human coronaviruses. Many knowledge gaps remain concerning the coronavirus life cycle, especially for MERS-CoV, which emerged only in 2012. We decided to focus on the assembly step of the virus, occurring at the ERGIC after the entry and replication steps. For proper assembly of new virions, all structural proteins (S, E, M and N) must be addressed to the assembly site, and must interact with each other. Nevertheless, the specific targeting signals allowing trafficking of the proteins to the assembly site and the domains involved in interactions between the structural proteins are still poorly characterized.

As the M protein was previously demonstrated to be the motor of the viral assembly, we decided to focus on the intracellular trafficking of the M protein of MERS-CoV. The aim of this project was to identify motifs involved in the trafficking of MERS-M to the assembly site.

In order to do so, we chose to work with the protein expressed alone in cells, so that we could decipher the trafficking mechanisms of M alone, before looking at their impact on infection. To study MERS-M intracellular trafficking, we used confocal microscopy and looked at the subcellular localization of the protein when expressed alone in cells, using co-stainings of several compartment markers. We constructed tagged forms of the proteins, allowing detection with various anti-tag antibodies, and thus making the co-staining experiments easier with a wider choice of antibodies. Additionally, the level of colocalization of the M protein (or its mutants) with the compartment markers was analyzed using Pearson's correlation coefficient (PCC).

To identify trafficking motifs, we constructed deletion mutants of the M protein in order to rapidly identify regions of the protein that were important for its trafficking. Then, single amino acid mutations in these regions allowed us to identify the precise targeting motifs.

As we have to date no antiviral active against coronaviruses at our disposal, another project focusing on antivirals was initiated in the lab. This project aims to identify and characterize an antiviral compound active against HCoV-229E. For this

purpose, a high-content screening (HCS) on the PRESTWICK library was performed in the lab, and four drugs were selected for further characterization: trifluoperazine dihydrochloride, perhexiline maleate, astemizole and digoxigenin. The inhibitory activity of the drugs was characterized in dose-response experiments and in experiments with addition of the drug at different steps of infection with HCoV-229E. In addition, cell survival after long incubation with the drugs was measured.

Digoxigenin was selected for further characterization and was also tested against HCV infection. Further tests using particles pseudotyped with the S protein of MERS-CoV (mimicking the entry step) and cells stably expressing the HCV replicon (bypassing the entry step) allowed us to have insight into the step of the viral cycle inhibited by digoxigenin.

3. Material & Methods

3.1 Project I: Intracellular trafficking of M proteins

3.1.1 Plasmid construction

The coding sequence of the M protein was cloned in the pCDNA3.1(+) vector, with or without a sequence coding for different tags, including HA (YPYDVPDYA), VSVG (YTDIEMNRLGK), and V5 (GKPIPNPLLGLDST). Total RNA from blood samples of infected patients was extracted by using the *Nucleospin RNA kit* (Macherey-Nagel) according to the manufacturer's instructions. Then, reverse transcription was performed using the high capacity cDNA reverse transcription kit (Applied Biosystems) and the M protein sequence was amplified by two successive PCRs, with the Q5® High-Fidelity DNA Polymerase (New England Biolabs). First, the sequence was amplified by using the two following primers: 5'-gacgagtgggtttaacgaact-3' and 5'-ggggatgccataacaatgaaa-3'. Then, to insert the sequence in expression vectors, the sequence was amplified with 5'-tcggatccaccatgtctaatatgacgcaactcactg-3' (primer A) and 5'cagaattcctaagctcgaagcaatgcaa-3' (primer B ; untagged protein) or by combination of primer A and 5'-tagaattcagctcgaagcaatgcaagttcaat-3' (primer C ; C-terminal tagged protein) or with 5'-acggatccaatatgacgcaactcactgagg-3' (primer D) with primer B (N-terminal tagged protein). PCR products were inserted between the BamH1 and EcoR1 restriction sites of the different vectors.

M protein deletion mutants were generated by PCR by using either primer A or D in combination with a reverse primer annealing at different positions of the M sequence. Deletion mutants were constructed as the following: MΔ5, MΔ10, MΔ15, MΔ20, MΔ40, MΔ60, MΔ80, MΔ100, MΔ120, respectively lacking the last 5, 10, 15, 20, 40, 60, 80, 100, and 120 last C-terminal amino acids. Mutant M proteins were generated by PCR using site-directed mutagenesis on an M wild-type matrix. Mutant or deleted proteins were then inserted into pCDNA3.1 vectors containing HA, VSVG or V5 tags.

The sequence of IBV-M from the Beaudette strain was amplified using primer E (5'-ctaagcttcccaacgagacaaattgtac-3') and primer F (5'-cagaattcttatgtgtaaagactacttc-3') for the N-terminal HA-tagged protein and a combination of primer G (5'-ttaagctttccatgcccaacgagacaaattg-3') and primer H (5'-ccgaattctgttgtaaagactacttcctc-3') for the C-terminal V5-tagged protein. PCR products were inserted between the EcoR1

and HindIII restriction sites of the different vectors. The IBV-M-VSVG construct was generated from digestion of the IBV-M-V5 construct by EcoRI and HindII, and ligation into the pCDNA3.1-V5-Ct vector between EcoRI and HindIII restriction sites. For deletion mutants we used the same strategy as for MERS-M, using primer E or G in combination with reverse primers annealing at different positions (MΔ5, MΔ10, MΔ15, MΔ20). Deleted M proteins were generated by PCR using site-directed mutagenesis on an M wild-type matrix. Mutant or deleted protein sequences were then inserted into pCDNA3.1 vectors containing HA, or VSVG tags.

For the construction of chimeras, the sequence of interest of each protein was amplified by PCR, using internal primers with 10 bases complementarity with the sequence of the other protein. The PCR products were then fused using the external primers annealing with the corresponding protein sequences, and the resulting PCR product was gel-purified and inserted into pCDNA3.1 vectors containing HA, VSVG or V5 tags. All the constructs were verified by DNA sequencing.

3.1.2 Cells and transfection

HeLa cells were maintained in MEM (*Minimum Essential Medium*) supplemented with 10% fetal calf serum and 1% Glutamax, at 37°C and 5% CO_2. 24h before transfection, HeLa cells were plated in 24-well plates on coverslips or in 6 well plates. The next day, plasmids encoding wild-type M protein or M mutant protein were transfected into HeLa cells using *TransIT®-LT1* Transfection Reagent (Mirus Bio) according to the manufacturer's instructions.

3.1.3 Immunofluorescence

At 18h post-transfection, cells were rinsed with PBS, fixed with 3% PFA and processed for immunofluorescence analysis. Cells were permeabilized with 0,1% Triton X100 in PBS for 5 min and then blocked with 10% goat or horse serum in PBS for 10 min. M protein was detected using anti-M pAbs (rabbit, Proteogenix) or anti-tag antibodies: anti-HA mAbs (3F10, Sigma Aldrich), anti-VSVG mAbs (P5D4, produced in the lab) or anti-V5 mAbs (ThermoFisher Scientific). For co-localization experiments, cells were double-labeled for M-proteins and a cellular marker, anti-calreticulin pAbs (CRT) for endoplasmic reticulum (ER) and anti-TGN46 pAbs for TGN (Biorad). All

primary antibodies were diluted in blocking buffer. In some cases, intracellular compartments were stained using an expression vector for a marker fused with the green fluorescent protein (GFP), which was co-transfected with the M protein. For ERGIC and TGN compartments, cells were co-transfected with M proteins and expression vectors for ERGIC53 fused to GFP (ERGIC53-GFP), and transmembrane domain and cytosolic tail of *cation-independent mannose-6-phosphate receptor* fused to GFP (GFP-CI-MPR), respectively.

For cell surface staining, cells were transfected with a vector expressing CD4 fused to GFP. After a 30 min incubation with primary antibodies, cells were washed 3 times for 5 min with PBS. The cells were then incubated with fluorescent secondary antibodies (cyanine-3 conjugated goat anti-mouse IgG; cyanine-3 conjugated goat anti-rabbit IgG, alexa488 conjugated donkey anti-rat IgG; cyanine-3 conjugated donkey anti-sheep IgG; alexa488 conjugated donkey anti-mouse IgG; alexa555 conjugated goat anti-rat IgG) and 1 µg/ml of 4',6-diamidino-2-phenylindole (DAPI) to stain the nuclei, for 20min in the dark. The cells were washed again 3 times for 5min with PBS, and mounted on slides.

3.1.4 Confocal microscopy and image analysis

Images were acquired using a laser scanning confocal microscope LSM 880 (Zeiss) and using a 63x oil immersion objective. Signals were sequentially collected using single fluorescence excitation and acquisition settings to avoid crossover.
The extent of colocalization was quantified by calculating the *Pearson's correlation coefficient* (PCC) using the JACoP plugin of *ImageJ*. The PCC examines the relationship between the intensities of the pixels of two channels in the same image. For each calculation, at least 15 images were analyzed to obtain a PCC mean. A PCC of 1 indicates perfect correlation, 0 no correlation, and -1 perfect anti-correlation.

3.1.5 Biotinylation and internalization

HeLa cells were seeded in 6-well plates and transfected the next day with pCDNA3.1-V5-N3Q-M, pCDNA3.1-V5-N3Q-MΔ20, pCDNA3.1-V5-N3Q-M-D211A,E213A, pCDNA3.1-V5-N3Q-M-K199A,G201A,Y203A,R204A, PCDNA3.1-V5-N3Q-M-K199A, pCDNA3.1-V5-N3Q-M-G201A, pCDNA3.1-V5-N3Q-M-Y203A or pCDNA3.1-V5-N3Q-M-

R204A. At 24h post-transfection, cells were washed on ice with ice-cold PBS, and incubated twice with 250 µg/mL of EZ-Link™ Sulfo-NHS-SS-Biotin (Pierce) diluted in PBS for 15 minutes in order to label cell surface proteins. Unfixed biotin was then quenched by two sequential incubations of the cells for 10 minutes with 50 mM Glycine/PBS.

For internalization assays, cells were biotinylated at 48h post-transfection and then incubated at 37°C for 30 min. The biotin of non-endocytosed proteins was then cleaved upon three 20 min incubations with glutathione buffer (50mM reduced glutathione, 75 mM NaCl, 75 mM NaOH, 10% FCS) followed by two 15 min incubations with iodoacetamide buffer (50mM iodoacetamide, 1% BSA, PBS). Cells were then lysed with B1 buffer (50 mM Tris pH 7.5, 100 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS, protease inhibitors cocktail) on ice. Lysates were centrifuged at 14,000 rpm at 4°C for 5 min to remove cellular debris, and were then incubated with 30 µL of streptavidin-conjugated agarose beads (Sigma) for 2h. Beads were then washed serially with 1mL of buffers B1, B2 (50 mM Tris pH 7.5, 100 mM NaCl, 2 mM EDTA, 0.1% Triton X-100, 0.5% SDS, 0.5% DOC), B3 (50 mM Tris pH 7.5, 500 mM NaCl, 2 mM EDTA, 0.1% Triton X-100) and B4 (50 mM Tris pH 7.5, 100 mM NaCl, 2 mM EDTA). Proteins were resuspended in Laemmli loading buffer and detected by immunoblotting. Samples were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and proteins were transferred on a nitrocellulose membrane (Amersham). Membrane-bound M proteins were then detected using a monoclonal anti-V5 antibody and horseradish peroxidase-conjugated secondary antibody. Detection was carried out by chemiluminescence (Pierce).

3.1.6 Glycosidases treatment.

HeLa cells were transfected with vectors expressing V5-M, V5-M-DxE, V5-M-KGYR or V5-N3Q-M proteins. 24h later, cells were lysed in B1 buffer. Then, 30 µL of lysates were mock-treated or treated with PNGase F or Endoglycosidase H according to the manufacturers instructions. Proteins were then separated on SDS-PAGE and detected by immunoblotting.

3.1.7 M-M interactions assay

HeLa cells were seeded in 10 cm dishes and transfected with vectors expressing V5-N3Q-M or V5-N3Q-M-KGYR. The next day, cell surface proteins were biotinylated at 4 °C and cross-linked with 0,8% PFA in PBS for 10 minutes. Then, PFA was quenched by washing the cells with 50 mM NH_4CL/PBS twice. Cells were lysed with B1 buffer and lysates were processed for streptavidin precipitation as previously described. Proteins were resuspended in non-reducing Laemmli loading buffer without heating and detected by immunoblotting.

3.2 Project II: Characterization of an antiviral against HCoV-229E

3.2.1 Viruses

The HCoV-229E strain containing the luciferase reporter gene in place of the accessory 4a gene was a kind gift from *Volker Thiel.* The HCV JFH1 strain used was previously engineered in the lab to express the A4 epitope (JFH1 CSA4N6) (Goueslain et al., 2010).

3.2.2 Cells

Huh-7 cells were maintained in DMEM (*Dulbecco's Modified Eagle's Medium*) supplemented with 10% fetal calf serum, at 37°C and 5% CO_2. 24h prior to transfection/transduction/infection, cells were plated in 96-well plates, 6-well plates or in T75 flasks. Huh-7 cells stably expressing an assembly-defective HCV replicon with a GFP reporter were available in the lab.

3.2.3 Infections

Huh-7 cells were plated in 96-well plates or in 24-well plates on coverslips, 24h prior to infection. Serial dilutions of drugs in medium were prepared right before use, from a stock solution at 10 µM (in DMSO). A solution of DMSO diluted in medium was used as a control. The cells were pre-incubated for 1h in absence or presence of different drug concentrations, and then cells were infected for 1h with HCoV-229E-Rluc or for 2h with HCV in presence of different concentrations of drug. The virus was removed and replaced by medium or drug dilution for 6h for HCoV-229E-Rluc and 30h for HCV. After

this post-infection step, cells were lysed for HCoV-229E and fixed 5 min with ice-cold methanol for HCV.

For HCoV-229E the lysed cells were used for measuring luciferase activity using a luminometer. For HCV, fixed cells were processed for immunofluorescence analysis using an anti-E1 antibody (A4, produced in the lab) for staining of HCV-infected cells and DAPI for nuclei staining. The number of infected cells and total cells were measured using ImageJ.

3.2.4 Viability assay

An MTS assay was performed to assess cell viability. Huh-7 cells were plated in 96-well plates. They were incubated 24h, 48h or 72h with the following concentrations of trifluoperazine dihydrochloride, perhexiline maleate, Astemizole and digoxigenin: 10 µM; 5 µM; 2,5 µM; 1,25 µM; 0,625 µM; and 0,3125 µM. DMSO diluted in DMEM was used as a control. After incubation with the drug, the medium was replaced with MTS reagent (Promega) diluted in DMEM and cells were incubated at 37°C and the OD490nm was measured.

3.2.5 Pseudoparticles

HCoV-229E pseudoparticles (229Epp) were generated by transfection of three expression vectors into HEK293-T cells: one for the surface glycoprotein S (300ng), one for the capsid and polymerase proteins of murine leukemia virus (MLV) pTG-gag-pol, (300 ng) and one coding for the minigenome containing the firefly-luciferase reporter gene pTG-luc (400 ng). Transfected cells were then incubated for 72h at 33°C. The supernatant was harvested and filtered on 0,45 µm filters.

Huh-7 cells were plated in 96-well plates. The next day, cells were pre-incubated 1h with the drug or DMSO, and inoculated with 50 µL of pseudoparticles per well for 3h at 37°C. The medium was then changed and the cells incubated at 37°C for 48h. 48h post-infection cells were lysed, and the luciferase activity was measured using a luminometer.

4. Results

4.1 Project I: Intracellular trafficking of M proteins

Coronaviral structural proteins are produced in the ER, and need to be transported to the ERGIC assembly site, and properly interact with each other in order to be assembled into new virions. The M protein is of special importance in this regard, because of its ability to interact with all other structural proteins S, N and E. Furthermore, it was observed for many M proteins that they were capable of going beyond the assembly site in the secretory pathway, mostly to the Golgi and TGN. Hence, they must contain specific targeting signals that can be recognized by the cellular transport machinery. Identifying the signals involved and deciphering the mechanisms by which they specifically target viral proteins to or near the assembly site is of great importance to further understand the viral cycle. Nevertheless, very few M proteins were extensively studied, and most of the trafficking signals that mediate their transport are still unknown.

Importantly, the trafficking mechanisms of the M protein of the recently emerged coronavirus MERS-CoV, and more generally its assembly, were not studied in depth yet. The aim of our study was to study MERS-M intracellular trafficking and to identify signals involved in its specific localization.

4.1.1 MERS-M localizes in the TGN

To begin our characterization of the trafficking of MERS-M, we cloned the sequence of the M protein in a pCDNA3.1 vector, with or without different tags. We generated pCDNA3.1-MERS-M, pCDNA3.1-HA-MERS-M, pCDNA3.1-V5-MERS-M, pCDNA3.1-MERS-M-V5, and pCDNA3.1-MERS-M-VSVG. The use of different tags allowed us first to have more flexibility in terms of antibodies for the co-stainings, and also to use specific properties of certain tags. First, we transfected the pCDNA3.1-HA-MERS-M construct into HeLa cells, and analyzed its subcellular localization in immunofluorescence and confocal microscopy, using co-staining with different compartment markers: TGN46 (TGN), GFP-CI-MPR (TGN), ERGIC-53 (ERGIC), calreticulin CRT (RE) and CD4 (cell surface) (Figure 14A).). The GFP-CI-MPR is a

reporter that was shown to localize in TGN in HeLa cells (Waguri et al., 2003a). It is composed of the GFP fused to the transmembrane domain and cytosolic tail of the cation-independent mannose-6-phosphate receptor. The images were then analyzed to calculate the Pearson Correlation Coefficient (PCC) with ImageJ software, to quantify colocalization of the M protein with the compartment markers (Figure 14B). The PCC measures the pixel-by-pixel fluorescence signals of two images. A PCC of 1 means perfect correlation, 0 no correlation, and -1 perfect anti-correlation. For colocalization analysis, it is usually admitted that a PCC below 0,5 represents no or partial colocalization, a PCC between 0,5 and 0,7 suggest partial colocalization, and a PCC above 0,7 represents strong colocalization.

The HA-MERS-M protein showed a strong colocalization with both TGN46 and GFP-CI-MPR, as seen in the merges in Figure 14A. This was confirmed with a PCC of respectively 0.878 (+/-0.014) and 0.852 (+/-0.012) for the MERS-CoV M with TGN46 and GFP-CI-MPR markers (Figure 14B). On the other hand, PCCs of HA-MERS-M with calreticulin and CD4 were both below 0,4, and the PCC between HA-MERS-M and ERGIC-53 was slightly higher (0.524 +/-0.03). These results indicate that the HA-MERS-M protein localizes to the TGN when expressed alone in cells.

We then wanted to check whether or not the tags had an effect on MERS-M localization. We transfected pCDNA3.1-MERS-M (untagged), pCDNA3.1-HA-MERS-M, pCDNA3.1-V5-MERS-M, pCDNA3.1-MERS-M-V5, or pCDNA3.1-MERS-M-VSVG in HeLa cells, and analyzed their co-localization with TGN46 in immunofluorescence (Figure 14C). The control untagged M was detected with an anti-M antibody directed against the C-terminal part of the protein. This untagged M protein strongly colocalized with TGN46, as seen for HA-MERS-M, with a PCC of 0.857 (+/-0.015). All tagged proteins also strongly colocalized with TGN46 regardless of the type (V5, VSVG, HA) or position (N-terminal or C-terminal) of the tag, as confirmed with all PCCs >0,8 (Figure 14D).

Altogether, these results demonstrate that MERS-M protein expressed alone in cells is localized in the TGN, and that adding a tag does not alter this localization no matter its position and type.

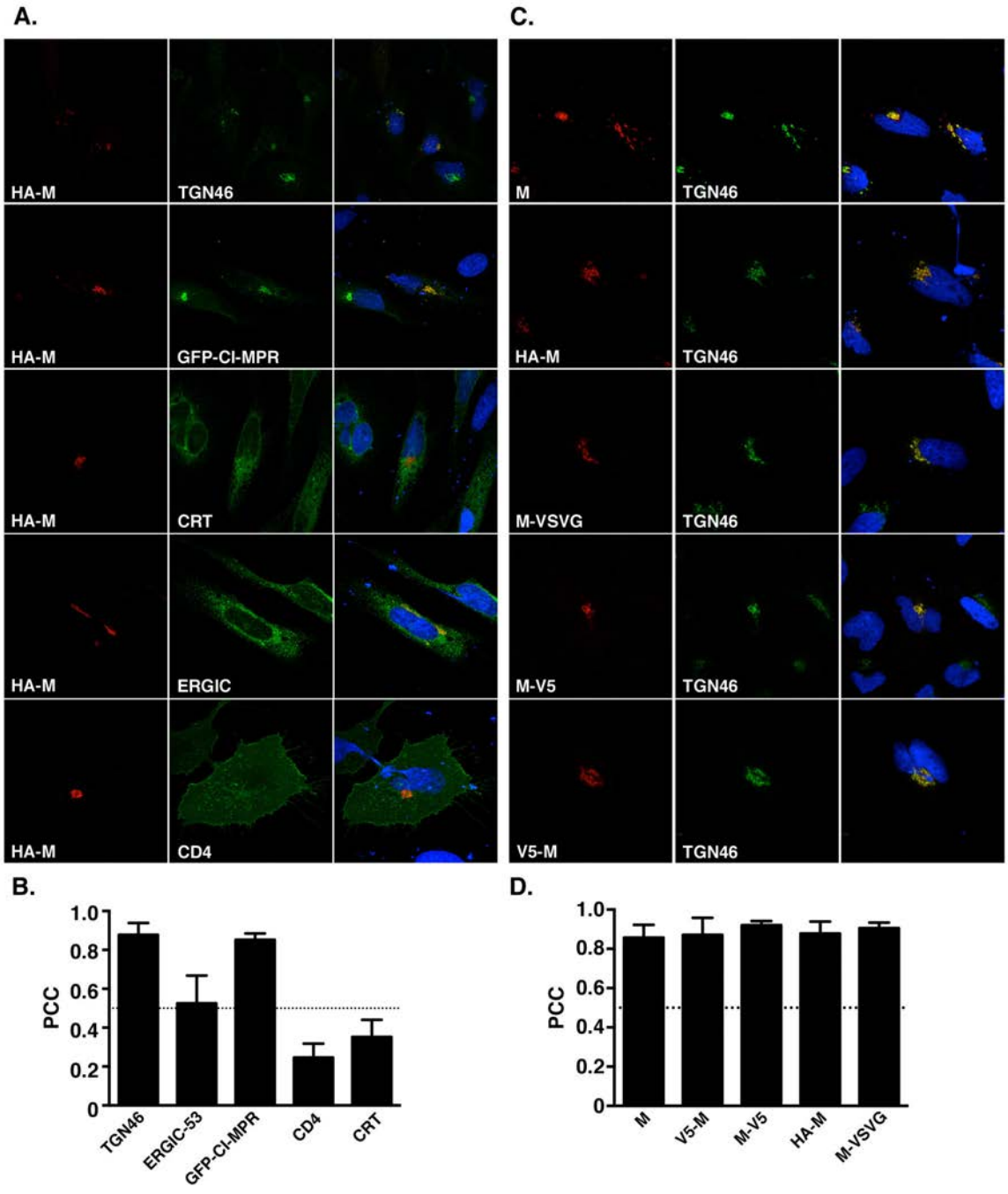


Figure 14- ***MERS-M subcellular localization***

(A, B) Cells expressing HA-MERS-M protein in combination with GFP-CI-MPR, ERGIC-53-GFP or CD4 fused with GFP were labeled with an anti-HA antibody. To detect the ER or the TGN compartments, cells were double-labeled for HA and calreticulin (CRT) or TGN46 using specific antibodies. Pearson's correlation coefficients were calculated for each combination of co-staining. (C, D) Cells expressing untagged MERS-M, HA-MERS-M, MERS-M-VSVG, MERS-M-V5 or V5-MERS-M were double-labeled with anti-M antibody together with an anti-TGN46 antibody. Pearson's correlation coefficients were calculated for each co-stainings.

4.1.2 Identification of an ER export signal

In order to identify motifs on MERS-M implicated in its traffic to the TGN, we started by constructing deletion mutants of the C-terminal domain. Indeed, this domain makes up more than half of the protein and is more exposed than the transmembrane segments, making it more accessible for recognition by cellular proteins. Furthermore, many known trafficking signals of the secretory pathway are located in the cytosolic-exposed domain of cargo proteins (Barlowe, 2003; Parmar et al., 2014).

Hence, we constructed mutants MERS-MΔ20, MERS-MΔ40, MERS-MΔ60, MERS-MΔ80, MERS-MΔ100, and MERS-MΔ120, lacking respectively the last 20, 40, 60, 80, 100, or 120 amino acids of the C-terminal domain. The Δ120 deletion is located right after the third transmembrane segment, at the beginning of the C-terminal domain, and the other deletions subdivide the C-terminal domain as depicted in Figure 15A. The constructs were transfected into HeLa cells and their subcellular localization was analyzed by immunofluorescence and confocal microscopy, with a TGN46 co-staining.

Interestingly, mutants MERS-MΔ20-V5, MERS-MΔ40-V5, MERS-MΔ60-V5, MERS-MΔ80-V5, and MERS-MΔ100-V5 all exhibited a cell surface localization, with a loss of retention into the TGN when compared to the wild type M (Figure 15B). On the other hand, mutant MERS-MΔ120-V5 had a very low expression rate, but positive cells showed a loss of localization to the TGN and displayed a more reticular pattern, likely corresponding to an ER localization. This would suggest that the Δ120 deletion impairs proper folding of the protein, resulting in a blockade in the ER. This is not surprising considering that this mutant lacks the whole C-terminal domain, including the amphipathic region.

Concerning the other mutants, as deletion of the last 20 amino acids was sufficient to disrupt MERS-M TGN localization, this region of the protein likely contains localization signals. In that light, we decided to further characterize the MΔ20 mutant and especially the role of the last 20 amino acids in ER export.

To that end, we constructed mutant MERS-MΔ20-VSVG, in order to compare its localization to the one of HA-MERS-MΔ20, and decipher the ER export mechanism. Indeed, some tags have an effect on ER export, whereas others don't : an N-terminal HA

tag has no effect on ER export, whereas the VSVG tag contains a known di-acidic (DxE) ER export signal. Additionally, in our hands, a C-terminal V5-tag was also able to induce ER export although it does not contain a known export signal.

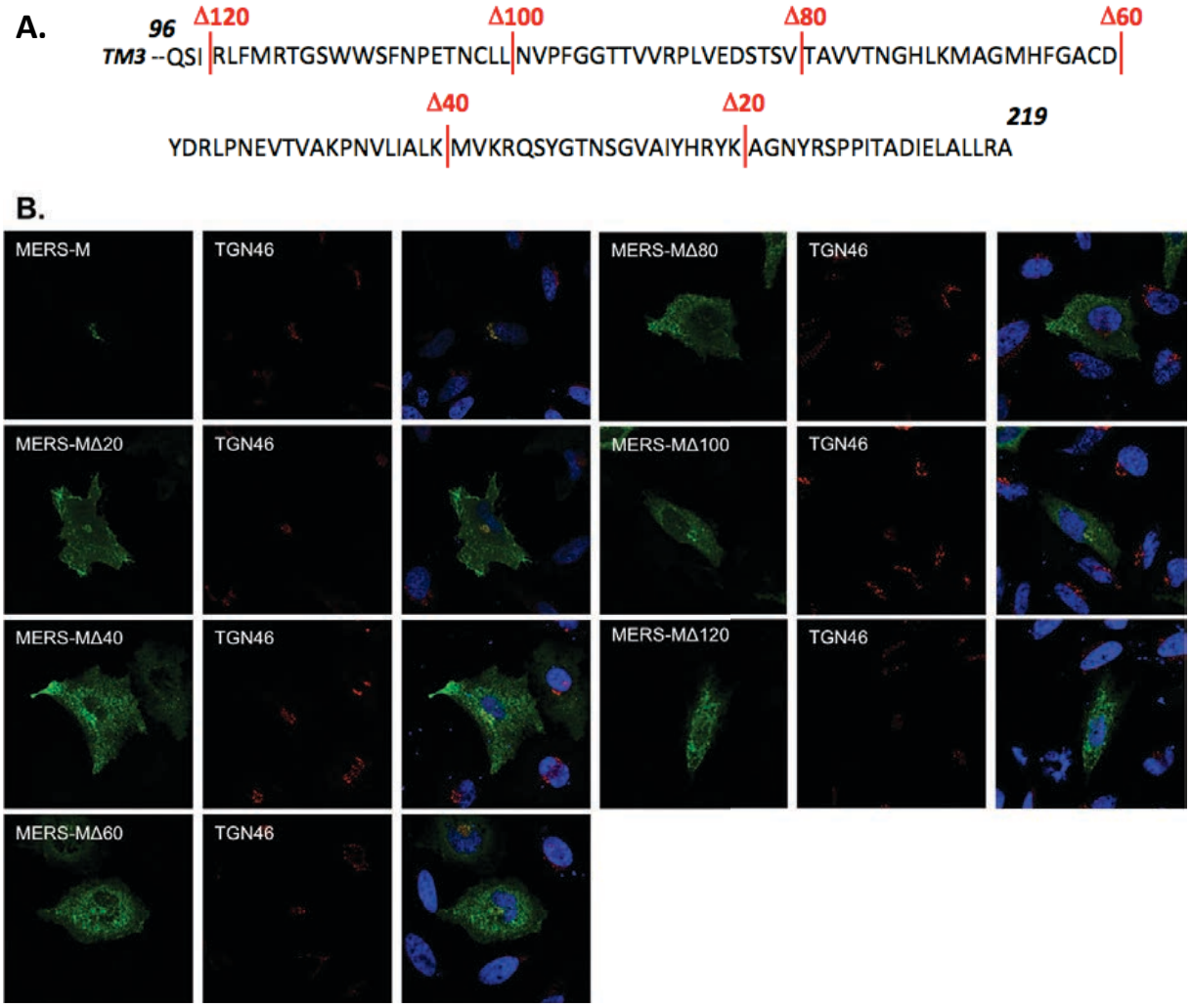


Figure 15 - ***C-terminal deletions of MERS-M disrupt TGN localization***

(A) Sequence of the C-terminal domain of MERS-M (96 to 219 amino acids), with position of the deletions marked in red. The Δ120 deletion is located 3 amino acids after the third transmembrane segment (TM3). (B) Subcellular localization of MERS-M deletion mutants. Cells expressing MERS-MΔ20-V5, MERS-MΔ40-V5, MERS-MΔ60-V5, MERS-MΔ80-V5, MERS-MΔ100-V5, or MERS-MΔ120-V5 were double-labeled with anti-V5 and anti-TGN46 antibodies, and their subcellular localization was analyzed by confocal microscopy.

HeLa cells were transfected with pCDNA3.1-HA-MΔ20 or pCDNA3.1-MΔ20-VSVG, and co-stained with an anti-tag (HA or VSVG) and a marker for the TGN (TGN46), the ER (CRT), or the cell surface (CD4) (Figure 16). As previously observed, the mutant protein showed a different localization depending on the tag used: the HA-MΔ20 protein

colocalizes with the ER marker CRT (Figure 16B), whereas the MΔ20-VSVG protein colocalizes with the cell surface marker CD4 (Figure 16C).

The PCCs were calculated for each construction with the three co-stainings (Figure 16D). For the wild type protein the PCCs were of 0.905 (+/-0.007) for TGN46 and of 0.296 (+/-0.014) and 0.239 (+/-0.017) for CRT and CD4 respectively. The PCC calculations confirmed colocalization of HA-MERS-MΔ20 with the ER marker (PCC of 0.855+/-0.011), accompanied by a loss of TGN localization (PCC of 0.369+/-0.019) and a slightly higher localization with the cell surface marker CD4 (PCC of 0.539 +/-0.018) when compared with the wild type. The PCC of MERS-MΔ20-VSVG with TGN46 also dramatically decreased when compared to the wild type (0.331+/-0.025), accompanied with a strong increase of the PCC with CD4 (0.79+/-0.016) and a slight increase of the PCC with CRT (0.475+/-0.018).

These results suggested the presence of important trafficking motifs in the last C-terminal 20 amino acids of the protein. More precisely, the fact that the HA-MERS-MΔ20 localizes to the ER and the MERS-MΔ20-VSVG to the cell surface suggested that the HA-MERS-MΔ20 was blocked in the ER because of a lack of export, that could be rescued by the presence of the DxE ER export signal present within the VSVG tag (YTDIEMNRLGK), confirming what we first observed with the V5-tagged protein.

Indeed, analysis of the sequence of the last 20 amino acids of the M protein showed the presence of a similar potential di-acidic signal $_{211}$DIE$_{213}$. We therefore mutated this signal by replacing both D211 and E213 residues by alanine on the full-length protein, generating the constructs HA-MERS-M-D211A;E213A (HA-M-DxE) and MERS-M-D211A;E213A-VSVG (M-DxE-VSVG), and analyzed their subcellular localization. Once again, we performed co-stainings of MERS-M with the ER, cell surface or TGN markers (Figure 16 A, B, C). Similarly to HA-MERS-MΔ20, HA-M-DxE localized in the ER (PCC with CRT >0,8), whereas M-DxE-VSVG localized in the TGN (PCC with TGN46 >0,8) as the MERS-M wild type (Figure 16 B, C, D). Taken together, these results show that the DxE signal located in the C-terminal part of MERS-M is a functional ER export signal, and that the VSVG tag is able to compensate for the D211A;E213A mutation.

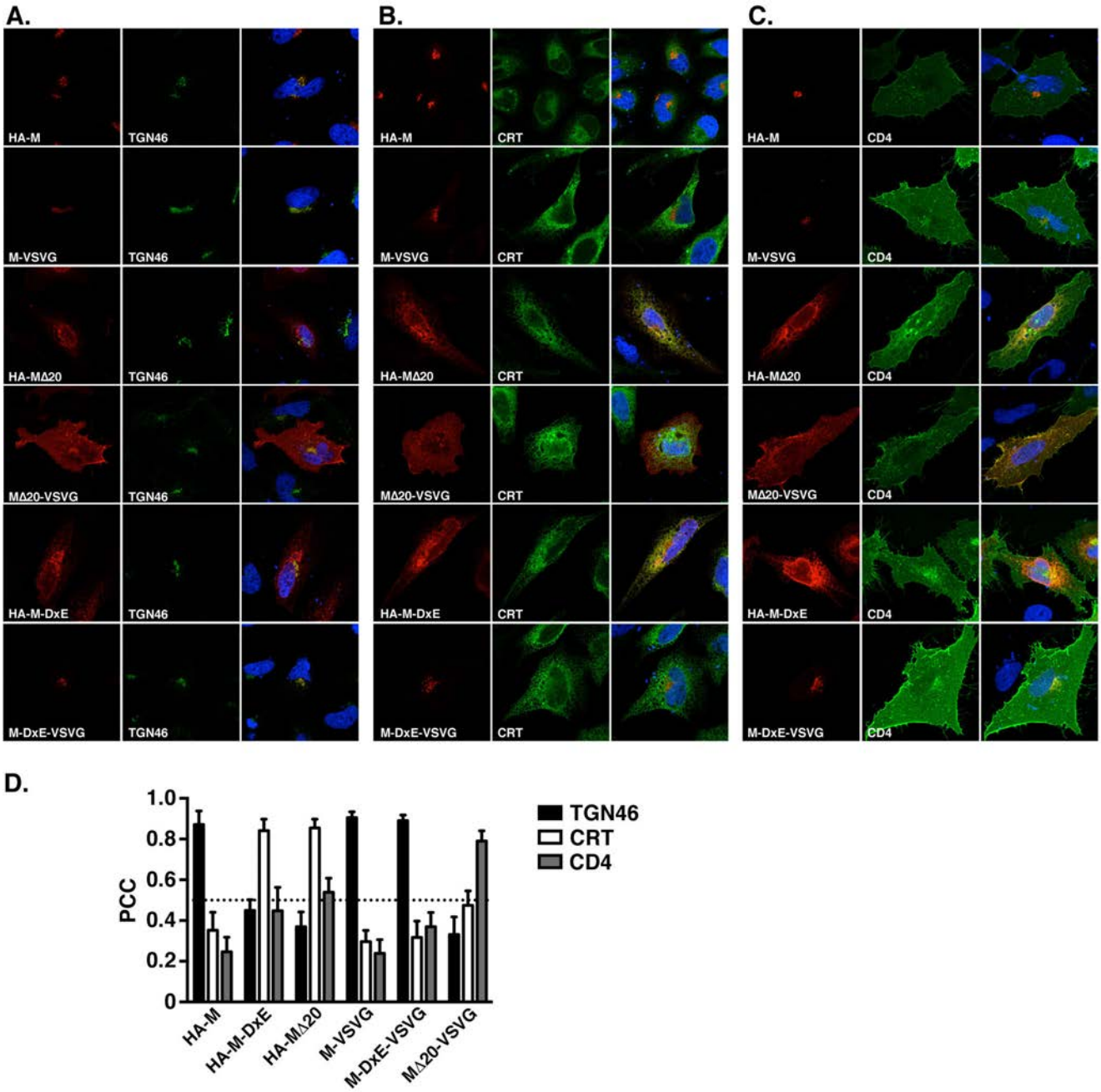


Figure 16 – ***The last 20 C-terminal amino acids of MERS-M contain motifs involved in its subcellular localization.***

(A, B, C) Cells expressing the HA-M, M-VSVG, HA-MΔ20, MΔ20-VSVG, HA-M-DxE or M-DxE-VSVG were stained with an anti-tag antibody, and a cellular marker. The TGN was stained with an anti-TGN46 antibody, the ER with an anti-CRT antibody, and the cell surface by co-transfection with a CD4-GFP construct. (D) PCCs were calculated for each co-staining.

The presence of the DxE signal is responsible for the exit of the M protein from the ER but not its retention in the TGN, since it is commonly accepted that in non-polarized cells, the constitutive secretory pathway leads to the plasma membrane by default (i.e in absence of specific addressing/retention signals). Considering that and the fact that the MERS-MΔ20-VSVG protein migrates to the cell surface and is not retained intracellularly, we looked for the presence of another signal in the last 20 amino acids of MERS-M, involved in its TGN retention.

4.1.3 Identification of a TGN retention signal

In order to identify the motif responsible for TGN retention in the last 20 residues of the C-terminal domain, we constructed smaller 5 residues deletion mutants: MERS-MΔ5-VSVG, MERS-MΔ10-VSVG and MERS-MΔ15-VSVG, lacking respectively the last 5, 10, or 15 residues of the C-terminal domain (Figure 17A). These mutants were constructed with a C-terminal VSVG tag in order to rescue the ER export of the proteins.

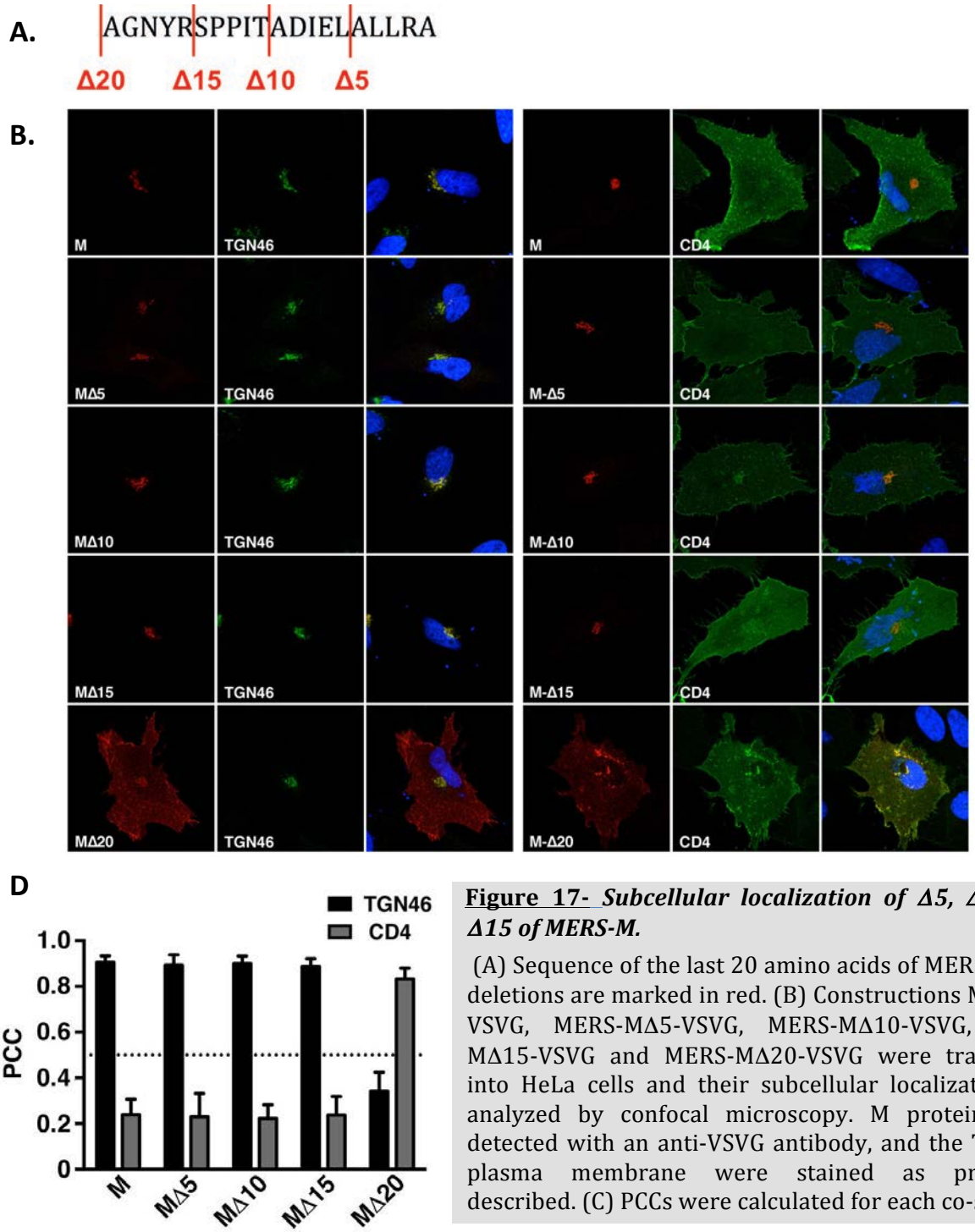


Figure 17- ***Subcellular localization of Δ5, Δ10 and Δ15 of MERS-M.***

(A) Sequence of the last 20 amino acids of MERS-M. The deletions are marked in red. (B) Constructions MERS-M-VSVG, MERS-MΔ5-VSVG, MERS-MΔ10-VSVG, MERS-MΔ15-VSVG and MERS-MΔ20-VSVG were transfected into HeLa cells and their subcellular localization was analyzed by confocal microscopy. M proteins were detected with an anti-VSVG antibody, and the TGN and plasma membrane were stained as previously described. (C) PCCs were calculated for each co-staining.

The subcellular localization of the resulting mutants was analyzed by immunofluorescence and confocal microscopy, and compared to the localization of MERS-M-VSVG and MERS-MΔ20-VSVG (Figure 17).

MERS-MΔ5-VSVG, MERS-MΔ10-VSVG and MERS-MΔ15-VSVG all strongly colocalized with TGN46 and did not colocalize with the cell surface marker CD4, similar to the wild type M protein (Figure 17 B). The calculation of the PCCs confirmed this with all PCC with TGN46 being >0,8 and PCC with CD4 being <0,4 (Figure 17C). This is in contrast with MERS-MΔ20-VSVG strongly localizing at the cell surface (PCC with CD4 > 0,8).

These results suggest that the TGN retention of the MERS-M protein is mediated by a signal located in between deletions Δ15 and Δ20. This five residues AGNYR sequence is showed in Figure 17A. In order to identify amino acid residues responsible for MERS-M retention in the TGN, we mutated each residue (except the alanine) individually to alanine on the MERS-MΔ15-VSVG protein. The constructs MERS-MΔ15-G201A-VSVG, MERS-MΔ15-N202A-VSVG, MERS-MΔ15-Y203A-VSVG and MERS-MΔ15-R204A-VSVG were transfected into HeLa cells and their subcellular localizations were analyzed by confocal microscopy. Cells were co-stained for TGN (Figure 18A) or plasma membrane (Figure 18B).

The mutants MERS-MΔ15-G201A-VSVG, MERS-MΔ15-Y203A-VSVG and MERS-MΔ15-R204A-VSVG showed an increased export at the cell surface when compared with MERS-MΔ15-VSVG, accompanied by a marked decrease of colocalization with the TGN marker TGN46. This was confirmed with calculation of the PCCs, with PCCs with the plasma membrane marker CD4 being all above 0,7 whereas PCCs with TGN marker TGN46 dropped below 0,5 (Figure 18C). On the other hand, mutant MERS-MΔ15-N202A-VSVG still strongly colocalized with TGN46 at a level similar to MERS-M-VSVG and MERS-MΔ15-VSVG (PCC> 0,8) and did not colocalize with the CD4 cell surface marker (PCC<0,4).

These results indicated that the residues G201, Y203 and R204 were involved in MERS-M retention to the TGN, as their mutation resulted in partial export of the mutant proteins to the cell surface. On the other hand, mutation of residue N202 had no effect on the localization of the protein, indicating that this residue is not involved in retention of MERS-M to the TGN.

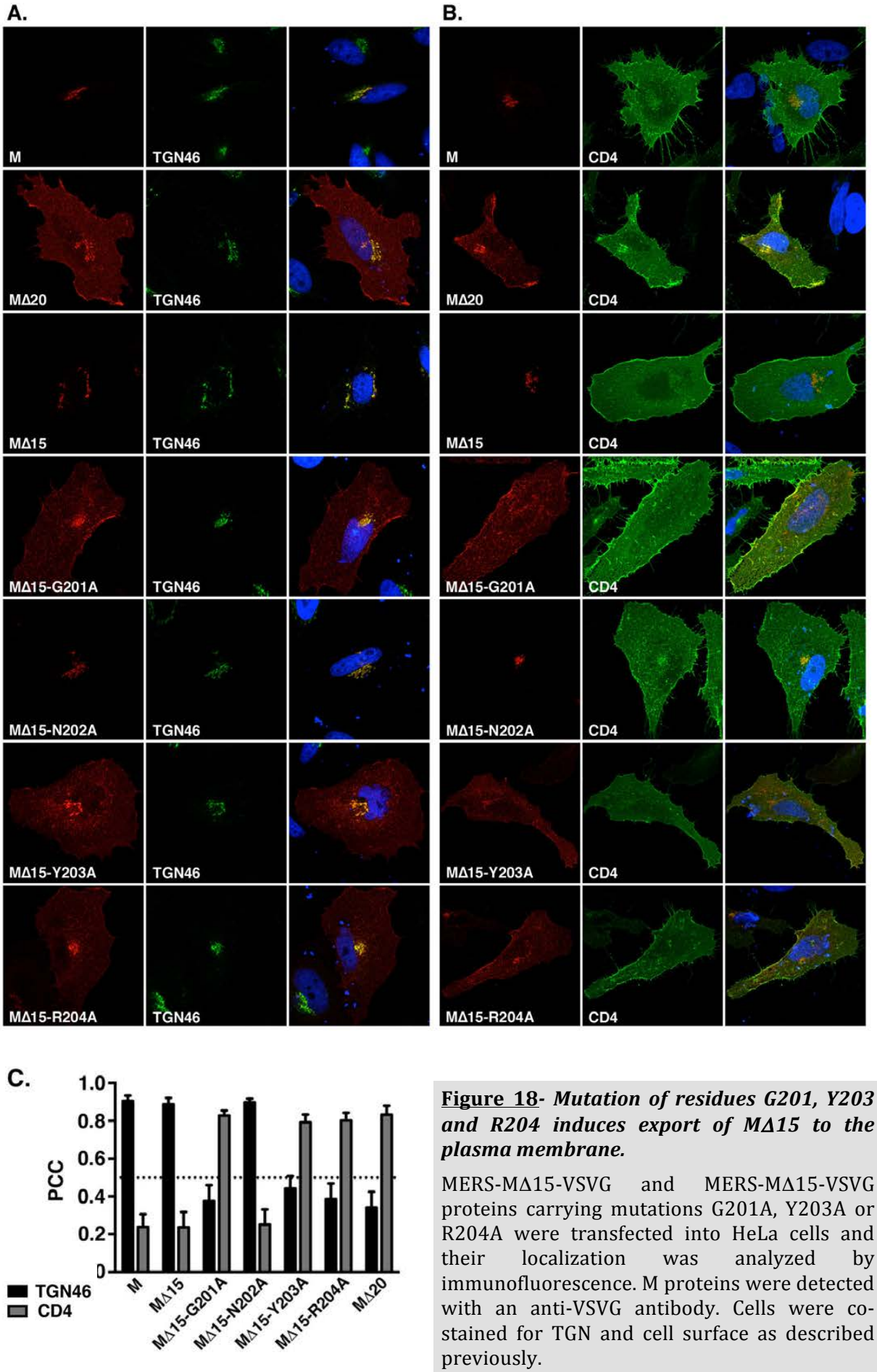


Figure 18- ***Mutation of residues G201, Y203 and R204 induces export of MΔ15 to the plasma membrane.***

MERS-MΔ15-VSVG and MERS-MΔ15-VSVG proteins carrying mutations G201A, Y203A or R204A were transfected into HeLa cells and their localization was analyzed by immunofluorescence. M proteins were detected with an anti-VSVG antibody. Cells were co-stained for TGN and cell surface as described previously.

To confirm these findings, we introduced G201, Y203 and R204 mutations individually in the full length MERS-M-VSVG protein.

Constructions MERS-M-G201A-VSVG, MERS-M-Y203A-VSVG and MERS-M-R204A-VSVG were transfected into HeLa cells and their subcellular localization was analyzed by immunofluorescence (Figure 19B, C). As expected, individual mutations of G201, Y203 and R204 into alanine led to partial export of the full-length protein to the cell surface, disrupting intracellular retention into the TGN, as was observed for the MΔ15 protein mutants. This was confirmed by the PCC calculations (Figure 19E) with all PCCs of the mutants with CD4 being above 0,7 and PCCs with TGN46 being under 0,5.

Additionally, in order to ensure that we identified the whole motif, we investigated if other residues upstream of the G201 could also be part of the signal. To that end, we looked for residues conserved among Betacoronaviruses upstream of the G201A residue (Figure 19A). As residues Y195A, R197A and K199A were quite conserved and close to G201 we decided to mutate those residues individually in the MERS-M-VSVG to check whether they were involved in trafficking of the protein to the TGN. Hence, MERS-M-Y195A-VSVG, MERS-M-R197A-VSVG and MERS-M-K199A-VSVG were transfected into HeLa cells and their subcellular localization was analyzed by immunofluorescence.

Mutations Y195A and R197A had no effect on the localization of the protein compared to the wild type (Figure 19D), both mutants still strongly colocalizing with the TGN marker TGN46.

Interestingly, the K199A mutation led to the export of the mutant protein to the cell surface, in a similar manner as for mutations G201A, Y203A and R204A (Figure 19B, C). This was confirmed with the PCC analysis, as the MERS-M-K199A mutant has a PCC with CD4 above 0,7 and a PCC with TGN46 below 0,5 (Figure 19E). This result indicated that the residue K199 was a part of the motif involved in the TGN retention of the MERS-M protein. Thus, we constructed a quadruple mutant MERS-M-K199A;G201A;Y203A;R204A-VSVG (M-KGYR) and analyzed its subcellular localization (Figure 20B,C). The quadruple mutant M-KGYR was also exported to the cell surface, and the extent of its colocalization with TGN46 and CD4 was comparable to those of the single mutants, as illustrated by the PCCs (Figure 19E).

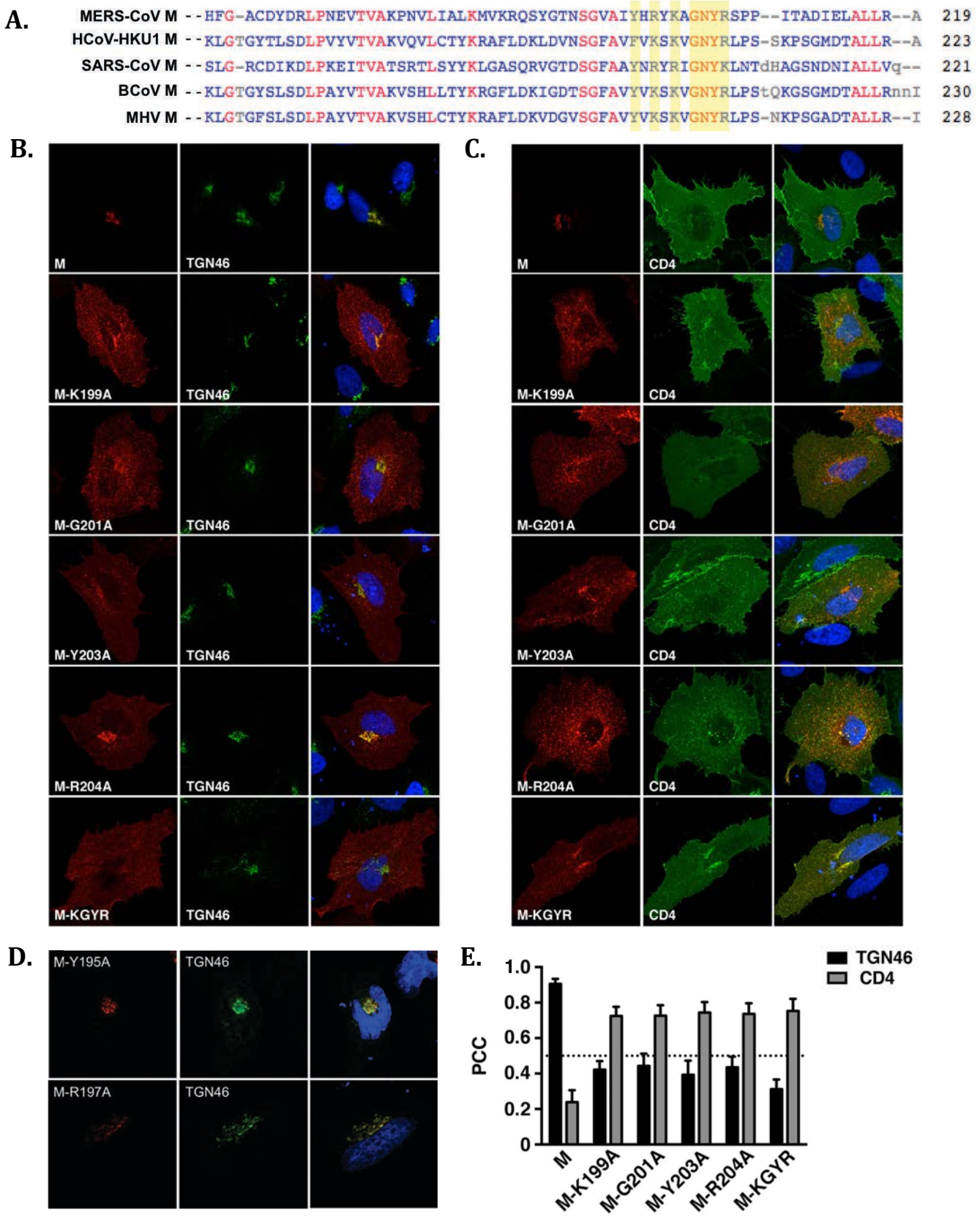


Figure 19- ***Identification of a KxGxYR motif involved in TGN localization of MERS-M***

(A) Alignment of the last 20aa of M protein of *betacoronaviruses.* Conserved residues that were mutated are indicated in yellow. (B, C, D) Lysates of cells expressing M protein wild type or carrying mutations Y195A, R197A, K199A, G201A, Y203A, R204A, or the four mutations were processed for immunofluorescence. M proteins were detected with an anti-VSVG antibody and the TGN and plasma membrane were stained as previously described. (E) PCCs were calculated for each co-staining.

We next wanted to confirm the role of the DxE and KGYR signals by investigating the glycosylation status of the M protein and of our mutants. Indeed, MERS-M contains an N-glycosylation site in its N-terminal domain (MSNMTQLTE), and as specific sugars are added to glycosylated proteins in the ER and the Golgi, the glycosylation status of a protein gives us insight on its localization and trafficking. For this purpose, we used EndoH, which cleaves N-linked mannose rich oligosaccharides that are not highly processed such as the ones added on proteins in the ER, and PNGase F, which is able to cleave N-linked high mannose, hybrid, and complex oligosaccharides, therefore cleaving almost any N-glycans.

The wild type protein migrated as three bands in western blot (Figure 20). The first band around 25 kDa corresponds to the unglycosylated form of the M, as confirmed by the migration profile of the N-glycosylation mutant containing only that band (N3Q-M). Similarly, this was confirmed by the same pattern arising from treatment of MERS-M with PNGase F. The second band around 30 kDa was sensitive to EndoH, indicating that it corresponds to M proteins glycosylated in the ER, that haven't reached the Golgi. The third band is more diffuse and seems to correspond to M proteins with sugars that were further modified in the Golgi, as this band was not sensitive to EndoH but was sensitive to PNGase F.

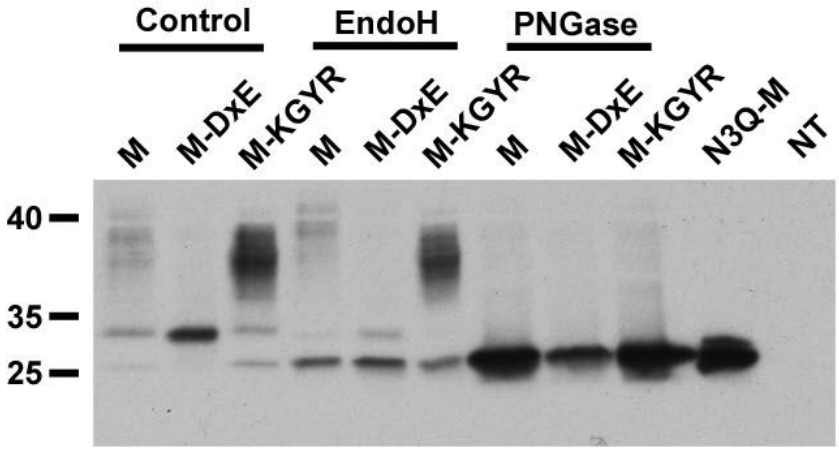


Figure 20- ***Glycosylation status of MERS-M mutants***.

V5-MERS-M, V5-MERS-M-DxE, V5-MERS-M-KGYR, or V5-MERS-N3Q-M were transfected into HeLa cells. Lysates were treated with EndoH or PNGase F, except the lysate of the V5-MERS-N3Q-M that was left untreated. The proteins were detected by immunoblot using an anti-V5 antibody.

In accordance with our previous results, the M-DxE mutant migrated as only one 30 kDa EndoH sensitive band, confirming that this protein is blocked and accumulated

within the ER (Figure 20). This result confirms the localization of the M-DxE mutant to the ER and therefore its role as a functional ER export signal for MERS-M. The M-KGYR mutant has a migration profile similar to the one of the wild type protein, but with an increased N-glycosylation, which is consistent with a better trafficking through the Golgi (Figure 20).

In order to confirm and quantify the cell surface localization of the mutant proteins M-K199A, M-G201A, M-Y203A, M-R204A and M-KGYR, we performed a cell surface biotinylation assay. As the migration profile of the MERS-M protein in western blot renders quantification of protein amounts difficult (see Figure 20), we decided to mutate the glycosylation site (mutation N3Q) on each construction, in order to have a clearer migration profile with only one unglycosylated band. All mutants were constructed with a V5 N-terminal tag, so that the intracellular trafficking of the proteins would not be modified by the tag. We first verified that the localization of the proteins was not affected by the mutation of the glycosylation site (Figure 21A).

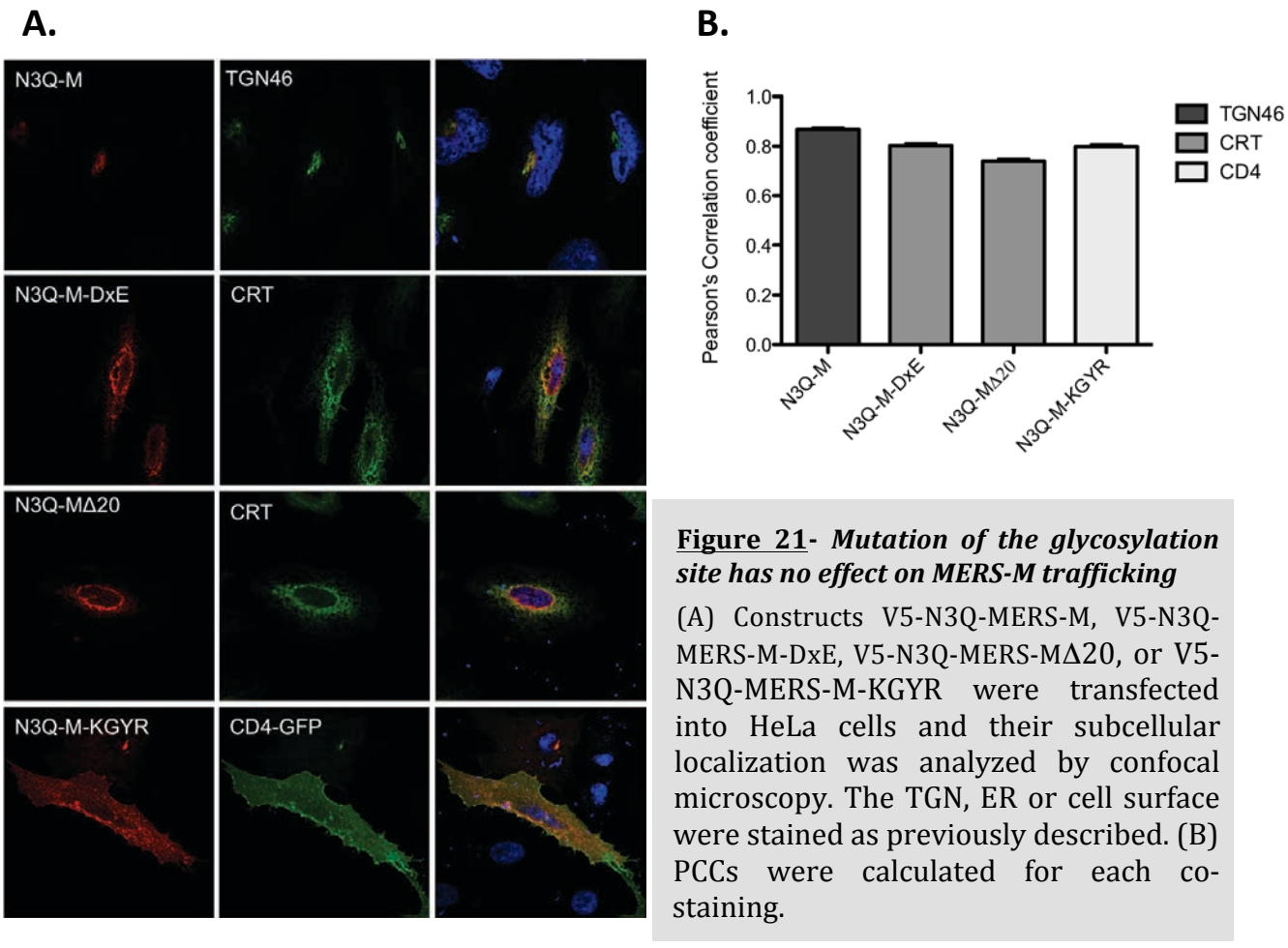


Figure 21- ***Mutation of the glycosylation site has no effect on MERS-M trafficking***

(A) Constructs V5-N3Q-MERS-M, V5-N3Q-MERS-M-DxE, V5-N3Q-MERS-MΔ20, or V5-N3Q-MERS-M-KGYR were transfected into HeLa cells and their subcellular localization was analyzed by confocal microscopy. The TGN, ER or cell surface were stained as previously described. (B) PCCs were calculated for each co-staining.

As expected, mutating the N-glycosylation site had no effect on the localization of the mutant proteins as confirmed with all PCCs being above 0,7: N3Q-M with TGN46, N3Q-M-DxE with CRT, N3Q-MERS-MΔ20 with CRT and N3Q-MERS-M-KGYR with CD4 (Figure 21B).

Hence, we used the constructions V5-N3Q-MERS-M, V5-N3Q-MERS-M-DxE, V5-N3Q-MERS-MΔ20, V5-N3Q-MERS-M-K199A, V5-N3Q-MERS-M-G201A, V5-N3Q-MERS-M-Y203A, V5-N3Q-MERS-M-R204A and V5-N3Q-MERS-M-KGYR for our biotinylation assays. Cells were transfected with these constructs, and plasma membrane proteins were labeled with non-permeable biotin. Then, biotinylated proteins were precipitated with streptavidin-conjugated agarose beads and analyzed in immunoblot (Figure 22).

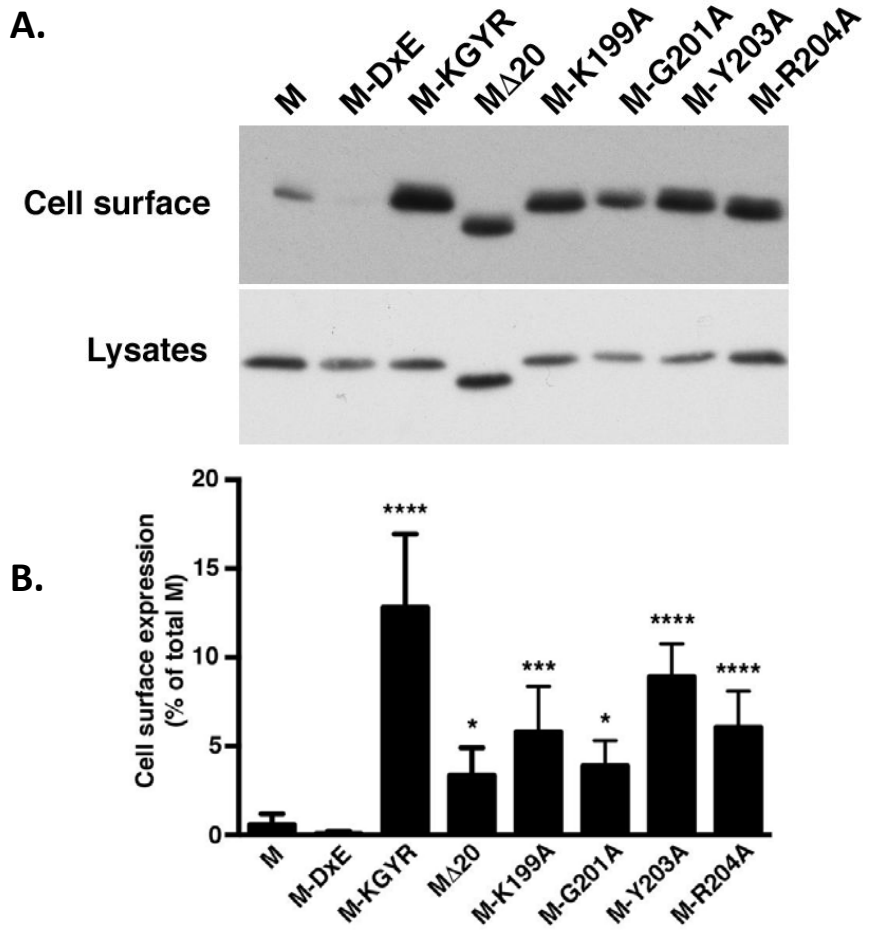


Figure 22- ***Cell surface expression of MERS-M mutants***

(A) Plasma membrane proteins of cells expressing the different M protein mutants were labeled with non-permeable biotin. Biotinylated proteins were purified using streptavidin-conjugated agarose beads. Biotinylated proteins and cell lysates were processed for immunoblot. M proteins were detected using an anti-V5 antibody. (B) Biotinylated M proteins and total M proteins in cell lysates were quantified in immunoblot. Results are expressed as the percentage of total M protein expressed at the cell surface and are expressed as the mean of five independent experiments. Error bars represent the standard error of the means (SEM). Results were analyzed by using an ANOVA test (* P<0.1; ** P<0.01; *** P<0.001; **** P< 0.0001).

The wild type M protein is only weakly exported to the cell surface, with less than 1% of the total amount of protein expressed at the cell surface. On the other hand, mutation of residues K199, G201, Y203 and R204 alone or in combination induced an increase in cell surface detection. Individual mutants displayed between 5 and 10% of cell surface expression, and approximately 13% of the total amount of the quadruple mutant M-KGYR was expressed at the cell surface. The N3Q-M-DxE was barely detected at the cell surface, consistent with the fact that this protein is blocked within the ER. Interestingly, as seen in immunofluorescent colocalization assay (Figure 16 C, D), the expression of the N3Q-MΔ20 mutant at the cell surface was slightly increased when compared to the wild type protein.

These results confirm that residues K199, G201, Y203 and R204 are involved in the specific localization of MERS-M to the TGN.

Nevertheless, intracellular trafficking is a dynamic process, and proteins can undergo cycles of cell surface expression and internalization. In that case at steady state, the localization of proteins in intracellular compartments results from equilibrium between anterograde intracellular trafficking and retrieval of protein by endocytosis. Any inhibition of endocytosis would result in protein accumulation at the cell surface. Hence, we wanted to test whether our KxGxYR motif was an endocytosis signal.

For this purpose, we analyzed the endocytosis of the M protein in a biotinylation assay. Cells were transfected with V5-N3Q-MERS-M or V5-N3Q-MERS-M-KGYR expression vectors, and cell surface proteins were labeled at 4°C with non-permeable cleavable biotin. Endocytosis was then allowed by incubating the cells at 37°C for 30 min. Afterwards, non-internalized biotin was cleaved with glutathione and internalized proteins were detected in western blot (Figure 23A) after streptavidin precipitation. The wild type and KGYR mutated proteins were internalized in comparable amounts (Figure 23B), indicating that the KxGxYR motif is not an internalization signal. Therefore, specific localization of the MERS-M protein to the TGN is probably mediated by a retention mechanism preventing export of the protein to the cell surface.

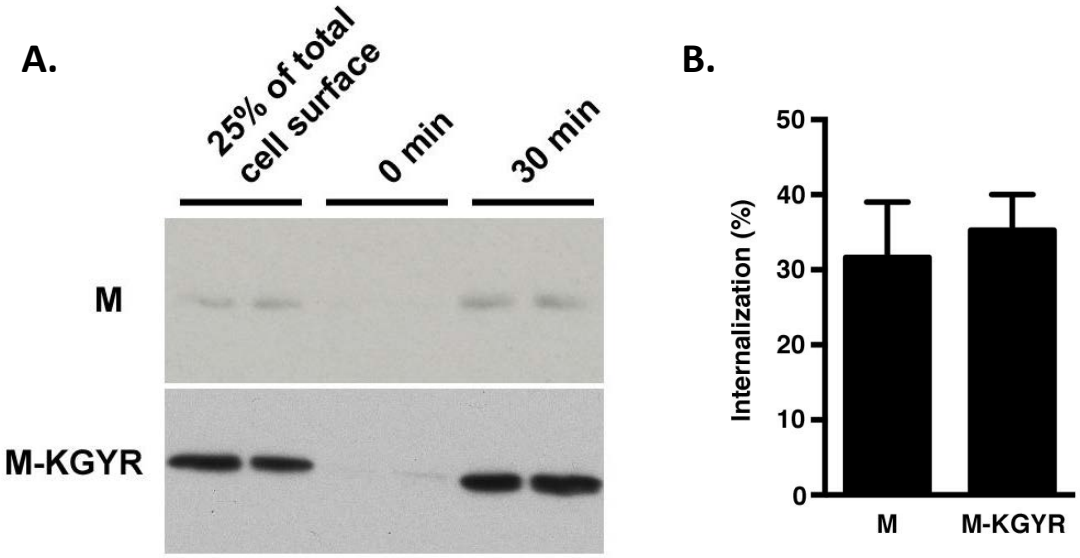


Figure 23- ***Endocytosis of M and M-KGYR mutant***

(A) *Cells were transfected with vectors expressing the M and M-K199A;G210A;Y203A;R204A (M-KGYR) proteins*. Cell surface proteins were then labeled with non-permeable biotin at 4°C. Endocytosis was allowed by incubating the cells for 30 min at 37°C. Biotin of non-internalized proteins was cleaved with glutathione. Internalized M proteins were detected after purification with streptavidin-conjugated agarose beads in western blot. (B) *Quantification of protein internalization*. Results are expressed as the percentage of cell surface-associated M protein and are expressed as the mean of three independent experiments. Error bars represent the standard error of the means (SEM).

Since it was demonstrated for MHV-M that oligomerization of the protein could be involved in its retention (Locker et al., 1995), we wanted to check whether the KxGxYR motif had any effect on this process. For this purpose, we compared oligomerization of MERS-M and MERS-M-KGYR in an M-M interaction assay. Cells were transfected with the constructs coding for V5-N3Q-M and V5-N3Q-M-KGYR. Cell surface proteins were biotinylated and cross-linked with PFA. The PFA was then quenched and the cells were lysed, precipitated with streptavidin beads and processed for immunoblot without heating.

Oligomers of the M protein were observed as a strong band at 40kDa (likely corresponding to dimers) and above it a smear of higher molecular weight bands. We observed no differences between MERS-M and MERS-M-KGYR, high molecular weight forms being equally abundant regardless of the presence of the KxGxYR motif (Figure 24). These multimers were more visible at the cell surface than in the lysates for both M and M-KGYR, probably because of the concentration of the sample. These results indicate that the KxGxYR motif does not play any role in oligomerization of the protein.

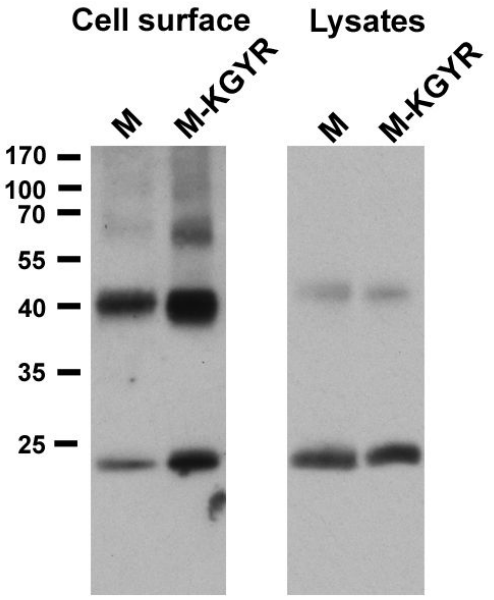


Figure 24- ***Oligomerization of MERS-M and MERS-M-KGYR proteins***

HeLa cells were seeded in 10 cm dishes and transfected with vectors expressing V5-N3Q-M or V5-N3Q-M-KGYR. Cell surface proteins were biotinylated at 4 °C and cross-linked with 0,8% PFA in PBS. Then, PFA was quenched by washing the cells with 50 mM NH4CL/PBS twice. Cells were lysed and lysates were processed for streptavidin precipitation as previously described. Proteins were resuspended in non-reducing Laemmli loading buffer without heating and detected by immunoblotting.

4.1.4 Transfer of the signal on another protein

In order to confirm the role of the KxGxYR motif as a TGN retention signal, we checked whether transferring the signal to a protein normally localized at the cell surface would induce its intracellular retention. To that end, we constructed chimeras between MERS-M and CD4, a cellular protein normally localized at the cell surface. Nevertheless, many of these chimeras exhibited folding default, which we could identify by using both an anti-CD4 (conformational antibody) and an anti-M antibody. All proteins detected with the anti-M located to the ER, whereas these forms could not be detected using the anti-CD4 antibody, indicating improper folding.

In light of the folding problems exhibited by our CD4-M chimeras, we decided to design other chimeras, but this time between MERS-M and IBV-M. That way, the three transmembrane segments structure of the M protein would be conserved, likely preventing folding impairment. We chose IBV-M because its trafficking has been extensively studied: when expressed alone in cells it localizes to the ERGIC compartment, and this specific localization is mediated by its first transmembrane segment (Machamer and Rose, 1987a; Machamer et al., 1993a). Therefore, IBV-M localization can be distinguished from MERS-M localization, allowing us to determine if our chimeras exhibit MERS-M-like or IBV-M-like localizations.

We constructed chimeras in which we switched the C-terminal domains of the proteins, MERS-M/IBV-M and IBV-M/MERS-M or IBV-M/MERS-M-KGYR. We also

constructed a chimera in which we replaced the first transmembrane segment of MERS-M-KGYR with the one of IBV-M to test if the first transmembrane segment of IBV-M can retain MERS-M-KGYR intracellularly (TM1-IBV/MERS-M-KGYR). Finally, we replaced the first transmembrane segment of IBV-M with the one of MERS-CoV M (TM1-MERS/IBV-M). These constructs were expressed in HeLa cells and their subcellular localization was analyzed by confocal microscopy. For each construct, we performed co-staining with both TGN46 and ERGIC53 (Figure 25 A, B).

Firstly, our results confirm that MERS-M and IBV-M expressed alone in cells have different localizations. As shown by our previous results MERS-M colocalizes with TGN46, whereas IBV-M colocalizes mainly with the ERGIC53 marker (PCC >0,7), confirming what was described in the literature (Figure 25 C).

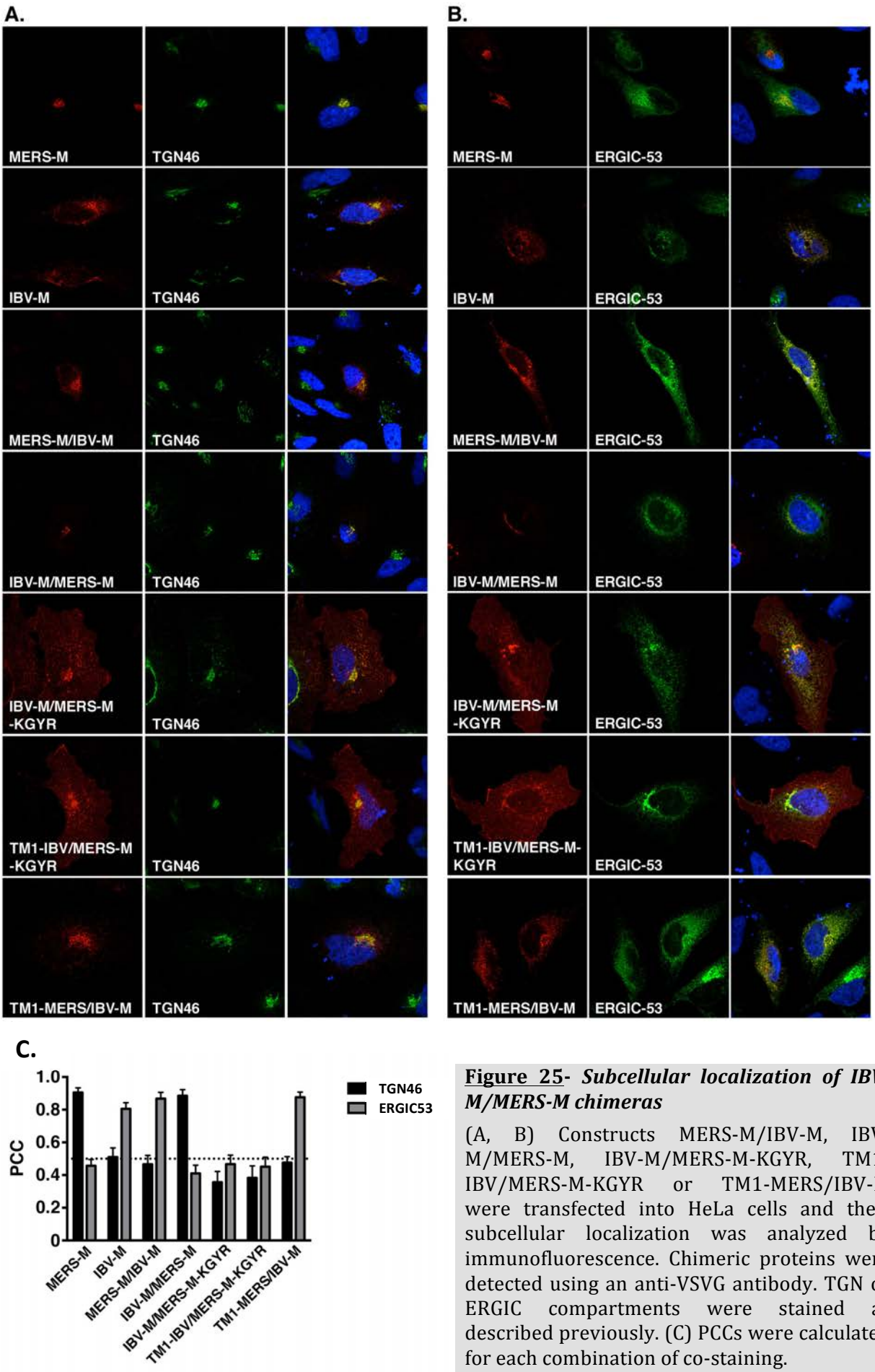


Figure 25- ***Subcellular localization of IBV-M/MERS-M chimeras***

(A, B) Constructs MERS-M/IBV-M, IBV-M/MERS-M, IBV-M/MERS-M-KGYR, TM1-IBV/MERS-M-KGYR or TM1-MERS/IBV-M were transfected into HeLa cells and their subcellular localization was analyzed by immunofluorescence. Chimeric proteins were detected using an anti-VSVG antibody. TGN or ERGIC compartments were stained as described previously. (C) PCCs were calculated for each combination of co-staining.

Interestingly, the IBV-M/MERS-M chimera colocalized with the TGN46 marker and not the ERGIC53 marker, and is thus localized in the TGN, whereas the MERS-M/IBV-M chimera colocalized with the ERGIC53 marker and not the TGN46 marker, indicating that it is localized in the ERGIC. In other words, switching the C-terminal domains of MERS-M and IBV-M led to the switch of their specific localizations. Interestingly, the IBV-M/MERS-KGYR protein localized to the cell surface, confirming the role of the KxGxYR signal in the specific localization of the MERS-M protein to the TGN. Surprisingly considering the literature on IBV-M, this result also suggests that the first transmembrane segment of IBV-M is not sufficient to retain the chimeric protein intracellularly. Consistently with this result, the chimera TM1-IBV/MERS-M-KGYR was also located at the cell surface and the chimera TM1-MERS/IBV-M was located in the ERGIC compartment.

Taken together, these results suggest that for both MERS-M and IBV-M the C-terminal domain is critical to their specific localization, and that the first transmembrane segment on the other hand is not able to induce specific intracellular targeting. Importantly, these results also confirm the role of the KxGxYR motif in the specific retention of MERS-CoV M in the TGN, even in the chimeric context.

4.1.5 Trafficking of IBV-M

Considering what was demonstrated in the literature concerning IBV-M trafficking and the conflicting results we obtained with our chimeras, we decided to further investigate IBV-M trafficking.

Analysis of the last 20 amino acids sequence of IBV-M showed the presence of a potential di-acidic ER export signal ExE (Figure 26A). Although di-acidic signals usually contain an aspartic acid (D), we decided to mutate both glutamic acids (E) in order to see if it could be a functional ER export motif. The construct coding for IBV-M-ExE-HA was transfected into HeLa cells, and its subcellular localization was analyzed by immunofluorescence. Co-stainings of ERGIC and ER compartments were performed. IBV-M-ExE-HA colocalized strongly with the ERGIC53 marker, and not with the CRT marker, similarly to the wild type IBV-M protein (Figure 26B). This result demonstrates that ExE is not a functional ER export motif for IBV-M.

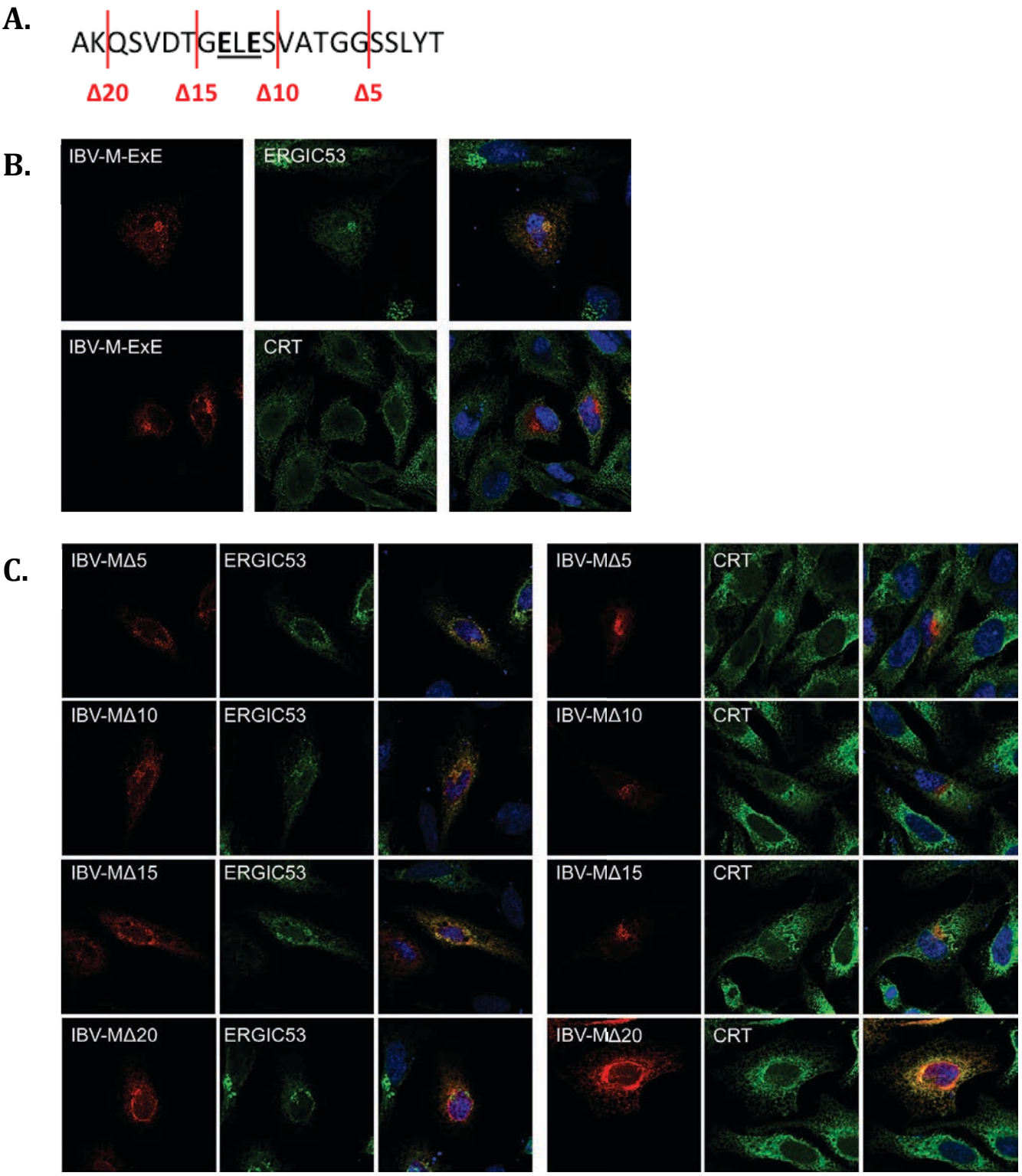


Figure 26- ***Involvement of the last 20 amino acids of IBV-M in its trafficking***

(A) Sequence of the last 20 residues of IBV-M. Positions of the deletions are indicated in red, and the putative ER export ExE signal is underlined. (B) Subcellular localization of IBV-ExE mutant. HeLa cells were transfected with pCDNA3.1-IBV-M-ExE-HA and processed for immunofluorescence. M proteins were detected using an anti-HA antibody and ER and ERGIC compartment were stained as previously described. (C) Subcellular localization of IBV-M deletion mutants. Constructs coding for IBV-MΔ5, IBV-MΔ10, IBV-MΔ15 and IBV-MΔ20 were transfected into HeLa cells, and the subcellular localization of the proteins was analyzed by immunofluorescence.

Consequently, to further investigate the ER export of IBV-M protein and identify the responsible signal, we constructed deletion mutants of the last 20 amino acids of IBV-M C-terminal domain: IBV-MΔ5, IBV-MΔ10, IBV-MΔ15 and IBV-MΔ20 (Figure 26A). All mutants were constructed with an HA C-terminal tag. These constructs were transfected in HeLa cells and their subcellular localization was analyzed in immunofluorescence and confocal microscopy (Figure 26C).

Mutants IBV-MΔ15, IBV-MΔ10 and IBV-MΔ5 all strongly colocalized with ERGIC53, regardless of the tag used, and did not colocalize with the ER marker CRT, similarly to wild type IBV-M (Figure 26C). On the other hand, IBV-MΔ20-HA colocalized with the CRT marker and not the ERGIC53 marker, indicating that the protein is blocked within the ER.

Taken together, these results indicate that there is a potential ER export signal located in the 5 amino acid sequence QSVDT between deletions Δ15 and Δ20.

4.2 Project II: Characterization of an antiviral against HCoV-229E

As previously mentioned, coronaviruses are an emerging family of pathogens, as illustrated by the recent emergence of SARS-CoV and MERS-CoV, and we do not have yet any specific antiviral or vaccine. This is why research on coronaviruses in general, and especially research for specific anti-coronaviruses compounds is of major importance. In that light, the laboratory initiated a project to identify molecules with an inhibitory effect on coronaviruses, using HCoV-229E as a model. Indeed, as this virus can be manipulated in a BSL-2 facility, it is easier and faster to perform experiments on it when compared with BSL-3 SARS-CoV and MERS-CoV. Hence, a high content screening of HCoV-229E infection was performed in the lab with the PRESTWICK library of 1280 compounds. The screen was performed at two different concentrations, 10 and 5µM during infection.

The aim of this second project was to validate the antiviral effect of four drugs identified in the screens. We selected four drugs with high inhibitory effect: digoxigenin, trifluoperazine dihydrochloride (TH), perhexiline maleate (PM), and astemizole, to confirm their antiviral properties and identify the mechanism of action of one of them.

Trifluoperazine is a dopamine antagonist with anti-psychotic properties used in the treatment of schizophrenia. It blocks central dopamine receptors, and thus prevents effects of an excess of dopamine. It can also function as a calmodulin inhibitor, leading to an elevation of intracellular calcium (Howland, 2016).

Perhexiline maleate is used in the treatment of angina pectoris. It acts through binding to the mitochondrial enzyme carnitine palmitoyltransferase (CPT)-1, which results in an increased ATP production for the same O_2 consumption and therefore increased myocardial efficiency (Horowitz and Mashford, 1979).

Astemizole is a reversible competitive inhibitor of histamine H1 receptors. It has anti-allergic properties and is thus used in the treatment of allergic rhinitis, asthma or allergic conjunctivis (Janssens, 1993).

After the first tests, digoxigenin was the most promising compound, and was thus selected for further characterization.

Digoxigenin is a steroid found in the leaves and flowers of *Digitalis lanata* (Figure 27). Digoxigenin is a derivative of digoxin, a cardiac glycoside used in the treatment of heart failure. The family of cardiac glycosides (also called cardiotonic steroïds) comprises several molecules all exhibiting cardiac activity, and being extracted from diverse plants including: *Strophanthus gratus, Acokanthera oblongifolia* and *Acokanthera schimperi* (ouabain), *Digitalis lanata* and *Digitalis purpurea* (digoxin, digitoxin), *Scilla maritima* (proscillaridin A), *Nerium oleander* (oleandrin, oleandrigenin); but also in some poisonous frogs' poisons (bufalin, marinobufagenin) (Winnicka et al.).

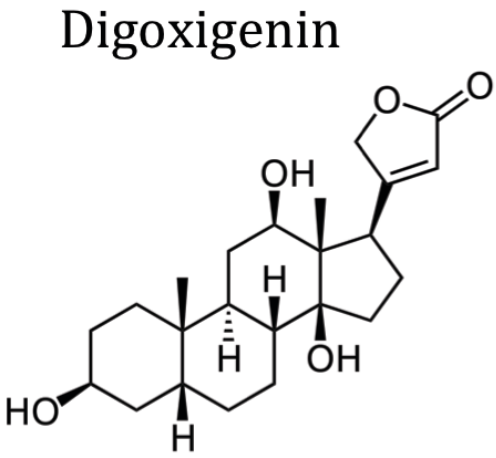


Figure 27- *Digoxigenin is found in the digitalis plants*

A picture of *Digitalis lanata* is shown on the left. It is one of the plants in which cardiac glycosides are produced, along with other plants and some poisonous frogs. On the right, the chemical structure of digoxigenin is depicted.

Interestingly, although digoxin was only isolated in the 1930s, digitalis plants have been used as heart drugs for centuries, and their toxic properties at high doses were also already known (Bessen, 1986). Indeed, although still widely used in the treatment of heart failure and arrhythmia (digoxin is on WHO's list of essential medicine), digoxin has an extremely narrow therapeutic index, and can be toxic if administered at high doses.

The cellular target of cardiac glycosides is the Na^+/K^+ ATPase pump, and more precisely they are capable of binding to the α subunit of the pump, thereby inhibiting its activity (Figure 28). Indeed, the pump consists of two functional subunits (α and β). The α subunit is a catalytic membrane protein containing 10 transmembrane segments. There are four different isoforms of the α subunit encoded by ATP1A1 to ATP1A4, and the α1 isoform is expressed in almost all tissues. The β subunit is a type II membrane protein responsible for translocation of the α subunit into the ER and its delivery to the cell surface.

The main role of the Na^+/K^+ ATPase pump is to maintain electrolyte homeostasis in cells by pumping K^+ cations in and Na^+ cations out of the cell against the concentration gradient, using the energy obtained from the hydrolysis of ATP (Figure 28, upper panel). Hence, inhibition of the pump by cardiac glycosides' binding (at the site of K^+ binding) leads to an increased intracellular Na^+ concentration, which in turn increases Ca^{2+} intracellular concentration because of the Na^+/Ca^{2+} exchanger (Figure 28, lower panel). This is the basis of treatment of heart failure by cardiac glycosides. But the function of the Na^+/K^+ ATPase is not limited to cation exchange. Indeed, it is also capable of working as a classical receptor, with binding of cardiac glycosides inducing many signaling cascades, the main signaling targets being phosphatidylinositol 3-kinase (PI3K), Src kinase, inositol 1,4,5-triphosphate receptor (IP3R), and phospholipase C (PLC). Activation of these targets could lead to downstream signaling effects involved in apoptosis, cell-cell interaction, gene expression… etc. (Liu et al., 2007; Xie, 2003; Zhang et al., 2008).

Importantly, this signaling is mediated by the α subunit, seems to be independent of changes in intracellular ions, but could differ depending on the α subunit isoforms.
It is also worth noting that cardiotonic steroïds can trigger the signaling pathways of the Na^+/K^+ ATPase at low concentrations that do not affect the pump function (Aizman et al., 2001).

Furthermore, the potential of cardiac glycosides as antiviral or anti-cancer drugs was previously suggested by various studies (Amarelle and Lecuona, 2018; Laird et al., 2014; Winnicka et al.).

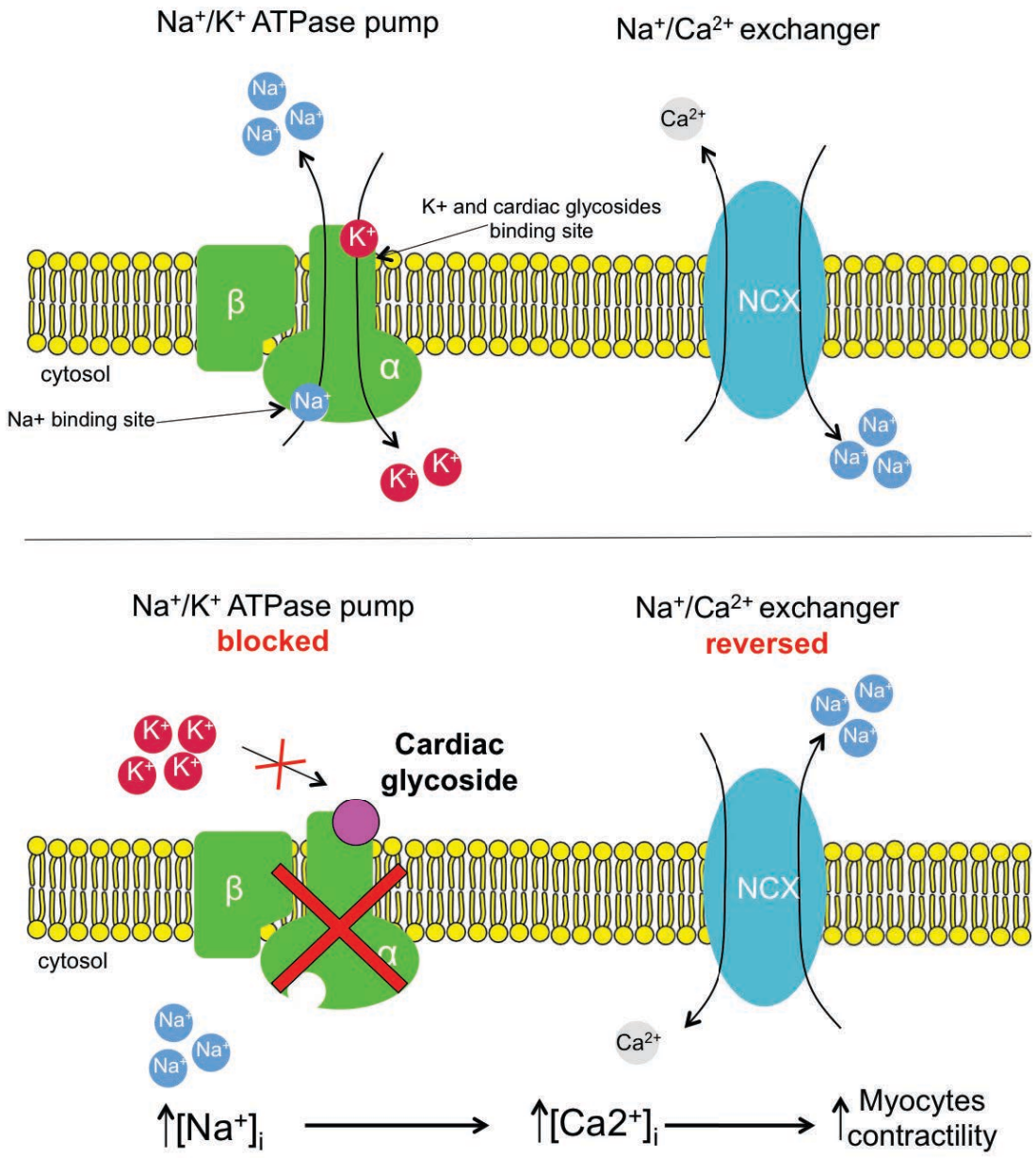


Figure 28- ***Mechanism of action of cardiac glycosides on the Na+/K+ ATPase pump***

Upper panel – Activity of the pump in absence of cardiac glycosides. The pump mediates the transport of $2K^+$ ions inside and $3Na^+$ ions outside of the cell, against the concentration gradients using the energy from ATP hydrolysis. The Na^+/Ca^{2+} exchanger mediates transport of $3Na^+$ cations outside the cell and of one Ca^{2+} cation inside the cell.

Lower panel – Binding of cardiac glycosides to the pump. Cardiac glycosides bind to the α subunit K^+ binding site, thereby blocking fixation of the K^+ ions on the pump. Hence, the activity of the pump is blocked, and Na^+ ions accumulate intracellularly. This increase in Na^+ concentration leads to a reversed action of the Na^+/Ca^{2+} exchanger that starts transporting Na^+ ions out and Ca^{2+} ions in the cell, leading to an increased intracellular Ca^{2+} concentration. This increased concentration of Ca^{2+} in turn increases myocytes contractility.

In our study, we looked at the antiviral potential of digoxigenin, trifluoperazine dihydrochloride (TH), perhexiline maleate (PM), and astemizole against infection with coronaviruses *in vitro*.

4.2.1 Dose dependent inhibition

After screening, four drugs were selected for further analysis: digoxigenin, TH, PM and astemizole. First, we confirmed their effect in dose-response assays. Huh-7 cells were infected with a recombinant HCoV-229E expressing the renilla luciferase, in presence of increasing concentration of the drugs. The drugs were present during every step: pre-incubation with the cells (1h), infection (1h) and post-infection (6h) (Figure 29A). We used concentrations going from 0,3 to 5µM for TH, PM and astemizole, and concentrations going from 0,039 to 2,5µM for digoxigenin, as first results showed that is seemed to inhibit infection at low concentrations.

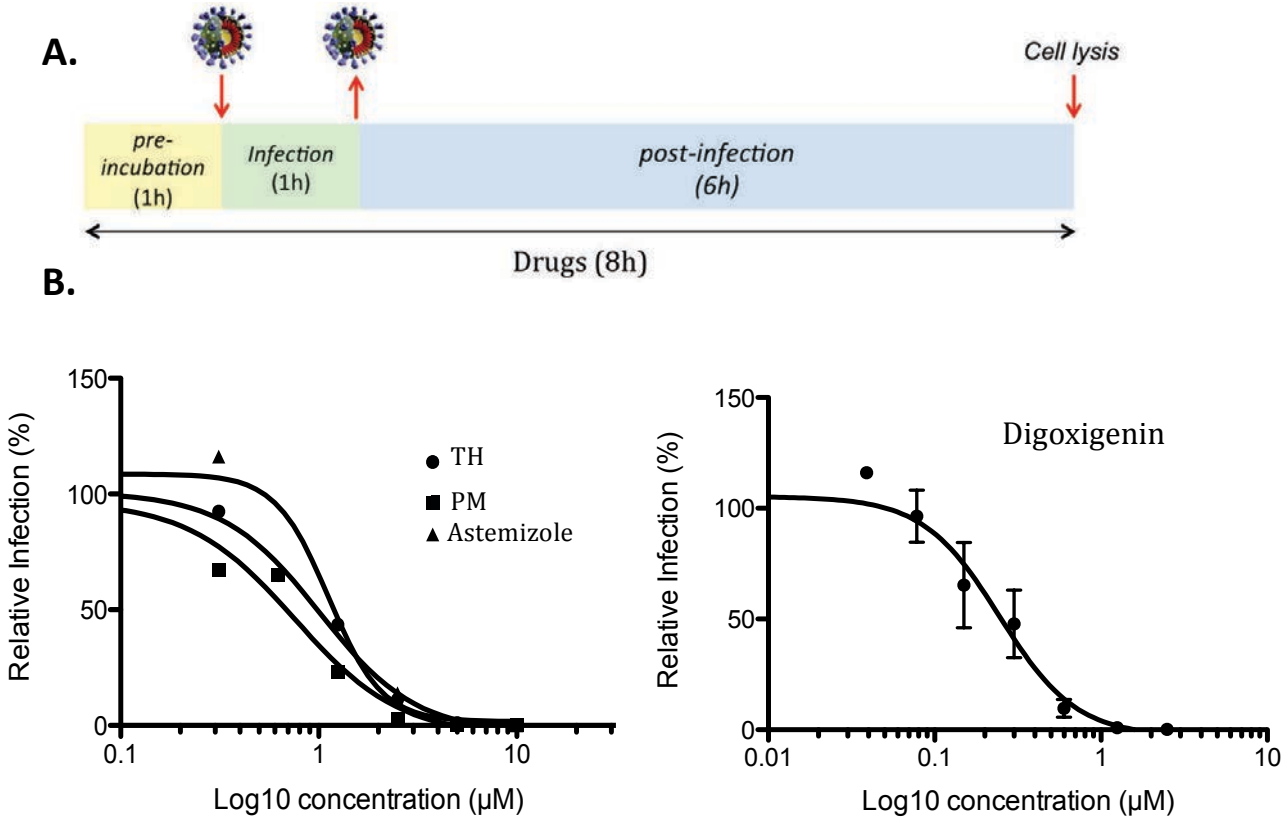


Figure 29- ***Digoxigenin, TH, PM and astemizole inhibit HCoV-229E infection***

(A) *Infection steps and time of addition of the drugs*. (B) *Dose response curves of digoxigenin, TH, PM and Astemizole*. Huh-7 cells were plated in 96-well plates. The next day, the medium was replaced by serial dilutions of digoxigenin, TH, PM or astemizole, and the cells were incubated 1h at 37°C. The cells were then infected with HCoV-229E-luc in presence of different drug concentrations for 1h. The medium was then removed and replaced with fresh dilutions of drugs, and the cells were incubated for 6h. Finally, the cells were lysed and luciferase activity was measured by a luminometer.

All drugs showed a strong inhibitory effect on HCoV-229E infection, as shown by the dose response curves in Figure 29B. TH, PM and astemizole have IC50s of respectively 1,012 μM, 756 nM and 1,124 μM. Interestingly, the IC50 of digoxigenin is much lower, at 245,7 nM, making it the most potent drug of the four.

4.2.2 Toxicity

In order to verify that the observed inhibitory effect on infection was not due to cell toxicity, we tested the drugs' toxicity in an MTS assay, after incubation of Huh-7 cells with the drugs for 24h, 48h or 72h at different concentrations (Figure 30).

For TH, no toxicity was observed, regardless of the concentration of the drug or the time of incubation. For PM and astemizole, a sudden drop of viability was observed when cells were incubated with 10 μM of drug, with almost no viable cells left after 72h of incubation. For digoxigenin, a progressive decrease in cell viability was observed after incubation with the drugs, but this is moderate at 24h and more marked after 72h. Importantly, no decrease of cell viability was observed after 24 and 48h for concentrations under 1,25 μM, which corresponds to five times the drug's IC50.

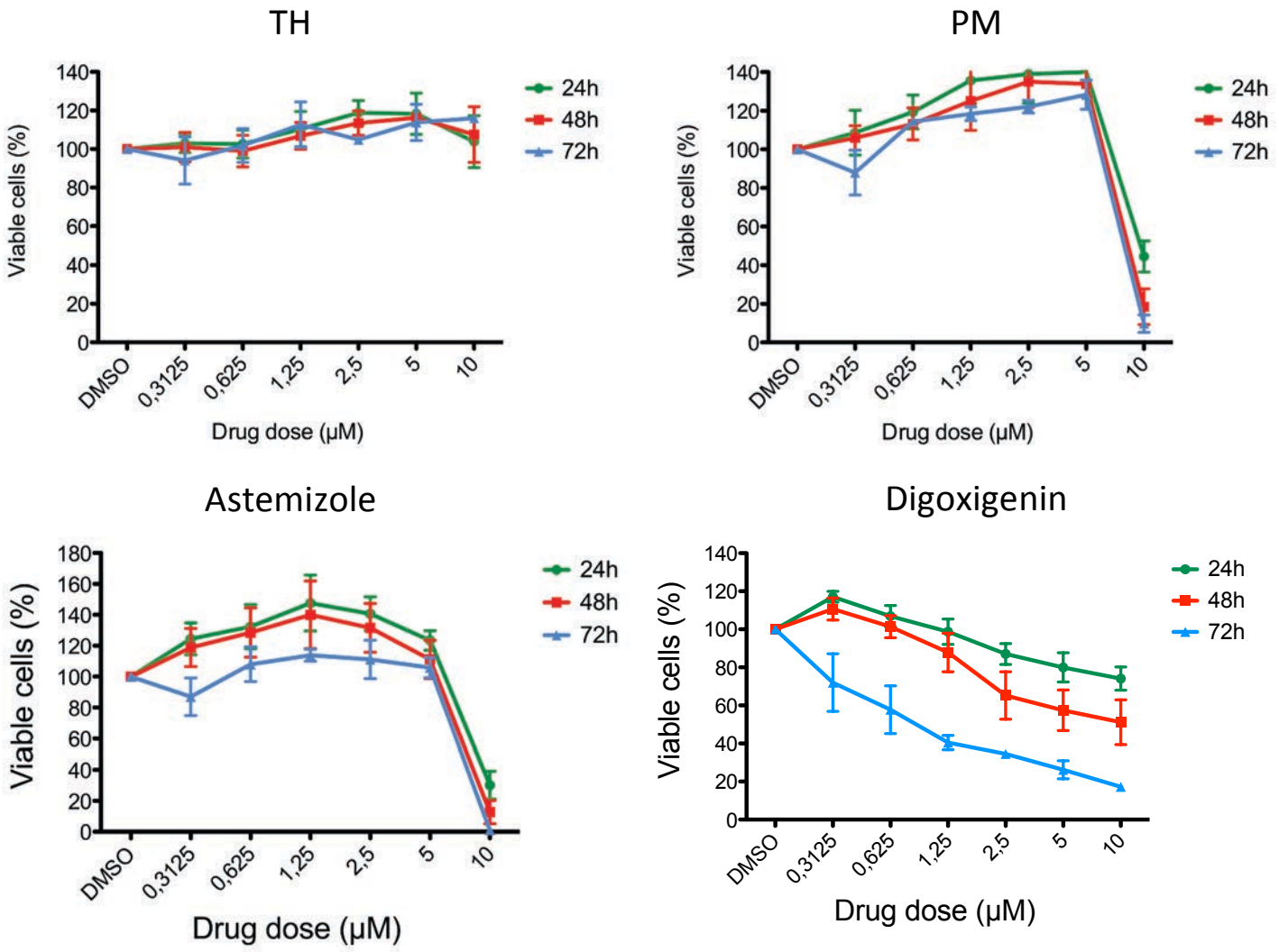


Figure 30- ***Cell toxicity of selected drugs***

Huh-7 cells were plated in 96-well plates, and incubated with concentrations of digoxigenin, TH, PM or astemizole going from 0,3 µM to 10 µM. After 24h, 48h or 72h of incubation with the drug, the medium was replaced with an MTS reagent diluted in DMEM, cells were incubated at 37°C and the OD490nm was measured. DMSO diluted in DMEM was used as a control.

Interestingly, none of the drugs was toxic at low concentrations, so the inhibitory effects previously observed were not due to cell toxicity. Hence, we decided to further characterize the mechanism by which these drugs inhibit HCoV-229E infection, by first identifying the step of the viral cycle that they inhibit.

4.2.3 Infection step inhibited

Then, we analyzed if the compounds act on the virus or on the cells. In parallel, we analyzed the step of the viral cycle that is inhibited. To do so, the compounds were added either during inoculation of the virus or after the entry step during the replication (Figure 31A). A dose of 2,5µM was used for TH, PM and astemizole, and a dose of 1,25µM was used for digoxigenin. In parallel either the cells or the virus were pre-incubated with the drugs. For the pre-incubation of the virus, the virus was incubated with drugs at 2,5µM or DMSO diluted in DMEM for 1h. Then, this mix was diluted 10 times with DMEM to obtain a 0,25 µM (for TH, PM and astemizole) or 0,125 µM (for digoxigenin) concentration, and this mix was added on cells for 1h. The concentrations used after dilution are not inhibitory. The medium was then changed and 6h post-infection the cells were lysed. Other conditions correspond to different adding times of the drugs, as shown in Figure 31A.

TH and PM seemed to have an effect on various steps of the viral cycle, making it difficult to interpret (Figure 31B). On the other hand, Astemizole demonstrated an effect when added during inoculation, or in pre-incubation with the cells. This indicates a possible effect on the entry step of HCoV-229E. Furthermore, astemizole did not act on the virus but rather on the cells. Interestingly, digoxigenin had a clear antiviral effect on a post-inoculation step, luciferase activity being reduced to almost 0% of the control when the drug is added during this step.

Thus, considering that digoxigenin has the lower IC50 (250nM), does not exhibit toxicity around this concentration, and has a clear effect on the post infection step, we decided to further characterize the mechanism by which it inhibits HCoV-229E infection.

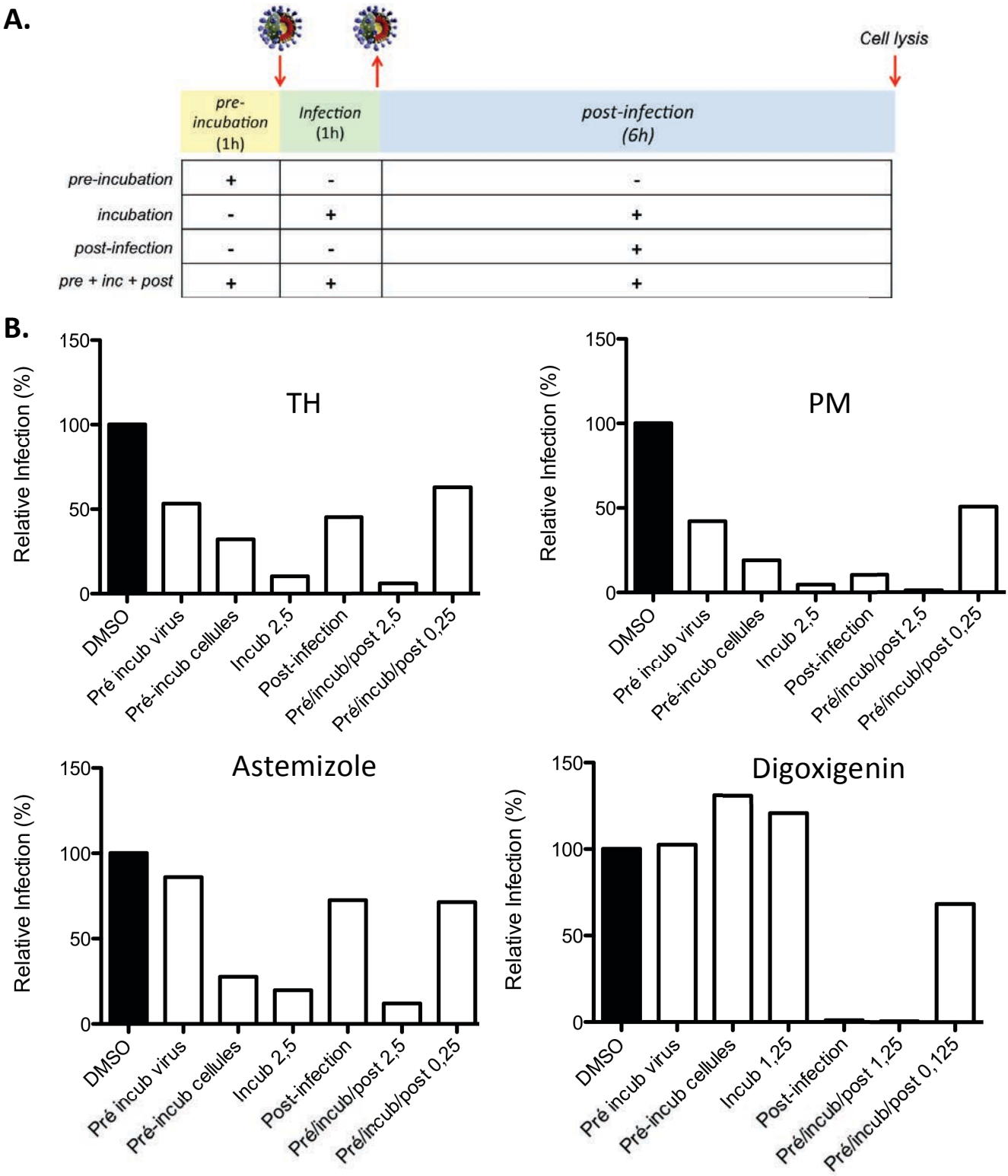


Figure 31- ***Step of the viral cycle inhibited by selected drugs***

(A) *Infection steps and time of addition of the drugs.* (B) Huh-7 cells were plated into 96-well plates and infected the next day with HCoV-229E-luc. TH, PM, astemizole or digoxigenin were added at different time points: during pre-incubation (with cells or virus), during infection, during post-infection, or during all experiment. TH, PM and astemizole were used at 2,5µM or 0,25 µM and digoxigenin at 1,25µM or 0,125µM. 6h post-infection the cells were lysed and luciferase activity was measured using a luminometer. Displayed results are from only one experiment.

4.2.4 Digoxigenin's inhibitory effect

First, we performed a new dose response experiment on HeLa cells, to confirm the drug's IC50 and to see whether the inhibitory effect is also exhibited in another cell type (Figure 32A). The dose response curve is similar to the one in Huh-7 cells (see Figure 29), and so is the IC50 (255nM). We then wanted to confirm that digoxigenin has an effect on the post-infection step, with the same protocol as before, and three independent experiments (Figure 32B). As before, the results demonstrate a strong inhibitory effect during the post-infection step, confirming what we first observed.

To further confirm that digoxigenin inhibits only the post-entry step, we tested whether it could inhibit infection with HCoV-229E pseudoparticles (229Epp). Pseudoparticles are retroviral cores containing a minigenome with a luciferase reporter gene and carrying a viral protein of interest in their envelope. Pseudoparticles fully mimic the entry step of the viral cycle, and the level of infection can be measured by quantifying luciferase activity. We produced particles pseudotyped with HCoV-229E spike protein (229Epp) and also particles pseudotyped with the G protein of VSV (VSVGpp) as a control. Huh-7 cells were then transduced with the pseudoparticles in presence or absence of digoxigenin. After 3h the medium was changed, and 48h post-infection cells were lysed and luciferase activity was measured (Figure 32C).

Surprisingly, digoxigenin inhibited infection of both 229Epp and VSVGpp, at similar levels. Nevertheless, the IC50 is higher than for HCoV-229E or HCV viruses, which suggests that the effect we observe here is different than the one we observed in the context of the whole virus. Therefore, we believe that the drug does not inhibit the entry step of the viral cycle, but it is more probable that it acts on the transcription/retrotranscription steps that occur after entry of the pseudoparticles in the cells.

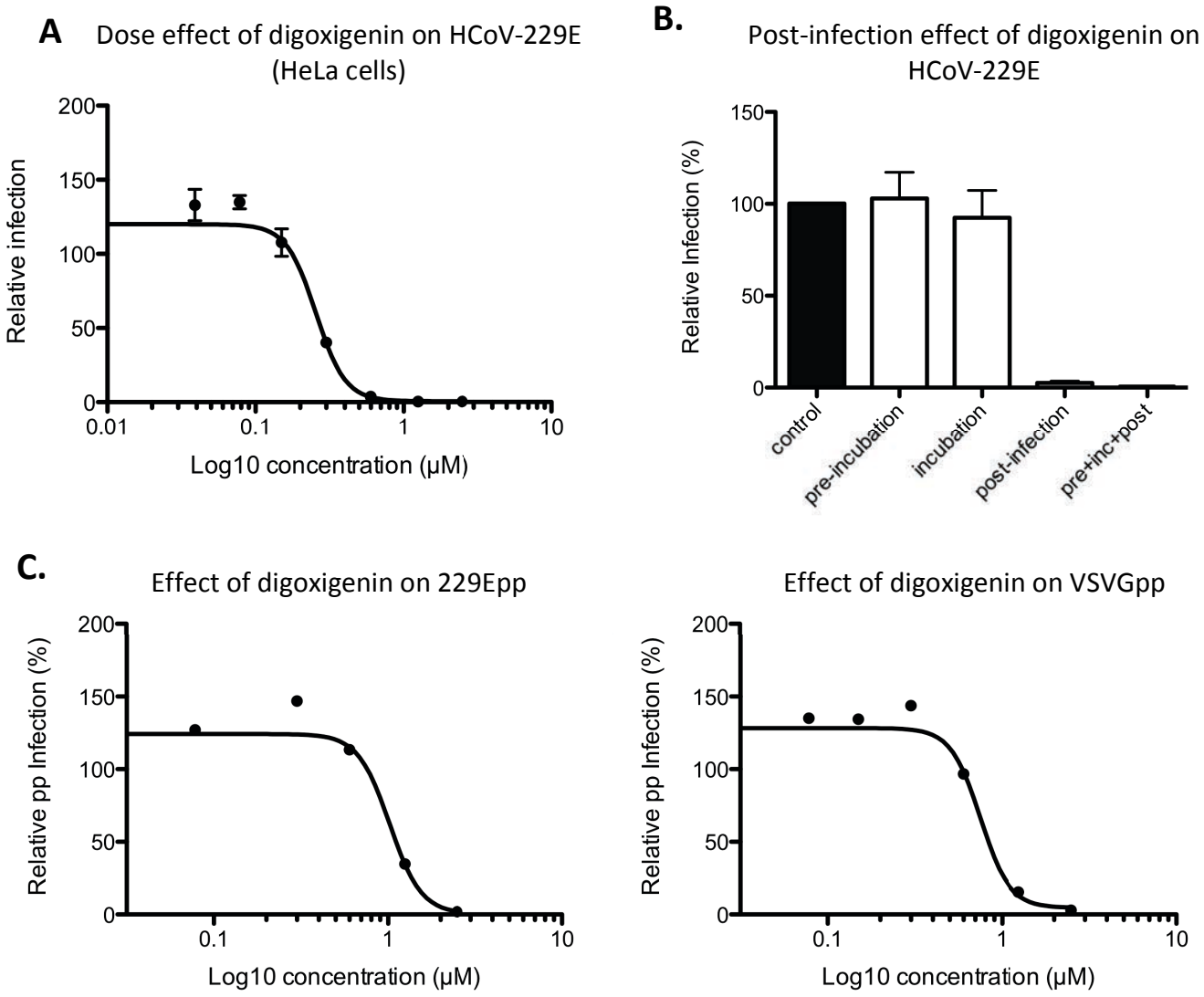


Figure 32- ***Digoxigenin inhibits the post-infection step of the viral cycle***

(A) *Dose response of digoxigenin in HeLa cells.* HeLa cells were plated in 96-well plates, and infection experiment was performed as described in Figure 28. (B) *Step of the viral cycle inhibited by digoxigenin.* Huh-7 cells were plated in 96-well plates and infected with HCoV-229E-luc. Digoxigenin at 1,25µM was added at different time points as described in Figure 30.
(C) *Effect of digoxigenin on pseudoparticles.* Huh-7 cells were infected with particles pseudotyped with S protein of HCoV-229E (229Epp) or G protein of VSV (VSVGpp) for 3h in presence or absence of digoxigenin (concentrations from 0,078 to 2,5µM). The media was changed and after 48h cells were lysed and luciferase activity was measured.

4.2.5 Digoxigenin also inhibits HCV infection

Since digoxigenin has a post-infection inhibitory effect, we tested if it could inhibit another RNA (+) virus. We therefore tested its effect on infection with *hepatitis C virus* (HCV). First, we performed a dose response experiment with concentrations of digoxigenin going from 0,078 µM to 2,5µM. Huh-7 cells were infected with HCV in presence of the drug during every step: pre-incubation with the cells (1h), infection (2h)

and post-Infection (30h). Interestingly, digoxigenin also inhibited HCV infection (Figure 33A).

Since digoxigenin also inhibited HCV, we further looked at which step of HCV cycle was inhibited. We proceeded as before, by adding and removing the digoxigenin at different time points during infection (Figure 33B).

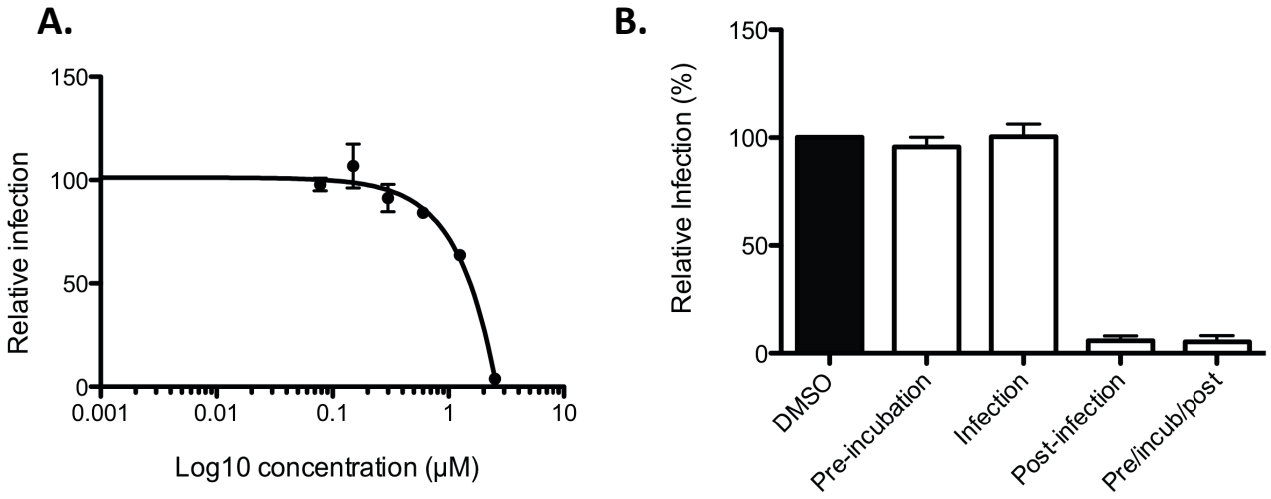


Figure 33- ***Digoxigenin's effect on HCV infection***

(A) Huh-7 cells were plated on coverslips in 24-well plates. The next day, cells were infected with HCV for 2h. Digoxigenin was added during pre-incubation with the cells (1h), during infection (2h) and during post-infection (30h). 30h post-infection cells were fixed and infected cells were stained in immunofluorescence with an anti-E1 antibody. (B) Huh-7 cells were plated in 24-well plates on coverslips and infected with HCV. Digoxigenin was added at different time points: pre-incubation with cells (1h), infection (2h) post-infection (30h), or during all experiment. 30h post-infection cells were fixed and infected cells were stained in immunofluorescence with an anti-E1 antibody.

Interestingly, digoxigenin also inhibited HCV infection when added during the post-infection step, further supporting an inhibition of replication (Figure 33B). In order to confirm the effect of the drug on replication, we tested the effect of the drug on an HCV replicon. A subgenomic replicon is a modified form of HCV genome that replicates autonomously in cells, but does not produce infectious particles. We used a cell line we had at our disposal in the lab, which stably expresses an HCV replicon with a GFP reporter gene (Sahuc et al., 2019) (Figure 34A). This replicon is assembly defective since there are no structural proteins, and it allows us to study replication only, and to bypass other steps of the viral cycle.

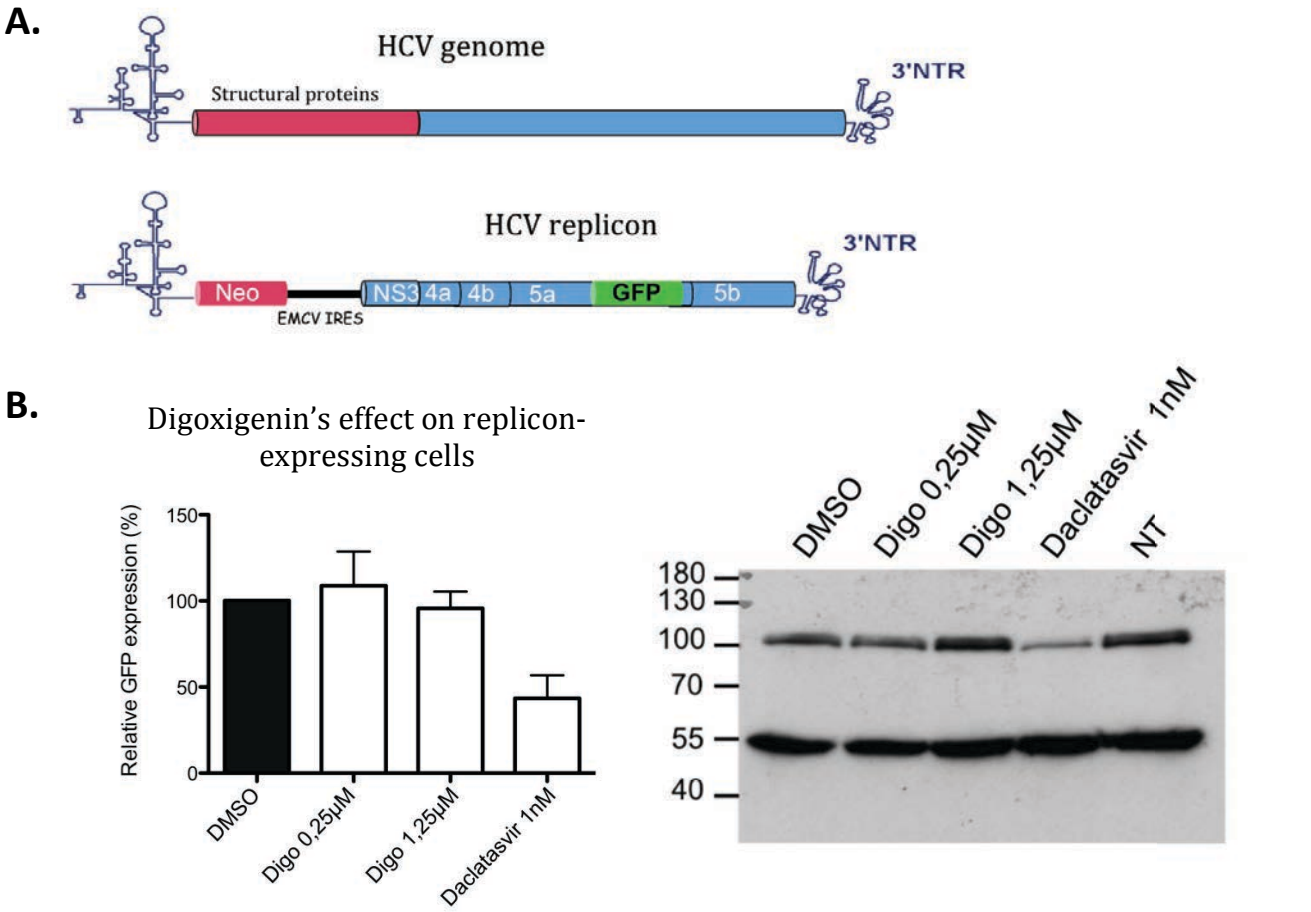


Figure 34- *Digoxigenin does not inhibit stably expressed HCV-replicon*

(A) *Structure of HCV genome and GFP-replicon.* The GFP reporter gene is fused to NS5a, and structural proteins were replaced by a Neomycin resistance gene. (B) *Effect of digoxigenin on Huh-7 cells stably expressing HCV GFP-replicon* (G13.2). G13.2 cells were plated in 24-well or 6-well plates and incubated for 24h with 0,25 µM or 1,25 µM or digoxigenin. The replication inhibitor daclatasvir was used as a control at 1nM. The levels of replicon replicating in each condition were measured by microscopy, by counting GFP positive cells (left panel) or by Western Blot using an anti-GFP antibody (right panel). The lower band in the Western Blot corresponds to tubulin detected with an anti-tubulin antibody.

The replicon-expressing cells were incubated for 24h with 0,25 µM or 1,25 µM of digoxigenin, and the level of replication in each condition was measured either by both microscopy and western blotting. The replication inhibitor daclatasvir was used as a control. GFP expression in presence of digoxigenin was comparable to the DMSO control when measured in both immunofluorescence and western blot (Figure 34B), indicating that digoxigenin has no effect on the replication of the replicon in these cells. This suggests that the post-infection inhibitory effect of the drug occurs at an early stage, likely before or at the onset of viral replication.

5. Discussion

5.1 Project I: Intracellular trafficking of M proteins

As coronaviruses structural proteins are produced in the ER and assemble in the ERGIC, specific targeting of viral structural proteins to the assembly site in the cell is crucial for viral egress and spreading. Additionally, because viruses divert many cellular processes, including the intracellular trafficking machinery, studying the trafficking of viral proteins is a useful way of deciphering protein sorting mechanisms and motifs.

In this study, we focused on the intracellular trafficking of the M protein of coronaviruses, which is considered to be the motor of viral assembly through interactions with all structural proteins. Interestingly, it was previously demonstrated that when expressed alone in cells, M proteins of coronaviruses could go beyond the assembly site of the virus (ERGIC) in the secretory pathway. Our project aimed at identifying motifs involved in the trafficking of the M protein of MERS coronavirus, that was reported to localize to the TGN (Yang et al., 2013a).

We confirmed this localization to the TGN and were able to identify two motifs involved in trafficking of MERS-M protein: a well-know di-acidic ER export signal, which we proved to be functional for MERS-M, and a novel KxGxYR TGN retention motif.

After their production in the ER, the first transport step for proteins in the secretory pathway is ER exit. This process occurs at specific ER-exit sites (ERES), and is mediated by COPII coat proteins. ER export signals, including LxxLE, diacidic DxE, YNNSNP or triple R, can interact directly with Sec24 or Sar1 and lead to incorporation of the proteins in the COPII carriers, and further transport towards the Golgi.

The DxE signal we identified in the C-terminal domain of MERS-M was first described on the G protein of VSV (Nishimura and Balch, 1997) and is present on many proteins of diverse origins. The signal was shown to be necessary for protein concentration into the COPII vesicles, but not for their recruitment to the pre-budding site (Nishimura et al., 1999). Interestingly, this DxE signal is also present in the C-terminal tail of PEDV and of many bat coronaviruses, including HKU4 and HKU5 which are closely related to MERS-CoV.

Some *Alpha*- and *Gammacoronavirus* M proteins, such as HCoV-229E-M, HCoV-NL63-M, FIPV-M or IBV-M also contain a di-acidic ExE motif. Nevertheless, such a signal was never demonstrated to be functional for ER export. Furthermore, in the yeast protein Sys1p, ExE cannot compensate for the DxE signal (Votsmeier, 2001a), and in our hands the mutation of the ExE motif to AxA on IBV-M did not disrupt ER export of the protein. Alternatively, a tri-acidic motif DxDxE was identified on the KAT1 ion channel from *Arabidopsis thaliana,* and mutations of the amino acids of the motif revealed a gradual reduction of ER export depending on the number of mutated acidic residues (Mikosch et al., 2006).

After its export from the ER, we observed that MERS-M protein localized to the TGN compartment at steady state. We identified the KxGxYR TGN retention motif in MERS-M cytosolic tail, which was previously unknown to literature. Mutation of any residue of this motif resulted in export of the mutated protein to the cell surface. Interestingly, this motif is highly conserved among *Betacoronaviruses* (Figure 19A). As other M proteins of *Betacoronaviruses* were reported to localize to the TGN, it is possible that the KxGxYR motif would be involved in this retention mechanism.

Interestingly in another study, a mutant of the MHV-M protein lacking the last 18 amino acids of the C-terminal tail was constructed and shown to localize to the cell surface, in contrast with wild type MHV-M which localizes to the TGN (Armstrong and Patel, 1991a). This deletion occurs in the middle of the KxGxYR signal, leaving only the K and G residues on the truncated protein. In another study, a 22 amino acids deletion of the C-terminal tail of MHV-M (removing the whole KxGxYR signal) also led to the export of the mutant protein to the cell surface (Locker et al., 1994a). In another study on MHV-M protein, it was mentioned that a mutant containing a mutation of the KxGxYR motif (Y211G) was exported at the cell surface (de Haan et al., 1999b).

Taken together, these observations strongly suggest that the KxGxYR motif we identified may be a functional TGN retention signal for MHV-M, and also more generally for M proteins of *Betacoronaviruses*.

Furthermore, SARS-CoV M was reported to localize to the Golgi, although its precise localization within the Golgi remains unclear. But interestingly, a SARS-M mutant lacking the N-terminal domain and the three transmembrane domains was reported to localize to the Golgi region (Tseng et al., 2010a), suggesting that the C-terminal domain of the protein is involved in this intracellular retention.

It would therefore be interesting to investigate the involvement of the KxGxYR motif in TGN retention of M proteins of other *Betacoronaviruses*, especially MHV-M, by mutational analysis.

Protein transport in the secretory pathway is a highly dynamic process. Hence, the localization of proteins in specific intracellular compartments generally results from equilibrium between anterograde and retrograde movements of the proteins. This is usually due to a crosstalk between targeting signals located in different domains of the protein.

For example, the TGN46 protein localization to the TGN results from a transport of the protein to the plasma membrane and a recycling back to the TGN after internalization and sorting via the endosomal network (Bos et al., 1993a). This mechanism is mediated by an endocytic motif located in the C-terminal domain, and an ER retrieval motif located in the transmembrane domain.

Similarly, it was demonstrated that the ERGIC-53 protein contains ER-export, ER-retention and ER-retrieval motifs, which mediate a continuous recycling of the protein from the Golgi to the ER, resulting in localization of the protein to the ERGIC and *cis*-Golgi at steady-state (Kappeler et al., 1997).

For TGN localization, one option was that the M protein was exported at the cell surface and retrieved back to the TGN through endocytosis, in a mechanism similar to what was demonstrated for TGN46 (Bos et al., 1993a; Reaves et al., 1993). Indeed, as a small proportion of wild type MERS-M protein was detected at the cell surface, we tested whether KxGxYR could be an endocytic motif. The M protein was indeed retrieved by endocytosis, but we saw no difference in internalization levels when comparing the mutant and wild-type M, indicating that KxGxYR acts as a retention motif rather than as an endocytic motif.

As discussed previously, in their study of MHV-M, *Locker et al.* generated an M protein mutant with a 22 amino acids C-terminal deletion that was exported at the cell surface and not retained in the TGN. They also hypothesized that this region could contain an endocytic motif, but their results showed that MHV-M protein does not cycle between the cell surface and the TGN, but rather acts as a Golgi resident protein (Locker et al., 1994a). Additionally, their results suggest that both the transmembrane domain and the C-terminal domain are involved in this retention. Nevertheless, this last interpretation was obtained through deletions of two transmembrane domains, resulting in a mutant protein containing only one transmembrane domain, which could drastically modify its trafficking, as discussed further.

In another paper, *Locker et al.* looked at the oligomerization of MHV-M as a possible determinant for its TGN retention (Locker et al., 1995). They followed oligomerization of the M protein along the secretory pathway, and used the O-glycosylation as a marker of transport, to identify in which compartment each form of the protein was located. They were able to demonstrate that the Golgi and TGN forms were heterogeneous complexes, whereas the ER and ERGIC forms migrated as monomers. Interestingly, the mutant lacking the C-terminal 22 amino acids also migrated as complexes, but of smaller size.

Hence, the model proposed by *Locker et al.* is that the TGN retention of MHV-M protein is mediated by both its oligomerization (likely through the transmembrane domains) and by an interaction with TGN protein(s) mediated by its cytoplasmic domain, resulting in retention into the TGN.

Similar results were obtained with the IBV-M/VSV-G chimeric protein generated by *Machamer et al.* This protein, in which the transmembrane domain of VSV-G was replaced by the first transmembrane segment of IBV-M, was retained intracellularly in the Golgi region. It was subsequently demonstrated that this protein formed stable, large oligomers, whereas mutants of this protein exported at the cell surface were detected in monomers.

In the case of MERS-M, we were able to show that the KxGxYR motif in the cytoplasmic tail of the protein is necessary for its TGN retention. Nevertheless, the underlying mechanism of this retention is unknown. As it was demonstrated in the

literature that oligomerization of M proteins could play a role in their retention to the Golgi region, investigating the role of the KxGxYR motif in oligomerization of the protein was interesting.

For this purpose we compared the formation of oligomers of MERS-M and MERS-M-KGYR proteins detected both at the cell surface and in lysates. We detected the presence of oligomers of the M protein, visible in western blot as a strong band at 40 kDa (likely corresponding to dimers) and a smear of higher molecular weights bands. Interestingly, those forms of the proteins were detected both at the cell surface and in the lysates, and in comparable amounts for MERS-M and MERS-M-KGYR.

Hence, the KxGxYR signal does not seem to have an effect on oligomerization of the M protein. Furthermore, oligomerization of MERS-M does not seem to mediate TGN retention, as high molecular weight oligomers were detected at the cell surface.

Other motifs, possibly located in other domains of MERS-M protein, could be involved in its specific localization to the TGN. In that light, we cannot exclude that there could be cooperation between the motifs of the C-terminal domain that we identified, and other yet unknown motifs located elsewhere in the protein.

To confirm the role of the KxGxYR motif as a TGN retention signal, we transferred it on a protein normally localized at the cell surface, in order to test if it could retain the chimeric protein intracellularly. For this purpose, we constructed MERS-M chimeras with the CD4 protein. Unfortunately, most of the chimeric proteins that we constructed displayed folding problems, and were largely blocked within the ER when expressed in cells. We hypothesized that this was due to the differences of protein structures, as CD4 is type I transmembrane protein containing only one transmembrane segment, whereas MERS-M is a triple membrane-spanning protein.

Hence, we constructed new chimeras, but this time between MERS-M and the M protein of another coronavirus, IBV. That way, we could keep the three transmembrane segment architecture of the protein, and supposedly avoid folding defaults. Importantly, MERS-M and IBV-M wild-type proteins display different localizations at steady state: MERS-M localizes to the TGN and IBV-M to the *cis*-Golgi and ERGIC.

Indeed, intracellular trafficking of IBV-M was previously studied and a role of the first transmembrane domain of the protein in ERGIC localization was reported. This was

demonstrated first by deleting transmembrane segments of the protein two by two: the authors found that a protein devoided of its second and third transmembrane segments (Δm2,3) was still retained intracellularly (Machamer and Rose, 1987a). Then, this was further confirmed by constructing chimeras in which the first transmembrane segment of IBV-M replaced the transmembrane domain of the G protein of VSV (Gm1). This chimeric protein was retained intracellularly (Swift and Machamer, 1991). Furthermore, four uncharged polar residues in the first transmembrane segment were identified as being critical for Golgi retention of the Gm1 protein (Asn465, Thr469, Thr476, and Gln480) as mutation of these residues on the chimeric protein led to its export to the cell surface (Machamer et al., 1993a). These mutations were also inserted into the full length M protein, but all mutants were blocked within the ER, and this was interpreted as being due to folding defaults (Swift and Machamer, 1991).

However, the results we obtained with our MERS/IBV chimeras argue in favor of a role of the C-terminal domain of IBV-M in specific localization to the ERGIC compartment. Indeed, switching the C-terminal domains of the proteins led to a switch of their specific localizations, whereas neither the TM1 of MERS-M nor the TM1 of IBV-M were able to induce specific intracellular retention. Specifically, when we replaced the first transmembrane segment of the MERS-M protein with the one of IBV-M in the context of the KxGxYR mutant (TM1-IBV/MERS-M-KxGxYR), the chimeric protein was exported to the cell surface.

As this was in contrast with what had been demonstrated by *Machamer et al.,* we decided to further investigate the role of the C-terminal domain of IBV-M in its retention in the ERGIC.

For this purpose, we constructed deletion mutants of the C-terminal domain of IBV-M. Our results suggest a role of the last 20 amino acids and especially of the QSVDT sequence, in ER export. We could not yet identify the motif involved in ER export, but individual mutations of these five amino acids will be necessary to identify the residues involved.

Actually, it is likely that the IBV-M protein contains several localization signals located in different parts of the protein, that together confer the specific ERGIC localization. The mechanism could resemble the one that was described for ERGIC53, with the protein containing both ER-export and ER-retrieval signals, and thus being

constantly exported from and recycled to the ER, resulting in ERGIC localization at steady state. This is especially likely considering what was demonstrated by *Machamer et al.* concerning the implication of the first transmembrane segment in ERGIC localization.

Nevertheless, their results were obtained using deletions of two transmembrane segments at a time or chimeras between IBV-M and VSV-G, so in all cases the proteins had only one transmembrane segment. This could be an issue both because of possible folding defaults (as we observed for our CD4/MERS-M chimeras) and because of the importance of transmembrane domains for localization. Additionally, in their chimeras, the presence of the C-terminal tail of the VSV-G protein containing the DxE ER export signal could also have altered the localization of the protein.

Sorting determinants of the transmembrane segments are usually not conserved amino acid sequences but rather physical properties (Cosson et al., 2013). Thus, transmembrane domains alone can influence localization depending on their length, their hydrophilicity, the type of residues they contain (basic, polar) or on which face of the helix they are exposed… etc.

There are a few examples of transmembrane domains determining Golgi localization, such as for proteins GOLPH2, alpha-2,6-sialyltransferase, or syntaxin-5 (Cosson et al., 2013), but it is more frequent that transmembrane domains mediate ER retention. Importantly, ER retention mediated by the transmembrane domains was demonstrated for several viral proteins, including HCV E2 (Ciczora et al., 2005), Yellow fever virus prME heterodimer (Ciczora et al., 2010) or Dengue virus envelope protein (Hsieh et al., 2010).

In that light, and considering previous results, cooperation between sorting signals located in different domains of the M protein is probable for IBV-M and MERS-M, and maybe more generally for coronavirus M proteins. Consequently, it would be interesting for both proteins to continue looking for specific targeting signals in the different domains, in order to decipher precisely the targeting mechanisms. In addition, identifying host factors involved in recognition of the KxGxYR motif would help to understand the TGN retention mechanism it mediates.

To go further in the characterization of the new KxGxYR motif, investigating its role in M-S, M-E and M-N interactions in co-immunoprecipitation assays would be

interesting, as the MERS-M protein supposedly interacts with all other structural proteins upon assembly, likely through its C-terminal domain. Using reverse genetics to insert the mutation into a full-length genome would also allow observing the effects of the mutation in the context of infection.

5.2 Project II: Characterization of an antiviral against HCoV-229E

Coronaviruses are emergent pathogens of growing interest to human medicine, and yet we do not have any specific antiviral agent against this family of viruses. Similarly, experimental vaccines are under development but none is available at the moment. Therefore, treatments are mostly symptomatic and the epidemics were only contained thanks to very strict quarantine and hygiene measures. Hence, there is an urgent need for specific antivirals directed against coronaviruses.

In our study, we used a screening approach to identify drugs with antiviral activity against coronavirus HCoV-229E. We used this coronavirus as a model because key steps of the coronavirus viral cycle are conserved in all species, it is a human pathogen, and HCoV-229E can be handled in BSL-2 level facilities, making experimentation easier than for SARS-CoV or MERS-CoV in BSL-3 facilities.

From that screen, we selected four drugs with antiviral activity against HCoV-229E for further characterization: digoxigenin, trifluoperazine dihydrochloride (TH), perhexiline maleate (PM) and astemizole. All drugs inhibited HCoV-229E infection, but after the first tests, digoxigenin was selected for further characterization as it had the lowest IC50 (250nM), had a clear post-infection effect, and exhibited no toxicity at low concentrations.

Notably, in our hands, digoxigenin induced a decrease in cell proliferation upon long incubations, but no toxicity. Importantly, cardiac glycosides in general are known to be toxic, and the therapeutic index of digoxin for its use in treatment of heart failure is extremely narrow: recommended concentration of digoxin in serum is 0.7 to 0.9 ng/ml, with concentrations above 1-2ng/mL causing dangerous adverse effects (Goldberger and Goldberger, 2012).

Nevertheless, at the concentrations we used, we did not observe toxicity, but rather a marked decrease of cell proliferation after incubation with the drug. Interestingly, it was previously demonstrated that several cardiac glycosides are capable of inhibiting proliferation in cancer cell lines (López-Lázaro et al., 2005), which suggests that this could be what we observed in our Huh-7 cells (derived from hepatocarcinoma) upon incubation with digoxigenin. More generally the anti-cancer potential of cardiac

glycosides has been demonstrated on various cell lines and in patient treatment (Winnicka et al.).

Since we only used the MTS assay in our study, which does not allow us to distinguish between a decreased cell proliferation and cell mortality, it would be interesting to confirm that we indeed observed a decrease in cell proliferation rather than toxicity using both MTS assay and trypan blue staining.

Regardless, we saw no decrease of proliferation or toxicity for the doses and incubation times that we used in our assays. Hence, the post-infection inhibitory effect we observed on HCoV-229E is not due to cell toxicity.

To further confirm that its inhibitory effect was only on a post-entry step, we tested digoxigenin's effect on pseudoparticles. As pseudoparticles are retroviral cores containing a minigenome with a luciferase reporter gene, and carrying a viral protein (HCoV-2229E spike or VSV-G) of interest in their envelope, they fully mimic the entry step of the viral cycle.

Surprisingly, digoxigenin inhibited infection with both 229Epp and VSVGpp, at comparable IC50s. The inhibitory effect could thus be on both HCoV-229E and VSV entry: this would be in line with what was demonstrated by *Burkard et al*, namely an effect of cardiac glycosides on MHV and VSV entry step (Burkard et al., 2015).

Indeed in this study, ATP1A1 gene silencing showed that the $\alpha 1$ subunit of the pump is essential to infection with MHV and FIPV. This absence of ATP1A1 did not affect virus binding to host cells but inhibited entry of MHV. Virus particles accumulate in pre-endosomal invaginations, and fusion seemed to be blocked. Consistently, low concentrations of ouabain and bufalin also inhibited infection with MHV, FIPV and MERS-CoV, when present during virus inoculation. Addition of Src kinase inhibitors blocked the antiviral effect of ouabain, suggesting that this signaling pathway is necessary for ATP1A1 mediated inhibition of MHV entry (Burkard et al., 2015).

However, in another study demonstrating the effect of cardiac glycosides on TGEV infection, *Yang et al.* observed no effect of these compounds on MHV infection (Yang et al., 2017).

Nevertheless, our experiments with different times of addition of the drug showed that adding digoxigenin before or during infection with HCoV-229E or HCV had no

effect, which does not correlate with an effect on entry. Hence, although MHV and HCoV-229E are both coronaviruses, the different receptors they use (respectively CEACAM1 and APN), or their differences in host and target cell lines could explain the different effect of cardiac glycosides on their infection. In addition, even though replication mechanisms are conserved across coronaviruses genera, these results suggest that there may be different host factors involved depending on the virus.

Thus, considering our own results indicating a post-infection effect on HCoV-229E, we believe it is more probable that the inhibition of pseudoparticles infection we observed was due to an inhibition of the transcription/retrotranscription steps occurring after entry of the pseudoparticles into the cells, rather than an effect on the entry step. This is corroborated by the fact that an effect of ouabain on the retrotranscription of Murine Leukemia Virus (MLV) was previously demonstrated in the literature (Tomita and Kuwata, 1978), which could account for the inhibition of pseudoparticles infection.

Interestingly, digoxigenin also inhibits HCV infection during the post-infection step of the viral cycle, with a slightly higher IC50 than for HCoV-229E. Hence, we decided to further investigate the inhibition mechanism of digoxigenin on HCV, as we had at our disposal in the lab Huh-7 cells stably expressing an HCV replicon with a GFP reporter gene. A subgenomic replicon is a modified form of HCV genome that replicates autonomously in cells, but does not produce infectious particles. This replicon is assembly defective and allows us to study replication only, and to bypass other steps of the viral cycle. These cells stably produce the replicon, and the level of expression can be quantified using the reporter gene GFP. After incubation of these cells with digoxigenin, we saw no decrease of GFP expression when compared to the control. This indicated that digoxigenin could not inhibit replication in that setting. Thus, we hypothesize that the drug acts on an early post-entry step, before replication onset.

This could be further investigated on coronaviruses using electroporation of viral RNA into cells, and incubation of these cells with digoxigenin. In that setting, we would be able to test the effect of digoxigenin on replication only, bypassing the entry step. Furthermore, replicon systems were developed for several coronaviruses, using different approaches (reviewed in (Almazán et al., 2014). Interestingly, an HCoV-229E

replicon with a GFP reporter gene was previously engineered by Thiel et al. (Thiel et al., 1998), and was previously used for identification of replicase inhibitors (Hertzig, 2004).

Taken together, these results suggest that digoxigenin has an effect on an early step of replication, which could be common to RNA (+) viruses, as suggested by its similar effect on both HCoV-229E and HCV.

In order to further characterize the inhibition mechanism of digoxigenin on HCoV-229E, it would be useful to first confirm that the effect is mediated by its cellular target, the Na^+/K^+ ATPase pump. For this purpose, RNA interference experiments with ATP1A1 and/or other isoforms of the α subunit would be necessary.

Additionally, testing the effect of digoxigenin on other coronaviruses, especially on the highly pathogenic MERS-CoV would be interesting to confirm its potential as an anti-CoV compound, and also to decipher its inhibition mechanism.

To go further, we tried to raise resistance mutants of HCoV-229E by incubating infected cells with increasing concentrations of digoxigenin upon serial passages. The objective was to isolate and then sequence these resistance mutants, and to identify mutations in their genome, in non-structural proteins. This could have given us insight into the mechanism of inhibition of the drug, and the precise step of viral replication that was impacted. Unfortunately, we were unable yet to raise mutants that were resistant to digoxigenin.

Interestingly, HCV and HCoV-229E share many of their replication mechanisms. First of all, their positive (+) ssRNA genome is translated into polyprotein(s) that is/are then cleaved at many sites by viral or cellular proteases. For HCV there is only one polyprotein that is cleaved by both viral proteases (non-structural proteins) and signal peptidases of the ER (structural proteins) (Dubuisson, 2007), whereas for coronaviruses the two polyproteins are cleaved by autoprocessing of the viral proteases. The replication complex then assembles, containing the RNA-dependent-RNA polymerase, along with many viral and cellular proteins. There is then the formation of negative genomic RNA intermediates serving as templates for replication of the viral genome. For HCV, all proteins are translated directly from the genomic RNA and then cleaved, whereas coronavirus produce subgenomic RNAs that act as mRNAs coding for structural and accessory proteins. Importantly, both HCV and coronaviruses induce important membrane rearrangements during their replication, and the replication complex

localizes to these "replication factories". For HCV, there is formation of a membranous web composed of DMVs, ER derived membranes and lipid droplets. For coronaviruses, DMVs, convoluted membranes or zippered-ER were observed depending on the species.

One of these common mechanisms could be targeted by digoxigenin, thus inhibiting replication of both HCV and HCoV-229E viruses.

More generally, the antiviral potential of cardiac glycosides was already demonstrated in several studies, and on many different families of RNA and DNA viruses (summed up in Figure 35 and reviewed in Amarelle and Lecuona, 2018). Interestingly, for many viruses, an effect of cardiac glycosides on a post-entry step or more precisely an early step of replication was described.

An inhibitory effect of digoxin was demonstrated on Chikungunya virus (CHIKV) infection, and on other *Togaviridae* (Ross river virus, Sindbis virus). An effect on vesicular stomatitis virus (VSV) and mammalian reovirus was also observed (Ashbrook et al., 2016). Importantly, the digoxin-mediated inhibition of CHIKV and reovirus occurred at a post-entry step, as it was not bypassed when viral fusion occurred directly at the cell surface. Interestingly, digoxin-resistant mutants of CHIKV were raised and sequencing of their genome allowed the identification of multiple mutations in non-structural proteins implicated in replication (Ashbrook et al., 2016).

A post-entry inhibitory effect of lanatoside C, another cardiac glycoside, was also observed on CHIKV, Sindbis Virus, and Dengue virus infection, and the results indicate that lanatoside C inhibits early replication processes of Dengue virus (Cheung et al., 2014). It is worth noting that Dengue virus is a member of the *Flaviviridae* family, as HCV, and that this effect on an early step of Dengue virus replication correlates with our results so far for HCV.

Digoxin is also an interesting anti-HIV drug candidate, as studies demonstrated its effect on HIV-1 infection, including clinical strains (Laird et al., 2014; Wong et al., 2013). Both studies demonstrated that cardiac glycosides inhibit HIV gene expression and therefore protein synthesis. Importantly, *Laird et al*. observed that this inhibition was dependent of the Na^+/K^+ ATPase pump, but independent of the increase in intracellular Ca^{2+} induced by cardiac glycosides (Laird et al., 2014). More recently, *Wong et al.* showed that other cardiac glycosides (including digoxigenin) are capable of

inhibiting HIV gene expression in a Ca^{2+} independent way, by modulating MEK/ERK signaling (Wong et al., 2018b). Furthermore, *Wong et al.* demonstrated that digoxin altered viral RNA splice site use, precisely by modification of SR proteins, resulting in loss of the viral factor Rev (Wong et al., 2013).

Interestingly, a similar effect on RNA splicing was also observed on Adenoviruses upon incubation with digoxin or digitoxin (Grosso et al., 2016). This correlates with the fact that cardiac glycosides were previously identified as RNA splicing inhibitors (Stoilov et al., 2008).

It is worth noting that many viruses are inhibited by cardiac glycosides, and the effect is often on an early step of replication. Interestingly, it was demonstrated for both Adenoviruses and HIV that the RNA splicing was altered in presence of cardiac glycosides. Although this mechanism cannot be involved in inhibition of coronavirus or HCV replication, as there is no RNA splicing during their cycle, other steps of their replication could be inhibited by cardiac glycosides.

Importantly, many viruses for different families are inhibited by cardiac glycosides (Figure 35), suggesting that these drugs inhibit a process shared by many viruses, or that they are able to act on different mechanisms depending on the virus and cell line.

Cardiac glycosides are thus very promising antiviral compounds, with a large spectrum of activity. Nevertheless, despite this potential, the established toxicity of cardiac glycosides illustrated by the narrow therapeutic window of digoxin in humans, may prevent their systemic use as antiviral drugs. However, recent studies demonstrated that at a lower dosage, digoxin toxicity in patients treated for heart failure could be avoided: hence, the recommendation for digoxin serum concentration in patients (which was historically 0,8 to 2ng/mL) should now be <1.0 ng/ml and preferably 0.7 to 0.9 ng/ml (Goldberger and Goldberger, 2012; MacLeod-Glover et al., 2016). Cardiac glycosides toxicity could also be circumvented by their diversity, as some derived compounds may exhibit less toxicity or different pharmacokinetics than the digoxin/digitoxin that are classically used.

As discussed before, several compounds are currently in testing as antivirals against coronaviruses, and especially against SARS-CoV and MERS-CoV (*see Treatments, in Introduction*). However, and despite our results being preliminary, digoxigenin is still an interesting candidate for various reasons. First, although this needs confirmation in our study, digoxigenin supposedly acts on a cellular target (the Na^+/K^+ ATPase pump), which may reduce the apparition of resistance mutants, compared to drugs targeting viral components. Another advantage of digoxigenin is the fact that the drug acts on an early replication step. Hence, as replication mechanisms are well conserved in all coronaviruses, digoxigenin could be a pan-coronavirus inhibitor, in a similar manner as the promising K22 compound, which inhibits the formation of DMVs (Lundin et al., 2014). Finally, digoxin has been used for centuries in the treatment of heart failure, was FDA approved in 1954, and is still in use: thus, we have a lot of knowledge on cardiac glycosides administration, toxicity, possible adverse effects… etc. This could greatly accelerate the use of the compound as an antiviral agent, which is not negligible considering the time needed to obtain approval of a new drug.

Figure 35 – ***Table of described antiviral effects of cardiac glycosides***

VIRUS	COMPOUND	SUGGESTED MECHANISM	REFERENCE(S)
RNA viruses			
Chikungunya virus *(Togaviridae)*	Digoxin, lanatoside C	Post-entry, likely replication	*Ashbrook et al., 2016 ; Cheung et al., 2014*
Sindbis virus (*Togaviridae)*	Digoxin, lanatoside C	/	*Ashbrook et al., 2016; Cheung et al., 2014*
Ross river virus *(Togaviridae)*	Digoxin	/	*Ashbrook et al., 2016*
Dengue virus *(Flaviviridae)*	Lanatoside C	Inhibition of early replication processes	*Cheung et al., 2014*
TGEV *(Coronaviridae)*	Digoxin, digitoxin, oleandrin, ouabain…	/	*Yang et al., 2017*
MHV *(Coronaviridae)*	Ouabain, bufalin	Activation of Src, inhibition of virus fusion	*Burkard et al., 2015*
MERS-CoV *(Coronaviridae)*	Ouabain, bufalin	/	*Burkard et al., 2015*
FIPV *(Coronaviridae)*	Ouabain, bufalin	/	*Burkard et al., 2015*
Ebola virus *(Filoviridae)*	Ouabain	Competition with VP24 to bind ATP1A1	*García-Dorival et al., 2014*
HIV *(Retroviridae)*	Digoxin and 12 other	Alteration of RNA splicing	*Laird et al., 2014 ; Wong et al., 2013 ; Wong et al., 2018*
Murine Leukemia Virus *(Retroviridae)*	Ouabain	Inhibition of protein synthesis, K+ dependent	*Tomita and Kuwata, 1978*
Vesicular stomatitis virus *(Rhabdoviridae)*	Digoxin	Activation of Src, inhibition of virus entry after binding	*Ashbrook et al., 2016 ; Burkard et al., 2015*
Vaccinia virus *(Poxviridae)*	Digoxin, cymarin, neriifolin, ouabain, and lanatoside C	Inhibition of protein synthesis	*Deng et al., 2007*
Mammalian reovirus *(Reoviridae)*	Digoxin	Post-entry, likely replication	*Ashbrook et al., 2016*
Influenza A and B *(Orthomyxoviridae)*	Lanatoside C, strophanthidin, digoxin	Inhibition of PKC	*Hoffmann et al., 2008 Mi et al., 2010*
Sendai virus *(Paramyxoviridae)*	Ouabain	Inhibition of RNA and protein synthesis	*Nagai et al., 1972*
Respiratory Syncitial virus *(Paramyxoviridae)*	Digoxin, Digitoxin	Modification of intracellular K+ and Na+	*Norris et al., 2018*
DNA viruses			
Adenoviridae	Digoxin, Digitoxin	Alteration of RNA splicing,	*Grosso et al., 2016*
Cytomegalovirus *(Herpesviridae)*	Digitoxin	Induction of AMPK activity and autophagy	*Cohen et al., 2016 ; Mukhopadhyay et al., 2018*
Herpes simplex virus *(Herpesviridae)*	Ouabain, Digitoxin, Digoxin	HSV replication prior to viral gene expression.	*Dodson et al., 2007 ; Su et al., 2008*

5.3 Importance

As illustrated by the recent emergence of highly pathogenic human coronaviruses SARS-CoV and MERS-CoV in the last two decades, coronaviruses are a family of emergent pathogens. This is especially striking considering their high capacity to cross the species barrier, and the large spectrum of hosts and pathogenesis they exhibit. Furthermore, animal coronaviruses are still a burden to the pork and poultry industries worldwide, and new variants of avian and swine coronaviruses constantly emerge, causing new epidemics. Consequently, coronaviruses are now a major interest to both human and veterinary medicine.

Nevertheless, many knowledge gaps remain concerning the coronavirus life cycle, transmission routes, host jumping mechanisms… etc. It is therefore essential that more research effort be put both into further understanding of coronaviruses biology and in the search of specific antiviral compounds and the development of vaccines.

In that context, the results of my first project bring light into the trafficking mechanisms of coronavirus M proteins to the TGN. The results obtained during this study are of interest both to virology and to cell biology. Our work on the intracellular trafficking of MERS-M protein is a first step to bring insight into the mechanisms and motifs involved in the trafficking of the protein to the assembly site. Even though this work needs to be pursued in order to decipher the importance of the identified motifs in the context of infection, using expression of M alone in cells was necessary to the comprehension of its intracellular trafficking, and allowed us to identify two motifs essential to its specific targeting to the TGN. Moreover, the assembly step can be a target for antiviral compounds, as seen with the protease inhibitors for HIV that induce the secretion of immature particles (Adamson, 2012). Therefore, it is essential to have a better understanding of the assembly step, and the identification of trafficking signals is a first step towards it.

Importantly concerning cell biology, the identification of the KxGxYR motif is interesting to the field of intracellular trafficking in general as this motif is new to literature. Interestingly, the KxGxYR motif is conserved in many *Betacoronaviruses* M proteins and might be a more general mean of TGN retention in coronaviruses M proteins. Furthermore, targeting motifs such as KxGxYR can be conserved across many

living organisms, so this might not be limited to coronaviruses. Further investigation is needed in order to understand how the KxGxYR motif mediates retention in the TGN.

The second project that we initiated aimed at identifying and characterizing an antiviral active against HCoV-229E. Indeed, even though SARS-CoV and MERS-CoV epidemics were mostly contained thanks to strict hygiene and quarantine measures, we still have no specific antivirals directed against coronaviruses, and the treatments are mostly symptomatic. Furthermore, the MERS-CoV epidemic is still ongoing.

Thus, coronaviruses still pose a major health threat, as illustrated by the fact that both SARS-CoV and MERS-CoV are on WHO's blueprint list of priority diseases (WHO). In that context, finding antivirals active against coronaviruses is of major importance. Our results demonstrate that digoxigenin, a cardiac glycoside derived from the well-known digoxin, is active against HCoV-229E and HCV, and that it may act on an early post-infection step. Although the work is preliminary and further investigation is needed, the compound is promising as it acts at relatively low concentrations (IC50 of 250nM) and exhibits little toxicity around those concentrations. We chose HCoV-229E as a model for biosafety reasons, but our results may be extended to the coronaviruses family in general, although this remains to be tested.

Bibliography

Adamson, C.S. (2012). Protease-Mediated Maturation of HIV: Inhibitors of Protease and the Maturation Process. Mol Biol Int *2012*.

Addie, D.D., Toth, S., Murray, G.D., and Jarrett, O. (1995). Risk of feline infectious peritonitis in cats naturally infected with feline coronavirus. Am. J. Vet. Res. *56*, 429–434.

Adedeji, A.O., Singh, K., Calcaterra, N.E., DeDiego, M.L., Enjuanes, L., Weiss, S., and Sarafianos, S.G. (2012). Severe acute respiratory syndrome coronavirus replication inhibitor that interferes with the nucleic acid unwinding of the viral helicase. Antimicrob. Agents Chemother. *56*, 4718–4728.

Adney, D.R., van Doremalen, N., Brown, V.R., Bushmaker, T., Scott, D., de Wit, E., Bowen, R.A., and Munster, V.J. (2014). Replication and Shedding of MERS-CoV in Upper Respiratory Tract of Inoculated Dromedary Camels. Emerg Infect Dis *20*, 1999–2005.

Agrawal, A.S., Tao, X., Algaissi, A., Garron, T., Narayanan, K., Peng, B.-H., Couch, R.B., and Tseng, C.-T.K. (2016). Immunization with inactivated Middle East Respiratory Syndrome coronavirus vaccine leads to lung immunopathology on challenge with live virus. Hum Vaccin Immunother *12*, 2351–2356.

Aizman, O., Uhlén, P., Lal, M., Brismar, H., and Aperia, A. (2001). Ouabain, a steroid hormone that signals with slow calcium oscillations. PNAS *98*, 13420–13424.

Akimkin, V., Beer, M., Blome, S., Hanke, D., Höper, D., Jenckel, M., and Pohlmann, A. (2016). New Chimeric Porcine Coronavirus in Swine Feces, Germany, 2012. Emerg Infect Dis *22*, 1314–1315.

Alagaili, A.N., Briese, T., Mishra, N., Kapoor, V., Sameroff, S.C., Wit, E. de, Munster, V.J., Hensley, L.E., Zalmout, I.S., Kapoor, A., et al. (2014). Middle East Respiratory Syndrome Coronavirus Infection in Dromedary Camels in Saudi Arabia. MBio *5*, e00884-14.

Al-Amri, S.S., Abbas, A.T., Siddiq, L.A., Alghamdi, A., Sanki, M.A., Al-Muhanna, M.K., Alhabbab, R.Y., Azhar, E.I., Li, X., and Hashem, A.M. (2017). Immunogenicity of Candidate MERS-CoV DNA Vaccines Based on the Spike Protein. Sci Rep *7*, 44875.

Alexander, D.J., and Gough, R.E. (1977). Isolation of avian infectious bronchitis virus from experimentally infected chickens. Res. Vet. Sci. *23*, 344–347.

Alharbi, N.K., Padron-Regalado, E., Thompson, C.P., Kupke, A., Wells, D., Sloan, M.A., Grehan, K., Temperton, N., Lambe, T., Warimwe, G., et al. (2017). ChAdOx1 and MVA based vaccine candidates against MERS-CoV elicit neutralising antibodies and cellular immune responses in mice. Vaccine *35*, 3780–3788.

Almazán, F., DeDiego, M.L., Sola, I., Zuñiga, S., Nieto-Torres, J.L., Marquez-Jurado, S., Andrés, G., and Enjuanes, L. (2013). Engineering a replication-competent, propagation-defective Middle East respiratory syndrome coronavirus as a vaccine candidate. MBio *4*, e00650-00613.

Almazán, F., Sola, I., Zuñiga, S., Marquez-Jurado, S., Morales, L., Becares, M., and Enjuanes, L. (2014). CORONAVIRUS REVERSE GENETIC SYSTEMS: INFECTIOUS CLONES AND REPLICONS. Virus Res *189*, 262–270.

Amarelle, L., and Lecuona, E. (2018). The Antiviral Effects of Na,K-ATPase Inhibition: A Minireview. Int J Mol Sci *19*.

Angelini, M.M., Akhlaghpour, M., Neuman, B.W., and Buchmeier, M.J. (2013). Severe Acute Respiratory Syndrome Coronavirus Nonstructural Proteins 3, 4, and 6 Induce Double-Membrane Vesicles. MBio *4*.

Appenzeller-Herzog, C. (2006). The ER-Golgi intermediate compartment (ERGIC): in search of its identity and function. Journal of Cell Science *119*, 2173–2183.

Armstrong, J., and Patel, S. (1991a). The Golgi sorting domain of coronavirus E1 protein. Journal of Cell Science *98*, 567–575.

Armstrong, J., and Patel, S. (1991b). The Golgi sorting domain of coronavirus E1 protein. Journal of Cell Science *98 (Pt 4)*, 567–575.

Armstrong, J., Patel, S., and Riddle, P. (1990). Lysosomal sorting mutants of coronavirus E1 protein, a Golgi membrane protein. J. Cell. Sci. *95 (Pt 2)*, 191–197.

Arndt, A.L., Larson, B.J., and Hogue, B.G. (2010a). A Conserved Domain in the Coronavirus Membrane Protein Tail Is Important for Virus Assembly. Journal of Virology *84*, 11418–11428.

Arndt, A.L., Larson, B.J., and Hogue, B.G. (2010b). A Conserved Domain in the Coronavirus Membrane Protein Tail Is Important for Virus Assembly. The Journal of Virology *84*, 11418–11428.

Ashbrook, A.W., Lentscher, A.J., Zamora, P.F., Silva, L.A., May, N.A., Bauer, J.A., Morrison, T.E., and Dermody, T.S. (2016). Antagonism of the Sodium-Potassium ATPase Impairs Chikungunya Virus Infection. MBio *7*, e00693-16.

Assiri, A., Al-Tawfiq, J.A., Al-Rabeeah, A.A., Al-Rabiah, F.A., Al-Hajjar, S., Al-Barrak, A., Flemban, H., Al-Nassir, W.N., Balkhy, H.H., Al-Hakeem, R.F., et al. (2013). Epidemiological, demographic, and clinical characteristics of 47 cases of Middle East respiratory syndrome coronavirus disease from Saudi Arabia: a descriptive study. Lancet Infect Dis *13*, 752–761.

Athmer, J., Fehr, A.R., Grunewald, M., Smith, E.C., Denison, M.R., and Perlman, S. (2017). In Situ Tagged nsp15 Reveals Interactions with Coronavirus Replication/Transcription Complex-Associated Proteins. MBio *8*.

Azhar, E.I., El-Kafrawy, S.A., Farraj, S.A., Hassan, A.M., Al-Saeed, M.S., Hashem, A.M., and Madani, T.A. (2014). Evidence for camel-to-human transmission of MERS coronavirus. N. Engl. J. Med. *370*, 2499–2505.

Bande, F., Arshad, S.S., Bejo, M.H., Moeini, H., and Omar, A.R. (2015). Progress and challenges toward the development of vaccines against avian infectious bronchitis. J Immunol Res *2015*, 424860.

Bande, F., Arshad, S.S., Omar, A.R., Bejo, M.H., Abubakar, M.S., and Abba, Y. (2016). Pathogenesis and Diagnostic Approaches of Avian Infectious Bronchitis. Advances in Virology *2016*, 1–11.

Banfield, D.K. (2011). Mechanisms of protein retention in the Golgi. Cold Spring Harb Perspect Biol *3*, a005264.

Barlan, A., Zhao, J., Sarkar, M.K., Li, K., McCray, P.B., Perlman, S., and Gallagher, T. (2014). Receptor Variation and Susceptibility to Middle East Respiratory Syndrome Coronavirus Infection. Journal of Virology *88*, 9.

Barlough, J.E., Stoddart, C.A., Sorresso, G.P., Jacobson, R.H., and Scott, F.W. (1984). Experimental inoculation of cats with canine coronavirus and subsequent challenge with feline infectious peritonitis virus. Lab. Anim. Sci. *34*, 592–597.

Barlough, J.E., Johnson-Lussenburg, C.M., Stoddart, C.A., Jacobson, R.H., and Scott, F.W. (1985). Experimental inoculation of cats with human coronavirus 229E and subsequent challenge with feline infectious peritonitis virus. Can J Comp Med *49*, 303–307.

Barlowe, C. (2003). Signals for COPII-dependent export from the ER: what's the ticket out? Trends in Cell Biology *13*, 295–300.

Barlowe, C.K., and Miller, E.A. (2013). Secretory Protein Biogenesis and Traffic in the Early Secretory Pathway. Genetics *193*, 383–410.

Barthold, S.W. (1997). Mouse Hepatitis Virus Infection, Intestine, Mouse. In Digestive System, T.C. Jones, J.A. Popp, and U. Mohr, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 379–384.

Barthold, S.W., and Smith, A.L. (1984). Mouse hepatitis virus strain--related patterns of tissue tropism in suckling mice. Arch. Virol. *81*, 103–112.

Basler, C.F. (2015). Innate immune evasion by filoviruses. Virology *479–480*, 122–130.

Baudoux, P., Carrat, C., Besnardeau, L., Charley, B., and Laude, H. (1998a). Coronavirus Pseudoparticles Formed with Recombinant M and E Proteins Induce Alpha Interferon Synthesis by Leukocytes. J Virol *72*, 8636–8643.

Baudoux, P., Carrat, C., Besnardeau, L., Charley, B., and Laude, H. (1998b). Coronavirus Pseudoparticles Formed with Recombinant M and E Proteins Induce Alpha Interferon Synthesis by Leukocytes. J. VIROL. *72*, 8.

Beachboard, D.C., Anderson-Daniels, J.M., and Denison, M.R. (2015). Mutations across murine hepatitis virus nsp4 alter virus fitness and membrane modifications. J. Virol. *89*, 2080–2089.

Becker, M.M., Graham, R.L., Donaldson, E.F., Rockx, B., Sims, A.C., Sheahan, T., Pickles, R.J., Corti, D., Johnston, R.E., Baric, R.S., et al. (2008). Synthetic recombinant bat SARS-like coronavirus is infectious in cultured cells and in mice. Proc. Natl. Acad. Sci. U.S.A. *105*, 19944–19949.

Beigel, J.H., Voell, J., Kumar, P., Raviprakash, K., Wu, H., Jiao, J.-A., Sullivan, E., Luke, T., and Davey, R.T. (2018). Safety and tolerability of a novel, polyclonal human anti-MERS coronavirus antibody produced from transchromosomic cattle: a phase 1 randomised, double-blind, single-dose-escalation study. Lancet Infect Dis *18*, 410–418.

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. Proceedings of the National Academy of Sciences of the United States of America *106*, 5871.

Belouzard, S., Millet, J.K., Licitra, B.N., and Whittaker, G.R. (2012a). Mechanisms of Coronavirus Cell Entry Mediated by the Viral Spike Protein. Viruses *4*, 1011–1033.

Belouzard, S., Millet, J.K., Licitra, B.N., and Whittaker, G.R. (2012b). Mechanisms of coronavirus cell entry mediated by the viral spike protein. Viruses *4*, 1011–1033.

Belsham, G.J., Rasmussen, T.B., Normann, P., Vaclavek, P., Strandbygaard, B., and Bøtner, A. (2016). Characterization of a Novel Chimeric Swine Enteric Coronavirus from Diseased Pigs in Central Eastern Europe in 2016. Transbound Emerg Dis *63*, 595–601.

Benetka, V., Kübber-Heiss, A., Kolodziejek, J., Nowotny, N., Hofmann-Parisot, M., and Möstl, K. (2004). Prevalence of feline coronavirus types I and II in cats with histopathologically verified feline infectious peritonitis. Vet. Microbiol. *99*, 31–42.

Bergmann, C.C., Ramakrishna, C., Kornacki, M., and Stohlman, S.A. (2001). Impaired T Cell Immunity in B Cell-Deficient Mice Following Viral Central Nervous System Infection. The Journal of Immunology *167*, 1575–1583.

Bertram, S., Glowacka, I., Müller, M.A., Lavender, H., Gnirss, K., Nehlmeier, I., Niemeyer, D., He, Y., Simmons, G., Drosten, C., et al. (2011). Cleavage and activation of the severe acute respiratory syndrome coronavirus spike protein by human airway trypsin-like protease. J. Virol. *85*, 13363–13372.

Bessen, H.A. (1986). Therapeutic and toxic effects of digitalis: William Withering, 1785. J Emerg Med *4*, 243–248.

Bidokhti, M.R.M., Tråvén, M., Krishna, N.K., Munir, M., Belák, S., Alenius, S., and Cortey, M. (2013). Evolutionary dynamics of bovine coronaviruses: natural selection pattern of the spike gene implies adaptive evolution of the strains. J. Gen. Virol. *94*, 2036–2049.

Bodmer, B.S., Fiedler, A.H., Hanauer, J.R.H., Prüfer, S., and Mühlebach, M.D. (2018). Live-attenuated bivalent measles virus-derived vaccines targeting Middle East respiratory syndrome coronavirus induce robust and multifunctional T cell responses against both viruses in an appropriate mouse model. Virology *521*, 99–107.

Bolles, M., Deming, D., Long, K., Agnihothram, S., Whitmore, A., Ferris, M., Funkhouser, W., Gralinski, L., Totura, A., Heise, M., et al. (2011). A double-inactivated severe acute respiratory syndrome coronavirus vaccine provides incomplete protection in mice and induces increased eosinophilic proinflammatory pulmonary response upon challenge. J. Virol. *85*, 12201–12215.

Boncompain, G., and Perez, F. (2013). The many routes of Golgi-dependent trafficking. Histochemistry and Cell Biology *140*, 251–260.

Boniotti, M.B., Papetti, A., Lavazza, A., Alborali, G., Sozzi, E., Chiapponi, C., Faccini, S., Bonilauri, P., Cordioli, P., and Marthaler, D. (2016). Porcine Epidemic Diarrhea Virus and Discovery of a Recombinant Swine Enteric Coronavirus, Italy. Emerg Infect Dis *22*, 83–87.

Borschensky, C.M., and Reinacher, M. (2014). Mutations in the 3c and 7b genes of feline coronavirus in spontaneously affected FIP cats. Res. Vet. Sci. *97*, 333–340.

Bos, K., Wraight, C., and Stanley, K.K. (1993a). TGN38 is maintained in the trans-Golgi network by a tyrosine-containing motif in the cytoplasmic domain. EMBO J. *12*, 2219–2228.

Bos, K., Wraight, C., and Stanley, K.K. (1993b). TGN38 is maintained in the trans-Golgi network by a tyrosine-containing motif in the cytoplasmic domain. EMBO J *12*, 2219–2228.

Boscarino, J.A., Logan, H.L., Lacny, J.J., and Gallagher, T.M. (2008). Envelope protein palmitoylations are crucial for murine coronavirus assembly. J. Virol. *82*, 2989–2999.

Bosch, B.J., van der Zee, R., de Haan, C.A.M., and Rottier, P.J.M. (2003). The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. J. Virol. *77*, 8801–8811.

Bosch, B.J., de Haan, C.A.M., Smits, S.L., and Rottier, P.J.M. (2005). Spike protein assembly into the coronavirion: exploring the limits of its sequence requirements. Virology *334*, 306–318.

Bosch, B.J., Smits, S.L., and Haagmans, B.L. (2014). Membrane ectopeptidases targeted by human coronaviruses. Curr Opin Virol *6*, 55–60.

Bradburne, A.F., Bynoe, M.L., and Tyrrell, D.A. (1967). Effects of a "new" human respiratory virus in volunteers. Br Med J *3*, 767–769.

Brandão, P.E., Gregori, F., Richtzenhain, L.J., Rosales, C. a. R., Villarreal, L.Y.B., and Jerez, J.A. (2006). Molecular analysis of Brazilian strains of bovine coronavirus (BCoV) reveals a deletion within the hypervariable region of the S1 subunit of the spike glycoprotein also found in human coronavirus OC43. Arch. Virol. *151*, 1735–1748.

Braulke, T., and Bonifacino, J.S. (2009). Sorting of lysosomal proteins. Biochim. Biophys. Acta *1793*, 605–614.

Brian, D.A., and Baric, R.S. (2005). Coronavirus genome structure and replication. Curr. Top. Microbiol. Immunol. *287*, 1–30.

Burkard, C., Verheije, M.H., Haagmans, B.L., van Kuppeveld, F.J., Rottier, P.J.M., Bosch, B.-J., and de Haan, C.A.M. (2015). ATP1A1-mediated Src signaling inhibits coronavirus entry into host cells. J. Virol. *89*, 4434–4448.

Callebaut, P.E., and Pensaert, M.B. (1980). Characterization and isolation of structural polypeptides in haemagglutinating encephalomyelitis virus. J. Gen. Virol. *48*, 193–204.

Canton, J., Fehr, A.R., Fernandez-Delgado, R., Gutierrez-Alvarez, F.J., Sanchez-Aparicio, M.T., García-Sastre, A., Perlman, S., Enjuanes, L., and Sola, I. (2018). MERS-CoV 4b protein interferes with the NF-κB-dependent innate immune response during infection. PLoS Pathog. *14*, e1006838.

Cavanagh, D. (2005). Coronaviridae: a review of coronaviruses and toroviruses. 54.

Cavanagh, D. (2007). Coronavirus avian infectious bronchitis virus. Vet. Res. *38*, 281–297.

Cavanagh, D., Davis, P.J., Pappin, D.J., Binns, M.M., Boursnell, M.E., and Brown, T.D. (1986). Coronavirus IBV: partial amino terminal sequencing of spike polypeptide S2 identifies the sequence Arg-Arg-Phe-Arg-Arg at the cleavage site of the spike precursor propolypeptide of IBV strains Beaudette and M41. Virus Res. *4*, 133–143.

Centers for Disease Control and Prevention (CDC) (2003). Severe acute respiratory syndrome--Singapore, 2003. MMWR Morb. Mortal. Wkly. Rep. *52*, 405–411.

Chan, C.-M., Chu, H., Wang, Y., Wong, B.H.-Y., Zhao, X., Zhou, J., Yang, D., Leung, S.P., Chan, J.F.-W., Yeung, M.-L., et al. (2016). Carcinoembryonic Antigen-Related Cell Adhesion Molecule 5 Is an Important Surface Attachment Factor That Facilitates Entry of Middle East Respiratory Syndrome Coronavirus. J. Virol. *90*, 9114–9127.

Chang, H.-W., Egberink, H.F., Halpin, R., Spiro, D.J., and Rottier, P.J.M. (2012). Spike protein fusion peptide and feline coronavirus virulence. Emerging Infect. Dis. *18*, 1089–1095.

Cheever, F.S., Daniels, J.B., Pappenheimer, A.M., and Bailey, O.T. (1949). A MURINE VIRUS (JHM) CAUSING DISSEMINATED ENCEPHALOMYELITIS WITH EXTENSIVE DESTRUCTION OF MYELIN. J Exp Med *90*, 181–194.

Chen, F., Knutson, T.P., Rossow, S., Saif, L.J., and Marthaler, D.G. (2019). Decline of transmissible gastroenteritis virus and its complex evolutionary relationship with porcine respiratory coronavirus in the United States. Scientific Reports *9*, 3953.

Chen, H., Gill, A., Dove, B.K., Emmett, S.R., Kemp, C.F., Ritchie, M.A., Dee, M., and Hiscox, J.A. (2005). Mass Spectroscopic Characterization of the Coronavirus Infectious Bronchitis Virus Nucleoprotein and Elucidation of the Role of Phosphorylation in RNA Binding by Using Surface Plasmon Resonance. J Virol *79*, 1164–1179.

Chen, S.-C., Lo, S.-Y., Ma, H.-C., and Li, H.-C. (2009). Expression and membrane integration of SARS-CoV E protein and its interaction with M protein. Virus Genes *38*, 365–371.

Cheng, V.C.C., Hung, I.F.N., Tang, B.S.F., Chu, C.M., Wong, M.M.L., Chan, K.H., Wu, A.K.L., Tse, D.M.W., Chan, K.S., Zheng, B.J., et al. (2004). Viral replication in the nasopharynx is associated with diarrhea in patients with severe acute respiratory syndrome. Clin. Infect. Dis. *38*, 467–475.

Cheng, V.C.C., Lau, S.K.P., Woo, P.C.Y., and Yuen, K.Y. (2007). Severe acute respiratory syndrome coronavirus as an agent of emerging and reemerging infection. Clin. Microbiol. Rev. *20*, 660–694.

Cheng, Y., Wong, R., Soo, Y.O.Y., Wong, W.S., Lee, C.K., Ng, M.H.L., Chan, P., Wong, K.C., Leung, C.B., and Cheng, G. (2005). Use of convalescent plasma therapy in SARS patients in Hong Kong. Eur. J. Clin. Microbiol. Infect. Dis. *24*, 44–46.

Cheung, Y.Y., Chen, K.C., Chen, H., Seng, E.K., and Chu, J.J.H. (2014). Antiviral activity of lanatoside C against dengue virus infection. Antiviral Research *111*, 93–99.

Chi, H., Zheng, X., Wang, X., Wang, C., Wang, H., Gai, W., Perlman, S., Yang, S., Zhao, J., and Xia, X. (2017). DNA vaccine encoding Middle East respiratory syndrome coronavirus S1 protein induces protective immune responses in mice. Vaccine *35*, 2069–2075.

Cho, H., Excler, J.-L., Kim, J.H., and Yoon, I.-K. (2018). Development of Middle East Respiratory Syndrome Coronavirus vaccines – advances and challenges. Human Vaccines & Immunotherapeutics *14*, 304–313.

Choudhury, B., Dastjerdi, A., Doyle, N., Frossard, J.-P., and Steinbach, F. (2016). From the field to the lab — An European view on the global spread of PEDV. Virus Research *226*, 40–49.

Chu, D.K.W., Poon, L.L.M., Gomaa, M.M., Shehata, M.M., Perera, R.A.P.M., Abu Zeid, D., El Rifay, A.S., Siu, L.Y., Guan, Y., Webby, R.J., et al. (2014). MERS coronaviruses in dromedary camels, Egypt. Emerging Infect. Dis. *20*, 1049–1053.

Chu, V.C., McElroy, L.J., Chu, V., Bauman, B.E., and Whittaker, G.R. (2006). The avian coronavirus infectious bronchitis virus undergoes direct low-pH-dependent fusion activation during entry into host cells. J. Virol. *80*, 3180–3188.

Ciczora, Y., Callens, N., Montpellier, C., Bartosch, B., Cosset, F.-L., Op de Beeck, A., and Dubuisson, J. (2005). Contribution of the charged residues of hepatitis C virus glycoprotein E2 transmembrane domain to the functions of the E1E2 heterodimer. J. Gen. Virol. *86*, 2793–2798.

Ciczora, Y., Callens, N., Seron, K., Rouille, Y., and Dubuisson, J. (2010). Identification of a dominant endoplasmic reticulum-retention signal in yellow fever virus pre-membrane protein. Journal of General Virology *91*, 404–414.

Circella, E., Camarda, A., Martella, V., Bruni, G., Lavazza, A., and Buonavoglia, C. (2007). Coronavirus associated with an enteric syndrome on a quail farm. Avian Pathol. *36*, 251–258.

Cohen, J.R., Lin, L.D., and Machamer, C.E. (2011). Identification of a Golgi complex-targeting signal in the cytoplasmic tail of the severe acute respiratory syndrome coronavirus envelope protein. J. Virol. *85*, 5794–5803.

Comar, C.E., Goldstein, S.A., Li, Y., Yount, B., Baric, R.S., and Weiss, S.R. (2019). Antagonism of dsRNA-Induced Innate Immune Pathways by NS4a and NS4b Accessory Proteins during MERS Coronavirus Infection. MBio *10*.

Cook, J.K. (1971). Recovery of infectious bronchitis virus from eggs and chicks produced by experimentally inoculated hens. J. Comp. Pathol. *81*, 203–211.

Cook, J.K., and Garside, J.S. (1967). A study of the infectious bronchitis status of a group of chicks hatched from infectious bronchitis infected hens. Res. Vet. Sci. *8*, 74–82.

Cook, J.K.A., and Mockett, A.P.A. (1995). Epidemiology of Infectious Bronchitis Virus. In The Coronaviridae, S.G. Siddell, ed. (Boston, MA: Springer US), pp. 317–335.

Cook, J.K.A., Jackwood, M., and Jones, R.C. (2012). The long view: 40 years of infectious bronchitis research. Avian Pathology *41*, 239–250.

Corman, V.M., Ithete, N.L., Richards, L.R., Schoeman, M.C., Preiser, W., Drosten, C., and Drexler, J.F. (2014). Rooting the Phylogenetic Tree of Middle East Respiratory Syndrome Coronavirus by Characterization of a Conspecific Virus from an African Bat. J Virol *88*, 11297–11303.

Corman, V.M., Baldwin, H.J., Tateno, A.F., Zerbinati, R.M., Annan, A., Owusu, M., Nkrumah, E.E., Maganga, G.D., Oppong, S., Adu-Sarkodie, Y., et al. (2015). Evidence for an Ancestral Association of Human Coronavirus 229E with Bats. J. Virol. *89*, 11858–11870.

Corman, V.M., Eckerle, I., Memish, Z.A., Liljander, A.M., Dijkman, R., Jonsdottir, H., Ngeiywa, K.J.Z.J., Kamau, E., Younan, M., Masri, M.A., et al. (2016). Link of a ubiquitous human coronavirus to dromedary camels. PNAS *113*, 9864–9869.

Corman, V.M., Muth, D., Niemeyer, D., and Drosten, C. (2018a). Hosts and Sources of Endemic Human Coronaviruses. In Advances in Virus Research, (Elsevier), pp. 163–188.

Corman, V.M., Muth, D., Niemeyer, D., and Drosten, C. (2018b). Hosts and Sources of Endemic Human Coronaviruses. Adv. Virus Res. *100*, 163–188.

Corse, E., and Machamer, C.E. (2000). Infectious Bronchitis Virus E Protein Is Targeted to the Golgi Complex and Directs Release of Virus-Like Particles. J Virol *74*, 4319–4326.

Corse, E., and Machamer, C.E. (2003). The cytoplasmic tails of infectious bronchitis virus E and M proteins mediate their interaction. Virology *312*, 25–34.

Cosson, P., and Letourneur, F. (1994). Coatomer interaction with di-lysine endoplasmic reticulum retention motifs. Science *263*, 1629–1631.

Cosson, P., Perrin, J., and Bonifacino, J.S. (2013). Anchors Aweigh: Protein Traffic Mediated by Transmembrane Domains. Trends Cell Biol *23*, 511–517.

Cottam, E.M., Maier, H.J., Manifava, M., Vaux, L.C., Chandra-Schoenfelder, P., Gerner, W., Britton, P., Ktistakis, N.T., and Wileman, T. (2011). Coronavirus nsp6 proteins generate autophagosomes from the endoplasmic reticulum via an omegasome intermediate. Autophagy *7*, 1335–1347.

Cottam, E.M., Whelband, M.C., and Wileman, T. (2014). Coronavirus NSP6 restricts autophagosome expansion. Autophagy *10*, 1426–1441.

Cox, E., Pensaert, M.B., and Callebaut, P. (1993). Intestinal protection against challenge with transmissible gastroenteritis virus of pigs immune after infection with the porcine respiratory coronavirus. Vaccine *11*, 267–272.

Crossley, B., Mock, R.E., Callison, S.A., and Hietala, S.K. (2012). Identification and characterization of a novel alpaca respiratory coronavirus most closely related to the human coronavirus 229E. Viruses *4*, 3689–3700.

Crossley, B.M., Barr, B.C., Magdesian, K.G., Ing, M., Mora, D., Jensen, D., Loretti, A.P., McConnell, T., and Mock, R. (2010). Identification of a novel coronavirus possibly associated with acute respiratory syndrome in alpacas (Vicugna pacos) in California, 2007. J. Vet. Diagn. Invest. *22*, 94–97.

Cui, J., Li, F., and Shi, Z.-L. (2019). Origin and evolution of pathogenic coronaviruses. Nature Reviews Microbiology *17*, 181–192.

David-Ferreira, J.F., and Manaker, R.A. (1965). AN ELECTRON MICROSCOPE STUDY OF THE DEVELOPMENT OF A MOUSE HEPATITIS VIRUS IN TISSUE CULTURE CELLS. J Cell Biol *24*, 57–78.

De Haan, C.A., Vennema, H., and Rottier, P.J. (2000). Assembly of the coronavirus envelope: homotypic interactions between the M proteins. Journal of Virology *74*, 4967–4978.

Dea, S., and Tijssen, P. (1989). Detection of turkey enteric coronavirus by enzyme-linked immunosorbent assay and differentiation from other coronaviruses. Am. J. Vet. Res. *50*, 226–231.

DeDiego, M.L., Álvarez, E., Almazán, F., Rejas, M.T., Lamirande, E., Roberts, A., Shieh, W.-J., Zaki, S.R., Subbarao, K., and Enjuanes, L. (2007). A Severe Acute Respiratory Syndrome Coronavirus That Lacks the E Gene Is Attenuated In Vitro and In Vivo. J Virol *81*, 1701–1713.

DeDiego, M.L., Nieto-Torres, J.L., Jiménez-Guardeño, J.M., Regla-Nava, J.A., Álvarez, E., Oliveros, J.C., Zhao, J., Fett, C., Perlman, S., and Enjuanes, L. (2011). Severe Acute Respiratory Syndrome Coronavirus Envelope Protein Regulates Cell Stress Response and Apoptosis. PLoS Pathog *7*.

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.

Desmarets, L.M.B., Vermeulen, B.L., Theuns, S., Conceição-Neto, N., Zeller, M., Roukaerts, I.D.M., Acar, D.D., Olyslaegers, D.A.J., Van Ranst, M., Matthijnssens, J., et al. (2016). Experimental feline enteric coronavirus infection reveals an aberrant infection pattern and shedding of mutants with impaired infectivity in enterocyte cultures. Sci Rep *6*.

Doherty, G.J., and McMahon, H.T. (2009). Mechanisms of Endocytosis. Annu. Rev. Biochem. *78*, 857–902.

Domanska-Blicharz, K., Jacukowicz, A., Lisowska, A., Wyrostek, K., and Minta, Z. (2014). Detection and molecular characterization of infectious bronchitis-like viruses in wild bird populations. Avian Pathology *43*, 406–413.

Dong, B.Q., Liu, W., Fan, X.H., Vijaykrishna, D., Tang, X.C., Gao, F., Li, L.F., Li, G.J., Zhang, J.X., Yang, L.Q., et al. (2007). Detection of a novel and highly divergent coronavirus from asian leopard cats and Chinese ferret badgers in Southern China. J. Virol. *81*, 6920–6926.

Donnelly, C.A., Ghani, A.C., Leung, G.M., Hedley, A.J., Fraser, C., Riley, S., Abu-Raddad, L.J., Ho, L.-M., Thach, T.-Q., Chau, P., et al. (2003). Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. The Lancet *361*, 1761–1766.

van Doremalen, N., Bushmaker, T., and Munster, V. (2013). Stability of Middle East respiratory syndrome coronavirus (MERS-CoV) under different environmental conditions. Eurosurveillance *18*, 20590.

Doyle, L.P., and Hutchings, L.M. (1946). A transmissible gastroenteritis in pigs. Journal of the American Veterinary Medical Association *108*, 257–259.

Doyle, N., Neuman, B.W., Simpson, J., Hawes, P.C., Mantell, J., Verkade, P., Alrashedi, H., and Maier, H.J. (2018). Infectious Bronchitis Virus Nonstructural Protein 4 Alone Induces Membrane Pairing. Viruses *10*.

Drechsler, Y., Alcaraz, A., Bossong, F.J., Collisson, E.W., and Diniz, P.P.V.P. (2011). Feline Coronavirus in Multicat Environments. Veterinary Clinics of North America: Small Animal Practice *41*, 1133–1169.

Drosten, C., Günther, S., Preiser, W., van der Werf, S., Brodt, H.-R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R.A.M., et al. (2003). Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. *348*, 1967–1976.

Du, L., Yang, Y., Zhou, Y., Lu, L., Li, F., and Jiang, S. (2017). MERS-CoV spike protein: a key target for antivirals. Expert Opin. Ther. Targets *21*, 131–143.

Dubuisson, J. (2007). Hepatitis C virus proteins. World J Gastroenterol *13*, 2406–2415.

Dunn, K.W., Kamocka, M.M., and McDonald, J.H. (2011). A practical guide to evaluating colocalization in biological microscopy. American Journal of Physiology-Cell Physiology *300*, C723–C742.

Earnest, J.T., Hantak, M.P., Park, J.-E., and Gallagher, T. (2015). Coronavirus and influenza virus proteolytic priming takes place in tetraspanin-enriched membrane microdomains. J. Virol. *89*, 6093–6104.

Earnest, J.T., Hantak, M.P., Li, K., McCray, P.B., Perlman, S., and Gallagher, T. (2017). The tetraspanin CD9 facilitates MERS-coronavirus entry by scaffolding host cell receptors and proteases. PLoS Pathog *13*.

Eckerle, I., Corman, V.M., Müller, M.A., Lenk, M., Ulrich, R.G., and Drosten, C. (2014). Replicative Capacity of MERS Coronavirus in Livestock Cell Lines. Emerg Infect Dis *20*, 276–279.

Erles, K., Toomey, C., Brooks, H.W., and Brownlie, J. (2003). Detection of a group 2 coronavirus in dogs with canine infectious respiratory disease. Virology *310*, 216–223.

Escors, D., Ortego, J., Laude, H., and Enjuanes, L. (2001). The Membrane M Protein Carboxy Terminus Binds to Transmissible Gastroenteritis Coronavirus Core and Contributes to Core Stability. Journal of Virology *75*, 1312–1324.

Escriou, N., Callendret, B., Lorin, V., Combredet, C., Marianneau, P., Février, M., and Tangy, F. (2014). Protection from SARS coronavirus conferred by live measles vaccine expressing the spike glycoprotein. Virology *452–453*, 32–41.

Fehr, A.R., and Perlman, S. (2015). Coronaviruses: An Overview of Their Replication and Pathogenesis. Methods Mol Biol *1282*, 1–23.

Ferguson, N.M., and Kerkhove, M.D.V. (2014). Identification of MERS-CoV in dromedary camels. The Lancet Infectious Diseases *14*, 93–94.

Fischer, F., Stegen, C.F., Masters, P.S., and Samsonoff, W.A. (1998). Analysis of constructed E gene mutants of mouse hepatitis virus confirms a pivotal role for E protein in coronavirus assembly. Journal of Virology *72*, 7885–7894.

Fiscus, S.A., and Teramoto, Y.A. (1987). Antigenic comparison of feline coronavirus isolates: evidence for markedly different peplomer glycoproteins. J Virol *61*, 2607–2613.

Foley, J.E., Poland, A., Carlson, J., and Pedersen, N.C. (1997). Risk factors for feline infectious peritonitis among cats in multiple-cat environments with endemic feline enteric coronavirus. J. Am. Vet. Med. Assoc. *210*, 1313–1318.

Fouchier, R.A.M., Hartwig, N.G., Bestebroer, T.M., Niemeyer, B., Jong, J.C. de, Simon, J.H., and Osterhaus, A.D.M.E. (2004). A previously undescribed coronavirus associated with respiratory disease in humans. PNAS *101*, 6212–6216.

Freundt, E.C., Yu, L., Goldsmith, C.S., Welsh, S., Cheng, A., Yount, B., Liu, W., Frieman, M.B., Buchholz, U.J., Screaton, G.R., et al. (2010). The Open Reading Frame 3a Protein of Severe Acute Respiratory Syndrome-Associated Coronavirus Promotes Membrane Rearrangement and Cell Death. J Virol *84*, 1097–1109.

Gadila, S.K.G., and Kim, K. (2016). Cargo trafficking from the trans-Golgi network towards the endosome. Biology of the Cell *108*, 205–218.

Gallon, M., and Cullen, P.J. (2015). Retromer and sorting nexins in endosomal sorting. Biochemical Society Transactions *43*, 33–47.

Gao, C., Cai, Y., Wang, Y., Kang, B.-H., Aniento, F., Robinson, D.G., and Jiang, L. (2014). Retention mechanisms for ER and Golgi membrane proteins. Trends in Plant Science *19*, 508–515.

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.

Gelhaus, S., Thaa, B., Eschke, K., Veit, M., and Schwegmann-Weßels, C. (2014). Palmitoylation of the Alphacoronavirus TGEV spike protein S is essential for incorporation into virus-like particles but dispensable for S-M interaction. Virology *464–465*, 397–405.

Gerber, J.D., Ingersoll, J.D., Gast, A.M., Christianson, K.K., Selzer, N.L., Landon, R.M., Pfeiffer, N.E., Sharpee, R.L., and Beckenhauer, W.H. (1990). Protection against feline infectious peritonitis by intranasal inoculation of a temperature-sensitive FIPV vaccine. Vaccine *8*, 536–542.

Geva, Y., and Schuldiner, M. (2014). The Back and Forth of Cargo Exit from the Endoplasmic Reticulum. Current Biology *24*, R130–R136.

Gledhill, A.W., and Andrewes, C.H. (1951). A Hepatitis Virus of Mice. Br J Exp Pathol *32*, 559–568.

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.

Godeke, G.J., de Haan, C.A., Rossen, J.W., Vennema, H., and Rottier, P.J. (2000). Assembly of spikes into coronavirus particles is mediated by the carboxy-terminal domain of the spike protein. J. Virol. *74*, 1566–1571.

Goldberger, Z.D., and Goldberger, A.L. (2012). Therapeutic ranges of serum digoxin concentrations in patients with heart failure. Am. J. Cardiol. *109*, 1818–1821.

Goldstein, S.A., Thornbrough, J.M., Zhang, R., Jha, B.K., Li, Y., Elliott, R., Quiroz-Figueroa, K., Chen, A.I., Silverman, R.H., and Weiss, S.R. (2017). Lineage A Betacoronavirus NS2 Proteins and the Homologous Torovirus Berne pp1a Carboxy-Terminal Domain Are Phosphodiesterases That Antagonize Activation of RNase L. J Virol *91*.

Gong, L., Li, J., Zhou, Q., Xu, Z., Chen, L., Zhang, Y., Xue, C., Wen, Z., and Cao, Y. (2017). A New Bat-HKU2–like Coronavirus in Swine, China, 2017. Emerg Infect Dis *23*, 1607–1609.

Goueslain, L., Alsaleh, K., Horellou, P., Roingeard, P., Descamps, V., Duverlie, G., Ciczora, Y., Wychowski, C., Dubuisson, J., and Rouillé, Y. (2010). Identification of GBF1 as a cellular factor required for hepatitis C virus RNA replication. J. Virol. *84*, 773–787.

Graham, R.L., and Baric, R.S. (2010). Recombination, Reservoirs, and the Modular Spike: Mechanisms of Coronavirus Cross-Species Transmission. Journal of Virology *84*, 3134–3146.

Graham, R.L., Sims, A.C., Baric, R.S., and Denison, M.R. (2006). The NSP2 Proteins of Mouse Hepatitis Virus and Sars Coronavirus are Dispensable for Viral Replication. In The Nidoviruses, S. Perlman, and K.V. Holmes, eds. (Springer US), pp. 67–72.

Graham, R.L., Deming, D.J., Deming, M.E., Yount, B.L., and Baric, R.S. (2018). Evaluation of a recombination-resistant coronavirus as a broadly applicable, rapidly implementable vaccine platform. Communications Biology *1*, 179.

Greig, A.S., Mitchell, D., Corner, A.H., Bannister, G.L., Meads, E.B., and Julian, R.J. (1962). A Hemagglutinating Virus Producing Encephalomyelitis in Baby Pigs. Can J Comp Med Vet Sci *26*, 49–56.

Grosso, F., Stoilov, P., Lingwood, C., Brown, M., and Cochrane, A. (2016). Suppression of Adenovirus Replication by Cardiotonic Steroids. J. Virol.

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.

Guo, J.-P., Petric, M., Campbell, W., and McGeer, P.L. (2004). SARS corona virus peptides recognized by antibodies in the sera of convalescent cases. Virology *324*, 251–256.

de Haan, C.A.M., and Rottier, P.J.M. (2005). Molecular interactions in the assembly of coronaviruses. Adv Virus Res *64*, 165–230.

de Haan, C. a. M., Haijema, B.J., Schellen, P., Wichgers Schreur, P., te Lintelo, E., Vennema, H., and Rottier, P.J.M. (2008). Cleavage of group 1 coronavirus spike proteins: how furin cleavage is traded off against heparan sulfate binding upon cell culture adaptation. J. Virol. *82*, 6078–6083.

de Haan, C.A., Kuo, L., Masters, P.S., Vennema, H., and Rottier, P.J. (1998a). Coronavirus particle assembly: primary structure requirements of the membrane protein. The Journal of Virology *72*, 6838–6850.

de Haan, C.A., Smeets, M., Vernooij, F., Vennema, H., and Rottier, P.J. (1999a). Mapping of the coronavirus membrane protein domains involved in interaction with the spike protein. The Journal of Virology *73*, 7441–7452.

de Haan, C.A.M., Kuo, L., Masters, P.S., Vennema, H., and Rottier, P.J.M. (1998b). Coronavirus Particle Assembly: Primary Structure Requirements of the Membrane Protein. J Virol *72*, 6838–6850.

de Haan, C.A.M., Smeets, M., Vernooij, F., Vennema, H., and Rottier, P.J.M. (1999b). Mapping of the Coronavirus Membrane Protein Domains Involved in Interaction with the Spike Protein. J Virol *73*, 7441–7452.

de Haan, C.A.M., Vennema, H., and Rottier, P.J.M. (2000). Assembly of the Coronavirus Envelope: Homotypic Interactions between the M Proteins. The Journal of Virology *74*, 4967–4978.

Hagemeijer, M.C., Rottier, P.J.M., and de Haan, C.A.M. (2012). Biogenesis and dynamics of the coronavirus replicative structures. Viruses *4*, 3245–3269.

Hagemeijer, M.C., Monastyrska, I., Griffith, J., van der Sluijs, P., Voortman, J., van Bergen en Henegouwen, P.M., Vonk, A.M., Rottier, P.J.M., Reggiori, F., and de Haan, C.A.M. (2014). Membrane rearrangements mediated by coronavirus nonstructural proteins 3 and 4. Virology *458–459*, 125–135.

Hamre, D., and Procknow, J.J. (1966). A New Virus Isolated from the Human Respiratory Tract. Proceedings of the Society for Experimental Biology and Medicine *121*, 190–193.

Han, D.P., Lohani, M., and Cho, M.W. (2007). Specific asparagine-linked glycosylation sites are critical for DC-SIGN- and L-SIGN-mediated severe acute respiratory syndrome coronavirus entry. J. Virol. *81*, 12029–12039.

Han, M.G., Cheon, D.-S., Zhang, X., and Saif, L.J. (2006). Cross-protection against a human enteric coronavirus and a virulent bovine enteric coronavirus in gnotobiotic calves. J. Virol. *80*, 12350–12356.

Heald-Sargent, T., and Gallagher, T. (2012). Ready, set, fuse! The coronavirus spike protein and acquisition of fusion competence. Viruses *4*, 557–580.

Heeney, J.L., Evermann, J.F., McKeirnan, A.J., Marker-Kraus, L., Roelke, M.E., Bush, M., Wildt, D.E., Meltzer, D.G., Colly, L., and Lukas, J. (1990). Prevalence and implications of feline coronavirus infections of captive and free-ranging cheetahs (Acinonyx jubatus). J. Virol. *64*, 1964–1972.

Hepojoki, S., Lindh, E., Vapalahti, O., and Huovilainen, A. (2017). Prevalence and genetic diversity of coronaviruses in wild birds, Finland. Infection Ecology & Epidemiology *7*, 1408360.

Hertzig, T. (2004). Rapid identification of coronavirus replicase inhibitors using a selectable replicon RNA. Journal of General Virology *85*, 1717–1725.

Hirano, N., Tohyama, K., Taira, H., and Hashikawa, T. (2001). Spread of Hemagglutinating Encephalomyelitis Virus (HEV) in the CNS of Rats Inoculated by Intranasal Route. In The Nidoviruses: Coronaviruses and Arteriviruses, E. Lavi, S.R. Weiss, and S.T. Hingley, eds. (Boston, MA: Springer US), pp. 127–132.

Ho, Y., Lin, P.-H., Liu, C.Y.Y., Lee, S.-P., and Chao, Y.-C. (2004). Assembly of human severe acute respiratory syndrome coronavirus-like particles. Biochem. Biophys. Res. Commun. *318*, 833–838.

Hoek, L.V.D., Pyrc, K., and Berkhout, B. (2006). Human coronavirus NL63, a new respiratory virus. FEMS Microbiology Reviews *30*, 760–773.

van der Hoek, L., Pyrc, K., Jebbink, M.F., Vermeulen-Oost, W., Berkhout, R.J., Wolthers, K.C., Wertheim-van Dillen, P.M., Kaandorp, J., Spaargaren, J., and Berkhout, B. (2004). Identification of a new human coronavirus. Nature Medicine *10*, 368.

Hofmann-Lehmann, R., Fehr, D., Grob, M., Elgizoli, M., Packer, C., Martenson, J.S., O'Brien, S.J., and Lutz, H. (1996). Prevalence of antibodies to feline parvovirus, calicivirus, herpesvirus, coronavirus, and immunodeficiency virus and of feline leukemia virus antigen and the interrelationship of these viral infections in free-ranging lions in east Africa. Clin. Diagn. Lab. Immunol. *3*, 554–562.

Hogue, B.G. (1995). Bovine coronavirus nucleocapsid protein processing and assembly. Adv. Exp. Med. Biol. *380*, 259–263.

Holzworth, J. (1963). Infectious diseases of cats. Cornell Vet *53*, 131–143.

Homberger, F.R. (1997). Enterotropic mouse hepatitis virus. Laboratory Animals *31*, 97–115.

Hora, A.S., Tonietti, P.O., Taniwaki, S.A., Asano, K.M., Maiorka, P., Richtzenhain, L.J., and Brandão, P.E. (2016). Feline Coronavirus 3c Protein: A Candidate for a Virulence Marker? Biomed Res Int *2016*.

Horowitz, J.D., and Mashford, M.L. (1979). Perhexiline maleate in the treatment of severe angina pectoris. Med. J. Aust. *1*, 485–488.

Horzinek, M.C., Lutz, H., and Pedersen, N.C. (1982). Antigenic relationships among homologous structural polypeptides of porcine, feline, and canine coronaviruses. Infect. Immun. *37*, 1148–1155.

Hou, Y., Yue, X., Cai, X., Wang, S., Liu, Y., Yuan, C., Cui, L., Hua, X., and Yang, Z. (2012). Complete Genome of Transmissible Gastroenteritis Virus AYU Strain Isolated in Shanghai, China. Journal of Virology *86*, 11935–11935.

Hou, Y., Meulia, T., Gao, X., Saif, L.J., and Wang, Q. (2019). Deletion of both the Tyrosine-Based Endocytosis Signal and the Endoplasmic Reticulum Retrieval Signal in the Cytoplasmic Tail of Spike Protein Attenuates Porcine Epidemic Diarrhea Virus in Pigs. J. Virol. *93*.

Howland, R.H. (2016). Trifluoperazine: A Sprightly Old Drug. J Psychosoc Nurs Ment Health Serv *54*, 20–22.

Hsieh, L.-E., Huang, W.-P., Tang, D.-J., Wang, Y.-T., Chen, C.-T., and Chueh, L.-L. (2013). 3C protein of feline coronavirus inhibits viral replication independently of the autophagy pathway. Res. Vet. Sci. *95*, 1241–1247.

Hsieh, P.-K., Chang, S.C., Huang, C.-C., Lee, T.-T., Hsiao, C.-W., Kou, Y.-H., Chen, I.-Y., Chang, C.-K., Huang, T.-H., and Chang, M.-F. (2005). Assembly of Severe Acute Respiratory Syndrome Coronavirus RNA Packaging Signal into Virus-Like Particles Is Nucleocapsid Dependent. J Virol *79*, 13848–13855.

Hsieh, S.-C., Tsai, W.-Y., and Wang, W.-K. (2010). The length of and nonhydrophobic residues in the transmembrane domain of dengue virus envelope protein are critical for its retention and assembly in the endoplasmic reticulum. J. Virol. *84*, 4782–4797.

Hsieh, Y.-C., Li, H.-C., Chen, S.-C., and Lo, S.-Y. (2008). Interactions between M protein and other structural proteins of severe, acute respiratory syndrome-associated coronavirus. J Biomed Sci *15*, 707–717.

Hsin, W.-C., Chang, C.-H., Chang, C.-Y., Peng, W.-H., Chien, C.-L., Chang, M.-F., and Chang, S.C. (2018). Nucleocapsid protein-dependent assembly of the RNA packaging signal of Middle East respiratory syndrome coronavirus. J Biomed Sci *25*.

Hu, B., Zeng, L.-P., Yang, X.-L., Ge, X.-Y., Zhang, W., Li, B., Xie, J.-Z., Shen, X.-R., Zhang, Y.-Z., Wang, N., et al. (2017). Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLOS Pathogens *13*, e1006698.

Huang, C., Ito, N., Tseng, C.-T.K., and Makino, S. (2006). Severe acute respiratory syndrome coronavirus 7a accessory protein is a viral structural protein. J. Virol. *80*, 7287–7294.

Huang, C., Lokugamage, K.G., Rozovics, J.M., Narayanan, K., Semler, B.L., and Makino, S. (2011). SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs: viral mRNAs are resistant to nsp1-induced RNA cleavage. PLoS Pathog. *7*, e1002433.

Huang, X., Dong, W., Milewska, A., Golda, A., Qi, Y., Zhu, Q.K., Marasco, W.A., Baric, R.S., Sims, A.C., Pyrc, K., et al. (2015). Human Coronavirus HKU1 Spike Protein Uses O-Acetylated Sialic Acid as an Attachment Receptor Determinant and Employs Hemagglutinin-Esterase Protein as a Receptor-Destroying Enzyme. J. Virol. *89*, 7202–7213.

Huang, Y., Yang, Z., Kong, W., and Nabel, G.J. (2004). Generation of synthetic severe acute respiratory syndrome coronavirus pseudoparticles: implications for assembly and vaccine production. J. Virol. *78*, 12557–12565.

Hulswit, R.J.G., de Haan, C.A.M., and Bosch, B.-J. (2016). Chapter Two - Coronavirus Spike Protein and Tropism Changes. In Advances in Virus Research, J. Ziebuhr, ed. (Academic Press), pp. 29–57.

Huynh, J., Li, S., Yount, B., Smith, A., Sturges, L., Olsen, J.C., Nagel, J., Johnson, J.B., Agnihothram, S., Gates, J.E., et al. (2012). Evidence supporting a zoonotic origin of human coronavirus strain NL63. J. Virol. *86*, 12816–12825.

Ignjatovic, J., and Sapats, S. (2000). Avian infectious bronchitis virus. 16.

Ismail, M.M., Cho, K.O., Ward, L.A., Saif, L.J., and Saif, Y.M. (2001). Experimental bovine coronavirus in turkey poults and young chickens. Avian Dis. *45*, 157–163.

Ito, N., Mossel, E.C., Narayanan, K., Popov, V.L., Huang, C., Inoue, T., Peters, C.J., and Makino, S. (2005). Severe acute respiratory syndrome coronavirus 3a protein is a viral structural protein. J. Virol. *79*, 3182–3186.

Iwata-Yoshikawa, N., Okamura, T., Shimizu, Y., Hasegawa, H., Takeda, M., and Nagata, N. (2019). TMPRSS2 Contributes to Virus Spread and Immunopathology in the Airways of Murine Models after Coronavirus Infection. J. Virol. *93*.

Jacobse-Geels, H.E., and Horzinek, M.C. (1983). Expression of feline infectious peritonitis coronavirus antigens on the surface of feline macrophage-like cells. J. Gen. Virol. *64 (Pt 9)*, 1859–1866.

Janssens, M.M. (1993). Astemizole. A nonsedating antihistamine with fast and sustained activity. Clin Rev Allergy *11*, 35–63.

Jevšnik, M., Uršič, T., Žigon, N., Lusa, L., Krivec, U., and Petrovec, M. (2012). Coronavirus infections in hospitalized pediatric patients with acute respiratory tract disease. BMC Infect Dis *12*, 365.

Jonassen, C.M., Kofstad, T., Larsen, I.-L., Løvland, A., Handeland, K., Follestad, A., and Lillehaug, A. (2005). Molecular identification and characterization of novel coronaviruses infecting graylag geese (Anser anser), feral pigeons (Columbia livia) and mallards (Anas platyrhynchos). J. Gen. Virol. *86*, 1597–1607.

Jones, K.E., Patel, N.G., Levy, M.A., Storeygard, A., Balk, D., Gittleman, J.L., and Daszak, P. (2008). Global trends in emerging infectious diseases. Nature *451*, 990–993.

Jung, K., Hu, H., and Saif, L.J. (2016). Porcine deltacoronavirus infection: Etiology, cell culture for virus isolation and propagation, molecular epidemiology and pathogenesis. Virus Res. *226*, 50–59.

Kaksonen, M., and Roux, A. (2018). Mechanisms of clathrin-mediated endocytosis. Nature Reviews Molecular Cell Biology *19*, 313–326.

Kam, Y.W., Kien, F., Roberts, A., Cheung, Y.C., Lamirande, E.W., Vogel, L., Chu, S.L., Tse, J., Guarner, J., Zaki, S.R., et al. (2007). Antibodies against trimeric S glycoprotein protect hamsters against SARS-CoV challenge despite their capacity to mediate FcgammaRII-dependent entry into B cells in vitro. Vaccine *25*, 729–740.

Kappeler, F., Klopfenstein, D.R.C., Foguet, M., Paccaud, J.-P., and Hauri, H.-P. (1997). The Recycling of ERGIC-53 in the Early Secretory Pathway ERGIC-53 CARRIES A CYTOSOLIC ENDOPLASMIC RETICULUM-EXIT DETERMINANT INTERACTING WITH COPII. J. Biol. Chem. *272*, 31801–31808.

Kazi, L., Lissenberg, A., Watson, R., de Groot, R.J., and Weiss, S.R. (2005). Expression of hemagglutinin esterase protein from recombinant mouse hepatitis virus enhances neurovirulence. J. Virol. *79*, 15064–15073.

Kennedy, M., Kania, S., Stylianides, E., Bertschinger, H., Keet, D., and van Vuuren, M. (2003). Detection of feline coronavirus infection in southern African nondomestic felids. J. Wildl. Dis. *39*, 529–535.

Kim, Y., Liu, H., Galasiti Kankanamalage, A.C., Weerasekara, S., Hua, D.H., Groutas, W.C., Chang, K.-O., and Pedersen, N.C. (2016). Reversal of the Progression of Fatal Coronavirus Infection in Cats by a Broad-Spectrum Coronavirus Protease Inhibitor. PLoS Pathog *12*.

King, B., and Brian, D.A. (1982). Bovine coronavirus structural proteins. J Virol *42*, 700–707.

Kint, J., Dickhout, A., Kutter, J., Maier, H.J., Britton, P., Koumans, J., Pijlman, G.P., Fros, J.J., Wiegertjes, G.F., and Forlenza, M. (2015). Infectious Bronchitis Coronavirus Inhibits STAT1 Signaling and Requires Accessory Proteins for Resistance to Type I Interferon Activity. J. Virol. *89*, 12047–12057.

Kint, J., Langereis, M.A., Maier, H.J., Britton, P., van Kuppeveld, F.J., Koumans, J., Wiegertjes, G.F., and Forlenza, M. (2016). Infectious Bronchitis Coronavirus Limits Interferon Production by Inducing a Host Shutoff That Requires Accessory Protein 5b. J Virol *90*, 7519–7528.

Kipar, A., Meli, M.L., Baptiste, K.E., Bowker, L.J., and Lutz, H. (2010). Sites of feline coronavirus persistence in healthy cats. Journal of General Virology *91*, 1698–1707.

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.

Klumperman, J., Locker, J.K., Meijer, A., Horzinek, M.C., Geuze, H.J., and Rottier, P.J. (1994a). Coronavirus M proteins accumulate in the Golgi complex beyond the site of virion budding. Journal of Virology *68*, 6523–6534.

Klumperman, J., Locker, J.K., Meijer, A., Horzinek, M.C., Geuze, H.J., and Rottier, P.J. (1994b). Coronavirus M proteins accumulate in the Golgi complex beyond the site of virion budding. The Journal of Virology *68*, 6523–6534.

Klumperman, J., Schweizer, A., Clausen, H., Tang, B.L., Hong, W., Oorschot, V., and Hauri, H.P. (1998). The recycling pathway of protein ERGIC-53 and dynamics of the ER-Golgi intermediate compartment. J. Cell. Sci. *111 (Pt 22)*, 3411–3425.

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*.

Ko, J.-H., Seok, H., Cho, S.Y., Ha, Y.E., Baek, J.Y., Kim, S.H., Kim, Y.-J., Park, J.K., Chung, C.R., Kang, E.-S., et al. (2018). Challenges of convalescent plasma infusion therapy in Middle East respiratory coronavirus infection: a single centre experience. Antivir. Ther. (Lond.) *23*, 617–622.

Koetzner, C.A., Kuo, L., Goebel, S.J., Dean, A.B., Parker, M.M., and Masters, P.S. (2010). Accessory Protein 5a Is a Major Antagonist of the Antiviral Action of Interferon against Murine Coronavirus. J Virol *84*, 8262–8274.

Konno, H., Wakabayashi, M., Takanuma, D., Saito, Y., and Akaji, K. (2016). Design and synthesis of a series of serine derivatives as small molecule inhibitors of the SARS coronavirus 3CL protease. Bioorg. Med. Chem. *24*, 1241–1254.

Kopecky-Bromberg, S.A., Martínez-Sobrido, L., Frieman, M., Baric, R.A., and Palese, P. (2007). Severe Acute Respiratory Syndrome Coronavirus Open Reading Frame (ORF) 3b, ORF 6, and Nucleocapsid Proteins Function as Interferon Antagonists. J Virol *81*, 548–557.

Kumar, V., Shin, J.S., Shie, J.-J., Ku, K.B., Kim, C., Go, Y.Y., Huang, K.-F., Kim, M., and Liang, P.-H. (2017). Identification and evaluation of potent Middle East respiratory syndrome coronavirus (MERS-CoV) 3CLPro inhibitors. Antiviral Res. *141*, 101–106.

Kuo, L., and Masters, P.S. (2002). Genetic Evidence for a Structural Interaction between the Carboxy Termini of the Membrane and Nucleocapsid Proteins of Mouse Hepatitis Virus. J Virol *76*, 4987–4999.

Kuo, L., and Masters, P.S. (2003). The small envelope protein E is not essential for murine coronavirus replication. J. Virol. *77*, 4597–4608.

Kuo, L., and Masters, P.S. (2010). Evolved variants of the membrane protein can partially replace the envelope protein in murine coronavirus assembly. J. Virol. *84*, 12872–12885.

Kuo, L., Hurst, K.R., and Masters, P.S. (2007). Exceptional Flexibility in the Sequence Requirements for Coronavirus Small Envelope Protein Function. J Virol *81*, 2249–2262.

Kuo, L., Koetzner, C.A., Hurst, K.R., and Masters, P.S. (2014). Recognition of the murine coronavirus genomic RNA packaging signal depends on the second RNA-binding domain of the nucleocapsid protein. J. Virol. *88*, 4451–4465.

Kuo, L., Hurst-Hess, K.R., Koetzner, C.A., and Masters, P.S. (2016a). Analyses of Coronavirus Assembly Interactions with Interspecies Membrane and Nucleocapsid Protein Chimeras. Journal of Virology *90*, 4357–4368.

Kuo, L., Koetzner, C.A., and Masters, P.S. (2016b). A key role for the carboxy-terminal tail of the murine coronavirus nucleocapsid protein in coordination of genome packaging. Virology *494*, 100–107.

Lai, S.T. (2005). Treatment of severe acute respiratory syndrome. European Journal of Clinical Microbiology & Infectious Diseases *24*, 583–591.

Laird, G.M., Eisele, E.E., Rabi, S.A., Nikolaeva, D., and Siliciano, R.F. (2014). A novel cell-based high-throughput screen for inhibitors of HIV-1 gene expression and budding identifies the cardiac glycosides. J Antimicrob Chemother *69*, 988–994.

Lamirande, E.W., DeDiego, M.L., Roberts, A., Jackson, J.P., Alvarez, E., Sheahan, T., Shieh, W.-J., Zaki, S.R., Baric, R., Enjuanes, L., et al. (2008). A live attenuated severe acute respiratory syndrome coronavirus is immunogenic and efficacious in golden Syrian hamsters. J. Virol. *82*, 7721–7724.

Laude, H., Rasschaert, D., and Huet, J.-C. (1987). Sequence and N-terminal Processing of the Transmembrane Protein E1 of the Coronavirus Transmissible Gastroenteritis Virus. Journal of General Virology *68*, 1687–1693.

LAUDE, H., Van Reeth, K., and Pensaert, M. (1993). Porcine respiratory coronavirus: molecular features and virus-host interactions. Veterinary Research *24*, 125–150.

Laviada, M.D., Videgain, S.P., Moreno, L., Alonso, F., Enjuanes, L., and Escribano, J.M. (1990). Expression of swine transmissible gastroenteritis virus envelope antigens on the surface of infected cells: epitopes externally exposed. Virus Res. *16*, 247–254.

Le Poder, S. (2011). Feline and canine coronaviruses: common genetic and pathobiological features. Adv Virol *2011*, 609465.

Lee, S., and Lee, C. (2014). Outbreak-related porcine epidemic diarrhea virus strains similar to US strains, South Korea, 2013. Emerging Infect. Dis. *20*, 1223–1226.

Lee, S.S., and Wong, N.S. (2015). Probable transmission chains of Middle East respiratory syndrome coronavirus and the multiple generations of secondary infection in South Korea. International Journal of Infectious Diseases *38*, 65–67.

Leutenegger, C.M., Hofmann-Lehmann, R., Riols, C., Liberek, M., Worel, G., Lups, P., Fehr, D., Hartmann, M., Weilenmann, P., and Lutz, H. (1999). VIRAL INFECTIONS IN FREE-LIVING POPULATIONS OF THE EUROPEAN WILDCAT. Journal of Wildlife Diseases *35*, 678–686.

Li, H.S., Kuok, D.I.T., Cheung, M.C., Ng, M.M.T., Ng, K.C., Hui, K.P.Y., Peiris, J.S.M., Chan, M.C.W., and Nicholls, J.M. (2018). Effect of interferon alpha and cyclosporine treatment separately and in combination on Middle East Respiratory Syndrome Coronavirus (MERS-CoV) replication in a human in-vitro and ex-vivo culture model. Antiviral Res. *155*, 89–96.

Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J.H., Wang, H., Crameri, G., Hu, Z., Zhang, H., et al. (2005a). Bats are natural reservoirs of SARS-like coronaviruses. Science *310*, 676–679.

Li, W., Zhang, C., Sui, J., Kuhn, J.H., Moore, M.J., Luo, S., Wong, S.-K., Huang, I.-C., Xu, K., Vasilieva, N., et al. (2005b). Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. EMBO J. *24*, 1634–1643.

Li, W., Hulswit, R.J.G., Widjaja, I., Raj, V.S., McBride, R., Peng, W., Widagdo, W., Tortorici, M.A., Dieren, B. van, Lang, Y., et al. (2017). Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein. PNAS *114*, E8508–E8517.

Li, Y., Surya, W., Claudine, S., and Torres, J. (2014). Structure of a Conserved Golgi Complex-targeting Signal in Coronavirus Envelope Proteins. J Biol Chem *289*, 12535–12549.

Liang, J.Q., Fang, S., Yuan, Q., Huang, M., Chen, R.A., Fung, T.S., and Liu, D.X. (2019). N-Linked glycosylation of the membrane protein ectodomain regulates infectious bronchitis virus-induced ER stress response, apoptosis and pathogenesis. Virology *531*, 48–56.

Liao, Y., Yuan, Q., Torres, J., Tam, J.P., and Liu, D.X. (2006). Biochemical and functional characterization of the membrane association and membrane permeabilizing activity of the severe acute respiratory syndrome coronavirus envelope protein. Virology *349*, 264–275.

Licitra, B.N., Millet, J.K., Regan, A.D., Hamilton, B.S., Rinaldi, V.D., Duhamel, G.E., and Whittaker, G.R. (2013). Mutation in spike protein cleavage site and pathogenesis of feline coronavirus. Emerging Infect. Dis. *19*, 1066–1073.

Lim, K.P. (2001). The Missing Link in Coronavirus Assembly. RETENTION OF THE AVIAN CORONAVIRUS INFECTIOUS BRONCHITIS VIRUS ENVELOPE PROTEIN IN THE PRE-GOLGI COMPARTMENTS AND PHYSICAL INTERACTION BETWEEN THE ENVELOPE AND MEMBRANE PROTEINS. The Journal of Cell Biology *276*, 17515–17523.

Lim, K.P., and Liu, D.X. (2001). The Missing Link in Coronavirus Assembly RETENTION OF THE AVIAN CORONAVIRUS INFECTIOUS BRONCHITIS VIRUS ENVELOPE PROTEIN IN THE PRE-GOLGI COMPARTMENTS AND PHYSICAL INTERACTION BETWEEN THE ENVELOPE AND MEMBRANE PROTEINS. J. Biol. Chem. *276*, 17515–17523.

Lin, C.-M., Saif, L.J., Marthaler, D., and Wang, Q. (2016). Evolution, antigenicity and pathogenicity of global porcine epidemic diarrhea virus strains. Virus Research *226*, 20–39.

Lin, J.-T., Zhang, J.-S., Su, N., Xu, J.-G., Wang, N., Chen, J.-T., Chen, X., Liu, Y.-X., Gao, H., Jia, Y.-P., et al. (2007). Safety and immunogenicity from a phase I trial of inactivated severe acute respiratory syndrome coronavirus vaccine. Antivir. Ther. (Lond.) *12*, 1107–1113.

Lin, L., Han, Y., and Yang, Z.M. (2003). [Clinical observation on 103 patients of severe acute respiratory syndrome treated by integrative traditional Chinese and Western Medicine]. Zhongguo Zhong Xi Yi Jie He Za Zhi *23*, 409–413.

Lipsitch, M., Cohen, T., Cooper, B., Robins, J.M., Ma, S., James, L., Gopalakrishna, G., Chew, S.K., Tan, C.C., Samore, M.H., et al. (2003). Transmission dynamics and control of severe acute respiratory syndrome. Science *300*, 1966–1970.

Liu, D.X., Fung, T.S., Chong, K.K.-L., Shukla, A., and Hilgenfeld, R. (2014). Accessory proteins of SARS-CoV and other coronaviruses. Antiviral Research *109*, 97–109.

Liu, G., Lv, L., Yin, L., Li, X., Luo, D., Liu, K., Xue, C., and Cao, Y. (2013). Assembly and immunogenicity of coronavirus-like particles carrying infectious bronchitis virus M and S proteins. Vaccine *31*, 5524–5530.

Liu, L., Zhao, X., Pierre, S.V., and Askari, A. (2007). Association of PI3K-Akt signaling pathway with digitalis-induced hypertrophy of cardiac myocytes. American Journal of Physiology - Cell Physiology *293*, C1489–C1497.

Liu, W.J., Zhao, M., Liu, K., Xu, K., Wong, G., Tan, W., and Gao, G.F. (2017). T-cell immunity of SARS-CoV: Implications for vaccine development against MERS-CoV. Antiviral Res. *137*, 82–92.

Locker, J.K., Griffiths, G., Horzinek, M.C., and Rottier, P.J. (1992a). O-glycosylation of the coronavirus M protein. Differential localization of sialyltransferases in N- and O-linked glycosylation. J. Biol. Chem. *267*, 14094–14101.

Locker, J.K., Griffiths, G., Horzinek, M.C., and Rottier, P.J. (1992b). O-glycosylation of the coronavirus M protein. Differential localization of sialyltransferases in N- and O-linked glycosylation. J. Biol. Chem. *267*, 14094–14101.

Locker, J.K., Klumperman, J., Oorschot, V., Horzinek, M.C., Geuze, H.J., and Rottier, P.J. (1994a). The cytoplasmic tail of mouse hepatitis virus M protein is essential but not sufficient for its retention in the Golgi complex. J. Biol. Chem. *269*, 28263–28269.

Locker, J.K., Klumperman, J., Oorschot, V., Horzinek, M.C., Geuze, H.J., and Rottier, P.J. (1994b). The cytoplasmic tail of mouse hepatitis virus M protein is essential but not sufficient for its retention in the Golgi complex. J Biol Chem *269*, 28263–28269.

Locker, J.K., Opstelten, D.-J.E., Ericsson, M., Horzinek, M.C., and Rottier, P.J.M. (1995). Oligomerization of a trans-Golgi/ trans-Golgi Network Retained Protein Occurs in the Golgi Complex and May Be Part of Its Retention. J. Biol. Chem. *270*, 8815–8821.

Lontok, E., Corse, E., and Machamer, C.E. (2004). Intracellular Targeting Signals Contribute to Localization of Coronavirus Spike Proteins near the Virus Assembly Site. Journal of Virology *78*, 5913–5922.

López-Lázaro, M., Pastor, N., Azrak, S.S., Ayuso, M.J., Austin, C.A., and Cortés, F. (2005). Digitoxin inhibits the growth of cancer cell lines at concentrations commonly found in cardiac patients. J. Nat. Prod. *68*, 1642–1645.

Losev, E., Reinke, C.A., Jellen, J., Strongin, D.E., Bevis, B.J., and Glick, B.S. (2006). Golgi maturation visualized in living yeast. Nature *441*, 1002–1006.

Loutfy, M.R., Blatt, L.M., Siminovitch, K.A., Ward, S., Wolff, B., Lho, H., Pham, D.H., Deif, H., LaMere, E.A., Chang, M., et al. (2003). Interferon alfacon-1 plus corticosteroids in severe acute respiratory syndrome: a preliminary study. JAMA *290*, 3222–3228.

Lundin, A., Dijkman, R., Bergström, T., Kann, N., Adamiak, B., Hannoun, C., Kindler, E., Jónsdóttir, H.R., Muth, D., Kint, J., et al. (2014). Targeting membrane-bound viral RNA synthesis reveals potent inhibition of diverse coronaviruses including the middle East respiratory syndrome virus. PLoS Pathog. *10*, e1004166.

Machamer, C.E., and Rose, J.K. (1987a). A specific transmembrane domain of a coronavirus E1 glycoprotein is required for its retention in the Golgi region. The Journal of Cell Biology *105*, 1205–1214.

Machamer, C.E., and Rose, J.K. (1987b). A specific transmembrane domain of a coronavirus E1 glycoprotein is required for its retention in the Golgi region. The Journal of Cell Biology *105*, 1205–1214.

Machamer, C.E., Mentone, S.A., Rose, J.K., and Farquhar, M.G. (1990). The E1 glycoprotein of an avian coronavirus is targeted to the cis Golgi complex. Proc Natl Acad Sci USA *87*, 6944–6948.

Machamer, C.E., Grim, M.G., Esquela, A., Chung, S.W., Rolls, M., Ryan, K., and Swift, A.M. (1993a). Retention of a cis Golgi protein requires polar residues on one face of a predicted alpha-helix in the transmembrane domain. Molecular Biology of the Cell *4*, 695–704.

Machamer, C.E., Grim, M.G., Esquela, A., Chung, S.W., Rolls, M., Ryan, K., and Swift, A.M. (1993b). Retention of a cis Golgi protein requires polar residues on one face of a predicted alpha-helix in the transmembrane domain. Molecular Biology of the Cell *4*, 695–704.

Mackay, I.M., and Arden, K.E. (2015). Middle East respiratory syndrome: An emerging coronavirus infection tracked by the crowd. Virus Research *202*, 60–88.

MacLeod-Glover, N., Mink, M., Yarema, M., and Chuang, R. (2016). Digoxin toxicity. Can Fam Physician *62*, 223–228.

Madu, I.G., Chu, V.C., Lee, H., Regan, A.D., Bauman, B.E., and Whittaker, G.R. (2007). Heparan sulfate is a selective attachment factor for the avian coronavirus infectious bronchitis virus Beaudette. Avian Dis. *51*, 45–51.

Maeda, J., Maeda, A., and Makino, S. (1999). Release of coronavirus E protein in membrane vesicles from virus-infected cells and E protein-expressing cells. Virology *263*, 265–272.

Maier, H.J., Hawes, P.C., Keep, S.M., and Britton, P. (2014). Spherules and IBV. Bioengineered *5*, 288–292.

Manaker, R.A., Piczak, C.V., Miller, A.A., and Stanton, M.F. (1961). A Hepatitis Virus Complicating Studies With Mouse Leukemia. J Natl Cancer Inst *27*, 29–51.

Martelli, P., Lavazza, A., Nigrelli, A.D., Merialdi, G., Alborali, L.G., and Pensaert, M.B. (2008). Epidemic of diarrhoea caused by porcine epidemic diarrhoea virus in Italy. Vet. Rec. *162*, 307–310.

Masters, P.S. (2006a). The Molecular Biology of Coronaviruses. In Advances in Virus Research, (Elsevier), pp. 193–292.

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

Matsuura-Tokita, K., Takeuchi, M., Ichihara, A., Mikuriya, K., and Nakano, A. (2006). Live imaging of yeast Golgi cisternal maturation. Nature *441*, 1007–1010.

Matsuyama, S., and Taguchi, F. (2002). Receptor-induced conformational changes of murine coronavirus spike protein. J. Virol. *76*, 11819–11826.

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.

Matsuyama, S., Nagata, N., Shirato, K., Kawase, M., Takeda, M., and Taguchi, F. (2010). Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by the Transmembrane Protease TMPRSS2. J Virol *84*, 12658–12664.

Maxfield, F.R., and McGraw, T.E. (2004). Endocytic recycling. Nature Reviews Molecular Cell Biology *5*, 121.

Mayer, T., Tamura, T., Falk, M., and Niemann, H. (1988). Membrane integration and intracellular transport of the coronavirus glycoprotein E1, a class III membrane glycoprotein. J. Biol. Chem. *263*, 14956–14963.

McBride, C.E., and Machamer, C.E. (2010). A Single Tyrosine in the Severe Acute Respiratory Syndrome Coronavirus Membrane Protein Cytoplasmic Tail Is Important for Efficient Interaction with Spike Protein. The Journal of Virology *84*, 1891–1901.

McBride, C.E., Li, J., and Machamer, C.E. (2007). The Cytoplasmic Tail of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein Contains a Novel Endoplasmic Reticulum

Retrieval Signal That Binds COPI and Promotes Interaction with Membrane Protein. J. Virol. *81*, 2418–2428.

McBride, R., van Zyl, M., and Fielding, B.C. (2014). The Coronavirus Nucleocapsid Is a Multifunctional Protein. Viruses *6*, 2991–3018.

McIntosh, K., Dees, J.H., Becker, W.B., Kapikian, A.Z., and Chanock, R.M. (1967). Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. PNAS *57*, 933–940.

Mellman, I., and Warren, G. (2000). The Road Taken: Past and Future Foundations of Membrane Traffic. Cell *100*, 99–112.

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

Menachery, V.D., Mitchell, H.D., Cockrell, A.S., Gralinski, L.E., Yount, B.L., Graham, R.L., McAnarney, E.T., Douglas, M.G., Scobey, T., Beall, A., et al. (2017b). MERS-CoV Accessory ORFs Play Key Role for Infection and Pathogenesis. MBio *8*.

Mikosch, M., Hurst, A.C., Hertel, B., and Homann, U. (2006). Diacidic Motif Is Required for Efficient Transport of the K+ Channel KAT1 to the Plasma Membrane. Plant Physiol. *142*, 923–930.

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. (2018). Physiological and molecular triggers for SARS-CoV membrane fusion and entry into host cells. Virology *517*, 3–8.

Min, C.-K., Cheon, S., Ha, N.-Y., Sohn, K.M., Kim, Y., Aigerim, A., Shin, H.M., Choi, J.-Y., Inn, K.-S., Kim, J.-H., et al. (2016). Comparative and kinetic analysis of viral shedding and immunological responses in MERS patients representing a broad spectrum of disease severity. Scientific Reports *6*, 25359.

Mora-Díaz, J.C., Piñeyro, P.E., Houston, E., Zimmerman, J., and Giménez-Lirola, L.G. (2019). Porcine Hemagglutinating Encephalomyelitis Virus: A Review. Front Vet Sci *6*.

Mounir, S., and Talbot, P.J. (1992). Sequence analysis of the membrane protein gene of human coronavirus OC43 and evidence for O-glycosylation. J. Gen. Virol. *73 (Pt 10)*, 2731–2736.

Müller, M.A., Corman, V.M., Jores, J., Meyer, B., Younan, M., Liljander, A., Bosch, B.-J., Lattwein, E., Hilali, M., Musa, B.E., et al. (2014). MERS coronavirus neutralizing antibodies in camels, Eastern Africa, 1983-1997. Emerging Infect. Dis. *20*, 2093–2095.

Munson, L., Marker, L., Dubovi, E., Spencer, J.A., Evermann, J.F., and O'Brien, S.J. (2004). Serosurvey of viral infections in free-ranging Namibian cheetahs (Acinonyx jubatus). J. Wildl. Dis. *40*, 23–31.

Munster, V.J., Wells, D., Lambe, T., Wright, D., Fischer, R.J., Bushmaker, T., Saturday, G., van Doremalen, N., Gilbert, S.C., de Wit, E., et al. (2017). Protective efficacy of a novel simian adenovirus vaccine against lethal MERS-CoV challenge in a transgenic human DPP4 mouse model. NPJ Vaccines *2*, 28.

Muradrasoli, S., Bálint, Á., Wahlgren, J., Waldenström, J., Belák, S., Blomberg, J., and Olsen, B. (2010). Prevalence and Phylogeny of Coronaviruses in Wild Birds from the Bering Strait Area (Beringia). PLoS One *5*.

Muramatsu, T., Kim, Y.-T., Nishii, W., Terada, T., Shirouzu, M., and Yokoyama, S. (2013). Autoprocessing mechanism of severe acute respiratory syndrome coronavirus 3C-like protease (SARS-CoV 3CLpro) from its polyproteins. FEBS J. *280*, 2002–2013.

Muthumani, K., Falzarano, D., Reuschel, E.L., Tingey, C., Flingai, S., Villarreal, D.O., Wise, M., Patel, A., Izmirly, A., Aljuaid, A., et al. (2015). A synthetic consensus anti-spike protein DNA vaccine induces protective immunity against Middle East respiratory syndrome coronavirus in nonhuman primates. Sci Transl Med *7*, 301ra132.

Nal, B. (2005). Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. Journal of General Virology *86*, 1423–1434.

Nal, B., Chan, C., Kien, F., Siu, L., Tse, J., Chu, K., Kam, J., Staropoli, I., Crescenzo-Chaigne, B., Escriou, N., et al. (2005). Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. Journal of General Virology *86*, 1423–1434.

Narayanan, K., Chen, C.-J., Maeda, J., and Makino, S. (2003). Nucleocapsid-Independent Specific Viral RNA Packaging via Viral Envelope Protein and Viral RNA Signal. J Virol *77*, 2922–2927.

Narayanan, K., Ramirez, S.I., Lokugamage, K.G., and Makino, S. (2015). Coronavirus nonstructural protein 1: common and distinct functions in the regulation of host and viral gene expression. Virus Res *202*, 89–100.

Naslavsky, N., and Caplan, S. (2018). The enigmatic endosome – sorting the ins and outs of endocytic trafficking. J Cell Sci *131*, jcs216499.

Neuman, B.W., and Buchmeier, M.J. (2016). Chapter One - Supramolecular Architecture of the Coronavirus Particle. In Advances in Virus Research, J. Ziebuhr, ed. (Academic Press), pp. 1–27.

Neuman, B.W., Adair, B.D., Yoshioka, C., Quispe, J.D., Orca, G., Kuhn, P., Milligan, R.A., Yeager, M., and Buchmeier, M.J. (2006). Supramolecular Architecture of Severe Acute Respiratory Syndrome Coronavirus Revealed by Electron Cryomicroscopy. J Virol *80*, 7918–7928.

Neuman, B.W., Kiss, G., Kunding, A.H., Bhella, D., Baksh, M.F., Connelly, S., Droese, B., Klaus, J.P., Makino, S., Sawicki, S.G., et al. (2011). A structural analysis of M protein in coronavirus assembly and morphology. Journal of Structural Biology *174*, 11–22.

Neuman, B.W., Chamberlain, P., Bowden, F., and Joseph, J. (2014). Atlas of coronavirus replicase structure. Virus Research *194*, 49–66.

Niemeyer, D., Zillinger, T., Muth, D., Zielecki, F., Horvath, G., Suliman, T., Barchet, W., Weber, F., Drosten, C., and Müller, M.A. (2013). Middle East respiratory syndrome coronavirus accessory protein 4a is a type I interferon antagonist. J. Virol. *87*, 12489–12495.

Nieto-Torres, J., Verdiá-Báguena, C., Castaño-Rodriguez, C., Aguilella, V., and Enjuanes, L. (2015). Relevance of Viroporin Ion Channel Activity on Viral Replication and Pathogenesis. Viruses *7*, 3552–3573.

Nieto-Torres, J.L., DeDiego, M.L., Álvarez, E., Jiménez-Guardeño, J.M., Regla-Nava, J.A., Llorente, M., Kremer, L., Shuo, S., and Enjuanes, L. (2011). Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein. Virology *415*, 69–82.

Nieto-Torres, J.L., DeDiego, M.L., Verdiá-Báguena, C., Jimenez-Guardeño, J.M., Regla-Nava, J.A., Fernandez-Delgado, R., Castaño-Rodriguez, C., Alcaraz, A., Torres, J., Aguilella, V.M., et al. (2014). Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis. PLoS Pathog. *10*, e1004077.

Nishimura, N., and Balch, W.E. (1997). A di-acidic signal required for selective export from the endoplasmic reticulum. Science *277*, 556–558.

Nishimura, N., Bannykh, S., Slabough, S., Matteson, J., Altschuler, Y., Hahn, K., and Balch, W.E. (1999). A Di-acidic (DXE) Code Directs Concentration of Cargo during Export from the Endoplasmic Reticulum. J. Biol. Chem. *274*, 15937–15946.

Nukoolkarn, V., Lee, V.S., Malaisree, M., Aruksakulwong, O., and Hannongbua, S. (2008). Molecular dynamic simulations analysis of ritronavir and lopinavir as SARS-CoV 3CLpro inhibitors. Journal of Theoretical Biology *254*, 861–867.

Oguma, K., OHNO, M., YOSHIDA, M., and SENTSUI, H. (2018). Mutation of the S and 3c genes in genomes of feline coronaviruses. J Vet Med Sci *80*, 1094–1100.

Okba, N.M., Raj, V.S., and Haagmans, B.L. (2017). Middle East respiratory syndrome coronavirus vaccines: current status and novel approaches. Curr Opin Virol *23*, 49–58.

O'Keefe, B.R., Giomarelli, B., Barnard, D.L., Shenoy, S.R., Chan, P.K.S., McMahon, J.B., Palmer, K.E., Barnett, B.W., Meyerholz, D.K., Wohlford-Lenane, C.L., et al. (2010). Broad-spectrum in vitro activity and in vivo efficacy of the antiviral protein griffithsin against emerging viruses of the family Coronaviridae. J. Virol. *84*, 2511–2521.

Olsen, S.J., Chang, H.-L., Cheung, T.Y.-Y., Tang, A.F.-Y., Fisk, T.L., Ooi, S.P.-L., Kuo, H.-W., Jiang, D.D.-S., Chen, K.-T., Lando, J., et al. (2003). Transmission of the severe acute respiratory syndrome on aircraft. N. Engl. J. Med. *349*, 2416–2422.

Oostra, M., de Haan, C.A.M., de Groot, R.J., and Rottier, P.J.M. (2006). Glycosylation of the Severe Acute Respiratory Syndrome Coronavirus Triple-Spanning Membrane Proteins 3a and M. J Virol *80*, 2326–2336.

Opstelten, D.J., Raamsman, M.J., Wolfs, K., Horzinek, M.C., and Rottier, P.J. (1995). Envelope glycoprotein interactions in coronavirus assembly. The Journal of Cell Biology *131*, 339–349.

Ortego, J., Ceriani, J.E., Patiño, C., Plana, J., and Enjuanes, L. (2007). Absence of E protein arrests transmissible gastroenteritis coronavirus maturation in the secretory pathway. Virology *368*, 296–308.

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.

Palade, G. (1975). Intracellular aspects of the process of protein synthesis. Science *189*, 347–358.

Pan, Y., Tian, X., Qin, P., Wang, B., Zhao, P., Yang, Y.-L., Wang, L., Wang, D., Song, Y., Zhang, X., et al. (2017). Discovery of a novel swine enteric alphacoronavirus (SeACoV) in southern China. Veterinary Microbiology *211*, 15–21.

Park, J.-E., Li, K., Barlan, A., Fehr, A.R., Perlman, S., McCray, P.B., and Gallagher, T. (2016). Proteolytic processing of Middle East respiratory syndrome coronavirus spikes expands virus tropism. Proc. Natl. Acad. Sci. U.S.A. *113*, 12262–12267.

Parmar, H.B., Barry, C., and Duncan, R. (2014). Polybasic Trafficking Signal Mediates Golgi Export, ER Retention or ER Export and Retrieval Based on Membrane-Proximity. PLoS ONE *9*.

Paul-Murphy, J., Work, T., Hunter, D., McFie, E., and Fjelline, D. (1994). Serologic survey and serum biochemical reference ranges of the free-ranging mountain lion (Felis concolor) in California. J. Wildl. Dis. *30*, 205–215.

Pedersen, N.C. (2009). A review of feline infectious peritonitis virus infection: 1963–2008. Journal of Feline Medicine and Surgery *11*, 225–258.

Pedersen, N.C. (2014). An update on feline infectious peritonitis: virology and immunopathogenesis. Vet. J. *201*, 123–132.

Pedersen, N.C., Black, J.W., Boyle, J.F., Evermann, J.F., McKeirnan, A.J., and Ott, R.L. (1984). Pathogenic differences between various feline coronavirus isolates. Adv. Exp. Med. Biol. *173*, 365–380.

Pedersen, N.C., Liu, H., Dodd, K.A., and Pesavento, P.A. (2009). Significance of Coronavirus Mutants in Feces and Diseased Tissues of Cats Suffering from Feline Infectious Peritonitis. Viruses *1*, 166–184.

Pedersen, N.C., Liu, H., Scarlett, J., Leutenegger, C.M., Golovko, L., Kennedy, H., and Kamal, F.M. (2012). Feline infectious peritonitis: role of the feline coronavirus 3c gene in intestinal tropism and pathogenicity based upon isolates from resident and adopted shelter cats. Virus Res. *165*, 17–28.

Peiris, J.S.M., Lai, S.T., Poon, L.L.M., Guan, Y., Yam, L.Y.C., Lim, W., Nicholls, J., Yee, W.K.S., Yan, W.W., Cheung, M.T., et al. (2003). Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet *361*, 1319–1325.

Pensaert, M.B., and Martelli, P. (2016). Porcine epidemic diarrhea: A retrospect from Europe and matters of debate. Virus Res. *226*, 1–6.

Pensaert, M., Callebaut, P., and Vergote, J. (1986). Isolation of a porcine respiratory, non-enteric coronavirus related to transmissible gastroenteritis. Veterinary Quarterly *8*, 257–261.

Perlman, S., and Netland, J. (2009). Coronaviruses post-SARS: update on replication and pathogenesis. Nature Reviews Microbiology *7*, 439–450.

Peters, H.L., Jochmans, D., de Wilde, A.H., Posthuma, C.C., Snijder, E.J., Neyts, J., and Seley-Radtke, K.L. (2015). Design, synthesis and evaluation of a series of acyclic fleximer nucleoside analogues with anti-coronavirus activity. Bioorganic & Medicinal Chemistry Letters *25*, 2923–2926.

Pfefferle, S., Oppong, S., Drexler, J.F., Gloza-Rausch, F., Ipsen, A., Seebens, A., Müller, M.A., Annan, A., Vallo, P., Adu-Sarkodie, Y., et al. (2009a). Distant relatives of severe acute respiratory syndrome coronavirus and close relatives of human coronavirus 229E in bats, Ghana. Emerging Infect. Dis. *15*, 1377–1384.

Pfefferle, S., Krähling, V., Ditt, V., Grywna, K., Mühlberger, E., and Drosten, C. (2009b). Reverse genetic characterization of the natural genomic deletion in SARS-Coronavirus strain Frankfurt-1 open reading frame 7b reveals an attenuating function of the 7b protein in-vitro and in-vivo. Virol J *6*, 131.

Ponnambalam, S. (1994). The TGN38 glycoprotein contains two non-overlapping signals that mediate localization to the trans-Golgi network. The Journal of Cell Biology *125*, 253–268.

Popova, R., and Zhang, X. (2002). The Spike but Not the Hemagglutinin/Esterase Protein of Bovine Coronavirus Is Necessary and Sufficient for Viral Infection. Virology *294*, 222–236.

Poteryaev, D., Datta, S., Ackema, K., Zerial, M., and Spang, A. (2010). Identification of the Switch in Early-to-Late Endosome Transition. Cell *141*, 497–508.

Pritchett-Corning, K.R., Cosentino, J., and Clifford, C.B. (2009). Contemporary prevalence of infectious agents in laboratory mice and rats. Lab Anim *43*, 165–173.

Puthenveedu, M.A., and Linstedt, A.D. (2005). Subcompartmentalizing the Golgi apparatus. Current Opinion in Cell Biology *17*, 369–375.

Qiu, M., Shi, Y., Guo, Z., Chen, Z., He, R., Chen, R., Zhou, D., Dai, E., Wang, X., Si, B., et al. (2005). Antibody responses to individual proteins of SARS coronavirus and their neutralization activities. Microbes and Infection *7*, 882–889.

Rappe, J.C.F., de Wilde, A., Di, H., Müller, C., Stalder, H., V'kovski, P., Snijder, E., Brinton, M.A., Ziebuhr, J., Ruggli, N., et al. (2018). Antiviral activity of K22 against members of the order Nidovirales. Virus Res. *246*, 28–34.

Rasschaert, D., Duarte, M., and Laude, H. (1990). Porcine respiratory coronavirus differs from transmissible gastroenteritis virus by a few genomic deletions. J. Gen. Virol. *71 (Pt 11)*, 2599–2607.

Reaves, B., Horn, M., and Banting, G. (1993). TGN38/41 recycles between the cell surface and the TGN: brefeldin A affects its rate of return to the TGN. Mol. Biol. Cell *4*, 93–105.

Regan, A.D., and Whittaker, G.R. (2008). Utilization of DC-SIGN for entry of feline coronaviruses into host cells. J. Virol. *82*, 11992–11996.

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

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

Ritz, S., Egberink, H., and Hartmann, K. (2007). Effect of feline interferon-omega on the survival time and quality of life of cats with feline infectious peritonitis. J. Vet. Intern. Med. *21*, 1193–1197.

Rossignol, J.-F. (2016). Nitazoxanide, a new drug candidate for the treatment of Middle East respiratory syndrome coronavirus. J Infect Public Health *9*, 227–230.

Roth-Cross, J.K., Stokes, H., Chang, G., Chua, M.M., Thiel, V., Weiss, S.R., Gorbalenya, A.E., and Siddell, S.G. (2009). Organ-specific attenuation of murine hepatitis virus strain A59 by replacement of catalytic residues in the putative viral cyclic phosphodiesterase ns2. J. Virol. *83*, 3743–3753.

Rottier, P.J., and Rose, J.K. (1987). Coronavirus E1 glycoprotein expressed from cloned cDNA localizes in the Golgi region. J Virol *61*, 2042–2045.

Rottier, P., Brandenburg, D., Armstrong, J., van der Zeijst, B., and Warren, G. (1984). Assembly in vitro of a spanning membrane protein of the endoplasmic reticulum: the E1 glycoprotein of coronavirus mouse hepatitis virus A59. Proc Natl Acad Sci U S A *81*, 1421–1425.

Rottier, P.J., Locker, J.K., Horzinek, M.C., and Spaan, W.J. (1990). Expression of MHV-A59 M glycoprotein: effects of deletions on membrane integration and intracellular transport. Adv. Exp. Med. Biol. *276*, 127–135.

Ruch, T.R., and Machamer, C.E. (2011). The Hydrophobic Domain of Infectious Bronchitis Virus E Protein Alters the Host Secretory Pathway and Is Important for Release of Infectious Virus. Journal of Virology *85*, 675–685.

Ruch, T.R., and Machamer, C.E. (2012a). The Coronavirus E Protein: Assembly and Beyond. Viruses *4*, 363–382.

Ruch, T.R., and Machamer, C.E. (2012b). The coronavirus E protein: assembly and beyond. Viruses *4*, 363–382.

Sahuc, M.-E., Sahli, R., Rivière, C., Pène, V., Lavie, M., Vandeputte, A., Brodin, P., Rosenberg, A.R., Dubuisson, J., Ksouri, R., et al. (2019). Dehydrojuncusol, a Natural Phenanthrene Compound Extracted from Juncus maritimus, Is a New Inhibitor of Hepatitis C Virus RNA Replication. J. Virol. *93*.

Saif, L.J. (2004). ANIMAL CORONAVIRUSES: LESSONS FOR SARS (National Academies Press (US)).

Sakai, Y., Kawachi, K., Terada, Y., Omori, H., Matsuura, Y., and Kamitani, W. (2017). Two-amino acids change in the nsp4 of SARS coronavirus abolishes viral replication. Virology *510*, 165–174.

Salanueva, I.J., Carrascosa, J.L., and Risco, C. (1999). Structural maturation of the transmissible gastroenteritis coronavirus. The Journal of Virology *73*, 7952–7964.

Sánchez, C.M., Gebauer, F., Suñé, C., Mendez, A., Dopazo, J., and Enjuanes, L. (1992). Genetic evolution and tropism of transmissible gastroenteritis coronaviruses. Virology *190*, 92–105.

Sasseville, A.M., Gélinas, A.M., Sawyer, N., Boutin, M., and Dea, S. (2001). Biological and molecular characteristics of an HEV isolate associated with recent acute outbreaks of encephalomyelitis in Quebec pig farms. Adv. Exp. Med. Biol. *494*, 57–62.

Sato, K., and Nakano, A. (2007). Mechanisms of COPII vesicle formation and protein sorting. FEBS Letters *581*, 2076–2082.

Schaecher, S.R., Mackenzie, J.M., and Pekosz, A. (2007). The ORF7b Protein of Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) Is Expressed in Virus-Infected Cells and Incorporated into SARS-CoV Particles. J Virol *81*, 718–731.

Schalk, A.F., and Hawn, M.C. (1931). An Apparently New Respiratory Disease of Baby Chicks.

Schindewolf, C., and Menachery, V.D. (2019). Middle East Respiratory Syndrome Vaccine Candidates: Cautious Optimism. Viruses *11*.

Schultze, B., Gross, H.J., Brossmer, R., and Herrler, G. (1991a). The S protein of bovine coronavirus is a hemagglutinin recognizing 9-O-acetylated sialic acid as a receptor determinant. J. Virol. *65*, 6232–6237.

Schultze, B., Wahn, K., Klenk, H.D., and Herrler, G. (1991b). Isolated HE-protein from hemagglutinating encephalomyelitis virus and bovine coronavirus has receptor-destroying and receptor-binding activity. Virology *180*, 221–228.

Schultze, B., Cavanagh, D., and Herrler, G. (1992). Neuraminidase treatment of avian infectious bronchitis coronavirus reveals a hemagglutinating activity that is dependent on sialic acid-containing receptors on erythrocytes. Virology *189*, 792–794.

Schultze, B., Krempl, C., Ballesteros, M.L., Shaw, L., Schauer, R., Enjuanes, L., and Herrler, G. (1996). Transmissible gastroenteritis coronavirus, but not the related porcine respiratory coronavirus, has a sialic acid (N-glycolylneuraminic acid) binding activity. J. Virol. *70*, 5634–5637.

Schwegmann-Wessels, C., Al-Falah, M., Escors, D., Wang, Z., Zimmer, G., Deng, H., Enjuanes, L., Naim, H.Y., and Herrler, G. (2004). A Novel Sorting Signal for Intracellular Localization Is Present in the S Protein of a Porcine Coronavirus but Absent from Severe Acute Respiratory Syndrome-associated Coronavirus. Journal of Biological Chemistry *279*, 43661–43666.

Scott, C.C., Vacca, F., and Gruenberg, J. (2014). Endosome maturation, transport and functions. Seminars in Cell & Developmental Biology *31*, 2–10.

Sebesteny, A., and Hill, A.C. (1974). Hepatitis and brain lesions due to mouse hepatitis virus accompanied by wasting in nude mice. Lab. Anim. *8*, 317–326.

Senio, K. (2003). Recent Singapore SARS case a laboratory accident. The Lancet Infectious Diseases *3*, 679.

Shang, J., Zheng, Y., Yang, Y., Liu, C., Geng, Q., Luo, C., Zhang, W., and Li, F. (2018a). Cryo-EM structure of infectious bronchitis coronavirus spike protein reveals structural and functional evolution of coronavirus spike proteins. PLoS Pathog. *14*, e1007009.

Shang, J., Zheng, Y., Yang, Y., Liu, C., Geng, Q., Tai, W., Du, L., Zhou, Y., Zhang, W., and Li, F. (2018b). Cryo-Electron Microscopy Structure of Porcine Deltacoronavirus Spike Protein in the Prefusion State. J. Virol. *92*.

Shaw, M.L. (2009). Henipaviruses employ a multifaceted approach to evade the antiviral interferon response. Viruses *1*, 1190–1203.

Shen, S., Lin, P.-S., Chao, Y.-C., Zhang, A., Yang, X., Lim, S.G., Hong, W., and Tan, Y.-J. (2005). The severe acute respiratory syndrome coronavirus 3a is a novel structural protein. Biochem. Biophys. Res. Commun. *330*, 286–292.

Shi, Z., and Hu, Z. (2008). A review of studies on animal reservoirs of the SARS coronavirus. Virus Research *133*, 74–87.

Shirato, K., Maejima, M., Matsuyama, S., Ujike, M., Miyazaki, A., Takeyama, N., Ikeda, H., and Taguchi, F. (2011). Mutation in the cytoplasmic retrieval signal of porcine epidemic diarrhea virus spike (S) protein is responsible for enhanced fusion activity. Virus Res. *161*, 188–193.

Shirato, K., Kanou, K., Kawase, M., and Matsuyama, S. (2017). Clinical Isolates of Human Coronavirus 229E Bypass the Endosome for Cell Entry. J. Virol. *91*.

Shirato, K., Kawase, M., and Matsuyama, S. (2018). Wild-type human coronaviruses prefer cell-surface TMPRSS2 to endosomal cathepsins for cell entry. Virology *517*, 9–15.

Sid, H., Benachour, K., and Rautenschlein, S. (2015). Co-infection with Multiple Respiratory Pathogens Contributes to Increased Mortality Rates in Algerian Poultry Flocks. Avian Dis. *59*, 440–446.

Siu, K.-L., Yeung, M.L., Kok, K.-H., Yuen, K.-S., Kew, C., Lui, P.-Y., Chan, C.-P., Tse, H., Woo, P.C.Y., Yuen, K.-Y., et al. (2014). Middle east respiratory syndrome coronavirus 4a protein is a double-stranded RNA-binding protein that suppresses PACT-induced activation of RIG-I and MDA5 in the innate antiviral response. J. Virol. *88*, 4866–4876.

Siu, Y.L., Teoh, K.T., Lo, J., Chan, C.M., Kien, F., Escriou, N., Tsao, S.W., Nicholls, J.M., Altmeyer, R., Peiris, J.S.M., et al. (2008a). The M, E, and N Structural Proteins of the Severe Acute Respiratory Syndrome Coronavirus Are Required for Efficient Assembly, Trafficking, and Release of Virus-Like Particles. J Virol *82*, 11318–11330.

Siu, Y.L., Teoh, K.T., Lo, J., Chan, C.M., Kien, F., Escriou, N., Tsao, S.W., Nicholls, J.M., Altmeyer, R., Peiris, J.S.M., et al. (2008b). The M, E, and N Structural Proteins of the Severe Acute Respiratory Syndrome Coronavirus Are Required for Efficient Assembly, Trafficking, and Release of Virus-Like Particles. The Journal of Virology *82*, 11318–11330.

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

Spanakis, N., Tsiodras, S., Haagmans, B.L., Raj, V.S., Pontikis, K., Koutsoukou, A., Koulouris, N.G., Osterhaus, A.D.M.E., Koopmans, M.P.G., and Tsakris, A. (2014). Virological and serological analysis of a recent Middle East respiratory syndrome coronavirus infection case on a triple combination antiviral regimen. International Journal of Antimicrobial Agents *44*, 528–532.

Stadler, K., Roberts, A., Becker, S., Vogel, L., Eickmann, M., Kolesnikova, L., Klenk, H.-D., Murphy, B., Rappuoli, R., Abrignani, S., et al. (2005). SARS vaccine protective in mice. Emerging Infect. Dis. *11*, 1312–1314.

Stein, R.A. (2011). Super-spreaders in infectious diseases. International Journal of Infectious Diseases *15*, e510–e513.

Stern, D.F., and Sefton, B.M. (1982). Coronavirus proteins: structure and function of the oligosaccharides of the avian infectious bronchitis virus glycoproteins. J. Virol. *44*, 804–812.

Stertz, S., Reichelt, M., Spiegel, M., Kuri, T., Martínez-Sobrido, L., García-Sastre, A., Weber, F., and Kochs, G. (2007). The intracellular sites of early replication and budding of SARS-coronavirus. Virology *361*, 304–315.

Stevenson, G.W., Hoang, H., Schwartz, K.J., Burrough, E.R., Sun, D., Madson, D., Cooper, V.L., Pillatzki, A., Gauger, P., Schmitt, B.J., et al. (2013). Emergence of Porcine epidemic diarrhea virus in the United States: clinical signs, lesions, and viral genomic sequences. J. Vet. Diagn. Invest. *25*, 649–654.

Stoilov, P., Lin, C.-H., Damoiseaux, R., Nikolic, J., and Black, D.L. (2008). A high-throughput screening strategy identifies cardiotonic steroids as alternative splicing modulators. Proceedings of the National Academy of Sciences of the United States of America *105*, 11218.

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 in Microbiology *24*, 490–502.

Subissi, L., Posthuma, C.C., Collet, A., Zevenhoven-Dobbe, J.C., Gorbalenya, A.E., Decroly, E., Snijder, E.J., Canard, B., and Imbert, I. (2014). One severe acute respiratory syndrome

coronavirus protein complex integrates processive RNA polymerase and exonuclease activities. Proc. Natl. Acad. Sci. U.S.A. *111*, E3900-3909.

Sugiyama, K., and Amano, Y. (1981). Morphological and biological properties of a new coronavirus associated with diarrhea in infant mice. Archives of Virology *67*, 241–251.

Sun et al., R.S. et (2012). Outbreak of Porcine Epidemic Diarrhea in Suckling Piglets, China - Volume 18, Number 1—January 2012 - Emerging Infectious Diseases journal - CDC.

Surjit, M., Kumar, R., Mishra, R.N., Reddy, M.K., Chow, V.T.K., and Lal, S.K. (2005). The severe acute respiratory syndrome coronavirus nucleocapsid protein is phosphorylated and localizes in the cytoplasm by 14-3-3-mediated translocation. J. Virol. *79*, 11476–11486.

Swift and Machamer (1991). A Golgi retention signal in a membrane-spanning domain of coronavirus E1 protein. J Cell Biol *115*, 19–30.

Taguchi, F., Hirano, N., Kiuchi, Y., and Fujiwara, K. (1976). Difference in response to mouse hepatitis virus among susceptible mouse strains. Jpn. J. Microbiol. *20*, 293–302.

Taharaguchi, S., Soma, T., and Hara, M. (2012). Prevalence of feline coronavirus antibodies in Japanese domestic cats during the past decade. J. Vet. Med. Sci. *74*, 1355–1358.

Tai, W., Zhao, G., Sun, S., Guo, Y., Wang, Y., Tao, X., Tseng, C.-T.K., Li, F., Jiang, S., Du, L., et al. (2016). A recombinant receptor-binding domain of MERS-CoV in trimeric form protects human dipeptidyl peptidase 4 (hDPP4) transgenic mice from MERS-CoV infection. Virology *499*, 375–382.

Tan, Y.-J. (2005). The Severe Acute Respiratory Syndrome (SARS)-coronavirus 3a protein may function as a modulator of the trafficking properties of the spike protein. Virol J *2*, 5.

Tan, Y.-J., Teng, E., Shen, S., Tan, T.H.P., Goh, P.-Y., Fielding, B.C., Ooi, E.-E., Tan, H.-C., Lim, S.G., and Hong, W. (2004). A novel severe acute respiratory syndrome coronavirus protein, U274, is transported to the cell surface and undergoes endocytosis. J. Virol. *78*, 6723–6734.

Tao, Y., Shi, M., Chommanard, C., Queen, K., Zhang, J., Markotter, W., Kuzmin, I.V., Holmes, E.C., and Tong, S. (2017). Surveillance of Bat Coronaviruses in Kenya Identifies Relatives of Human Coronaviruses NL63 and 229E and Their Recombination History. J. Virol. *91*.

Tekes, G., and Thiel, H.-J. (2016). Chapter Six - Feline Coronaviruses: Pathogenesis of Feline Infectious Peritonitis. In Advances in Virus Research, J. Ziebuhr, ed. (Academic Press), pp. 193–218.

Thiel, V., Siddell, S.G., and Herold, J. (1998). Replication and transcription of HCV 229E replicons. p6. Adv. Exp. Med. Biol. *440*, 109–113.

Thiel, V., Herold, J., Schelle, B., and Siddell, S.G. (2001). Viral replicase gene products suffice for coronavirus discontinuous transcription. J. Virol. *75*, 6676–6681.

Thorp, E.B., Boscarino, J.A., Logan, H.L., Goletz, J.T., and Gallagher, T.M. (2006). Palmitoylations on Murine Coronavirus Spike Proteins Are Essential for Virion Assembly and Infectivity. J Virol *80*, 1280–1289.

Tomita, Y., and Kuwata, T. (1978). Suppression of murine leukaemia virus production by ouabain and interferon in mouse cells. J. Gen. Virol. *38*, 223–230.

Tooze, J., Tooze, S., and Warren, G. (1984). Replication of coronavirus MHV-A59 in sac- cells: determination of the first site of budding of progeny virions. Eur. J. Cell Biol. *33*, 281–293.

Torres, J., Parthasarathy, K., Lin, X., Saravanan, R., Kukol, A., and Liu, D.X. (2006). Model of a Putative Pore: The Pentameric α-Helical Bundle of SARS Coronavirus E Protein in Lipid Bilayers. Biophys J *91*, 938–947.

Tsang, K.W., Ho, P.L., Ooi, G.C., Yee, W.K., Wang, T., Chan-Yeung, M., Lam, W.K., Seto, W.H., Yam, L.Y., Cheung, T.M., et al. (2003). A cluster of cases of severe acute respiratory syndrome in Hong Kong. N. Engl. J. Med. *348*, 1977–1985.

Tseng, Y.-T., Wang, S.-M., Huang, K.-J., Lee, A.I.-R., Chiang, C.-C., and Wang, C.-T. (2010a). Self-assembly of Severe Acute Respiratory Syndrome Coronavirus Membrane Protein. J Biol Chem *285*, 12862–12872.

Tseng, Y.-T., Wang, S.-M., Huang, K.-J., Lee, A.I.-R., Chiang, C.-C., and Wang, C.-T. (2010b). Self-assembly of severe acute respiratory syndrome coronavirus membrane protein. Journal of Biological Chemistry *285*, 12862–12872.

Tsunemitsu, H., el-Kanawati, Z.R., Smith, D.R., Reed, H.H., and Saif, L.J. (1995). Isolation of coronaviruses antigenically indistinguishable from bovine coronavirus from wild ruminants with diarrhea. J. Clin. Microbiol. *33*, 3264–3269.

Tu, L., and Banfield, D.K. (2010). Localization of Golgi-resident glycosyltransferases. Cellular and Molecular Life Sciences *67*, 29–41.

Ujike, M., and Taguchi, F. (2015). Incorporation of Spike and Membrane Glycoproteins into Coronavirus Virions. Viruses *7*, 1700–1725.

Ujike, M., Huang, C., Shirato, K., Matsuyama, S., Makino, S., and Taguchi, F. (2012). Two palmitylated cysteine residues of the severe acute respiratory syndrome coronavirus spike (S) protein are critical for S incorporation into virus-like particles, but not for M–S co-localization. Journal of General Virology *93*, 823–828.

Ujike, M., Huang, C., Shirato, K., Makino, S., and Taguchi, F. (2016). The contribution of the cytoplasmic retrieval signal of severe acute respiratory syndrome coronavirus to intracellular accumulation of S proteins and incorporation of S protein into virus-like particles. J. Gen. Virol. *97*, 1853–1864.

Ulasli, M., Verheije, M.H., Haan, C.A.M. de, and Reggiori, F. (2010). Qualitative and quantitative ultrastructural analysis of the membrane rearrangements induced by coronavirus. Cellular Microbiology *12*, 844–861.

Ulferts, R., and Ziebuhr, J. (2011). Nidovirus ribonucleases: Structures and functions in viral replication. RNA Biology *8*, 295–304.

te Velthuis, A.J.W., van den Worm, S.H.E., and Snijder, E.J. (2012). The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. Nucleic Acids Res. *40*, 1737–1747.

Venditti, R., Wilson, C., and De Matteis, M.A. (2014a). Exiting the ER: what we know and what we don't. Trends in Cell Biology *24*, 9–18.

Venditti, R., Wilson, C., and De Matteis, M.A. (2014b). Exiting the ER: what we know and what we don't. Trends Cell Biol *24*, 9–18.

Venkatagopalan, P., Daskalova, S.M., Lopez, L.A., Dolezal, K.A., and Hogue, B.G. (2015). Coronavirus envelope (E) protein remains at the site of assembly. Virology *478*, 75–85.

Vennema, H., Godeke, G.J., Rossen, J.W., Voorhout, W.F., Horzinek, M.C., Opstelten, D.J., and Rottier, P.J. (1996a). Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. EMBO J *15*, 2020–2028.

Vennema, H., Godeke, G.J., Rossen, J.W., Voorhout, W.F., Horzinek, M.C., Opstelten, D.J., and Rottier, P.J. (1996b). Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. EMBO J *15*, 2020–2028.

Vennema, H., Poland, A., Foley, J., and Pedersen, N.C. (1998). Feline Infectious Peritonitis Viruses Arise by Mutation from Endemic Feline Enteric Coronaviruses. Virology *243*, 150–157.

Vijgen, L., Keyaerts, E., Moës, E., Thoelen, I., Wollants, E., Lemey, P., Vandamme, A.-M., and Van Ranst, M. (2005). Complete Genomic Sequence of Human Coronavirus OC43: Molecular Clock Analysis Suggests a Relatively Recent Zoonotic Coronavirus Transmission Event. J Virol *79*, 1595–1604.

Vijgen, L., Keyaerts, E., Lemey, P., Maes, P., Van Reeth, K., Nauwynck, H., Pensaert, M., and Van Ranst, M. (2006). Evolutionary history of the closely related group 2 coronaviruses: porcine hemagglutinating encephalomyelitis virus, bovine coronavirus, and human coronavirus OC43. J. Virol. *80*, 7270–7274.

Vlasak, R., Luytjes, W., Spaan, W., and Palese, P. (1988). Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses. Proc. Natl. Acad. Sci. U.S.A. *85*, 4526–4529.

Voss, D., Kern, A., Traggiai, E., Eickmann, M., Stadler, K., Lanzavecchia, A., and Becker, S. (2006). Characterization of severe acute respiratory syndrome coronavirus membrane protein. FEBS Lett. *580*, 968–973.

Votsmeier, C. (2001a). An acidic sequence of a putative yeast Golgi membrane protein binds COPII and facilitates ER export. The EMBO Journal *20*, 6742–6750.

Votsmeier, C. (2001b). An acidic sequence of a putative yeast Golgi membrane protein binds COPII and facilitates ER export. The EMBO Journal *20*, 6742–6750.

Waguri, S., Dewitte, F., Le Borgne, R., Rouillé, Y., Uchiyama, Y., Dubremetz, J.-F., and Hoflack, B. (2003a). Visualization of TGN to endosome trafficking through fluorescently labeled MPR and AP-1 in living cells. Mol. Biol. Cell *14*, 142–155.

Waguri, S., Dewitte, F., Borgne, R.L., Rouille, Y., Uchiyama, Y., Dubremetz, J.-F., and Hoflack, B. (2003b). Visualization of TGN to Endosome Trafficking through Fluorescently Labeled MPR and AP-1 in Living Cells□V. Molecular Biology of the Cell *14*, 14.

Walls, A.C., Tortorici, M.A., Frenz, B., Snijder, J., Li, W., Rey, F.A., DiMaio, F., Bosch, B.-J., and Veesler, D. (2016a). Glycan shield and epitope masking of a coronavirus spike protein observed by cryo-electron microscopy. Nat. Struct. Mol. Biol. *23*, 899–905.

Walls, A.C., Tortorici, M.A., Bosch, B.-J., Frenz, B., Rottier, P.J.M., DiMaio, F., Rey, F.A., and Veesler, 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. Journal of Infectious Diseases *208*, 1634–1642.

Wang (2014). Porcine Coronavirus HKU15 Detected in 9 US States, 2014 - Volume 20, Number 9—September 2014 - Emerging Infectious Diseases journal - CDC.

Wang, C., Zheng, X., Gai, W., Zhao, Y., Wang, H., Wang, H., Feng, N., Chi, H., Qiu, B., Li, N., et al. (2016). MERS-CoV virus-like particles produced in insect cells induce specific humoural and cellular imminity in rhesus macaques. Oncotarget *5*.

Wang, H., Yuan, X., Sun, Y., Mao, X., Meng, C., Tan, L., Song, C., Qiu, X., Ding, C., and Liao, Y. (2019a). Infectious bronchitis virus entry mainly depends on clathrin mediated endocytosis and requires classical endosomal/lysosomal system. Virology *528*, 118–136.

Wang, M., Yan, M., Xu, H., Liang, W., Kan, B., Zheng, B., Chen, H., Zheng, H., Xu, Y., Zhang, E., et al. (2005). SARS-CoV infection in a restaurant from palm civet. Emerging Infect. Dis. *11*, 1860–1865.

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.

Wang, Q., Vlasova, A.N., Kenney, S.P., and Saif, L.J. (2019b). Emerging and re-emerging coronaviruses in pigs. Current Opinion in Virology *34*, 39–49.

Wang, W., Lin, X.-D., Guo, W.-P., Zhou, R.-H., Wang, M.-R., Wang, C.-Q., Ge, S., Mei, S.-H., Li, M.-H., Shi, M., et al. (2015). Discovery, diversity and evolution of novel coronaviruses sampled from rodents in China. Virology *474*, 19–27.

Ward, J.M. (1970). Morphogenesis of a virus in cats with experimental feline infectious peritonitis. Virology *41*, 191–194.

Weisz, O.A., Swift, A.M., and Machamer, C.E. (1993). Oligomerization of a membrane protein correlates with its retention in the Golgi complex. J. Cell Biol. *122*, 1185–1196.

Weiwei, H., Qinghua, Y., Liqi, Z., Haofei, L., Shanshan, Z., Qi, G., Kongwang, H., and Qian, Y. (2014). Complete genomic sequence of the coronavirus transmissible gastroenteritis virus SHXB isolated in China. Arch Virol *159*, 2295–2302.

Westerbeck, J.W., and Machamer, C.E. (2015). A Coronavirus E Protein Is Present in Two Distinct Pools with Different Effects on Assembly and the Secretory Pathway. Journal of Virology *89*, 9313–9323.

Westerbeck, J.W., and Machamer, C.E. (2019a). The Infectious Bronchitis Coronavirus Envelope Protein Alters Golgi pH to Protect Spike Protein and Promote Release of Infectious Virus. J. Virol.

Westerbeck, J.W., and Machamer, C.E. (2019b). The Infectious Bronchitis Coronavirus Envelope Protein Alters Golgi pH to Protect Spike Protein and Promote Release of Infectious Virus. Journal of Virology.

White, J.M., and Whittaker, G.R. (2016). Fusion of Enveloped Viruses in Endosomes. Traffic *17*, 593–614.

de Wilde, A.H., Jochmans, D., Posthuma, C.C., Zevenhoven-Dobbe, J.C., van Nieuwkoop, S., Bestebroer, T.M., van den Hoogen, B.G., Neyts, J., and Snijder, E.J. (2014). Screening of an FDA-Approved Compound Library Identifies Four Small-Molecule Inhibitors of Middle East

Respiratory Syndrome Coronavirus Replication in Cell Culture. Antimicrob Agents Chemother *58*, 4875–4884.

Winnicka, K., Bielawski, K., and Bielawska, A. CARDIAC GLYCOSIDES IN CANCER RESEARCH AND CANCER THERAPY. 7.

Winter, C., Schwegmann-Wessels, C., Neumann, U., and Herrler, G. (2008). The Spike Protein of Infectious Bronchitis Virus Is Retained Intracellularly by a Tyrosine Motif. J Virol *82*, 2765–2771.

de Wit, E., van Doremalen, N., Falzarano, D., and Munster, V.J. (2016). SARS and MERS: recent insights into emerging coronaviruses. Nature Reviews Microbiology *14*, 523–534.

Wong, R.S.M., and Hui, D.S. (2004). Index Patient and SARS Outbreak in Hong Kong. Emerg Infect Dis *10*, 339–341.

Wong, G., Liu, W., Liu, Y., Zhou, B., Bi, Y., and Gao, G.F. (2015). MERS, SARS, and Ebola: The Role of Super-Spreaders in Infectious Disease. Cell Host & Microbe *18*, 398–401.

Wong, H.H., Fung, T.S., Fang, S., Huang, M., Le, M.T., and Liu, D.X. (2018a). Accessory proteins 8b and 8ab of severe acute respiratory syndrome coronavirus suppress the interferon signaling pathway by mediating ubiquitin-dependent rapid degradation of interferon regulatory factor 3. Virology *515*, 165–175.

Wong, R.W., Balachandran, A., Ostrowski, M.A., and Cochrane, A. (2013). Digoxin Suppresses HIV-1 Replication by Altering Viral RNA Processing. PLoS Pathogens *9*.

Wong, R.W., Lingwood, C.A., Ostrowski, M.A., Cabral, T., and Cochrane, A. (2018b). Cardiac glycoside/aglycones inhibit HIV-1 gene expression by a mechanism requiring MEK1/2-ERK1/2 signaling. Sci Rep *8*, 850.

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. (2005a). 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., Huang, Y., Tsoi, H.-W., Chan, K.-H., and Yuen, K.-Y. (2005b). Phylogenetic and recombination analysis of coronavirus HKU1, a novel coronavirus from patients with pneumonia. Arch Virol *150*, 2299–2311.

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.

Woo, P.C.Y., Lau, S.K.P., Lam, C.S.F., Tsang, A.K.L., Hui, S.-W., Fan, R.Y.Y., Martelli, P., and Yuen, K.-Y. (2014). Discovery of a Novel Bottlenose Dolphin Coronavirus Reveals a Distinct Species of Marine Mammal Coronavirus in Gammacoronavirus. J Virol *88*, 1318–1331.

Xia, L., Yang, Y., Wang, J., Jing, Y., and Yang, Q. (2018). Impact of TGEV infection on the pig small intestine. Virol J *15*.

Xie, Z. (2003). Molecular mechanisms of Na/K-ATPase-mediated signal transduction. Ann. N. Y. Acad. Sci. *986*, 497–503.

Xiong, X., Tortorici, M.A., Snijder, J., Yoshioka, C., Walls, A.C., Li, W., McGuire, A.T., Rey, F.A., Bosch, B.-J., and Veesler, D. (2018). Glycan Shield and Fusion Activation of a Deltacoronavirus Spike Glycoprotein Fine-Tuned for Enteric Infections. J. Virol. *92*.

Xu, K., Zheng, B.-J., Zeng, R., Lu, W., Lin, Y.-P., Xue, L., Li, L., Yang, L.-L., Xu, C., Dai, J., et al. (2009). Severe acute respiratory syndrome coronavirus accessory protein 9b is a virion-associated protein. Virology *388*, 279–285.

Xu, R.-H., He, J.-F., Evans, M.R., Peng, G.-W., Field, H.E., Yu, D.-W., Lee, C.-K., Luo, H.-M., Lin, W.-S., Lin, P., et al. (2004). Epidemiologic Clues to SARS Origin in China. Emerg Infect Dis *10*, 1030–1037.

Xu, Z., Zhang, Y., Gong, L., Huang, L., Lin, Y., Qin, J., Du, Y., Zhou, Q., Xue, C., and Cao, Y. (2019). Isolation and characterization of a highly pathogenic strain of Porcine enteric alphacoronavirus causing watery diarrhoea and high mortality in newborn piglets. Transboundary and Emerging Diseases *66*, 119–130.

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

Yamane, I., Yamazaki, H., Ishizeki, S., Watanabe, Y., Okumura, H., Okubo, M., Kure, K., Hayakawa, Y., Furukawa, M., Ooi, M., et al. (2016). Impact of a porcine epidemic diarrhea outbreak on swine productivity in Japan: a retrospective cohort study. J. Vet. Med. Sci. *78*, 1385–1389.

Yang, C.-W., Chang, H.-Y., Hsu, H.-Y., Lee, Y.-Z., Chang, H.-S., Chen, I.-S., and Lee, S.-J. (2017). Identification of anti-viral activity of the cardenolides, Na+/K+-ATPase inhibitors, against porcine transmissible gastroenteritis virus. Toxicol. Appl. Pharmacol. *332*, 129–137.

Yang, X.-L., Hu, B., Wang, B., Wang, M.-N., Zhang, Q., Zhang, W., Wu, L.-J., Ge, X.-Y., Zhang, Y.-Z., Daszak, P., et al. (2015). Isolation and Characterization of a Novel Bat Coronavirus Closely Related to the Direct Progenitor of Severe Acute Respiratory Syndrome Coronavirus. J. Virol. *90*, 3253–3256.

Yang, Y., Zhang, L., Geng, H., Deng, Y., Huang, B., Guo, Y., Zhao, Z., and Tan, W. (2013a). The structural and accessory proteins M, ORF 4a, ORF 4b, and ORF 5 of Middle East respiratory syndrome coronavirus (MERS-CoV) are potent interferon antagonists. Protein Cell *4*, 951–961.

Yang, Y., Zhang, L., Geng, H., Deng, Y., Huang, B., Guo, Y., Zhao, Z., and Tan, W. (2013b). The structural and accessory proteins M, ORF 4a, ORF 4b, and ORF 5 of Middle East respiratory syndrome coronavirus (MERS-CoV) are potent interferon antagonists. Protein & Cell *4*, 951–961.

Yang, Y., Du, L., Liu, C., Wang, L., Ma, C., Tang, J., Baric, R.S., Jiang, S., and Li, F. (2014). Receptor usage and cell entry of bat coronavirus HKU4 provide insight into bat-to-human transmission of MERS coronavirus. Proc. Natl. Acad. Sci. U.S.A. *111*, 12516–12521.

Yao, Q., Masters, P.S., and Ye, R. (2013). Negatively charged residues in the endodomain are critical for specific assembly of spike protein into murine coronavirus. Virology *442*, 74–81.

Yeh, K.-M., Chiueh, T.-S., Siu, L.K., Lin, J.-C., Chan, P.K.S., Peng, M.-Y., Wan, H.-L., Chen, J.-H., Hu, B.-S., Perng, C.-L., et al. (2005). Experience of using convalescent plasma for severe acute respiratory syndrome among healthcare workers in a Taiwan hospital. J. Antimicrob. Chemother. *56*, 919–922.

You, J.-H., Reed, M.L., and Hiscox, J.A. (2007). Trafficking motifs in the SARS-coronavirus nucleocapsid protein. Biochemical and Biophysical Research Communications *358*, 1015–1020.

Yount, B., Roberts, R.S., Lindesmith, L., and Baric, R.S. (2006). Rewiring the severe acute respiratory syndrome coronavirus (SARS-CoV) transcription circuit: Engineering a recombination-resistant genome. PNAS *103*, 12546–12551.

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.

Zhang, X.W., and Yap, Y.L. (2004). Old drugs as lead compounds for a new disease? Binding analysis of SARS coronavirus main proteinase with HIV, psychotic and parasite drugs. Bioorganic & Medicinal Chemistry *12*, 2517–2521.

Zhang, L., Zhang, Z., Guo, H., and Wang, Y. (2008). Na+/K+-ATPase-mediated signal transduction and Na+/K+-ATPase regulation. Fundamental & Clinical Pharmacology *22*, 615–621.

Zhang, L., Li, L., Yan, L., Ming, Z., Jia, Z., Lou, Z., and Rao, Z. (2018). Structural and Biochemical Characterization of Endoribonuclease Nsp15 Encoded by Middle East Respiratory Syndrome Coronavirus. J. Virol. *92*.

Zhang, Y., Buckles, E., and Whittaker, G.R. (2012). Expression of the C-type lectins DC-SIGN or L-SIGN alters host cell susceptibility for the avian coronavirus, infectious bronchitis virus. Vet. Microbiol. *157*, 285–293.

Zhao, J., Li, K., Wohlford-Lenane, C., Agnihothram, S.S., Fett, C., Zhao, J., Gale, M.J., Baric, R.S., Enjuanes, L., Gallagher, T., et al. (2014). Rapid generation of a mouse model for Middle East respiratory syndrome. PNAS *111*, 4970–4975.

Zhao, J., Perera, R.A.P.M., Kayali, G., Meyerholz, D., Perlman, S., and Peiris, M. (2015). Passive Immunotherapy with Dromedary Immune Serum in an Experimental Animal Model for Middle East Respiratory Syndrome Coronavirus Infection. Journal of Virology *89*, 6117–6120.

Zhou, H., Ferraro, D., Zhao, J., Hussain, S., Shao, J., Trujillo, J., Netland, J., Gallagher, T., and Perlman, S. (2010). The N-Terminal Region of Severe Acute Respiratory Syndrome Coronavirus Protein 6 Induces Membrane Rearrangement and Enhances Virus Replication. J Virol *84*, 3542–3551.

Zhou, P., Fan, H., Lan, T., Yang, X.-L., Shi, W.-F., Zhang, W., Zhu, Y., Zhang, Y.-W., Xie, Q.-M., Mani, S., et al. (2018). Fatal swine acute diarrhoea syndrome caused by an HKU2-related coronavirus of bat origin. Nature *556*, 255–258.

Zhou, Y.L., Ederveen, J., Egberink, H., Pensaert, M., and Horzinek, M.C. (1988). Porcine epidemic diarrhea virus (CV 777) and feline infectious peritonitis virus (FIPV) are antigenically related. Arch. Virol. *102*, 63–71.

Zumla, A., Chan, J.F.W., Azhar, E.I., Hui, D.S.C., and Yuen, K.-Y. (2016). Coronaviruses — drug discovery and therapeutic options. Nature Reviews Drug Discovery *15*, 327–347.

Züst, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B.W., Ziebuhr, J., Szretter, K.J., Baker, S.C., Barchet, W., Diamond, M.S., et al. (2011). Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. Nat. Immunol. *12*, 137–143.

Abstract

Coronaviruses are an important family of emerging pathogens, as shown by the recent emergence of pathogenic SARS-CoV (Severe acute respiratory syndrome coronavirus) and MERS-CoV (Middle-East respiratory syndrome coronavirus) in the last two decades. There are still some knowledge gaps concerning the biology of coronaviruses and we do not have any specific treatment or vaccine.

Among the four structural proteins of the virus, the M protein is considered to be the motor of viral assembly. Expressed alone in cells, M proteins can go beyond the assembly site of the virus (Endoplasmic reticulum-Golgi intermediate compartment, ERGIC) in the secretory pathway. We confirmed MERS-M localization in the Trans-Golgi network (TGN) and identified two signals involved in its intracellular trafficking in its C-terminal domain: a DxE ER export signal, and a KxGxYR TGN retention signal. The DxE signal was already identified on another viral protein, whereas the KxGxYR signal is a new motif. To confirm the role of KxGxYR signal in TGN retention, we constructed chimeras between MERS-M and the protein M of the *Infectious bronchitis virus* (IBV), located in the ERGIC. Our results suggest that for both MERS-M and IBV-M the C-terminal domain is determinant for the specific localization of the proteins.

We also initiated a project on the characterization of the antiviral activity of digoxigenin against HCoV-229E. Our results demonstrated that it inhibits HCoV-229E at a post-entry step, with an IC50 of 250nM, and that it is not toxic at this concentration. Digoxigenin also inhibits hepatitis C virus (HCV) and likely has an effect on an early step of replication of RNA (+) viruses.

Keywords: coronavirus – trafficking motifs – MERS – M protein – antivirals