Université du Québec
Institut National de la Recherche Scientifique
Centre Institut Armand-Frappier

LA PROTÉINE D'ENVELOPPE (E) DU CORONAVIRUS RESPIRATOIRE HUMAIN HCOV-OC43 EST NÉCESSAIRE POUR LA FORMATION DE VIRIONS INFECTIEUX DANS LES CELLULES ÉPITHÉLIALES ET NEURONALES

Par

Jenny Karen Stodola

Mémoire présentée pour l'obtention du grade de
Maître ès sciences (M.Sc.)
Virologie et immunologie

Jury d'évaluation

Président du jury et examinateur interne	Patrick Labonté INRS – Institut Armand-Frappier
Examinateur externe	Roger Lippé Département de pathologie et biologie cellulaire Université de Montréal
Directeur de recherche	Pierre Talbot INRS – Institut Armand-Frappier
Codirecteur de recherche	Marc Desforges INRS – Institut Armand-Frappier

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Pierre Talbot for first accepting me into his lab in September 2013, and allowing me the opportunity to undertake my Master's degree in French and his full support and mentorship during my studies. *J'ai beaucoup apprécié la communauté de l'INRS, mon équipe de laboratoire ainsi que tous les autres membres de l'institut, qui ont tous participé à mon apprentissage de la science en français. C'était une expérience unique et je suis très contente d'être maintenant bilingue et d'avoir bien intégré la culture québécoise. Merci beaucoup.*

I would like to thank the members of the laboratory during my stay: Élodie B, Mathieu MP, Alain LC, Guillaume D, Mathieu D and Julie G for their support, advice and corrections of my French these past two years and the great scientific environment of the lab which helped enormously in my studies and research. Importantly, I would like to thank my co-director, Marc Desforges for his continuous guidance, advice and support for my project; without his direction, I would not have been able to produce an article during my short stay. I would be in remiss to not thank the *Fondation Universitaire Armand-Frappier INRS* and Mr. and Mrs. Bellini for their support in granting me a bursary to support my studies.

I would also like to thank Karine and Constance for being science sounding-board and for their friendship during my French-immersion experience, something I truly appreciated. I would like to acknowledge the support of my parents and brothers during all my studies; always asking me if my viruses were "happy." And finally, a thank you to my love, Rob, who kept me motivated and happy even though we were always a province apart. Thank you.

RÉSUMÉ

La souche OC43 des coronavirus humains (HCoV-OC43) est un virus ubiquitaire infectant le tractus respiratoire détenant aussi des capacités neuroinvasives et neurotropes. Comme tous les coronavirus, HCoV-OC43 contient une petite protéine structurale de l'enveloppe (E) qui, bien que présente en petite quantité dans les virions, est plutôt importante dans la morphogenèse et l'assemblage dû à la présence de plusieurs motifs protéiques et/ou la formation des canaux homo-oligomériques. Selon la souche de coronavirus, la protéine E joue toujours un rôle important et peut même être essentielle pour la production de virions infectieux. L'importance de la protéine E d'HCoV-OC43 dans la production des virus infectieux n'a pas encore été étudiée et les études sur cette protéine sont peu abondantes. En modifiant un clone infectieux contenant le génome complet du virus, nous avons démontré que la protéine E d'HCoV-OC43 joue un rôle critique pour la production des virions infectieux ainsi que pour la propagation du virus dans les cellules épithéliales et neuronales. De plus, nous avons démontré que certains acides aminés dans le domaine transmembranaire potentiel sont impliqués dans la modulation de production de virions infectieux. Finalement, la région C-terminale de la protéine E contient un motif potentiel d'interaction protéine-protéine. En modifiant partiellement ou complètement ce motif potentiel, nous avons déterminé son importance significative dans la production et la propagation de particules infectieuses, indiquant une implication potentielle dans des neuropathologies.

Mots clés : coronavirus, protéine d'eveloppe, assemblage, propagation, virus, PBM, viroporine, neurones

ABSTRACT

The OC43 strain of human coronavirus (HCoV-OC43) is a ubiquitous respiratory tract pathogen for which we have shown neuroinvasive and neurotropic capacities. As with all coronaviruses, it comprises a small structural envelope (E) protein that, despite its scarcity on virions, is known to play a variety of important roles, including virion morphology/assembly and effects on cell response and/or virulence due to the presence of several protein motifs and/or homo-oligomeric ion channel formation. Depending on the strain of coronavirus, the presence of E protein has been shown to have various levels of importance, ranging from degrees of dispensability to absolutely essential for the production of progeny infectious virions. To date, the importance of the E protein for HCoV-OC43 infectious virion production has not been studied, and any literature directly on this protein is lacking. By modifying a full-length cDNA infectious clone, we demonstrated for the first time that the HCoV-OC43 E protein is critical for infectious virus production and propagation in both epithelial and neuronal cells. In addition, investigations of the transmembrane domain revealed key amino acids implicated in modulating infectious virus production and spread in the same cell types. Finally, the C-terminal end of the E protein has been previously predicted to contain a PDZ binding motif (PBM). By fully or partially abroagating this potential protein-protein interaction motif, we were able to discern its high importance in infectious virus production and spread in epithelial and neuronal cells; pointing to possible implications in neuropathogenesis which remain to be elucidated.

Key words : coronavirus, envelope protein, assembly, propagation, virus, PBM, viroporin, neurons

SOMMAIRE RÉCAPITULATIF

Les coronavirus humains (HCoVs) sont des virus d'ARN monocaténaire de polarité positive avec un génome d'approximativement 30 kb encapsidé dans une enveloppe pléomorphe de 120 à 160 nm qui représente le plus grand virus à génome ARN connu. Infectant les mammifères et les oiseaux, les coronavirus sont des virus connus pour leur tropisme respiratoire bien qu'ils démontrent aussi des tropismes entériques, hépatiques et du système nerveux central (SNC). Les souches de coronavirus respiratoires 229E (HCoV-229E) et OC43 (HCoV-OC43) étaient connus comme des virus bénins causant le rhume chez les humains depuis les années 1960. Cependant, l'épidémie récente en 2002 et 2003, du syndrome respiratoire aïgu sévère (SRAS), et l'identification de l'agent pathogène responsable comme étant un coronavirus émergent d'origine zoonotique (le coronavirus du syndrome respiratoire aïgu sévère (SARS-CoV)), a démontré l'impact de cette famille de virus au niveau économique ainsi que potentiellement sur la santé publique et vétérinaire. Il y a donc présentement un renouvellement d'intérêt concernant l'étude de cette famille de virus. Par conséquent, depuis 2003 les études moléculaires virologiques de surveillance ont découvert plus de 60 nouvelles espèces de coronavirus dont deux coronavirus respiratoires humains (HCoV-HKU1 et HCoV-NL63) et un coronavirus animal pouvant se transmettre à l'homme (le coronavirus du syndrome respiratoire du Moyen-Orient (MERS-CoV)) en 2012 (Figure 1.1). La diversité d'espèces des coronavirus pourrait s'expliquer par le fait que les coronavirus contiennent des enzymes, incluant les exoribonucléases, capables de modifier l'ARN directement, ce qui implique des mécanismes de vérification de séquence pour augmenter la tolérance pour les erreurs de séquence. Cela aide les coronavirus à s'adapter rapidement aux nouveaux hôtes et niches écologiques.

Quatre coronavirus humains connus (HCoV-OC43, -229E, -NL63, -HKU1) co-circulent dans la population et sont considérés comme étant des virus causant surtout des maladies mineures en infectant principalement les voies respiratoires supérieures. Cependant, chez des populations plus susceptibles telles que les patients

immunodéprimés, les enfants et les aînés, ces espèces virales peuvent mener à de plus graves complications. En marge de leur tropisme respiratoire, certaines souches, comme le HCoV-OC43, peuvent atteindre le système nerveux central (SNC) et y infecter diverses cellules présentes, notamment les neurones (démontré chez les souris). Notre laboratoire a détecté de l'ARN viral d'HCoV-OC43 dans les cerveaux des patients atteints de maladies neurologiques telles que la maladie d'Alzheimer et la sclérose en plaques ainsi que chez des patients sains. Présentement nous manquons de données afin de mettre en évidence une forte corrélation entre une infection aigue ou chronique d'HCoV-OC43 et le développement de pathologies neurologiques chez l'humain, mais il existe des cas cliniques qui lient fortement le développement de différents types d'encéphalites sévères et l'infection par HCoV-OC43. De plus, la détection de particules infectieuses de SARS-CoV dans le SNC chez des patients décédés du SRAS supporte d'avantage cette hypothèse. Le développement d'un modèle murin d'infection par le HCoV-OC43, au laboratoire du Dr Pierre Talbot, a permis de confirmer le neurotropisme, la neuroinvasion et la neurovirulence du virus chez la souris, et permet de suggérer une potentielle neurovirulence du virus chez l'humain. Toutefois, les mécanismes sous-jacents à la pathogenèse restent encore à élucider et le rôle précis de plusieurs protéines d'HCoV-OC43 pendant l'infection du SNC n'est pas encore bien connu.

Le génome à ARN monocaténaire des coronavirus varie entre 27 et 32 kb et est composé de 6 cadres de lecture ouverts (ORF) (Figure 2.1), dont les deux premiers (ORF1a et ORF1b) représentent les deux tiers du génome et encodent les gènes du complexe de la réplicase pour produire deux polyprotéines. Ces polyprotéines sont clivées en 15 ou 16 protéines non-structurales par des protéases virales. Les autres cadres de lectures produisent les protéines structurales du virus : la glycoprotéine de surface (S), la protéine d'enveloppe (E), la protéine de membrane (M), la protéine de la nucléocapside (N) et la protéine hémagglutinine-estérase (HE) pour certaines espèces de coronavirus, tel que HCoV-OC43. En fonction de l'espèce, on peut retrouver jusqu'à 8 gènes accessoires entre les gènes structuraux. Chez HCoV-OC43, seulement deux

gènes de ce type sont présents, ns2 et ns5, qui agissent potentiellement dans la réponse de la cellule hôte à l'infection et jouent un role dans la pathogenèse.

Au début du cycle de réplication des coronavirus, les particules infectieuses se lient soit au récepteur cellulaire protéique ou à un acide sialique, par l'intermédiaire de leur protéine S, pour fusionner et entrer dans la cellule (Figure 2.2). L'ARN polarité positive du génome produit et assemble sa propre ARN polymérase dans le cytoplasme pour produire un brin négatif complémentaire. Ce brin négatif sert lui-même de modèle pour transcrire de petits ARN sous-génomiques pour fabriquer toutes les autres protéines, notamment les protéines structurales. La protéine N se lie à l'ARN génomique et ce complexe bourgeonne à travers la région intermédiaire du réticulum endoplasmique pour être encapsidé par les protéines S, M, E et HE. La nouvelle particule virale est ensuite relâchée par exocytose. Cette dernière étape distingue les coronavirus des plupart des autres virus enveloppés qui s'assemblent à la membrane plasmique (Figure 2.3).

La protéine E est une petite protéine transmembranaire de 74 à 109 acides aminés présente chez tous les coronavirus. L'homologie de séquence des acides aminés entre espèces est minime, mais sa structure secondaire et ses propriétés biochimiques sont très conservées. La protéine E est aujourd'hui associée à plusieurs fonctions, mais était historiquement classée comme une protéine structurale importante pour la formation de nouvelles particules infectieuses. Récemment, des études portant sur plusieurs souches de coronavirus différents ont démontrées l'importance de la présence de cette protéine dans la voie de sécrétion où elle pourrait interagir avec la protéine M pour induire la courbure adéquate de la membrane virale, ainsi qu'avec elle-même ou d'autres protéines cellulaires à travers des motifs d'interaction protéine-protéine pour contribuer aux pathologies suivant l'infection. Ces études ont récemment permis d'établir la protéine E comme un nouveau facteur de virulence, en plus de la protéine S, notamment dans le cas du SARS-CoV.

Bien que le gène E du HCoV-OC43 ait été séquencé (des souches d'origine clinique et de laboratoire), jusqu'à présent il n'y a pas des données publiées concernant les propriétés ou fonctions spécifiques de la protéine E d'HCoV-OC43. Toutes les informations sont tirées d'études sur d'autres coronavirus, plus spécifiquement le coronavirus murin (MHV), le coronavirus aviaire (virus de la bronchite infectieuse; IBV) et le SARS-CoV. Ce mémoire contient un état des connaisances (Chapitres 1 à 3) concernant la protéine E d'autres coronavirus qui pourrait aider à expliquer les nouvelles données obtenues sont contenues dans le Chapitre 4 (article soumis). De plus, les recherches présentées dans ce mémoire démontrent pour la première fois l'importance de cette protéine dans la production et la propagation de particules infectieuses d'HCoV-OC43 ce qui sera décrit en plus détails après un court survol de certaines caractéristiques biochimiques connues de la protéine E des coronavirus.

Même si la protéine E des coronavirus est importante pour l'assemblage de l'enveloppe virale, la majorité de la produite par la cellule est localisée dans la voie sécrétoire des cellules entre les membranes du réticulum endoplasmique et le Golgi (ERGIC), ce qui a été démontré plus récemment par imagerie en temps réel à l'aide d'une protéine E de MHV-A59 portant à son extrémité C-terminale une étiquette tétracystéine. La forte localisation de cette protéine dans la zone ERGIC est attribuée à la présence d'un ou plusieurs motifs qui ciblent cette région sur la protéine E. Les protéines E d'IBV, MHV et SARS-CoV possèdent un signal de localisation du Golgi, sous la forme d'une structure secondaire «β-hairpin» dans leurs région C-terminale. De plus, en modifiant cette structure «β-hairpin» dans la protéine E de SARS-CoV, un deuxième signal ciblant la région Golgi de la cellule a été découvert. La littérature suggère cette redondance de signaux ciblant le complexe du Golgi pourrait représenter un mécanisme d'évolution assurant la bonne localisation de cette protéine critique durant l'assemblage des virions et/ou être utilisé dans deux différentes populations de protéine E qui pourraient ou non être intégrées dans les virions.

En fonction de l'espèce virale, la protéine E des coronavirus contient de 2 à 4 cystéines dans sa région cytoplasmique près de la région transmembranaire de la protéine (Figure 3.1). Ces cystéines peuvent être palmitoylées chez lBV, SARS-CoV et MHV. Cependant la capacité des cystéines de la protéine E de MHV à former des ponts disulfures n'a pas encore été établie, contrairement au SARS-CoV où cela semble être le cas. Les cystéines de la protéine E ne semblent pas être impliquées dans la localisation de la protéine vers le Golgi, mais seraient importantes pour la stabilité de la protéine et la production de pseudoparticules virales (VLPs) ou de virions. Il est suggéré que la palmitoylation des cystéines servirait à concentrer la protéine E vers la région ERGIC en augmentant l'association de la protéine à ces membranes ou en facilitant son interaction avec d'autre protéines présentes. La palmitoylation de la protéine E pourrait affecter le mouvement de cette protéine dans les membranes pendant l'assemblage des nouveaux virions afin d'aider à son interaction avec la protéine M, assurant ainsi une bonne courbure de la membrane et/ou un relâchement efficace de nouveaux virions.

Il a été démontré que la protéine E joue une rôle très important dans la production efficace des virions infectieux pour de nombreux coronavirus. La délétion de cette protéine dans plusieurs espèces virales abouti généralement à deux différents phénotypes. Dans certains cas, tel que le coronavirus du porc; nommé virus de la gastroentérite transmissible du porc (TGEV), et le MERS-CoV, la délétion de la protéine E fait qu'il n'y ait pas de virus infectieux détectable. Dans le cas de MHV et SARS-CoV, la délétion de la protéine mène à une réduction de l'efficacité de production de virus infectieux en fonction du type cellulaire utilisé, bien que les titres viraux soient encore détectables. À ce jour, il n'existe pas d'information concernant l'importance de la protéine E pour la production des virions infectieux pour la souche circulante HCoV-OC43. Étant donné que la protéine E du SARS-CoV a récemment été établie comme un facteur de virulence, l'étude présentée dans ce mémoire est une première étude de la protéine E d'HCoV-OC43 et de son importance dans la production virale. Ces données pourraient mettre en évidence différents facteurs jouant un rôle pour produire des souches soit circulantes (HCoV-OC43), soit plus pathogènes (SARS-CoV, MERS-

CoV), ou des facteurs impliqués dans les neuropathologies que nous avons déjà observé suite à l'infection des souris par HCoV-OC43.

A l'aide d'un clone infectieux d'ADNc du génome complet d'HCoV-OC43 déjà utilisé au laboratoire (pBAC-OC43^{FL}), nous avons introduit un codon stop vers le début du gène E afin de prévenir la production complète de la protéine E dans le virus recombinant correspondant (rOC/E-Stop) (Figure 4.1A). La transfection de ce clone infectieux muté dans des cellules BHK-21 n'a pas permit de détecter de nouveaux virus infectieux (Figure 4.1B). Cependant lorsqu'une co-transfection transitoire a été effectuée avec un plasmide contenant la protéine E sauvage, pcDNA(OC-E), la production de virions infectieux a pu être ramenée à un niveau détectable (Figure 4.1C-E). Afin de déterminer si ces virions initialement complémentés mais n'exprimant pas la protéine E seraient capables de produire des nouveaux virions, nous avons effectué trois amplifications successives sur les cellules épithéliales humaines HRT-18 (Figure 4.2A). Nos résultats démontrent qu'un virus rOC/E-Stop non-complémenté ne produit pas de virus infectieux détectables, alors que le virus rOC/E-Stop initialement complémenté pourrait produire des virions infectieux. Cependant l'efficacité de ce processus est largement inférieure au virus sauvage, et n'est plus détectable après trois amplifications sur les cellules HRT-18. Il est aussi intéressant de noter que des répétitions indépendantes de ce processus d'amplification, ont permis la détection de virus infectieux rOC/E-Stop (initialement complémentés) qui augmentent après plusieurs amplifications jusqu'au niveau du virus sauvage dans certains cas. Dans ces cas, l'analyse de la séquence du gène E a mis en évidence la présence de révertants aux positions où le codon stop a été introduit dans le clone infectieux (Figure 4.2).

Notre laboratoire s'intéresse principalement aux neuropathologies induites par HCoV-OC43, ainsi qu'aux facteurs cellulaires et viraux impliqués. Nous avons voulu déterminer si la protéine E est aussi importante dans le contexte du SNC que ce qui a été préalablement démontré au niveau des voies respiratoires pour d'autres coronavirus. Etant donné que notre laboratoire a déjà démontré que les neurones sont

les cibles principales du virus dans le SNC, nous avons voulu déterminer si l'absence de la protéine E peut engendrer une différence de tropisme après l'infection de cellules susceptibles. Dans ce but, nous avons infecté la lignée LA-N-5 (cellules neuronales différenciées humaines) ainsi que des cultures primaires mixtes du SNC murin avec notre virus rOC/E-Stop initialement complémenté (Figure 4.3). En infectant les cellules LA-N-5, on observe une importante diminution de production de virus infectieux, un phénomène qui est exacerbé lors de l'infection de cultures mixtes du SNC murin. De plus, dans le cas des cultures mixtes murines, nous n'avons pas été capables de détecter l'infection de cellules au cours des 72 heures post-infection, suggérant une importante inhibition de propagation virale (Figure 4.3, Figure 4.6).

L'ensemble de ces résultats nous permet de conclure pour la première fois que la protéine E d'HCoV-OC43 joue un rôle critique pour la production de virion infectieux. Comme pour le MERS-CoV et TGEV, la souche HCoV-OC43 nécessite la protéine E pour produire de nouveaux virions infectieux à un niveau détectable. Les cellules épithéliales et neuronales peuvent être infectées par les virus recombinant HCoV-OC43 manquant la protéine E mais la propagation virale subséquente est fortement diminuée.

De plus, l'importance d'une protéine E fonctionnelle est soulignée par le fait que l'amplification du virus recombinant rOC/E-Stop intialement complémenté sur les cellules épithéliales HRT-18 introduit une forte sélection pour des populations virales révertantes ramenant la production de virions infectieux à un niveau similaire au virus sauvage après très peu de passages. Ce phénomène a déjà été décrit pour MHV et le SARS-CoV où après quelques passages, les deux souches virales ont subies des mutations dans le gène M permettant de remplacer la protéine E préalablement supprimée. Nos résultats suggèrent fortement qu'une pression sélective s'exerce sur la protéine E elle-même. Enfin, la protéine E d'HCoV-OC43 joue un rôle critique pour la production des virions infectieux et l'importance de cette protéine est amplifiée dans les cellules originaires du SNC, peut-être en impliquant d'autres motifs dans la protéine elle-même et en interagissant avec d'autres facteurs viraux et cellulaires.

Plusieurs virus avec un génome d'ARN (et quelques virus à ADN) sont connus pour posséder des protéines transmembranaires pouvant former des canaux ioniques ; protéines nommées viroporines (Table 3.1). Généralement les viroporines sont de petites protéines virales transmembranaires de 60 à 120 acides aminés possédant la capacité de s'homo-oligomériser afin de perméabiliser les membranes. Cette caractéristique facilite la fonction de ce type de protéines dans la réplication virale, dans l'entrée/sortie de la particule virale, dans l'assemblage des virions et dans la perturbation des fonctions normale de la cellule. Les protéines 3a et 8a de SARS-CoV, la protéine 4a de HCoV-229E et la protéine ns5 d'HCoV-OC43 ont récemment été caractérisées comme étant de potentielles viroporines et comme étant importantes pour le cycle de réplication des coronavirus. De plus, au cours des dix dernières années, les études sur la protéine E de différents coronavirus ont mis en évidence ce comportement caractéristique de viroporine. Il a été démontré que les protéines E d'IBV, d'HCoV-229E, de MERS-CoV et de SARS-CoV peuvent former des canaux ioniques dans des membranes synthétiques perméables aux ions. En effet, la protéine E d'IBV pourrait exister sous la forme d'homo-pentamères et celle de MERS-CoV et de SARS-CoV peut former des homo-pentamères dans des membranes lipidiques artificielles (Figure 3.2).

Etant donné que la majorité de la protéine E produite par la cellule est localisé dans les membranes de la zone ERGIC, ce comportement potentiel de viroporine de la protéine E des coronavirus pourrait avoir un impact majeur sur la voie de sécrétion de la cellule et/ou le cycle de réplication virale. En effet, l'inhibition chimique de la fonction de canaux ioniques, la destruction de l'intégrité ou le remplacement de la région transmembranaire de la protéine avec celui d'une autre espèce virale a permis d'observer une réduction des titres de virus infectieux chez plusieurs espèces de coronavirus. Plus spécifiquement, il a été démontré pour SARS-CoV et IBV que deux acides aminés sont très importants pour la stabilité de ce domaine lors de la formation de viroporines. Le premier, l'asparagine à la position 15 (N15) dans la région transmembranaire chez le SARS-CoV (équivalent à T16 chez l'IBV) est supposé être

important pour la sélectivité d'ions prêt de l'ouverture du canal, tandis que la valine à la position 25 (V25) dans la région transmembranaire (équivalent à A26 chez IBV) est supposé être à l'interface des interactions des monomères et serait donc important pour la formation de viroporines (Figure 3.3). De plus, la fonction de canaux ioniques de la protéine E de SARS-CoV a été mise en lien avec les capacités adaptatives du virus, modulant la production virale et la pathogénicité chez les souris. Récemment, la protéine E d'IBV a été décrite comme étant présente sous deux formes, soit sous forme monomérique pouvant perturber la voie de sécrétion, soit sous forme oligomérique pouvant moduler l'assemblage de virions.

Présentement, il n'y a aucune étude démontrant que la protéine E d'HCoV-OC43 contient véritablement une région transmembranaire. Cependant, même s'il y a peu de similarité de séquence entre les différentes souches de coronavirus, il a été démontré que la similarité de structures secondaires et de caractéristiques biochimiques est plus importantes. Dans l'étude présentée ici, nous avons déterminé que la mutation des deux acides aminés équivalents à ceux déjà étudiés dans la région transmembranaire du SARS-CoV et d'IBV pourrait avoir un effet sur la production des particules virales chez HCoV-OC43 et pourrait nous donner des indices concernant la fonction de cette région de la protéine E d'HCoV-OC43.

En utilisant notre clone infectieux d'ADNc, pBAC-OC43^{FL}, nous avons produit des mutants dans la région transmembranaire de la protéine E. L'acide aminé polaire, la glutamine à la position 17 (Q17) de la région transmembranaire potentielle d'HCoV-OC43 (à la position équivalente à N15 en SARS-CoV) a été remplacé par mutagenèse dirigée par une alanine plus petite et non-chargée. Théoriquement, ce changement devrait diminuer la fonction associée au passage d'ions à travers un potentiel canal ionique. Une deuxième mutation a été effectuée à la position 27 contenant une leucine (L27) dans la région transmembranaire potentielle d'HCoV-OC43 (équivalente au V25 de SARS-CoV). Ce petit acide aminé a été muté en une phénylalanine (plus grosse) afin de bloquer ou déstabiliser la région transmembranaire potentielle. Dans le cas de la

protéine E de SARS-CoV, cet acide aminé est à l'interface des interactions de monomères dans la section la plus mince du canal ionique.

Suite à la transfection des clones infectieux mutés dans des cellules BHK-21, nous avons pu évaluer la production de virus infectieux de rOC/E-TM$_{Q17A,}$ mais cette production était sous le seuil de détection pour le mutant rOC/E-TM$_{L27F}$ (Figure 4.4A). Dans les deux cas, l'ARN du gène E a bien été détecté par PCR, suggérant que la transfection a bien fonctionné (Figure 4.4B). Le stock viral de rOC/E-TM$_{Q17A}$ produit après transfection a ensuite été amplifié sur des cellules épithéliales HRT-18, donnant des titres de virus infectieux moins importants comparé au virus sauvage. Quand nous avons infecté des cellules humaines neuronales LA-N-5 ou des cultures primaires mixtes du SNC murin avec le même mutant, nous avons pu observer une diminution similaire des titres de virus infectieux dans les fractions de virus non-associés aux cellules (Figure 4.4C et D). Cependant les fractions de virus associés aux cellules montraient des titres comparables au virus sauvage pour les deux types cellulaires, suggérant un défaut dans le relâchement du virus. Ces résultats sont supportés par des études d'immunofluorescence sur cultures primaires mixtes du SNC murin démontrant un délai de propagation de rOC/E-TM$_{Q17A}$ comparé au virus sauvage à 48 heures post-infection (Figure 4.6). Ces résultats soulignent l'importance du potentiel domaine transmembranaire de la protéine E d'HCoV-OC43 pour la production virale et l'efficacité de propagation du virus dans les cellules neuronales.

Dans le cas de notre mutant de la protéine E L27F, nous n'étions pas capables de détecter des virions infectieux après transfection. Un phénomène similaire a été observé pour IBV, où une mutation à la position équivalente, A26, ne permettait pas de produire des VLPs, probablement à cause de l'abolition des contacts lors de l'oligomérisation. Chez le SARS-CoV, des études par résonance magnétique nucléaire (RMN) suggèrent également que c'est la région la plus mince du canal ionique. Notre hypothèse est donc que l'acide aminé L27 d'HCoV-OC43 représente un acide aminé très important pour l'infection, soit pour l'homo-oligomérisation et/ou l'activité du canal

ionique, est c'est pourquoi nous ne pouvons pas détecter de virions infectieux. Nous avons pu observer une diminution importante du titre viral du mutant transmembranaire Q17A dans les cellules épithéliales et neuronales. Plus précisément, une comparaison des fractions associées et non-associées aux cellules suite à l'infection de cellules LA-N-5 et cultures primaire mixtes du SNC murin suggère un défaut dans le relâchement viral au niveau de la voie de sécrétion. Cela a déjà été démontré avec un mutant d'IBV (à la position équivalente T16) n'exhibant pas une conductance ionique associée aux canaux formés par la région transmembranaire de sa protéine E. Les auteurs ont démontré que cet acide aminé n'est pas important pour la formation de virions, mais qu'il joue un rôle dans la perturbation de la voie de sécrétion. Comparativement, les positions N15 et V25 de SARS-CoV sont importantes pour l'activité ionique des canaux et l'adaptabilité du virus, mais n'ont pas d'effet sur la production virale comparé au virus sauvage. Cependant, l'observation de plages de lyse moins développées après infection avec ces mutants suggèrent que les canaux ioniques pourraient induire des effets cytopathiques et jouer un rôle dans la dissémination virale.

En conclusion, nos études suggèrent que la position Q17 de la potentielle région transmembranaire de la protéine E pourrait servir à moduler le relâchement de virus dans les cultures cellulaires et que la position L27 semble être une position nécessaire pour la production de particules infectieux de virus. Ces résultats représentent les premiers indices en lien avec la présence d'une région transmembranaire de la protéine E d'HCoV-OC43 et de son important rôle dans le cycle de réplication du virus.

Un dernier thème émergent dans la littérature portant sur les interactions virus-cellule est celui portant à propos les interactions protéine-protéine à l'aide de motifs liant des domaines PDZ (les PBMs). Les domaines PDZ sont des domaines de reconnaissance protéine-protéine répandus dans le protéome humain, qui agissent comme un système d'échafaudage pour faciliter l'assemblage de plusieurs composants intracellulaires. Ces domaines sont composés d'approximativement 80 à 90 acides aminés formant six feuillets bêta et deux hélices alpha établissant une structure en baril

bêta avec un point d'interaction protéique de fixation (Figure 3.4). Les domaines PDZ sont impliqués dans d'importantes fonctions dans la cellule, dont la formation de complexes protéiques, la formation des jonctions cellulaires, le developement de polarité cellulaire ou dans les voies de transduction de signaux. Des mutations retrouvées dans des protéines contenant les domaines PDZ ou leurs ligands, et l'expression différentielle de protéines contenant des domaines PDZ ont été liés au cancer et associés aux troubles neurologiques chez l'humain. La majorité des interactions protéiques concernant les domaines PDZ reconnaissent un motif spécifique à l'extrémité C-terminale d'un ligand. Ce motif nommé motif liant un domaine PDZ (*PDZ-binding motif*; PBM) peut être représenté par une séquence canonique.

La majorité des PBMs peuvent être classifiés selon trois types dépendamment des acides aminés présents aux positions -0 et -2 de la région C-terminale de la protéine. Ces types sont : PBM de type I (-X-S/T-X-Φ_{COOH}), PBM de type II (X-Φ-X-Φ_{COOH}) et PBM de type III (-X-D/E-X-Φ_{COOH}), où X représente n'importe quel acide aminé et Φ est un résidu hydrophobe (Table 3.2). Les interactions entre les domaines PDZ et leur ligand contenant un PBM sont hautement spécifiques et sont fortement dépendants de la séquence du PBM et de la localisation intracellulaire du ligand. L'étude des motifs PBM dans les protéines virales est un domaine en émergence (Table 3.3). Les PBMs viraux perturbent et modifient les processus réguliers de la cellule pour, entre autre, moduler la production virale, la propagation entre les cellules, ou pour augmenter la pathogénicité.

Récemment les études sur la protéine E du SARS-CoV ont démontré que cette protéine contient un PBM de type II et que cette région peut lier la protéine nommée *protéine associée à Lin Seven 1* (PALS1). Les auteurs ont démontré que pendant l'infection de cellules de poumons, cette interaction perturbe la localisation normale de cette protéine cellulaire qui est importante pour la formation des jonctions serrées, pour le développement de la polarisation cellulaire et pourrait expliquer les dommages alvéolaires dans les poumons suite à l'infection. Une deuxième interaction du PBM de

la protéine E a été observée décrivant un lien direct entre une interaction du PBM viral et sa fonction dans la pathologie du virus du SRAS. En effet, le PBM de la protéine E du SARS-CoV interagit avec la protéine synténine activant directement la MAP kinase p38 pour exacerber la réponse inflammatoire de la maladie. Quand la région PBM a été supprimée, les pathologies associées à l'infection ont disparu dans un modèle murin, démontrant un nouveau mécanisme de modulation de la pathogenèse des coronavirus. La présence d'un motif PBM dans la protéine E d'HCoV-OC43 et de plusieurs autres coronavirus a été prédite par analyse bio-informatique, mais à ce jour, aucune étude n'a été publiée à propos de ce sujet, hormis ceux sur SARS-CoV. Dans ce mémoire nous avons effectué une étude préliminaire afin de déterminer si en modifiant le potentiel motif PBM de la protéine E d'HCoV-OC43, il y a un effet sur la production et la propagation virale, ce qui pourrait nous donner des premiers indices concernant la fonctionnalité de ce potentiel motif durant l'infection.

Nous avons modifié notre clone infectieux d'ADN complémentaire pBAC-OC43^{FL} pour modifer le motif PBM potentiel de type II potentiel dans la région C-terminale de la protéine. Pour ce faire, nous avons changé les acides aminés soit au position -0 (acide aminé numéro 84) ou -2 (acide aminé numéro 82), ou soit les deux acides aminés ensemble, en alanines pour abroger une interaction protéique potentielle dans la région C-terminale. Les simples mutants (pBAC-OC-E-PBM-D82A et pBAC-OC-E-PBM-V84A) et le double mutant (pBAC-OC-E-PBM-D82A-V84A) ont été transfectés sur les cellules BHK-21 et amplifiés sur les cellules épithéliales HRT-18. L'amplification sur les cellules épithéliales a démontré que l'abrogation complète du domaine PBM potentiel (rOC/E-PBM$_{D82A-V84A}$) engendre une diminution de virions infectieux en comparaison aux mutants simples (rOC/E-PBM$_{D82A}$ ou rOC/E-PBM$_{V84A}$) et au virus sauvage (Figure 4.5A). Afin de déterminer si le motif PBM potentiel pourrait jouer un rôle dans l'infection ou la propagation dans le contexte du SNC, nous avons effectué des infections de cellules neuronales humaines LA-N-5 et de cultures primaires mixtes du SNC murin. La production de virus infectieux suite à l'infection des cellules neuronales humaines par le double mutant a été encore plus réduite que dans les cellules épithéliales lorsque comparé aux simples mutants et au virus sauvage (Figure 4.5B). La tendance est

encore plus marquée dans le contexte des cultures primaires où il n'y a pas de détection de virus infectieux du double mutant rOC/E-PBM$_{D82A-V84A}$ en comparaison au virus sauvage ou aux virus arborant un seule mutation (Figure 4.5C). La détection de protéines virales par immunofluorescence a démontré que les deux types de cellules neuronales sont également infectées indépendamment du virus mutant utilisé. Dans les deux modèles, l'infection avec le mutant arborant un domaine PBM potentiel abrogé, rOC/E-PBM$_{D82A-V84A}$, mène à une réduction de propagation virale importante (Figure 4.5D, Figure 4.6).

En abrogeant le motif PBM potentiel dans la région C-terminale de la protéine E d'HCoV-OC43, nous avons démontré son importance dans la production de virus infectieux et la dissémination efficace du virion. Nous avons observé une diminution significative de la propagation virale dans le contexte de cellules neuronales. Des études similaires avec la protéine E du SARS-CoV ont cependant démontré que la délétion du motif PBM menait à une légère diminution des titres infectieux dans certains types cellulaires et aucune différence dans d'autres. Nous avons également démontré que la capacité initiale des virus à infecter les cellules n'était pas affectée par des mutations dans la région PBM potentiel, ce qui correspond à ce qui a été démontré chez le SARS-CoV. Notre étude semble indiquer que le motif PBM de la protéine E est fonctionnel, avec de possibles fonctions plus importantes dans le contexte du SNC, ce qui devra être confirmé par le biais d'études sur le neuro-interactome de la protéine. Nous pouvons par contre formuler l'hypothèse qu'en abrogeant le motif PBM de la protéine E, des interactions cellulaires et/ou virales nécessaires pour l'assemblage des virions infectieux seraient rendues inefficaces. Une autre possibilité est que le motif PBM de la protéine E pourrait être lié aux neuropathologies ayant déjà été observées chez la souris suite à l'infection d'HCoV-CoV telles que l'excitotoxicité glutamatergique en interagissant avec les protéines contenant des domaines PDZ présentes dans les cellules neuronales.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LA PROTÉINE D'ENVELOPPE (E) DU CORONAVIRUS RESPIRATOIRE HUMAIN HCOV-OC43 EST NÉCESSAIRE POUR LA FORMATION DE VIRIONS INFECTIEUX DANS LES CELLULES ÉPITHÉLIALES ET NEURONALES

I - LITERATURE REVIEW

1 CORONAVIRINAE – A UBIQUITOUS VIRUS SUBFAMILY

Coronaviruses, members of the the *Coronavirinae* subfamily, of the greater *Coronaviridae* family and *Nidovirales* order are the largest RNA viruses identified to date (de Groot *et al.*, 2012). Coronaviruses are morphologically pleomorphic enveloped viruses of approximately 120-160nm across and containing a single-stranded positive-sense RNA genome of approximately 30kb (de Groot *et al.*, 2012). Coronaviruses are generally respiratory pathogens of many mammalian and avian species but are also known to cause enteric, hepatic and central nervous system diseases (de Groot *et al.*, 2012, Desforges *et al.*, 2014). The virus target is usually epithelial cells, while transmission usually follows fomite, aerogenic or fecal-oral route depending on the species. Coronaviruses that infect humans usually infect the upper respiratory tract however infections in immunocompromised individuals as well as elderly and infant patients can lead to more severe clinical outcomes (Gaunt *et al.*, 2010, Vabret *et al.*, 2009).

The emerging diversity of coronaviruses (Figure 1.1), delimited by research spurred by the severe acute respiratory syndrome (SARS) outbreak associated with the SARS coronavirus (SARS-CoV) in 2003, underline the extensive adaptative capacity of the virus. Previously the size limit of RNA genomes was normally thought to be between 10^3 and 10^4 base pairs, linked to the inherent low fidelity of the RNA-dependent RNA polymerases (RdRps) used in viral RNA synthesis (10^{-3} to 10^{-4}mutations/nucleotide/round of replication) (Steinhauer *et al.*, 1992). While bioinformatic models suggest that RNA genome size is constrained to approximately 30kb for stability and precise replication, coronaviruses with genomes of 27-32kb do not conform to this theoretical limit (Smith *et al.*, 2012). Recent work has demonstrated that coronaviruses contain many RNA-modifying enzymes, including an exoribonuclease domain that could act as a proofreading mechanism and could help explain the high tolerance for fidelity variation of sequence compared to other RNA viruses (reviewed in (Smith *et al.*, 2012)). These characteristics allow coronaviruses to not only establish

intra-species diversity but to also adapt quickly to new hosts and ecological niches. Indeed, since the SARS epidemic, molecular surveillance and virus discovery studies have yielded evidence for at least 60 novel coronaviruses, two being human respiratory coronaviruses (HCoV-HKU1 and HCoV-NL63) and one virus capable of zoonotic transmission to humans (Middle-Eastern respiratory syndrome coronavirus; MERS-CoV) (de Groot *et al.*, 2012, Zaki *et al.*, 2012).

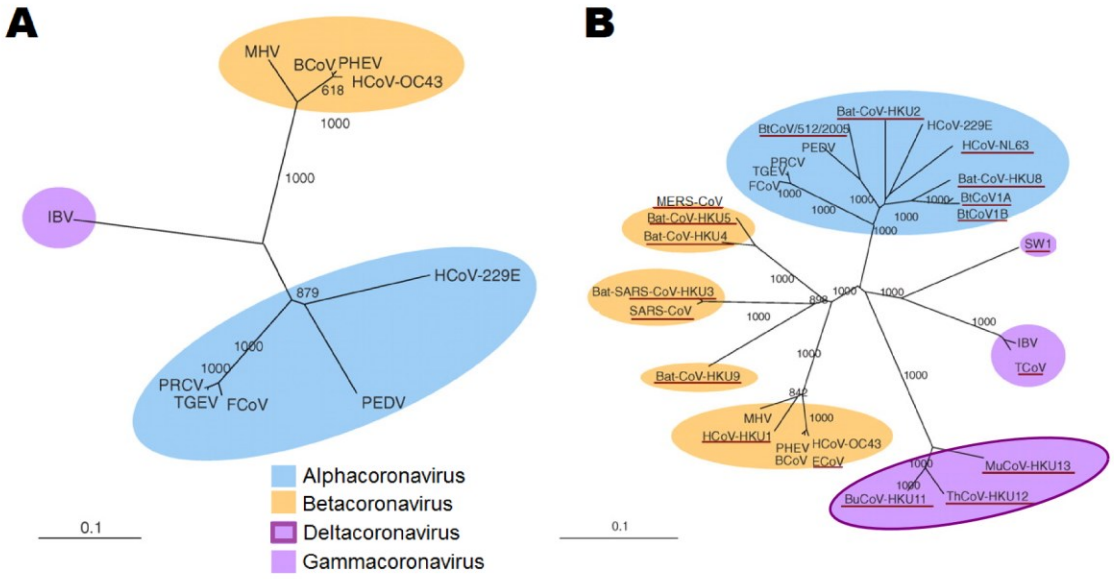


Figure 1.1 **Phylogenetic analysis of RNA-dependent RNA polymerases (RdRp) of coronaviruses. The complete genome sequences of the ten species available before SARS-CoV discovery in 2002 (A), and those available by the end of 2008 (with the 2012 MERS-CoV added to the Bat-CoV-HKU4/U5 cluster for reference) (B).** Colour differentiation represents the different *Coronavirinae* genera with Deltacoronavirus as a fourth proposed genus. Scale bars represent the estimated number of substitutions per 10 amino acids. SARS-CoV, severe acute respiratory syndrome coronavirus; MERS-CoV, Middle-East respiratory syndrome virus; HCoV, human coronavirus; BCoV, bovine coronavirus; MHV, murine hepatitis virus; IBV, infectious bronchitis virus; PEDV, porcine epidemic diarrhea virus; TGEV, porcine transmissible gastroenteritis virus; FCoV, feline coronavirus, PRCoV, porcine respiratory coronavirus; bat-CoV, bat coronavirus; PHEV, porcine hemagglutinating encephalomyelitis virus; ECoV, equine coronavirus; TCoV, turkey coronavirus; SW1, beluga whale coronavirus; BuCoV, bulbul coronavirus; ThCoV, thrush coronavirus; MuCoV, munia coronavirus. Modified from (Woo *et al.*, 2009) with additional information from (de Groot *et al.*, 2012, Graham *et al.*, 2013).

1.1 Coronaviruses infecting humans

Human coronavirus (HCoV) strain B814 was the first discovered species of coronaviruses to infect humans in the 1960s by Tyrell and Bynoe (Tyrrell *et al.*, 1965). First passaged in human embryonic tracheal organ cultures and then inoculated intranasally into human volunteers, common cold-like symptoms were produced. Hamre and Procknow recovered the now named HCoV-229E from subjects with mild respiratory illness and grew the virus on human tissue cultures (Hamre *et al.*, 1966). HCoV-OC43 was originally isolated in 1967 by Chanock and colleagues from human embryonic tracheal organ cultures (origin of “OC” designation) from patients with upper respiratory tract disease and passaged once in suckling mice (McIntosh *et al.*, 1967a, McIntosh *et al.*, 1967b, Tyrrell *et al.*, 1965). Infections of healthy volunteers demonstrated the HCoV-OC43 and -229E viruses to be the cause of the common cold but the number of published studies in the field remained limited until the mid 1990s.

The 2002-2003 SARS epidemic caused by the causative agent SARS-CoV, discovered in 2002 (Drosten *et al.*, 2003, Ksiazek *et al.*, 2003), represented the first instance of a coronavirus inducing high morbidity and mortality disease as well as the first detected instance of zoonotic transmission of coronaviruses to humans; proposed to be transmitted via palm civets from a bat reservoir (Ge *et al.*, 2013, W. Li *et al.*, 2005). There was an enormous public health response to limit and stop the spread of the virus, and it is no longer circulating after the last reported case in July 2003. This epidemic resulted in more than 8000 reported cases with a mortality rate of 10-15% (Peiris *et al.*, 2004). The renewed focus on the coronavirus family led to the discovery of the human coronavirus NL63 (HCoV-NL63) isolated from children with bronchiolitis and pneumonia in the Netherlands in 2004 (Fouchier *et al.*, 2004, van der Hoek *et al.*, 2004) and HKU1 (HCoV-HKU1) from an adult with chronic pulmonary disease in Hong Kong in 2005 (Woo *et al.*, 2005).

In 2012, ten years after the SARS outbreak, a new coronavirus, eventually named Middle East respiratory syndrome coronavirus (MERS-CoV), was isolated from the sputum of a 60-year-old man in Saudi Arabia who presented acute pneumonia and subsequent renal failure and later died (Zaki *et al.*, 2012). This new virus, proposed to have a bat reservoir and transmitted to humans through dromedary camels, causes significantly higher mortality (at approximately 40%) but decreased human-to-human transmission compared to SARS-CoV (Zumla *et al.*, 2015). Since 2012, it has become endemic to the Arabian peninsula with outbreak cases reported in many countries across the world, but which are always linked to recent travels to the incubation area (Zumla *et al.*, 2015). New cases are continuously reported and as of February 2016 there are more than 1,600 cases reported by the World Health Organization (http://www.who.int/emergencies/mers-cov/en/; accessed 13-02-2016). The section below will focus on the non-zoonotic coronaviruses known to infect humans, this being the focus of this thesis work.

1.2 Human coronavirus epidemiology and associated clinical outcomes to infection

The four human coronaviruses, HCoV-OC43, HCoV-229E, HCoV-NL63 and HCoV-HKU1 currently co-circulate worldwide and have seasonal infection peaks throughout the year (Gaunt *et al.*, 2010, Lau *et al.*, 2006, Vabret *et al.*, 2009, Vabret *et al.*, 2008), causing infections of the upper and lower respiratory tract. Human coronaviruses have been found in co-infection with other respiratory viruses, especially human respiratory syncytial virus and human rhinovirus/enterovirus (Gaunt *et al.*, 2010, Hara *et al.*, 2015, Lepiller *et al.*, 2013). Current knowledge regarding infection rates and clinical manifestations of coronaviruses, although growing in recent years, is variable and limited given that: virus prevalence can fluctuate by year and location, surveillance studies often target differing populations, are in limited length and the contribution of coronavirus co-infection with other viruses in the respiratory tract is not widely investigated (Gaunt *et al.*, 2010, Lepiller *et al.*, 2013). Van der Hoek and colleagues

suggested that in hospitalized children under the age of 2, HCoV-OC43 and HCoV-NL63 are the most common coronavirus infections in the Netherlands, followed by HCoV-HKU1 and HCoV-229E as determined by detection and seroconversion rates (Dijkman *et al.*, 2012), corroborating previous work done on a wider age group in the United Kingdom (Gaunt *et al.*, 2010). Pellier *et al.* observed the same prevalence of HCoV-OC43 and HCoV-NL63 in their study with children under the age of 2 in France, however the distribution changed to predominant HCoV-229E and HCoV-HKU1 infections in children 3 years or older and adults, suggesting variation of species between age groups and between different geographical regions of the world (Lepiller *et al.*, 2013). Based on the fact that coronaviruses are sorted into their genera based on a combination of serology and genetic sequence, the authors hypothesize that (fully or partially) neutralizing antibodies against the spike protein of HCoV-OC43 may cross-react with the spike protein of HCoV-HKU1 (which is in the same genus) to protect, or partially protect infection by this second virus, explaining its lower relative infection rate to HCoV-OC43 (Dijkman *et al.*, 2012). It is also hypothesized that the same type of immunity may be raised against HCoV-229E after HCoV-NL63 infection; however these concepts have yet to be experimentally confirmed (Dijkman *et al.*, 2012).

Although the four human coronaviruses are all considered to be common cold viruses with associated mild upper respiratory tract disease symptoms, they can cause more severe clinical symptoms in vulnerable populations including children, the elderly and immunocompromised individuals after the infection of the lower respiratory tract (Lepiller *et al.*, 2013, Vabret *et al.*, 2009). HCoV-NL63 was first described in young children with severe lower respiratory track infections (Fouchier *et al.*, 2004, van der Hoek *et al.*, 2004) and one elderly Canadian patient is reported to have died 5 days after onset of disease (Bastien *et al.*, 2005). An association has been suggested between HCoV-NL63 infection and croup (laryngotrachyitis) in young children (van der Hoek *et al.*, 2005, Wu *et al.*, 2008) but in other studies no such association was made (Gaunt *et al.*, 2010). HCoV-HKU1 is consistently identified in vulnerable patients, either elderly or young children with major underlying disease, especially of the respiratory and cardiovascular systems, leading to more severe outcomes such as bronchiolitis and

pneumonia, reviewed extensively in (Pyrc *et al.*, 2007). HCoV-HKU1 has also been suggested to play a role in gastrointestinal disease (Vabret *et al.*, 2006) and has been associated with febrile seizures in children with acute respiratory tract infections (Lau *et al.*, 2006). HCoV-229E is mainly associated with infections of immunocompromised individuals and has low detection rates in otherwise healthy individuals (Gaunt *et al.*, 2010, Gerna *et al.*, 2006, Pene *et al.*, 2003).

1.3 Neurovirulent potential of human coronaviruses

As indicated previously, clinical manifestations of HCoV-OC43 and HCoV-229E infections are most often presented as symptoms of the common cold. However, when HCoV-OC43 was first isolated, it was shown to be neurovirulent and cause disease (tremors, rigidity, lethargy) in suckling mice after one passage and encephalitis after 2-4 passages (McIntosh *et al.*, 1967a). Neurotropism of the virus was confirmed in later studies showing that HCoV-OC43 and HCoV-229E were able to infect human and murine neural cells (Bonavia *et al.*, 1997, Jacomy *et al.*, 2006) with the principal (or main) target of HCoV-OC43 in the central nervous system of mice being neurons (Jacomy *et al.*, 2006). Further investigations showed the virus to be naturally neuroinvasive in humans, with viral HCoV-OC43 and HCoV-229E RNA detected in the brains of both healthy controls and patients with neurological disease (Arbour *et al.*, 2000, Stewart *et al.*, 1992).

It is difficult to correlate acute and/or chronic HCoV-OC43 human central nervous system (CNS) infections with neurological pathologies, however there is a case where HCoV-OC43 was detected in the cerebrospinal fluid of a child presumed to have acute disseminated encephalomyelitis (Yeh *et al.*, 2004) and there is a recent case strongly linking HCoV-OC43 infection with the development of fatal encephalitis of a toddler in the United Kingdom (Morfopoulou *et al.*, 2016). In addition, invasion of the CNS by SARS-CoV has been recorded in deceased SARS patients (Gu *et al.*, 2005, Xu *et al.*,

2005) and recent observations of patients presenting severe (and fatal) neurological syndromes concurrent with positive MERS-CoV RNA detection in the respiratory tract suggest a neurological pathology (Arabi *et al.*, 2015).

In recent years, a mouse model of HCoV-OC43 has demonstrated the neurovirulent capacities of the virus (Brison *et al.*, 2011, Brison *et al.*, 2014, Jacomy *et al.*, 2006, Jacomy *et al.*, 2010, Jacomy *et al.*, 2003) and may provide insights on certain human pathologies. A significant portion of mice infected with HCoV-OC43 which survived initial encephalitis, had detectable levels of viral RNA in the CNS for at least a year, showed abnormal limb clasping reflexes and had a smaller hippocampus associated with loss of neurons (Jacomy *et al.*, 2006). Furthermore, mutations within the spike (S) protein which serves for viral attachment to susceptible cells led to the development of other severe neuronal dysfunctions in addition to previously observed encephalitis, including paralysis induced by glutamate excitotoxicity of neurons as well as demyelination of the same cell type (Brison *et al.*, 2011, Jacomy *et al.*, 2010). Given that first infections of HCoV-OC43 occur early in life, that persistent natural infection is possible and that the demonstration of a wide range of neuronal pathologies, it is conceivable that invasion of HCoV-OC43 into the brain could trigger and/or exacerbate the development of neurodegenerative conditions such as multiple sclerosis or Parkinson’s disease (Arbour *et al.*, 2000).

2 CORONAVIRUS GENOME ORGANIZATION AND REPLICATION CYCLE

Coronavirus genomes are composed of positive-sense single-stranded infectious RNA between 27 and 32 kb in length (de Groot *et al.*, 2012). The genome is composed of at least 6 open reading frames (ORFs) in a conserved order and flanked by 5' and 3' untranslated regions (UTRs) (Figure 2.1). Two-thirds of the genome are taken up by the ORF1a and ORF1b encoding replication-transcriptase (replicase) genes which are synthesized into two polyproteins; pp1a from ORF1a and pp1a/b from ORF1a/1b by programmed ribosomal frameshifting which are subsequently proteolytically cleaved by viral proteases into 15 or 16 non-structural proteins (Sawicki *et al.*, 2007). Downstream of ORF1a/1b, are the ORFS for the structural proteins spike (S), envelope (E), membrane (M), nucleocapsid (N) and hemagglutinin-esterase (HE), the latter found in lineage A betacoronavirus such as HCoV-OC43. Up to eight accessory genes are interspersed in between the structural genes which are group or species specific and the products of which yield protein products that are generally dispensable for replication in cell culture and are often associated with viral pathogenesis and virulence (de Groot *et al.*, 2012, D. X. Liu *et al.*, 2014). HCoV-OC43 contains two accessory proteins, ns2 and ns5 (formerly ns12.9) which may be involved in the host cell response to infection. The murine coronavirus MHV (murine hepatitis virus) ns2 homologue participates in the inhibition of type I interferon response (Zhao *et al.*, 2012) while ns5 has been reported to antagonise this response during infection (Koetzner *et al.*, 2010) as well as act as a viroporin to influence virion morphogenesis and pathogenesis (R. Zhang *et al.*, 2015). Recent data strongly suggest that ns2 may localize in the nucleus of HCoV-OC43 infected neurons in the CNS where it could work as a DNA-binding protein (Levros *et al.*, unpublished data).

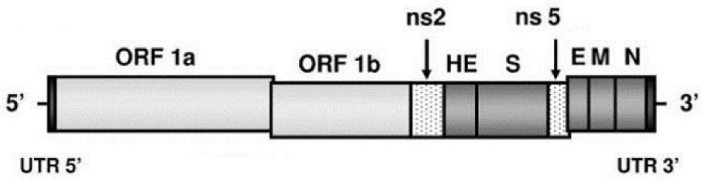


Figure 2.1 **General coronavirus genome organization based on the HCoV-OC43 lineage A betacoronavirus genome.** Replicase genes includes in ORF1a/b (light grey shading), structural genes in medium shade grey and accessory proteins are represented as speckled boxes. UTR, untranslated region; ORF, open-reading frame; ns, non-structural; HE, hemagglutinin-esterase; S, spike; E, envelope; M, membrane; N, nucleocapsid. Modified from (St-Jean *et al.*, 2006).

At the beginning of the coronavirus replication cycle, infectious virions undergo membrane fusion via a host cell receptor or a sialic acid mediated by the spike (S) protein (Graham *et al.*, 2013). Upon entry, the positive-sense, single-stranded RNA genome is released from the enveloped virus into the cytoplasm where replication takes place (Figure 2.2). The host machinery translates the overlapping open reading frames ORF1a and ORF1b by a ribosomal frame-shifting mechanism to produce two polyproteins. Cleavage of these polyproteins by virally encoded proteinases yields the components that are necessary to assemble the viral replication complex which then synthesizes the full-length, negative-strand as a template for genomic RNA (Sawicki *et al.*, 2007). Subgenomic negative-strand templates are synthesized from discontinuous transcription on the plus-strand genome and serve as templates for mRNA synthesis (Sawicki *et al.*, 2007). Nucleocapsid (N) protein and genomic RNA assemble in the cytoplasm and then bud into the lumen of the ERGIC to form the helical nucleocapsid. The envelope structure is created when the M, E, HE (if present) and S proteins bud through the intracellular membranes of the ER and Golgi apparatus and where they trigger assembly, likely through M proteins, to produce virions (Stadler *et al.*, 2003). Virion-containing vesicles travel via exocytosis towards the plasma membrane and fuse with it to release virions from the cell (Stadler *et al.*, 2003).

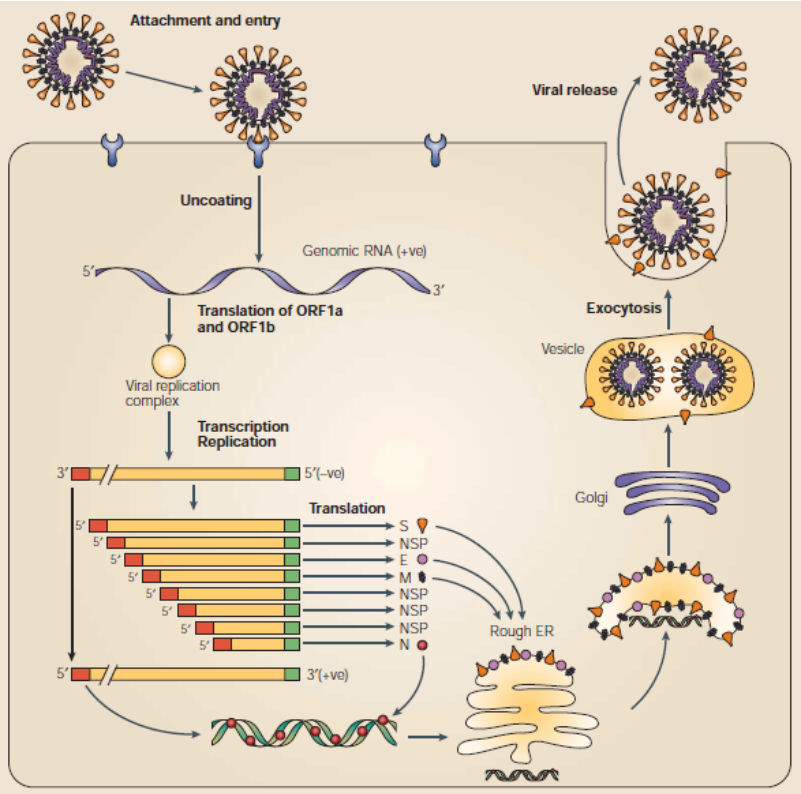


Figure 2.2 **Coronavirus replication cycle.** The coronavirus replication cycle is distinct from other enveloped viruses as virions are assembled at and bud into the lumen of the endoplasmic reticulum (ER)-Golgi intermediate compartment (ERGIC), followed by release by exocytosis rather than assembly occurring at the plasma membrane. Image taken from (Stadler *et al.*, 2003).

2.1 Coronavirus structural proteins

The four or five structural proteins found in any given coronavirus species come together to create a recoverable infectious virion of pleomorphic shape (120 to 160nm across), and all have essential roles critical for an efficient viral replication cycle, described briefly below (Figure 2.3).

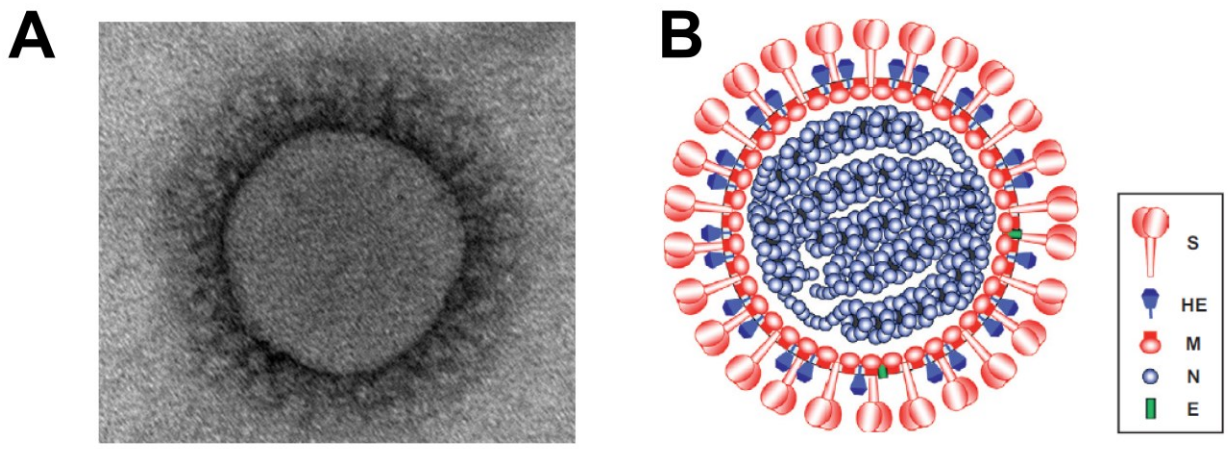


Figure 2.3 **Virion morphology and details on coronavirus virion structure.** (A) Electron micrograph showing HCoV-OC43 morphology, typical of all coronaviruses (unpublished image, Alain LeCoupanec, P.J. Talbot Laboratory). (B) Schematic representation of a betacoronavirus virion. S, spike protein; HE, hemagglutinin-esterase protein; M, membrane protein; N, nucleocapsid protein; E, envelope protein. Image taken from (de Groot *et al.*, 2012).

2.1.1 Spike protein

All coronaviruses contain a structural spike (S) protein which is well-described as the principal virulence determinant. S proteins of coronaviruses are large (approximately 1100 to 1500 amino acids in length) homo-trimeric glycoproteins that give coronaviruses their characteristic spiky appearance, or “crown” (Belouzard *et al.*, 2012). Being a Type 1 transmembrane protein, the principal function is to mediate receptor binding and membrane fusion during virion entry into target cells. The S protein is highly

glycosylated and cleaved into S1 and S2 domains in many coronaviruses where the S1 domain contains a binding domain for a host cell receptor and the S2 domain is responsible for the viral entry and cell fusion (Belouzard *et al.*, 2012, Ujike *et al.*, 2015).

2.1.2 Membrane protein

Membrane (M) proteins (approximately 230 amino acids) are the most abundant structural protein of the coronavirus envelope. Though sequence homology is low across species, the characteristic structure is retained. The M protein has a short N-terminal ectodomain, followed by three transmembrane domains (triple spanning the membrane) and a long C-terminal CT domain that is partially membrane associated thus giving an overall *N*-terminal ecto- and a *C*-terminal endodomain (Ujike *et al.*, 2015). This CT domain is hypothesized to associate with the inner leaflet of the membrane, giving coronaviruses their characteristic thick envelopes (de Groot *et al.*, 2012). All coronaviruses M protein N-terminal ectodomains contain either an O- or N-glycosylation site depending on the species. The M protein plays important roles in virus morphogenesis and is critical for assembly within the secretory pathway and budding, especially in tandem with coronavirus envelope (E) protein (Ujike *et al.*, 2015).

2.1.3 Nucleocapsid protein

The nucleocapsid protein (N) principal role is to interact with and encapsulate the coronavirus RNA genome via its phosphoprotein properties. This 349 to 470 amino acid long protein has been recently confirmed to be multifunctional. N protein has roles in regulating the virus replication cycle and modulating the host cell response such as interacting with the viral M protein during virion assembly, important for RNA synthesis, translation and transcription, having RNA chaperone activity (Zuniga *et al.*, 2010, Zuniga *et al.*, 2007) and acting as a type I interferon antagonist, among others (de Groot *et al.*, 2012, McBride *et al.*, 2014).

2.1.4 Hemagglutinin-Esterase protein

In some Betacoronaviruses (of the lineage A subgroup, including HCoV-OC43), the hemagglutinin-esterase (HE) protein represents an additional structural protein. HE proteins are approximately 400 amino acids in length, is N-glycosylated and form small surface projections in virion electron micrographs that forms a small layer in the coronavirus “crown” (de Groot, 2006). This homo-dimeric type I transmembrane envelope protein can induce reversible virion attachment to O-acetylated sialic acids by acting as a sialate-O-acetylesterase (de Groot *et al.*, 2012, X. Huang *et al.*, 2015). HE has recently been implicated as a potential virulence determinant in HCoV-OC43 as it is essential for the efficient production of infectious virions (Desforges *et al.*, 2013) and is involved in the MHV fitness of natural hosts (Lissenberg *et al.*, 2005). The HE protein is present and functional in bovine coronavirus (BCoV), HCoV-HKU1 and HCoV-OC43; in MHV the HE gene is present and expressed in some strains such as the JHM strain, while in others, such as MHV-A59, it is not expressed and acts as a pseudogene (de Groot, 2006, X. Huang *et al.*, 2015, Yokomori *et al.*, 1991).

2.1.5 Envelope protein

The envelope (E) protein is a small type I transmembrane protein of 74-109 amino acids in length found in all members of the C*oronavirinae* subfamily, showing only some amino acid sequence agreement; though it has been suggested that it is the conserved secondary structure predicted between species rather than sequence conservations which is more important for function (Cohen *et al.*, 2011, Kuo *et al.*, 2007). The protein is demonstrated to have a wide range of functions; generally, coronavirus E proteins are classified as structural proteins and possess crucial roles in virion formation and morphology of the viral envelope (de Groot *et al.*, 2012), but at the same time only a small proportion of E protein generated during infection, approximately 20 copies (Godet *et al.*, 1992), are incorporated into virions (Maeda *et al.*, 2001, Raamsman *et al.*, 2000). Coronavirus E proteins are increasingly found to perform important and significant roles within the infected cell; they are demonstrated to have

viroporin activity, are implicated within the secretory pathway where they interact with the M structural protein (Corse *et al.*, 2003, Y. C. Hsieh *et al.*, 2008), as well as mediate cellular protein-protein interactions, contributing to infection pathologies (DeDiego *et al.*, 2014). These finding have recently established the coronavirus E protein as an additional virulence factor to the S protein, especially demonstrated for SARS-CoV (DeDiego *et al.*, 2014).

3 CORONAVIRUS ENVELOPE PROTEIN CHARACTERIZATION

3.1 Protein Characteristics

Although the HCoV-OC43 E gene has been sequenced (both from reference strains and clinical samples), to date there is no published data focusing on the protein's characteristics. All present knowledge is inferred from studies done on other coronavirus species such as MHV and SARS-CoV. The aim of the present literature review is to contextualize current information on the coronavirus E protein, potentially providing insights for initial HCoV-OC43 E characterization studies which have been conducted as part of the Masters thesis work presented herein.

3.1.1 Coronavirus E protein localization

Coronaviruses assemble at intracellular membranes in the endoplasmic reticulum intermediate compartment (ERGIC) and Golgi, where they bud into the lumen and are subsequently transported out of the cell by exocytosis in cargo vesicles (Klumperman *et al.*, 1994, Krijnse-Locker *et al.*, 1994, Tooze *et al.*, 1984, Tooze *et al.*, 1985) which is in contrast to most enveloped viruses that assemble and bud at the plasma membrane. Although the E protein is found in minute quantities in the virion envelope, a large quantity is present within the secretory pathway. Real-time imaging of carboxy-terminal tetracysteine tagged MHV-A59 with the utilization of a panel of exocytic pathway cellular markers has recently shown E protein localization in the ERGIC and Golgi membranes (Venkatagopalan *et al.*, 2015) while prior work had concluded that MHV E accumulated in pre-Golgi membranes based on co-localization with Rab-1, a marker of ER and ERGIC compartments (Raamsman *et al.*, 2000). The avian infectious bronchitis virus (IBV) E protein has been reported to localize in the Golgi or pre-Golgi complexes (Corse *et al.*, 2000, Lim *et al.*, 2001).

Many studies of epitope-tagged SARS-CoV E protein have been conducted and demonstrated the presence of the protein in the ER, ERGIC or Golgi (Alvarez *et al.*, 2010, Liao *et al.*, 2006, Lopez *et al.*, 2006, Nal *et al.*, 2005, Nieto-Torres *et al.*, 2011, Teoh *et al.*, 2010) with the slight variability in localization suggested to be an artefact of epitope-tag interference; N-terminal tags could interfere with translocations into ER lumen during biosynthesis or tags on either end could affect ER export and trafficking by blocking protein-protein interaction motifs (Cohen *et al.*, 2011). Indeed, Machamer and colleagues have demonstrated that exogenously expressed, untagged SARS-CoV E protein in HeLa cells localizes to the cis-Golgi area (Cohen *et al.*, 2011). Taken together, this data shows a general localization of coronavirus E protein to the ER, ERGIC and Golgi subcellular compartments with small variations depending on the specific viral species and cell type utilized.

It must be noted that there has been some confusion within the literature regarding the localization of the coronavirus E protein, with early studies reporting that it traffics to the cell surface (Godet *et al.*, 1992, Yu *et al.*, 1994, Yuan *et al.*, 2006). Studies involving surface immunofluorescence staining, cellular permeabilization by E protein, and whole cell patch clamp measurement of ion channel activity in cells expressing SARS-CoV E protein were also taken as evidence of expression at the plasma membrane (Liao *et al.*, 2004, Liao *et al.*, 2006, Madan *et al.*, 2005, Pervushin *et al.*, 2009). Conversely, neither Altmeyer or Liu and colleagues were able to provide evidence of SARS-CoV E localization at the plasma membrane (Nal *et al.*, 2005, Yuan *et al.*, 2006). To this effect, it has recently been clearly demonstrated that E protein is not present at the plasma membrane in cells transfected with SARS-CoV or MHV E protein nor in SARS-CoV or MHV infected cells using more sensitive methods (such as immunoelectron microscopy, live cell imaging and time-course confocal microscopy) in more biologically relevant systems (Nieto-Torres *et al.*, 2011, Venkatagopalan *et al.*, 2015). The authors of these more recent studies suggest that phenomena attributed to E protein plasma membrane localization previously observed could have been due to indirect effects such as using E proteins with very large tags, utilization of different cell lines or the strength of expression systems rather than actual E protein presence (Nieto-

Torres *et al.*, 2011). Furthermore, using either untagged SARS-CoV E protein (Cohen *et al.*, 2011) or a small tetracysteine tagged MHV-A59 (Venkatagopalan *et al.*, 2015), two studies were unable to detect biotinylation nor detect the protein using immunofluorescence staining at the cell surface of infected cells. Thus presently coronavirus E protein is not considered to traffic to the cell surface.

Localization of coronavirus E proteins to the Golgi area can be attributed to the presence of one or more Golgi complex targeting signals present on the protein. The cytoplasmic domain of the IBV E protein has been found to be sufficient for Golgi targeting with the targeting information identified to be between amino acids 13 and 63 (Corse *et al.*, 2000, Corse *et al.*, 2002). Similarly, a study of SARS-CoV E protein using a VSV-G glycoprotein/SARS-CoV E cytoplasmic tail chimeric protein demonstrated that Golgi complex-targeting signals are also present in the cytoplasmic tail portion of the protein (Cohen *et al.*, 2011). In this region, a predicted β-hairpin secondary structure is formed around a strongly conserved proline residue which is critical for Golgi targeting. However, abrogation of the secondary structure motif within a full length SARS-CoV E protein did not disrupt Golgi targeting; by using a reciprocal chimeric protein, this led to the discovery of a second Golgi targeting signal in the N-terminal end of the E protein (Cohen *et al.*, 2011). Sequence analysis of the β-hairpin motif region suggests that the β-hairpin structure is conserved among beta and gamma coronaviruses, but not alpha; indeed, MHV and IBV E proteins were found to utilize this predicted secondary structure for Golgi targeting in the same study (Cohen *et al.*, 2011).

Machamer and colleagues suggested that redundancy in Golgi complex targeting information may be an evolutionary mechanism which ensures proper localization of this critical protein during assembly, or alternatively, may be used in two different pools of E protein, depending on whether or not the E protein is incorporated into virions (Cohen *et al.*, 2011) (tested and elaborated by (Westerbeck *et al.*, 2015); see Section 3.7 below). Furthermore, a recent study where small sections of the C-terminal end of the SARS-CoV E protein were deleted independently of each other, producing recombinant

viruses that did not contain the demonstrated Golgi complex targeting signal, nonetheless showed no difference in localization compared to wild-type virus, further suggesting redundancy within the protein (Regla-Nava *et al.*, 2015). A recent biochemical study, the first to determine the structure of the transmembrane and C-terminal of a coronavirus E protein, determined that this conserved C-terminal domain forms a β-coil-β motif structure on its own, but in the context of full-length SARS-CoV E protein, adopts an α-helical structure (Y. Li *et al.*, 2014). The authors suggest that the β-structure may exist in dynamic equilibrium in smaller proportions compared to the α-helical form and affect processes in the infected cell such as membrane scission, protein binding, or localization (Y. Li *et al.*, 2014).

3.1.2 Coronavirus E protein palmitoylation

Coronavirus E proteins contain two to four cysteine residues, depending on the viral species, located in the cytoplasmic tail more proximally to the transmembrane domain side rather than the C-terminal end (Figure 3.1). These cytoplasmic residues have been found to be post-translationally cysteine palmitoylated in IBV (Corse *et al.*, 2002), SARS-CoV (Liao *et al.*, 2006, Tseng *et al.*, 2014) and MHV (Boscarino *et al.*, 2008, Lopez *et al.*, 2008, Yu *et al.*, 1994) however it is not known if these cysteine residues are modified at the same time or undergo dynamic palmitoylation/depalmitoylation cycling during infection (Venkatagopalan *et al.*, 2015). Results from Lopez *et al.* demonstrate that disulphide linkages are not present between cysteine residues of the MHV E protein (Lopez *et al.*, 2008), however Wang and colleagues demonstrated that intracellular SARS-CoV E protein was in a monomeric form and released E protein was dimeric, suggesting the formation of disulphide bridges (Tseng *et al.*, 2014). In MHV (Lopez *et al.*, 2008) and IBV (Corse *et al.*, 2002), cysteines and their palmitoylation do not appear to play a role in Golgi targeting. However, the former is demonstrated to be important in protein stability and virus-like particle (VLP)/virus production. Conversely, SARS-CoV E protein cysteine residue

palmitoylation does not appear to affect VLP release or protein assembly (Tseng *et al.*, 2014) but is still important for membrane association (Liao *et al.*, 2006).

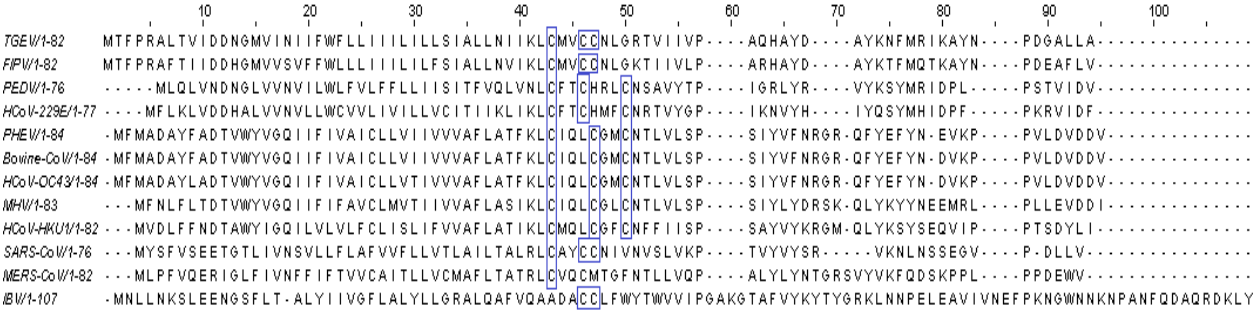


Figure 3.1 **Sequence alignment of representative coronavirus E protein sequences with cysteine residues able to be palmitoylated indicated by blue outline.** Sequence alignment produced using ClustalW2. Modified from data presented in (Ruch *et al.*, 2012a, Ruch *et al.*, 2012b).

Palmitoylation is suggested to concentrate the coronavirus E protein to its primary subcellular localization site by increasing membrane association that would facilitate roles in assembly and/or interactions with resident Golgi complex proteins or membranes (Liao *et al.*, 2006, Lopez *et al.*, 2008, Venkatagopalan *et al.*, 2015). For example, given that the SARS-CoV E protein cysteine residues are just upstream of the predicted β-hairpin Golgi complex-targeting motif (discussed in the Section 3.1), Machamer and colleagues hypothesized that E protein targeting resident Golgi complex membrane proteins could be mediated by the proximally-located palmitoylated residues of the E protein (Cohen *et al.*, 2011). It has also been observed that the E protein undergoes a range of mobilities at the site of assembly early in infection (Venkatagopalan *et al.*, 2015); the number of palmitoylated residues on the protein could affect this movement. It has been suggested that lateral movement of the protein in membranes at the site of assembly could facilitate the membrane curvature and scission properties previously observed for MHV E protein and facilitate the budding process (Venkatagopalan *et al.*, 2015, Vennema *et al.*, 1996). Gallagher and colleagues suggest that palmitoylation could affect how the E protein interacts with the M protein at

the site of assembly, allowing for proper virus morphogenesis and efficient release (Boscarino *et al.*, 2008).

3.1.3 Coronavirus E protein topology

The emerging picture surrounding coronavirus E protein localization within infected cells and potential membrane-association palmitoylation sites in this small protein can help explain the relative range of coronavirus E protein topology descriptions found in the literature. At the time of the discovery of the first coronavirus E protein in transmissible gastroeneteritis virus (TGEV), it was determined that the protein follows a C_{exo}-N_{endo} orientation (amino terminus oriented towards the lumen of intracellular membranes and carboxy terminus facing the cell cytoplasm) following a Type I transmembrane protein classification (Godet *et al.*, 1992). This orientation has been confirmed more recently in studies of E proteins of MHV-A59 (Venkatagopalan *et al.*, 2015) and SARS-CoV (Nieto-Torres *et al.*, 2011) using untagged or proteins with small tags in transfected and infected cells. Conversely the cytoplasmic tail of MHV E was previously determined to reside in the cytoplasm of infected cells, but with a FLAG tag placed at the amino end it was suggested that the protein rather spans the lipid bilayer twice, with both ends in the cytoplasm (Maeda *et al.*, 2001); a similar hairpin topology was suggested in C- and N-terminally FLAG tagged SARS-CoV E protein (Yuan *et al.*, 2006) a concept originally suggested by Arkin and colleagues (Arbely *et al.*, 2004). These latter studies highlight the potential pitfalls of a topological studies with large-tagged proteins but do not necessarily rule out two topological confirmations for coronavirus E proteins. Furthermore, palmitoylation of the coronavirus E protein could render topological study interpretation more difficult due to possible close and strong associations of the C-terminal of the protein with membranes within the cell; extensive structural studies have been hampered by problems in expression, purification and stabilization of E proteins.

To help explain these conflicting studies, work on IBV by Ruch and Machamer suggests that the E protein may adopt multiple topologies during infection (Ruch *et al.*,

2012b). Versions of IBV E protein that were forced into either a transmembrane or hairpin conformation were produced; the former conformation could alter the secretory pathway while the latter was more efficient in forming VLP particles (Ruch *et al.*, 2012b). Further confirmation of two distinct populations with distinct functions of IBV E protein was recently demonstrated in the context of infection (Westerbeck *et al.*, 2015). Furthermore, a recent structural study of a truncated form of the SARS-CoV E monomer that includes the transmembrane domain and C-terminal tail containing a Golgi-complex targeting signal demonstrated a dynamic equilibrium between an α and β secondary structure form, with each structure possibly facilitating different processes within an infected cell (Y. Li *et al.*, 2014).

3.1.4 Other post-translational modifications of coronavirus E protein

Other post-translational modifications, other than palmitoylation, of coronavirus E proteins have been observed, however their functional roles have not been extensively investigated. Transiently expressed flag-tagged SARS-CoV E protein was shown by Liu and colleagues to be N-glycosylated at the asparagine at position 66 and it is hypothesized that it could influence the topology of the protein (Yuan *et al.*, 2006). It has also been noted by the same authors that this asparagine is not conserved in the E protein of other coronaviruses, and may be a unique feature of SARS-CoV (Yuan *et al.*, 2006). A previous study by Nal *et al.* of the subcellular distribution and maturation of SARS-CoV surface proteins determined that, as opposed to S, M, and HE proteins, C-terminally tagged E protein was not glycosylated (Nal *et al.*, 2005). This lack of detection could be due to the methods employed in this study, however, the E protein of other coronaviruses are also thought not to be glycosylated (de Groot *et al.*, 2012). Altmeyer and colleagues also demonstrated that E protein is quickly degraded by pulse-chase labelling experiments (Nal *et al.*, 2005), which could be explained by later findings that SARS-CoV E protein is ubiquitinated (Alvarez *et al.*, 2010, Keng *et al.*, 2011), however further investigation is merited.

3.2 Importance of coronavirus E protein in virus production

As observed by E protein deletions in a wide variety of coronavirus species, it is clear that this protein is a very important component of efficient viral production. Different coronavirus species show what can generally be classified as two different phenotypes when taking into account E protein dispensability: completely indispensible (essential for virion production; without E protein no virus can be detected) and partially dispensible (very important for virion production; without E protein a decreased titer can be detected). In the case of TGEV (α genus) (Curtis *et al.*, 2002, Ortego *et al.*, 2007, Ortego *et al.*, 2002) and MERS-CoV (β genus) (Almazan *et al.*, 2013), E protein deletion leads to replication-competent but propagation deficient phenotypes; no viruses can be detected upon deletion. Alternatively, deletion of the E protein results in a more than 1000-fold decrease in MHV titers (Kuo *et al.*, 2007, Kuo *et al.*, 2003) and a more modest decrease of 20-200-fold in SARS-CoV (DeDiego *et al.*, 2007); both β genus coronaviruses. SARS-CoV without E protein has been demonstrated to be attenuated *in vivo* (DeDiego *et al.*, 2007, DeDiego *et al.*, 2008). To date, there is no published information regarding the necessity of HCoV-OC43 E protein for infectious virion production and this will represent an important section of the presented Masters thesis work found in Chapter 4.

3.2.1 E protein flexibility in virus assembly

The E protein secondary structure appears to be flexible enough to be able to accept the E protein from certain species of coronaviruses, suggesting primary sequence specificity has reduced importance in virion formation (Kuo *et al.*, 2007). Substitution mutants, where MHV E proteins were switched with different coronavirus species E proteins, demonstrated that γ- and other β coronaviruses were able to produce recombinant viruses indistinguishable to wild-type, whereas α coronavirus E protein substitution into MHV (β genus) was inactive, but could recover virus production with one amino acid substitution (Kuo *et al.*, 2007). Interestingly, passage of a MHV E

protein-deleted virus showed that after serial passage, revertants arose with partial duplication of the M gene containing its three transmembrane domains but lacking a cytoplasmic tail (Kuo *et al.*, 2010).

Early coronavirus studies suggested that E and M proteins were sufficient for VLP formation in IBV, MHV, TGEV and bovine coronavirus (BCoV) (Baudoux *et al.*, 1998, Corse *et al.*, 2000, Vennema *et al.*, 1996) and the cytoplasmic tails of the same proteins were shown to interact in IBV (Corse *et al.*, 2003, Lim *et al.*, 2001) and SARS-CoV (S. C. Chen *et al.*, 2009, Y. C. Hsieh *et al.*, 2008). However more recent results using SARS-CoV appear contradictory to these studies as they report M and N (Yue Huang *et al.*, 2004), M and E (Ping-Kun Hsieh *et al.*, 2005) or even M alone (Almazan *et al.*, 2013) is sufficient for VLP formation. A possible explanation is that the expression systems used in VLP production (e.g. overexpression of viral membrane particles and/or vaccinia based-expression systems) could produce artefacts and/or would represent non-recoverable particles in the context of infection (Ruch *et al.*, 2012a). This hypothesis is supported by the findings that transient transfections to express proteins from plasmids in the presence of N protein can greatly increase VLP yield in MHV (Boscarino *et al.*, 2008) and SARS-CoV (Siu *et al.*, 2008). Ruch and Machamer suggest that N protein may play a role in producing a complete virion that can be detected during the infection of cells but not necessarily important for envelope formation (Ruch *et al.*, 2012a). A second explanation would be that the role of E protein in assembly is not to directly interact with the M protein, but to induce membrane curvature by allowing enough space between aggregates of M protein during virion formation (Kuo *et al.*, 2007). Another function independent of M protein interaction could be the induction of scission events at the ERGIC; mutations in the C-terminal tail of MHV E protein have been shown to produce elongated and unstable virions, possibly due to failed membrane scission (Fischer *et al.*, 1998).

3.3 Viroporins

Many RNA viruses have been found to have common ion channel properties through their respective small transmembrane proteins, designated viroporins (Table 3.1). Viroporins were first predicted after enhanced membrane permeability to ions and small molecules was noted in several virus-cell systems of poliovirus with the viral agent itself later implicated in this destabilization (Carrasco, 1978, Carrasco *et al.*, 1976). After further studies on poliovirus 2B and 3B proteins (Aldabe *et al.*, 1996, Doedens *et al.*, 1995, Lama *et al.*, 1992) and coxsackievirus 2B protein (van Kuppeveld *et al.*, 1997), the idea of a viroporin family of proteins modulating the viral replication cycle and host cell response was solidified (Carrasco, 1995). Since this initial characterization many additional viruses have been identified as modifying membrane permeability and studied for viroporin activity including: human respiratory syncytial virus (HRSV) SH protein (Perez *et al.*, 1997), Semliki forest virus (SFV) 6K protein (Sanz *et al.*, 1994), influenza A virus M2 protein (Pinto *et al.*, 1992, Pinto *et al.*, 2006, Sakaguchi *et al.*, 1996), hepatitis C virus (HCV) p7 (Carrere-Kremer *et al.*, 2002, Griffin *et al.*, 2003, Pavlovic *et al.*, 2003), and human immunodeficiency virus (HIV-1) viral protein U (Vpu) (Ewart *et al.*, 1996, Gonzalez *et al.*, 1998); the last three examples being the most intensely studied along with poliovirus 2B and 3A proteins. In addition, viroporins have been found in small DNA viruses such as JC polyomavirus, Simian virus 40 and human papillomavirus type 16, summarized in review done by Nieva *et al.* (Nieva *et al.*, 2012)). Interestingly, SARS-CoV 3a (Lu *et al.*, 2006) and 8a (C. C. Chen *et al.*, 2011) proteins, HCoV-229E 4a protein (R. Zhang *et al.*, 2013) as well as the HCoV-OC43 ns5 protein (R. Zhang *et al.*, 2015), have also recently been proposed as likely viroporins, showing the importance of the implicated function in the general coronavirus replication cycle. In the coronavirus field, the E protein across several species has recently been implicated as an ion channel/viroporin and will be discussed in more detail below.

Table 3.1 **List of selected viroporins indicating their RNA or DNA family, virus species origin and size.** Adapted from (Gonzalez *et al.*, 2003) with information from (C. C. Chen *et al.*, 2011, Lu *et al.*, 2006, Nieva *et al.*, 2012, R. Zhang *et al.*, 2013, R. Zhang *et al.*, 2015).

	Virus Family	Virus (viroporin protein)	Amino acid residues
RNA	*Picornaviridae*	Poliovirus (2B)	97
		Poliovirus (3A)	87
	Togaviridae	Semliki Forest Virus (6K)	60
	Retroviridae	HIV-1 (Vpu)	81
	Paramyxoviridae	HRSV (SH)	64
	Orthomyxovirus	Influenza (M2)	97
	Flaviviridae	Hepatitis C virus (p7)	63
	Coronaviridae	SARS-CoV (E)	76
		SARS-CoV (3a)	274
		SARS-CoV (8a)	39
		HCoV-229E (4a)	133
		HCoV-OC43 (ns5)	109
DNA	*Polyomaviridae*	JC polyomavirus (agnoprotein)	71
		Simian virus 40 (VP4)	125
	Hepadnaviridae	Hepatitis B virus 16 (E5)	83

Generally viroporins are small transmembrane proteins of approximately 60 to 120 amino acids in length and have the capacity to oligomerize, thereby permeating membranes (Table 3.1). This property allows viroporins to participate in a wide range of functions including genome replication, virus particle entry/release, and assembly of virus particles in infected cells while also interrupting a number of important physiological processes within the host cell. The importance of viroporins to the viral replication cycle and pathogenesis is underlined by the fact that deletion of a viroporin-

encoded gene greatly diminishes the formation of viral progeny, that viral infectivity can be rescued in *trans* by complementation of their own viroporin genes and in some cases can be rescued by other viroporins containing biochemical and/or structural properties of sufficient similarity (Brohm *et al.*, 2009, Gonzalez *et al.*, 2001, Wozniak *et al.*, 2010).

Viroporins have the ability to form ion channels and/or size-limited pores to allow the passage of small solutes and/or ions. Upon oligomerization, the influenza A M2 protein is able to produce tetrameric channels that conduct protons (Pinto *et al.*, 1997) and HCV p7 has a similar conductivity when forming heptameric channels (Griffin *et al.*, 2003), while HIV-1 Vpu oligomers can function as ion-conducting channels that open in a voltage-dependent or voltage-independent manner (Mehnert *et al.*, 2008, Schubert *et al.*, 1996). This weak ion selectivity, and the additional capacity to induce leakage of solutes into and out of the cell across a sealed membranes, can affect cellular and viral processes (Mehnert *et al.*, 2008). In the case of poliovirus 2B protein, it can cause the entry of extracellular Ca^{2+} through the plasma membrane and/or Ca^{2+} leakage from intracellular stores (Aldabe *et al.*, 1996, Aldabe *et al.*, 1997) which could induce apoptosis, as described by other proteins such as HIV-1 protein R (Chami *et al.*, 2003, Jacotot *et al.*, 2000).

Some viroporins modify intracellular membranes in order to produce favourable conditions for genome replication, best described for poliovirus protein 2B and 2C precursor, 2B2C. Transgenic expression of 2B2C remodels the ER membranes leading to the formation of a double-membrane multivesicular body, termed ‘viroplasm’, required for viral RNA replication (Suhy *et al.*, 2000), although the exact mechanisms leading to this extensive remodelling are not known. Trafficking can also be influenced by viroporins as seen with influenza A virus M2 or HCV p7 which delay trafficking of cellular and viral glycoproteins by proton redistribution via formed ion channels, thereby preventing the acidification of intracellular compartments (vesicles and the Golgi and trans-Golgi network); this alteration has an effect on the production of infectious particles (Sakaguchi *et al.*, 1996, Wozniak *et al.*, 2010). Influenza A M2 protein ion

channel activity equilibrates the pH gradient between the trans Golgi network and the cytosol to prevent premature maturation of the viral protein hemagglutinin (Pielak *et al.*, 2011, Pinto *et al.*, 2006). p7 of HCV is needed for polyprotein processing (Brohm *et al.*, 2009) and is essential in a late step of viral assembly and release of infectious virions (Jones *et al.*, 2007, Steinmann *et al.*, 2007).

In the production of viral particles of HCV, p7 is not only important in its viroporin activity, but also on its interactions with other proteins (possibly NS2, a non-structural HCV protein), however the mechanism for this relationship is not yet known (Steinmann *et al.*, 2007). In the same vein, the transmembrane domain of Vpu is known to interact with cellular proteins, including CD4 and tetherin. Vpu antagonizes the innate immune response to facilitate morphogenesis and release by degrading newly synthesized CD4 thereby downregulating its concentration at the cell surface and allowing increased viral budding, as well as by counteracting the interferon-induced host restriction factor, tethrin, to promote the release of new virions (reviewed in (Nieva *et al.*, 2012)). Functions outside the classical ion channel activity described for viroporins such as interactions with other viral or cellular components is currently an emerging concept in the field.

3.3.1 Coronavirus E protein – a viroporin

Shortly after the discovery of SARS-CoV, its E protein was found to be capable of altering the permeability of *E. coli* cells (Liao *et al.*, 2004) and mammalian cells (Liao *et al.*, 2006). A similar study was conducted with the MHV E protein with corroborative results (Madan *et al.*, 2005). This membrane permeabilization ability was subsequently mapped to the E protein transmembrane domain (Liao *et al.*, 2006) and led to the suggestion that coronavirus E proteins could participate in pore formation. Molecular simulation of the transmembrane domain of 13 coronavirus species and subsequent *in vitro* oligomerization studies with a synthetic SARS-CoV E protein transmembrane domain peptide in SDS micelles, determined dimeric, trimeric, and pentameric

aggregates as possible conformations (Torres *et al.*, 2005). The ability of SARS-CoV E protein to produce pentameric pores in lipid membranes was subsequently confirmed by various biochemical methods, including NMR (Figure 3.2) (Parthasarathy *et al.*, 2008, Pervushin *et al.*, 2009, Torres *et al.*, 2006). Biochemical analysis has demonstrated that IBV E purified from bacteria can form an oligomer consistent with a homopentamer conformation (Parthasarathy *et al.*, 2012). More recently, purified MERS-CoV E protein was found to form homo-oligomeric pentameric channels in lipid bilayers via its α-helical transmembrane domain (Surya *et al.*, 2015).

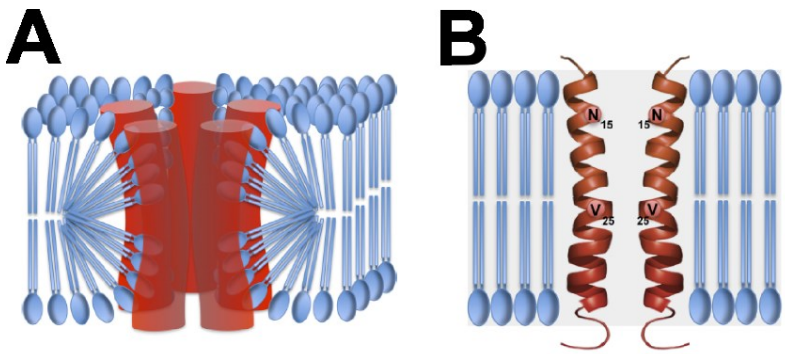


Figure 3.2 **SARS-CoV E protein homopentameric ion channel structure based on NMR study.** (A) General structure of homo-oligomeric pore made of E protein monomers represented as red cylinders with surrounding phosophlolipids in blue (DeDiego *et al.*, 2014). (B) Simplified cross-section of the SARS-CoV E protein ion channel with phospholipids in blue and E protein transmembrane α-helix secondary structure detailed in red. Amino acids important for SARS-CoV ion channel activity are indicated, N15 and V25, with equivalent positions in HCoV-OC43 being Q17 and L27 respectively (Nieto-Torres *et al.*, 2014). Original NMR study conducted by Torres and colleagues (Pervushin *et al.*, 2009).

The discovery of the pentameric conformation of SARS-CoV E protein channels led to the hypothesis that these conformations could serve as ion channels. Indeed, even before the first study determining the structure of the coronavirus E protein channels was conducted, a study was done wherein planar lipid bilayers demonstrated ion channel activity (small cation selectivity) mediated by a synthetic peptide corresponding to full-length SARS-CoV E protein (Wilson *et al.*, 2004) and which was later confirmed (Torres *et al.*, 2007). Chemical inhibition screens have demonstrated that HCoV-229E, MHV and IBV have cation ion channel activity in planar lipid bilayers

(Wilson *et al.*, 2006). Ion channel selectivity of E proteins appears to be related to the genus of the virus with α genus coronaviruses (e.g. HCoV-229E) having preferential selection for K^+ over Na^+, whereas β and γ genus coronaviruses (e.g. MHV/SARS-CoV and IBV, respectively) prefer the inverse (Wilson *et al.*, 2006, Wilson *et al.*, 2004). However, more recent results in lipid membranes broadly mimicking ERGIC/Golgi membrane composition, the general localization of coronavirus E protein during infection, cation channel selectivity and conductance has been further shown to be regulated by the lipid composition and charge of the membrane surrounding the channel composed of either full length SARS-CoV E protein or HCoV-229E E protein derived peptide (Verdia-Baguena *et al.*, 2013, Verdia-Baguena *et al.*, 2012). Surprisingly, very recently is has also been demonstrated in the same ERGIC/Golgi-like membranes that SARS-CoV E protein is permeable to Ca^{2+} (Nieto-Torres *et al.*, 2015). Taken together, these results strongly suggest that E protein ion channel activity is a general characteristic of coronaviruses in the context of infection and it is now considered that coronavirus E proteins act as viroporins.

3.3.2 Importance of coronavirus E protein ion channel activity

As described above, the coronavirus E protein is important for virion biogenesis, however, more specifically, the transmembrane domain is responsible for ion channel activity and is an important domain implicated in virus maturation and production. A series of alanine substitutions within the MHV transmembrane domain were found to produce crippled viruses that eventually reverted to restore the structural integrity of α-helix structure within the transmembrane domain to restore virus production (Y. Ye *et al.*, 2007). The importance of the transmembrane domain was further demonstrated with substitution mutants created by replacing the genus β CoV MHV E protein transmembrane domain with E protein transmembrane domain of other coronavirus genera. Domains belonging to other genus β and γ but not α coronaviruses where able to functionally replace the MHV E transmembrane domain when measuring for viral production; an effect suggested to be due to the different ion selectivity of these domains (Kuo *et al.*, 2007). In another case, replacement of the IBV E protein

transmembrane domain with vesicular stomatitits virus (VSV) G protein, that did not contain ion channel activity, interfered with efficient trafficking and release of virions in infected cells (Ruch *et al.*, 2011).

In order to further elucidate the possible function(s) of coronavirus E protein transmembrane domain, structural information was needed. An NMR structural study of the SARS-CoV E protein transmembrane domain by Torres and colleagues provided much of the basis for recent ion channel mutagenesis work (Figure 3.2) (Pervushin *et al.*, 2009). From the resolved structure, the authors speculated that the asparagine at position 15, being located at the top of the transmembrane channel, acts as an ion selectivity filter. Based on consensus sequence analysis (Figure 3.3), the authors suggested that larger, polar amino acids at this position in other coronavirus species (e.g. MHV), could explain the observed higher selectivity for ions (Pervushin *et al.*, 2009). This same amino acid in SARS-CoV E protein was subsequently confirmed to be important for ion channel activity in planar lipid bilayers and mammalian cells, with the valine at position 25 also determined to be important for this function (Verdia-Baguena *et al.*, 2012). This latter residue was found in the NMR study to be the most restricted portion of the SARS-CoV E protein channel (Pervushin *et al.*, 2009). With their SARS-CoV E protein transmembrane domain NMR structure, Pervushin *et al.* were also able to map hexamethylene amiloride sensitivity/amiloride insensitivity to the biochemical characteristics of the residue at position 15, among several others (Pervushin *et al.*, 2009). Interestingly, targeted mutagenesis studies of IBV showed that mutating the threonine at position 16, equivalent to the amino acid 15 postion in SAR-CoV and predicted to inhibit ion channel activity in the E protein of the virus, did not affect virus-like particle formation. This is hypothesized by the authors to be due to the lack of this amino acid being charged in this virus species, preventing it from acting as a selectivity filter for the ion channel (Ruch *et al.*, 2012b).

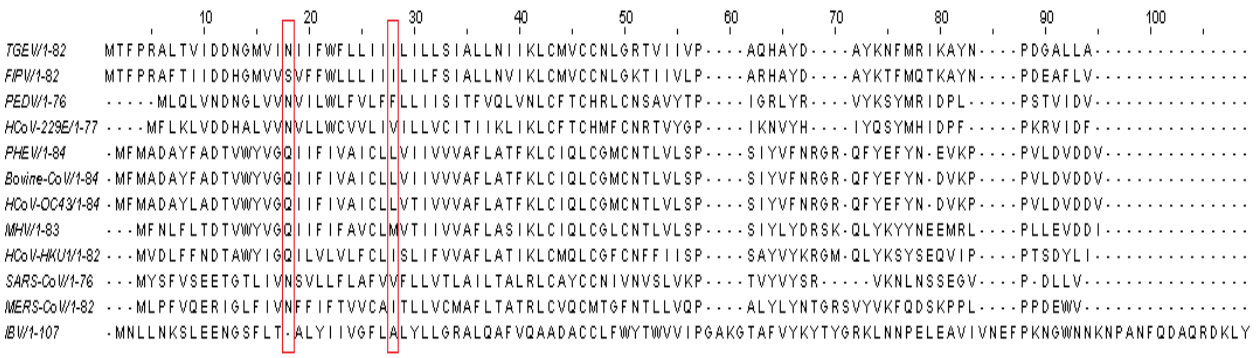


Figure 3.3 **Sequence alignment of coronavirus E proteins with amino acids implicated in ion channel activity highlighted in red.** Sequence alignment produced using ClustalW2. Modified from data presented in (Ruch *et al.*, 2012a, Ruch *et al.*, 2012b).

The implication of coronavirus E protein ion channel activity in the host cell response and pathogenesis is an area of active research. Single point mutations (N15A and V25F) were introduced into the transmembrane domain of SARS-CoV E protein in order to determine the contribution of ion channel activity to pathogenesis (Nieto-Torres *et al.*, 2014). Recombinant viruses with ion selectivity prevented (A substitution for N) or ion passage blocked (F in place of V) were deficient in ion channel activity and showed a tendency to evolve to restore E ion channel structure and activity after serial passaging in cell cultures. Ion channel activity showed a clear role in viral fitness, as viruses with ion channel activity quickly outcompeted those without during a co-infection experiment. In mice, infection with viruses deficient in ion channel activity correlated with decreased inflammatory response and increased survival (Nieto-Torres *et al.*, 2014). The authors conclude that ion channel activity of SARS-CoV E protein is a virulence determinant.

3.3.3 Impact of coronavirus infection on the secretory pathway

It has long been observed that viroporins disrupt the host secretory pathway to their own benefit, and the same has been observed in various coronavirus species for some time, specifically MHV (David-Ferreira *et al.*, 1965, Ruebner *et al.*, 1967, Svoboda *et al.*, 1962) and more recently for SARS-CoV (Snijder *et al.*, 2006) and IBV (Corse *et*

al., 2000). It has been observed that during MHV infection, the Golgi complex becomes fragmented and dissociated from its usual localization adjacent to the nucleus (Lavi *et al.*, 1996). The discovery of SARS-CoV and subsequent implication of the E protein as a possible viroporin lead to a re-evaluation of previous observations of transmembrane domain modification studies and to new studies involving various coronavirus E proteins. Deletion of E protein showed that TGEV virions were assembled in infected cells, but appeared arrested in the secretory pathway and were non-infectious (Ortego *et al.*, 2007, Ortego *et al.*, 2002) while SARS-CoV lacking E gene showed accumulation of aberrant virions intracellularly, suggesting an inability to be sorted without E protein (DeDiego *et al.*, 2007).

A recent, in-depth study has determined, through immuno-electron microscopy, that extensive membrane reorganization plays a major role in the MHV viral replication cycle, with Golgi cisternae enlarged and fragmented during infection possibly in order to aid the trafficking of virions (Venkatagopalan *et al.*, 2015). Subsequently, it was determined that the IBV E protein can disrupt the Golgi structure (Ruch *et al.*, 2011). Even though the same authors did not observe this phenomenon for IBV (Corse *et al.*, 2000), they suggest this to be due to differences in cell types and expression systems used between experiments (Ruch *et al.*, 2011). This modification of the host secretory pathway was specifically mapped to the hydrophobic domain of the IBV E protein (Ruch *et al.*, 2011) and later to the threonine at amino acid residue position 16 in this domain (Ruch *et al.*, 2012b). These results have led to speculation that the E protein acts as an ion channel in the secretory pathway, driving the rearrangement of secretory organelles, modifying the luminal environment and playing a role in trafficking of virions (Ruch *et al.*, 2012a).

3.4 Importance of coronavirus E protein ion channel activity in virus production

As discussed previously, the coronavirus E protein as a whole has been demonstrated to be very important for infectious virion formation among multiple species (Almazan *et al.*, 2013, DeDiego *et al.*, 2007, Kuo *et al.*, 2003, Ortego *et al.*, 2007, Ortego *et al.*, 2002). Several research groups have engaged in studies in order to determine which domain(s) of the E protein could be implicated in viral production. One series of hypothesis implicated the transmembrane domain of the E protein, where the integrity of the E protein α-helix is critical for virion formation. Ye and Machamer demonstrated that alanine substitution mutational studies done with the transmembrane domain of MHV, resulted in significantly reduced titers in cell culture and correlated to disruption of the helical pitch of the domain (Y. Ye *et al.*, 2007). This suggests that the structural integrity of the transmembrane domain is important for virus assembly where transmembrane domain disruption could prevent pore formation, and by extension, ion channel activity, either of which could influence virion biogenesis. Indeed, in an experiment where the transmembrane domain of IBV E protein was replaced with a heterologous version that did not have ion conductance, resultant infectious virus was poorly excreted into the external media and remained within the cells (Ruch *et al.*, 2011).

Attempts to link ion channel activity to viral production have been made with further targeted transmembrane domain mutational studies. Introduction of amino acid substitutions at position N15A and V25F in the transmembrane domain of SARS-CoV E protein, suggested to block ion channel activity, resulted in marginally reduced titers (less than 5-fold) in Vero E6 cells (Verdia-Baguena *et al.*, 2012). Studies where the threonine at position 16 in the IBV E protein transmembrane domain (equivalent to N15 in SARS-CoV) is mutated to an alanine do not affect VLP production in cell culture, however a link to secretory pathway modification is seen (Ruch *et al.*, 2012b). This suggests that while ion channel activity mediated by the N15/T16 positions in SARS-

CoV and IBV respectively, could play a role in secretory pathway disruption, they do not necessary play a role in virus production. Furthermore, in-depth studies with ion channel activity-deficient mutants N15A and V25F in SARS-CoV showed only minor differences in growth rates in human Vero E6 cells and mouse DBT-mACE2 cells compared to parental viruses, nor do titers vary in intranasally infected mice compared to wild-type infected and N15A infected mice (Nieto-Torres *et al.*, 2014), indicating this specific property of coronavirus E protein is likely not a major factor in virion production, thought could have greater implications in cellular membrane reorganization or pathology in mice. It may be that the structure of the transmembrane domain may be more important in terms of mediating virus-virus or virus-host protein interactions which could directly affect virus assembly and release through the exocytic pathway, as suggested by Ye and Hogue (Y. Ye *et al.*, 2007). At the same time, interactions between proteins are often based on particular specific domains or sequences.

3.5 PDZ domain-containing proteins

PDZ domains are modular protein-protein recognition domains acting as scaffolding complexes that mediate the assembly of numerous intracellular components and cellular functions. PDZ domains, an acronym of the first three discovered proteins of this class, PSD-95 (postsynaptic density-95), the *Drosophila* tumour suppressor protein Dlg-1 (discs large 1) and the tight junction protein ZO-1 (zonula occludens 1) (Cho *et al.*, 1992, Kennedy, 1995, Willott *et al.*, 1993, Woods *et al.*, 1993) take on a characteristic structure composed of approximately 80-90 amino acid residues with six β-strands and two α-helices forming an antiparallel β-barrel which contains a cellular protein interaction/binding cleft between one of the β-strands and α-helices (Figure 3.4). PDZ domains are very abundant and found in bacteria, fungi and metazoans with 250 to 900 PDZ proteins reported within the human proteome (Javier *et al.*, 2011, Letunic *et al.*, 2012, Spaller, 2006). However, the higher end estimations were since shown to be in erratum with independent bioinformatic analyses converging at approximately 270 PDZs in the human proteome (Luck *et al.*, 2012). The abundance of this class of protein interaction domain is not surprising given that PDZ containing proteins are often

cytoplasmic or membrane adapter proteins concentrated at specialized subcellular sites, such as epithelial tight junctions (Humbert *et al.*, 2003), neuronal post-synaptic densities (Feng *et al.*, 2009, Kim *et al.*, 2004), and immunological synapses of T cells (Ludford-Menting *et al.*, 2005). These localizations logically support the findings that PDZ domains are involved in important cellular interactions including: components of multi-domain scaffolding structures, the maintenance of cell-cell junctions or cellular polarity, and in signal transduction pathways (F. Ye *et al.*, 2013). Mutations in cellular PDZ-containing proteins and their interaction ligands, as well as differential expression of PDZ-containing proteins, have been linked to cancer and associated with neurological disorders in humans (Facciuto *et al.*, 2012, Toyooka *et al.*, 2002, F. Ye *et al.*, 2013).

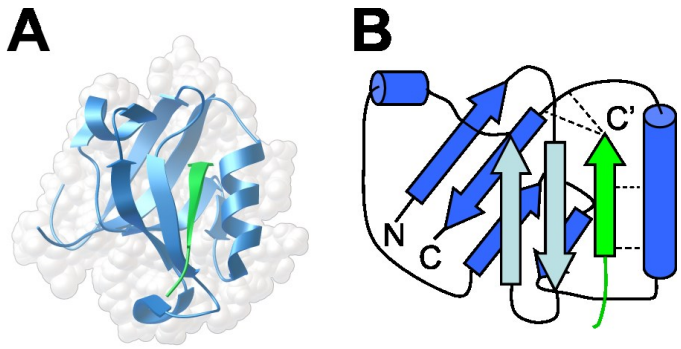


Figure 3.4 **Representative PDZ domain-containing protein interacting with the canonical C-terminal end PDZ binding motif present on the target ligand.** Schematic representation in ribbon diagram with α-helices as spirals and β-sheets as arrows (A) and simplified 2-dimensional topological diagram with α-helices as tubes and β-sheets as arrows. PDZ domain-containing protein represented in shades of blue and C-terminal end of ligand represented in green. Panel A taken from the PSI Structural Biology Knowledgebase (Gabanyi *et al.*, 2011) and Panel B inspired from (F. Ye *et al.*, 2013).

Domain organization analysis has provided general characteristics of PDZ domain containing proteins; PDZ domains often co-exist with a wide variety of other protein-protein interaction elements such as SH3, L27 or leucine-rich repeat (LRR) domains or other PDZ domains (Figure 3.5). Additionally, PDZ domains can contain multiple repeats and can also form large structures while generally lacking any intrinsic catalytic activities, thereby acting largely as scaffolding proteins (Javier *et al.*, 2011, F. Ye *et al.*, 2013). PDZ domains act as platforms for several types of protein-protein interactions within the cell, with interactions with phospholipids or internal protein motifs being less common. The majority of studied PDZ domain interactions recognize a specific, extreme C-terminal sequence on a target cellular protein and follow a canonical amino acid residue pattern, called a PDZ-binding motif, or PBM.

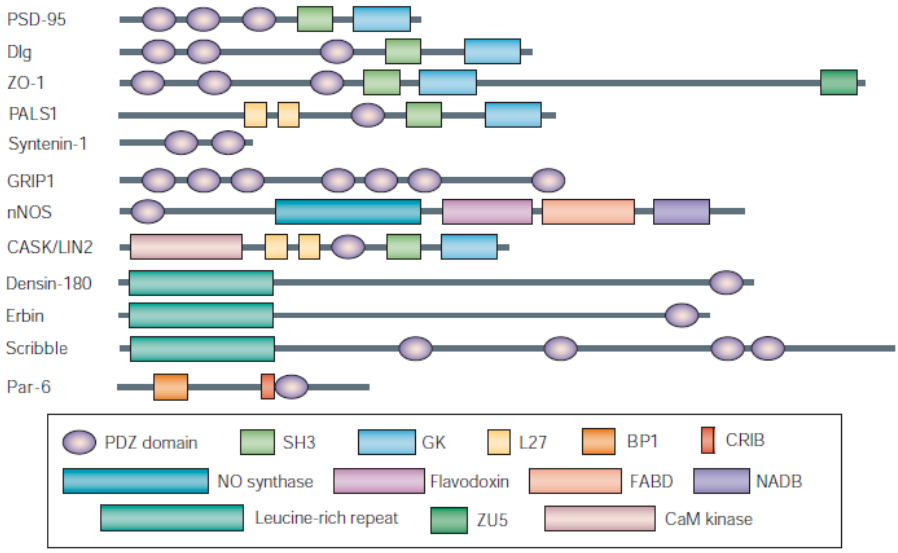


Figure 3.5 **Schematic representations of selected PDZ proteins.** PDZ domains are often found in scaffold proteins as multiple tandem arrays and/or linked to other kinds of modular protein-interaction domains. PDZ domains are indicated in purple elipses. Proteins: PDS-95, post-synaptic density 95; Dlg, discs large; ZO-1, zona occludens protein 1 ; PALS1, membrane-associated palmitoylated protein 5; Syntenin, syndecan binding protein; GRIP1, glutamate-receptor-interacting protein; nNOS, neuronal nitric oxide synthase; CASK/LIN2, Calcium/Calmodulin-dependent serine protein kinase; Par, partitioning defective homologue. Domains: SH3, Src homology 3 domain; GK, guanylate kinase domain; BP1, Phox and Bem1p domain; CRIB, Cdc42/Rac-interactive binding domain; NO, nitric oxide; FABD, FAD-binding domain; NADB, NAD-binding domain; Zu5, domain present in ZO-1; CaM kinase, calmodulin-dependent kinase. Adapted from (Kim *et al.*, 2004) with information on Par-6 and Scribble structure included from (F. Ye *et al.*, 2013) and (Penkert *et al.*, 2004), respectively.

3.5.1 PDZ-binding motifs (PBMs)

The majority of studied PDZ-binding motifs can be classified into three canonical specificity classes based on the residues at the -0 and -2 positions at the C-terminal of the peptide ligands, designated: type I PBM (-X-S/T-X-Φ_{COOH}), type II PBM (X-Φ-X-Φ_{COOH}) and type III PBM (-X-D/E-X-Φ_{COOH}), where X is any residue and Φ is a hydrophobic residue (Table 3.2) (Songyang *et al.*, 1997, Stricker *et al.*, 1997). Although the -0 and -2 positions of the peptide ligand are the key residues of interaction, it has been shown that residues -1, -3, -4 and -5 contribute to binding specificities; the PDZ binding cleft has been further shown to interact specifically with up to ten C-terminal ligand residues (Stiffler *et al.*, 2007, Tonikian *et al.*, 2008, F. Ye *et al.*, 2013, Y. Zhang *et al.*, 2006).

Table 3.2 Canonical C-terminal PDZ-binding motifs (PBMs).

Designation	Amino Acid Motif
Type I	**-X-S/T-X-Φ_{COOH}**
Type II	**-X-Φ-X-Φ_{COOH}**
Type III	**-X-D/E-X-Φ_{COOH}**

X = any amino acid; Φ = hydrophobic amino acid

An emerging theme of PDZ binding proteins and their ligands is the importance of context for an interaction. The context involves two concepts; firstly the sequence context, whereby both the PDZ domain and the ligand require additional amino acid extensions beyond the canonical PDZ domain fold and ligand PBM in order to mediate an interaction of high specificity. Secondly, the context of the cellular environment is considered where many factors within the cell (ie. temporal, spatial distribution, local concentration) dictate the moment of physical interaction between the two proteins (Luck *et al.*, 2012). A recent review conducted by Feng and Zhang on PDZ domains in the postsynaptic density shows that when many PDZ-containing proteins within this specialized subcellular compartment come together in supramodules which not only mediate distinct and highly specific target-binding properties, but are multi-PDZ-containing complexes can act as regulatory switches integrated into the PDZ scaffolds (Feng *et al.*, 2009). This concept suggests that while individual PDZ domains may be catalytically inactive as previously mentioned, PDZ domain supramodules may have a dynamic function in signalling events when found in PDZ-complexes (Feng *et al.*, 2009).

3.5.2 Viral PBMs

Not long after the discovery of the PDZ domains, PBMs were identified within viral oncoproteins of human adenovirus E4-ORF1, human T-lymphotropic virus type 1 (HTLV-1) Tax and human papillomavirus (HPV) E6 (Kiyono *et al.*, 1997, Lee *et al.*, 1997, Rousset *et al.*, 1998). Recent work has shown that non-transforming viruses such as SARS-CoV, rabies virus, influenza A virus and tick-borne encephalitis virus (TBEV), also present functional PBMs targeting cellular PDZ domains (Table 3.3) (Javier *et al.*, 2011). Viral PBMs interfere and modulate regular cellular processes to modulate viral replication, dissemination in the host or transmission to new hosts, and can enhance pathogenesis.

Table 3.3 Known cellular PDZ protein targets of viral proteins, circa 2011.

Adenoviridae	*Hepadnaviridae*	*Orthomyxoviridae*	*Papillovaviridae*	*Retroviridae*
Human Ad E4-ORF1 proteins *PBMs: ASLI, ATMI, ASNV, ATLV* PDZ targets • Dlg1 • MAGI-1 • MUPP1 • PATJ • ZO-2 (Ad9 only)	**HBV** Core protein *PBM: SQAR* PDZ target • TIP-2/GIPC	**Influenza A** NS1 proteins *Avian PBM: ESEV* PDZ targets • Dlg1 • MAGI-1 • MAGI-2 • MAGI-3 • Scribble • PDLIM2	**HPV** E6 proteins *HPV-16 PBM: ETQL* *HPV-18 PBM: ETQV* PDZ targets • CAL/GOPC • Dlg1 • Dlg4 • MAGI-1 • MAGI-2 • MAGI-3 • MUPP1 • PATJ • PTPN3 • PTPN13 • Scribble • TIP-1 • TIP-2/GIPC **RhPV** E7 protein *PBM: ASRV* PDZ target • Par3 **CRPV** LE6 and SE6 proteins *PBM: Not identified* PDZ target • Dlg1	**HTLV I** Tax protein *PBM: ETEV* PDZ targets • β1-syntrophin • Dlg1 • Dlg4 • Erbin • MAGI-3 • pro-IL-16 • Scribble • TIP-1 • TIP-2/GIPC • PDLIM2* Env protein *PBM: ESSL* PDZ target • Dlg1 **HIV** Gag protein *PBM: None* PDZ target • Dlg1 • PDZD8
Flaviviridae	***Rhabdoviridae***	***Coronaviridae***		
TBEV NS5 protein *PBM: internal* PDZ targets • RIMS2 • Scribble • ZO-1 **DV** NS5 protein *PBM: internal* PDZ target • ZO-1	**Rabies** G proteins *PBM: ETRL* PDZ targets • Dlg1 • MAST2 • MUPP1 • PTPN4	**SARS** E protein *PBM: DLLV* PDZ target • PALS1 • Syntenin		

Legend: Human Ad, human adenovirus; HBV, hepatitis B virus; DV, dengue virus; HPV, human papillomavirus; RhPV, rhesus papillomavirus; CRPV, cottontail rabbit papillomavirus; HIV, human immunodeficiency virus. Taken from (Javier *et al.*, 2011) and adapted with more recent information from (Jimenez-Guardeno *et al.*, 2014).

Association of viral PBMs with cellular PDZ proteins can result in loss of function, either by aberrant sequestration in cellular structures or by proteosome-mediated degradation, although some associations result in gain of functions for cellular proteins. For example, the NS1 protein PBM of influenza A respiratory virus of birds is an important virulence determinant with key functions in counteracting the antiviral mechanisms of the innate immune system and has been shown to re-localize the cellular protein Scribble from the plasma membrane into cytoplasmic puncta to block its pro-apoptotic function during infection (H. Liu *et al.*, 2010). Furthermore, it has been recently demonstrated that the NS1 PBM plays a role in transmission efficiency in a guinea pig model demonstrated by interchanging the PBM of a highly pathogenic H5N1 avian influenza strain to a pandemic H1N1 strain (Toyooka *et al.*, 2002). In the case of the oncogenic human papillomavirus types 16 and 18 (HPV-16, -18) of the mucosal

epithelium, the PBM of their E6 proteins have been shown to bind to the protein MAGI-1, targeting it for degradation thereby promoting tight junction disruption in cells (Kranjec *et al.*, 2011). In a final example, the neurotropic encephalitic rabies virus demonstrates the specificity of function of PBMs. Wild-type rabies virus is propagated quickly through neurons that survive infection, thus killing the host swiftly while attenuated rabies virus causes neuronal cell death upon infection, slowing the infection rate and allowing the immune system to slow and limit the infection, albeit still leading to eventual death. It has been shown that this difference in disease phenotype is mediated by different PBM sequences on the C-terminal of the rabies envelope glycoprotein G and the resulting cellular protein it interacts with mediates either neuronal cell survival or death (Prehaud *et al.*, 2010).

3.5.3 Coronavirus E protein PBMs

SARS coronavirus infection is pathologically characterized by diffuse alveolar damage and occasionally extensive lung epithelium damage (Teoh *et al.*, 2010). Early work by Enjuanes and colleagues suggested that the E protein was a virulence factor influencing replication level, virus dissemination, and pathogenicity (DeDiego *et al.*, 2007, DeDiego *et al.*, 2008), however the underlying molecular mechanisms were not known. Yeast-two hybrid studies, followed by confirmation by immunoprecipitation in mammalian cells done by Nal and colleagues, first demonstrated an interaction between the viral SARS E protein PBM and Lin Seven 1 (PALS1), a tight-junction-associated protein, mediated by the former's extreme C-terminus PBM (Teoh *et al.*, 2010). SARS coronavirus E protein was demonstrated to contain a type II PBM (-D-L-L-V$_{COOH}$). When present during viral infections, SARS-CoV E interacts with PALS1, redistributing it to the endoplasmic reticulum-Golgi intermediate compartment (ERGIC) and Golgi region in Vero E6 cells and causing delayed tight junction formation and polarity establishment in MDCK cells suggesting a role in the disease pathology (Teoh *et al.*, 2010).

The functionality of the SARS coronavirus E protein PBM was further confirmed as a molecular determinant of virulence by Enjuanes and colleagues using *in vivo*

models of infection (Jimenez-Guardeno *et al.*, 2014). In this study, recombinant SARS coronaviruses lacking the E protein PBM were produced, leading to a complete lack of disease phenotype (complete mouse survival and significant decrease in lesions and inflammation found within the lungs) after intranasal infection of mice but without affecting viral titers in the same organ. Following a yeast two-hybrid analysis and co-immunoprecipitation assays, it was determined that syntenin (syndecan binding protein) also associates with the SARS coronavirus E protein PBM. In Vero E6 cells, the interaction of the viral protein was shown to redistribute syntenin from the nucleus to the perinuclear region and regions close to the plasma membrane. Furthermore, syntenin was shown to participate in the activation of p38 mitogen-activated protein kinase (MAPK), which mediates proinflammatory cytokines; deletion of the SARS coronavirus E protein PBM caused a significant decrease in the amount of detected proinflammatory cytokines in infected mice compared to wild-type, elucidating a novel mechanism of modulation of pathogenesis in SARS coronavirus. Interestingly enough, a follow-up study where a series of small deletions in the C-terminal end of the SARS-CoV protein were created to evaluate the potential roles of different domains of the E protein in replication and pathogenesis, demonstrated that an 11 amino acid deletion of the C-terminal end, which contains the PBM, did not result in an attenuated disease phenotype as expected (Regla-Nava *et al.*, 2015). The authors suggest that this may be due to the fact that the resultant C-terminal amino acids actually form an alternative PBM which could rescue the original E protein PBM interaction to produce a wild-type disease phenotype (Regla-Nava *et al.*, 2015).

The human OC43 coronavirus viral E protein along with many other human and non-human coronavirus species such as MERS-CoV, HCoV-229E, HCoV-NL63 and HCoV-HKU1, have been recently identified via bioinformatic analyses as containing a PBM at their respective C-terminal ends (Figure 3.6) (Jimenez-Guardeno *et al.*, 2014). Furthermore, a large scale interaction profiling of PDZ domains using proteomic peptide-phage display of human viral peptidomes suggests an interaction between Erbin, a protein involved in the maintenance of epithelial cell architecture, and a bat coronavirus E protein (Ivarsson *et al.*, 2014). These preliminary observations suggest coronavirus E proteins in general may have a modulatory effect on cellular interactions or pathogenesis via their putative C-terminal PBMs.

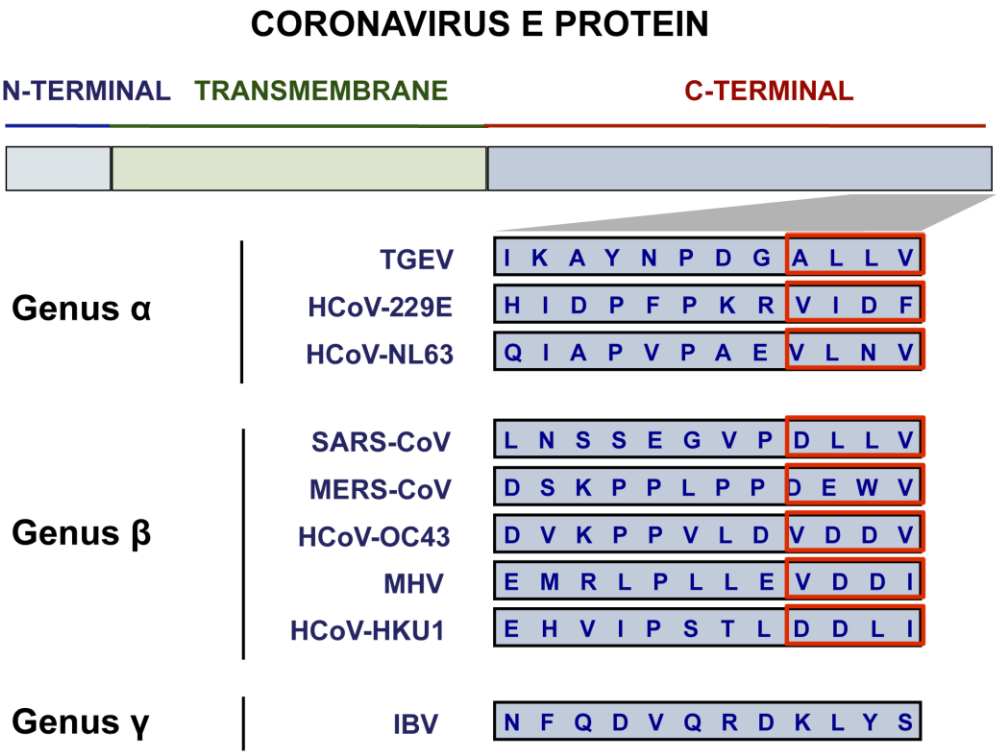


Figure 3.6 **Putative C-terminal PBMs of select coronaviruses identified by bioinformatic analysis.** Red box delimits putative PBM. Modified from (Jimenez-Guardeno *et al.*, 2014).

3.5.4 HCoV-OC43 E protein PBM

The putative OC43 coronavirus E protein PBM is composed of the amino acid sequence (-V-D-D-V$_{COOH}$) following the type II PBM canonical pattern (-X-D/E-X-Φ$_{COOH}$) (Table 3.2). It was noted by the present author that downstream of the canonical PBM at the -8 and -10 positions from the C-terminus, the residues showed similarity to a type III C-terminal PBM (-X-D/E-X-Φ$_{COOH}$) (Figure 3.6) and (Table 3.2). Non-canonical internal binding PBM-containing proteins have been described in the literature with varying binding modes (F. Ye *et al.*, 2013). For example, although the extreme C-terminus carboxylic acid is a major contributor to ligand affinity and selectivity for canonical PBMs, comparisons between the second PDZ domain of PSD-95 have shown it to interact both with a C-terminal ligand or an internal binding site ligand, indicating that an extreme terminal carboxylic acid is not essential for PDZ domain/ligand interaction (Pedersen *et al.*, 2014). In another example, partitioning-defective-6 (Par-6) PDZ domains has been shown to bind an internal fragment of PALS1 with a side-chain aspartic acid mimicking a carboxy group of the canonical C-terminal peptide ligand, a relevant example supporting the case for a non-canonical internal fragment binding of the hypothetical internal HCoV-OC43 E protein PBM described above. However, these types of internal PBMs are found much more distally from the C-terminal end (e.g. on an adjacent β-strand to the binding cleft). This hypothetical OC43 E protein PBM is more likely an extension of the canonical type II PBM identified, given that up to ten residues upstream of a canonical PBM have been found to be important for PDZ targets binding to cognate PDZ domains (F. Ye *et al.*, 2013) and the identified residues are found at the -8 and -10 positions. In either case, comprehensive biochemical, *in vitro* and *in vivo* studies are needed in order to determine the physiological relevance of these residues.

3.6 Importance of coronavirus E protein PBM in virus production

This principle function of PDZ binding motif-containing proteins is to bind to their corresponding PDZ-domain and, in the case of viral PBMs, interrupt regular cellular processes; this could include an effect on viral replication. Enjuanes and colleagues studied the growth kinetics of viruses containing full-length SARS-CoV E protein and those containing an E protein with the PBM deleted, and found slight replication defects at 24h but not at 72h post-infection on monkey Vero E6 and mouse DBT-mACE2 cells, suggesting a functional PDZ domain is not required for replication in cell cultures (Jimenez-Guardeno *et al.*, 2014). Intranasal inoculation of mice showed a significant decrease in lung titers (and disease pathology) of those infected with virus lacking the entire E protein or deleted C-terminal end containing the PBM compared to wild-type virus. Furthermore, a recombinant virus lacking a functional E protein PBM, but conserving a full length E protein, grew to similar levels as wild-type virus while still having an attenuated phenotype, suggesting a direct role for PBMs in replication is limited *in vivo* for SARS-CoV (Jimenez-Guardeno *et al.*, 2014). In a follow-up study by the same group, a series of small deletions were introduced to sections of the C-terminal end of a mouse-adapted SARS-CoV E protein, and the recombinant virus with the last 11 amino acids deleted did not affect the viral titer in Vero E6 and Huh7.5.1 cells nor the titer in the lungs of intranasally infected BALB/c mice (Regla-Nava *et al.*, 2015). The same study potentially maps replication to regions upstream of the PMB-containing region of the SARS-CoV E protein. Taken together, the literature suggests that deletion of the PBM of SARS-CoV can significantly decrease disease pathology, by a mechanism not yet fully elucidated, which could be independent of virus replication.

3.7 Explaining coronavirus E protein functionality

The coronavirus E protein is a small structural viral protein but appears to play a multitude of important roles during coronavirus infection that is only starting to be appreciated. Very recently, an interesting study has been conducted by Machamer and colleagues where they explored the possibility of distinct populations of IBV E proteins to help explain its demonstrated multifunctionality (Westerbeck *et al.*, 2015). Biochemical studies from IBV E transfections and infections of cells demonstrated the existence of two distinct and independent pools of E protein, representing monomeric and multimeric (likely homopentameric) states. More importantly, it is suggested that each pool supports a distinct function. Mutation of the alanine at position 26 of IBV E protein (equivalent to SARS-CoV E protein V25) within the transmembrane domain demonstrated a requirement for this amino acid for the robust formation of the multimeric E protein pool; mutation to a phenylalanine resulted in severely inhibited formation of this population. It is also suggested that this higher order oligomerization is needed for VLP production, given that with A26F mutation, no VLPs are produced (Westerbeck *et al.*, 2015). Threonine 16 IBV E protein was previously suggested to regulate ion channel activity (Ruch *et al.*, 2012b); the study by Machamer and colleagues demonstrated an additional role of this amino acid in generating the monomeric pool of E proteins that need to travel slowly through the secretory pathway for proper virion maturation (Westerbeck *et al.*, 2015). The authors suggested that in this monomeric state, the E protein may interact with other host proteins to mediate secretory pathway interactions (Westerbeck *et al.*, 2015); it is imaginable that the putative PBM at the C-terminal could mediate a majority of these proposed interactions.

The multiple properties of coronavirus E proteins described above have not yet been fully investigated or explained, and are at times contradictory. The phenomena observed during coronavirus infection, such as extensive secretory pathway modification within the cell as well as trafficking, assembly and release of virions, has been mapped to this protein, however it is still unclear of the exact domains contributing to these observations. In many cases, the mechanism of action mediated via E protein is not known or contradictory between species. The concept of two pools of E proteins with their own distinct functions has been hypothesized for some time, and has only recently begun to be studied. Regardless, the proposed viroporin activity as well as the ability to interact with viral and/or cellular components mediated via a specific domain (such as a C-terminal PDZ-binding motif), are likely two major contributors to the observed effects on viral replication cycle and the host-cell response to infection in relation to the coronavirus E protein.

This thesis will represent an initial characterization study (submitted manuscript) of the HCoV-OC43 E protein for which there is no published literature on its importance in viral replication. This study will describe the first efforts to determine the relative contribution of full-length E protein to infectious HCoV-OC43 production, through the production of a recombinant virus lacking the E protein in epithelial and neuronal cell culture. This study will elucidate if predicted amino acids suggested to be important for the formation/function a putative HCoV-OC43 E protein viroporin are relevant for virus replication by constructing recombinant viruses with amino acids Q17 and L27 of this protein mutated and observing the corresponding replication rates of in cell culture. A final portion of this project will describe the first efforts to characterize the functionality of putative HCoV-OC43 PBMs by constructing recombinant viruses with partially or fully abrogated motifs and investigating the corresponding replication rates in cell culture.

LA PROTÉINE D'ENVELOPPE (E) DU CORONAVIRUS RESPIRATOIRE HUMAIN HCOV-OC43 EST NÉCESSAIRE POUR LA FORMATION DE VIRIONS INFECTIEUX DANS LES CELLULES ÉPITHÉLIALES ET NEURONALES

II - ARTICLE

4 A FULLY FUNCTIONAL HUMAN CORONAVIRUS OC43 ENVELOPE (E) PROTEIN IS NECESSARY FOR PRODUCTION OF INFECTIOUS VIRUS AND EFFICIENT VIRAL SPREAD IN CELL CULTURE

Authors : Jenny K. Stodola, Marc Desforges, Pierre J. Talbot

Affiliation of all authors : Laboratory of Neuroimmunovirology, INRS-Institut Armand-Frappier, Laval, Québec, Canada

Contribution of Jenny K. Stodola : JKS conducted all the set up and execution of all of the experiments found within the article (except preparation of the primary murine central nervous system cultures, done by Alain Le Coupanec and Mathieu Messen-Pinard). JKS undertook the majority of the preparation of the manuscript as well as creation of all the images found within. Experiments were advised and overseen by MD and PJT, and the manuscript was also extensively reviewed by these individuals.

Article submittted in : Journal of Virology (February 11, 2016)

Titre français: Une protéine d'envelope (E) complètement fonctonnelle est nécessaire pour la production de virions infectieux du coronavirus humain OC43 et sa propagation efficace en culture cellulaire.

Résumé français: Le coronavirus humain (OC43) est un pathogène ubiquitaire du tractus respiratoire dont notre laboratoire a démontré des propriétés neurotropiques. Comme tout coronavirus, il comporte une petite protéine structurale d'enveloppe (E) qui, malgré sa rareté dans les virions, est connu pour jouer plusieurs rôles importants notamment dans la morphologie/assemblage du virion, dans la réponse cellulaire à l'infection et dans la virulence. Ces fonctions sont associées à la présence de plusieurs motifs protéiques présents sur la protéine elle-même et/ou à la formation de canaux

ioniques causés par l'homo-oligomérisation de cette protéine. En fonction de la souche étudiée, la protéine E des coronavirus a montré plusieurs niveaux d'importance dans la production des nouveaux virions infectieux, allant graduellement de dispensable jusqu'à absolument critique. En utilisant notre clone infectieux d'ADNc, nous avons produit des virus recombinants HCoV-OC43 dont la protéine E a été supprimée ou mutée dans des domaines spécifiques. En transfectant des cellules BHK-21 avec ces clones infectieux d'ADNc mutants, nous avons démontré qu'une protéine E complètement fonctionnelle est nécessaire pour la production optimale de virus recombinants infectieux, production qui est rétablie suite à l'introduction de la protéine E sauvage en co-transfection. De plus, l'infection des cellules épithéliales et neuronales d'origine humaine et des cultures primaires du système nerveux central murin ont démontré que la protéine E est essentielle pour la production et propagation efficace des virions. Nos résultats indiquent que la protéine E pourrait représenter une cible thérapeutique d'intérêt pour limiter la neuroinvasion des coronavirus.

4.1 Abstract

The OC43 strain of human coronavirus (HCoV-OC43) is a ubiquitous respiratory tract pathogen for which we have shown neurotropic capacities. As with all coronaviruses, it comprises a small structural envelope (E) protein that, despite its scarcity on virions, is known to play a variety of important roles, including in virion morphology/assembly and in cell response to infection and/or virulence. This has been associated to the presence of several protein motifs and/or homo-oligomeric ion channel formation by the protein. Depending on the strain of coronavirus, the E protein has been shown to have various levels of importance, ranging from degrees of dispensability to absolutely essential for production of progeny infectious virions. Making use of our cDNA infectious clone, we have generated recombinant viruses either devoid of the E protein or harboring mutations in specific domains of the protein. Transfection of BHK-21 cells by mutant cDNA infectious clones revealed that a fully functional HCoV-OC43 E protein is needed for optimal production of recombinant infectious viruses, wherein wild-type E protein provided in trans was sufficient to rescue infectious virus production. Furthermore, HCoV-OC43 infection of human epithelial and neuronal cell lines and of mixed murine primary cultures from the central nervous system showed that the E protein is critical for efficient virus replication and spread. Taken together, our results indicate that the E protein could represent an interesting therapeutic target to limit neuroinvasive coronavirus infection.

4.2 Importance of study

The OC43 strain of human coronavirus (HCoV-OC43) represents a circulating respiratory virus with neurotropic and potential neurovirulent properties. The E protein of different coronaviruses has recently emerged as a virulence factor, which plays different roles in virus production and pathogenesis. Herein, we confirm the importance of this multifunctional protein in HCoV-OC43 infection and spread in both epithelial and neuronal cells. Our studies point to the presence of a true transmembrane domain and underlines the critical role of a protein-protein interaction motif present at the C-terminal end of the protein, especially in neuronal cells. Given that HCoV-OC43 is related to the highly pathogenic severe acute respiratory syndrome coronavirus (SARS-CoV) and recently emerged Middle-East respiratory syndrome coronavirus (MERS-CoV), these studies can help elucidate which factors could contribute to differences in disease outcome between coronavirus species and the possibility that HCoV-OC43 could exacerbate or modulate neurologic disease.

4.3 Introduction

Coronaviruses are widespread RNA viruses of the *Nidovirales* order, *Coronaviridae* family, most often associated with human and veterinary respiratory infections (de Groot *et al.*, 2012). Of the six human-infecting coronavirus strains, four (HCoV-229E, HCoV-NL63, HCoV-HKU1 and HCoV-OC43) are currently co-circulating and elicit respiratory illnesses (Vabret *et al.*, 2009). Coronaviruses also represent a significant public health concern due to the recent zoonotically emerged, highly pathogenic species, SARS corovonavirus (SARS-CoV) (Drosten *et al.*, 2003, Ksiazek *et al.*, 2003) in 2002-2003 and, since 2012, Middle-East respiratory syndrome coronavirus (MERS-CoV) (Zaki *et al.*, 2012), localized to the Arabian Peninsula, but with sporadic travel-related outbreaks worldwide. In addition to their respiratory tropism, human coronaviruses have been detected concurrently with severe and acute neurological symptoms (Arabi *et al.*, 2015, Yeh *et al.*, 2004) and shown to naturally infect the central nervous system (CNS) (Arbour *et al.*, 2000, Gu *et al.*, 2005, Xu *et al.*, 2005) with neurons demonstrated as the main target of infection in HCoV-OC43 (Bonavia *et al.*, 1997, Favreau *et al.*, 2012, Jacomy *et al.*, 2006, Jacomy *et al.*, 2003) and SARS-CoV (Gu *et al.*, 2005, Xu *et al.*, 2005).

Coronaviruses represent the largest known enveloped RNA (single-stranded positive sense) viruses with a genome of approximately 30 kb (de Groot *et al.*, 2012). The viral envelope is composed of four or five proteins, the spike (S), membrane (M), envelope (E) and hemagglutinin-esterase protein (HE), the latter in some β genus coronaviruses, such as HCoV-OC43. Coronavirus E proteins are 74-109 amino acids in length, 84 amino acids for HCoV-OC43, and share only a small amount of sequence identity between coronavirus species. However, its secondary structure, composed of a short N-terminal domain followed by single hydrophobic transmembrane (TM) domain and hydrophilic cytoplasmic tail, remains overall conserved and is suggested to be more important than sequence for function (Kuo *et al.*, 2007, Torres *et al.*, 2005). The importance of the presence of the E protein in the viral envelope is emphasized by the

fact that there are only about twenty E molecules incorporated within the virion structure (Godet *et al.*, 1992, D. X. Liu *et al.*, 1991, Yu *et al.*, 1994) and deletion of the protein can either completely prevent the production of detectable infectious virions (Almazan *et al.*, 2013, Curtis *et al.*, 2002, Ortego *et al.*, 2007, Ortego *et al.*, 2002) or significantly reduce infectious virus titers (DeDiego *et al.*, 2007, DeDiego *et al.*, 2008, Kuo *et al.*, 2007, Kuo *et al.*, 2003).

The majority of the coronavirus E protein in the infected cell is localized within the secretory pathway between the membranes of the endoplasmic reticulum (ER), Golgi and intermediate compartment between them (ERGIC) (Cohen *et al.*, 2011, Nieto-Torres *et al.*, 2011, Venkatagopalan *et al.*, 2015). It is in this intracellular region that additional functions mediated by various domains of the coronavirus E proteins are proposed to occur. Homo-pentameric oligomerization of the E protein TM domain in membranes to form ion channels, called viroporins, has been predicted for several coronaviruses (Torres *et al.*, 2005) and extensively studied for species such as SARS-CoV (Nieto-Torres *et al.*, 2014, Pervushin *et al.*, 2009) or avian infectious bronchitis virus (IBV) (Ruch *et al.*, 2012b, Westerbeck *et al.*, 2015). Another domain found at the extreme C-terminal end of the E protein, a PDZ-domain binding motif (PBM), has also been predicted for several coronavirus species (Jimenez-Guardeno *et al.*, 2014). This protein-protein interaction motif capable of interrupting normal cellular functions, has been demonstrated in other viruses to play important roles in replication, dissemination in the host and pathogenesis (Javier *et al.*, 2011). The multiple properties of coronavirus E proteins have not yet been fully investigated or explained, and are at times apparently contradictory between coronavirus species. The multifunctionality of the E protein could be explained by the presence of two distinct pools (monomeric versus homo-oligomeric states) present in the infected cell (Westerbeck *et al.*, 2015). Furthermore, the different motifs found within the protein (Jimenez-Guardeno *et al.*, 2015) could mediate different specific functions.

HCoV-OC43 represents a circulating strain of human coronavirus causing respiratory illness, which is naturally capable of invading the CNS where neurons are preferentially targeted for infection. In this study, we demonstrate that the fully functional HCoV-OC43 E protein is critical in infectious virus production and dissemination in epithelial and neuronal cells, with strong selective pressure to recover functional E protein production. In addition, modification of key positions of the putative E protein TM domain and C-terminal PBM have significant effects on infectious virus production suggesting a role of E protein during infection, a first demonstration for this coronavirus species.

4.4 Materials and Methods

Cell lines and mixed primary murine CNS cells. The BHK-21 cell line (ATCC-CCL10) was cultured in minimal essential medium alpha (MEM-α; Life Technologies) supplemented with 10% (vol/vol) fetal bovine serum (FBS; PAA GE Healthcare) and used for transfection. The HRT-18 cell line (a gift from the late David Brian, University of Tennessee) was cultured in the same medium and used for virus infections/amplifications. The LA-N-5 cell line (a kind gift of Stephan Ladisch, George Washington University School of Medicine) was cultured in RPMI medium supplemented with 15% (vol/vol) fetal bovine serum (FBS), 10 mM HEPES, 1 mM sodium pyruvate, and 100 µM non-essential amino acids (Gibco - Invitrogen). The LA-N-5 cells were differentiated into human neurons as previously described (Hill *et al.*, 1998). Briefly, cells were seeded in 24-well plates (1.25×10^3 cells/well) in RPMI medium supplemented with 10% (vol/vol) FBS, 10 mM HEPES, 1 mM sodium pyruvate, and 100 µM non-essential amino acids. The next day and every 2 days for 6 days, the medium was replaced with the same medium supplemented with 10% (vol/vol) FBS and 10 µM all-trans retinoic acid (Sigma-Aldrich).

Mixed primary cultures of mouse CNS cells were prepared as previously described (Le Coupanec *et al.*, 2015). Briefly, embryos at 14 to 16 days of gestation were removed from pregnant anesthetized CD1 mice and their cortex and hippocampus were

harvested and placed in Hanks balanced salt solution (HBSS) medium, without Ca^{2+} and Mg^{2+}, supplemented with 1 mM sodium pyruvate and 10 mM HEPES buffer. Tissues were gently pipetted up and down with a Pasteur pipette to dissociate the cells. After a decantation step of 5 min at room temperature, supernatants were transferred into a 50-mL tube with 36 mL of Neurobasal Medium (Invitrogen) supplemented with 0.5 mM GlutaMAX-I (Life Technologies), 10 mM HEPES buffer, B27 supplement (Life Technologies), gentamycin and 10% (vol/vol) of horse serum (Life Technologies). Cells were then seeded at 1×10^5 cells/cm^2 and grown on collagen+poly-D-lysine (3:1 for a final concentration at 50 µg/mL for both)-treated 12-well plates containing glass coverslips (for immunofluorescence) or not (for evaluation of infectious virus production) in the same medium, which was replaced by fresh Neurobasal Medium without horse medium the next day. The medium was changed every 2 days after and the cultures were ready for infection after 7 days in culture.

Site-directed mutagenesis. Using our full-length, cDNA infectious clone pBAC-OC43^{FL} (St-Jean *et al.*, 2006) the recombinant HCoV-OC43 virus (rOC/ATCC) was generated. In parallel, a series of recombinant mutant viruses were produced by site-directed mutagenesis using the QuikChange Multi Site-Directed Mutagenesis Kit (Stratagene) and a variety of primers (Table 4.1, Section A) to introduce nucleotide substitutions in the E gene. These substitutions in the cDNA clone introduced were: nucleotide mutations at positions 24 and 25, corresponding to a change of amino acid position 9 into a stop codon (pBAC-OC43-E-Stop, plasmid; rOC/E-Stop, recombinant virus) or mutations at nucleotide positions 49 and 50 corresponding to amino acid 17 (pBAC-OC43-E-TM-Q17A; rOC/E-TM$_{Q17A}$), nucleotide position 81; amino acid 27 (pBAC-OC43-E-TM-L27F; rOC/E-TM$_{L27F}$), nucleotide position 245 and 246; amino acid 82 (pBAC-OC43-E-PBM-D82A; rOC/E-PBM$_{D82A}$) or nucleotide position 251; amino acid 84 (pBAC-OC43-E-PBM-V84A; rOC/E-PBM$_{V84A}$). A double mutant cDNA clone pBAC-OC43-E-PBM-D82A-V84A (rOC/E-PBM$_{D82A\text{-}V84A}$) was also produced using the pBAC-OC43-E-PBM-D82A as a DNA template for a second-round of mutagenesis reaction with OC-E-PBM-D82A-V84A primer to introduce a second mutation at nucleotide position 251; amino acid 84. Prior to transfection of BHK-21 cells, all samples were

sequenced to make sure that only the introduced mutations were present and that no other mutations appeared.

Table 4.1 **Primers used to introduce nucleotide substitutions in the E gene of pBAC-OC43^{FL} (St-Jean *et al.*, 2006) for recombinant HCoV-OC43 virus production with amino acid modifications within the E protein (Section A). Primers used to verify sequences of the full E HCoV-OC43 E and partial M gene of recombinant viruses (Section B).** Bold and underlined sections represent newly introduced nucleotide substitutions. Underlined section in mutant E-PBM 82-84 represents previously introduced nucleotide substitutions.

A - Primers for site-directed mutagenesis

Recombinant Virus	Primer Name	Primer sequence
rOC/E-Stop	OC-E-Stop	5'-CT GAT GCT TAT CTT **TG**A GAC ACT GTG TGG-3'
rOC/E-TM$_{Q17A}$	OC-E-TM-Q17A	5'-GTG TGG TAT GTG GGG **GC**A ATA ATT TTT ATA GTT GCC-3'
rOC/E-TM$_{L27F}$	OC-E-TM-L27F	5'-GTT GCC ATT TGT TTA TT**C** GTT ACA ATA GTT GTA GTG GC-3'
rOC/E-PBM$_{D82A}$	OC-E-PBM-D82A	5'-CA GTC CTT GAT GTG G**CC** GAC GTT TAG GTA ATC-3'
rOC/E-PBM$_{V84A}$	OC-E-PBM-V84A	3'-GT TTG GAT TAC CTA A**G**C GTC ATC CAC ATC AAG-5'
rOC/E-PBM$_{D82A-V84A}$	OC-E-PBM-D82A-V84A	5'-CTT GAT GTG GCC GAC G**C**T TAG GTA ATC CAA AC-3'

B - Primers for sequence verification

Primer Name	Primer sequence
OC-ns5-116-E-For	5'-GTA GAG TCC TAG TCA TGC TTG-3'
OC-M127-E-Rev	5'-ACA TAC TGC GAC TTG TAT AGC C-3'
OC-M241-E-Rev	5'-CTA TAG AAA GGC CAA GAT ACA C-3'
GAPDH-For	5'-CGG AGT CAA CGG ATT TGG TCG TAT-3'
GAPDH-Rev	5 '-AGC CTT CTC CAT GGT GGT GAA GAC-3'

Plasmid for transient co-transfection rescue. In order to insert the HCoV-OC43 E gene into the pcDNA3.1(+) expression vector (pcDNA; Invitrogen) and allow for E protein expression upon transient co-transfection with pBAC-OC43 infectious clones in BHK-21 cells, restriction enzymes *Nhe*I and *Bam*HI were added to the 5' (primer: 5'-GCTAGC ATG TTT ATG GCT GAT GCT TA-3') and 3' (primer: 5'-GGATCC CTA AAC GTC ATC CAC AT-5') ends of the E gene respectively. The E gene with added restriction enzyme sites was PCR-amplified from cDNA originating from an HCoV-OC43 reference strain (ATCC) infection on HRT-18 cells using Accuprime *Pfx* Supermix (Life Technologies) with 1 cycle at 95°C for 5 min, followed by 35 cycles at 94°C for 15 s, 48°C for 30 s and 68°C for 1 min and 1 cycle at 68°C for 4 min and then introduced into the pcDNA plasmid.

Recombinant virus production, transient co-transfection and virus amplification. The BHK-21 cells were cultured in MEM-α supplemented with 10% (vol/vol) FBS and used for transfection of pBAC-OC43 cDNA infectious clones with Lipofectamine 3000 Reagent (Life Technologies) according to the manufacturer's instructions. Briefly, for production of recombinant viruses, BHK-21 cells were seeded in 6-well cell culture plates at $6x10^5$ cells/well. The next day, when cells were 70-90% confluent, the medium was replaced and cells were transfected with 7.5µl Lipofectamine 3000 Transfection Reagent, 10µl P3000 Reagent, 5µg of pBAC-$OC43^{FL}$ or other modified pBAC DNA, and 2µg of pcDNA(OC-E) or empty pcDNA plasmid per well. For semi-quantitative determination of transfection efficiency by immunofluorescence assay (IFA), BHK-21 cells were seeded at 5×10^4 cells/well onto glass coverslips in 24-well plates and transfected with 1.5µl Lipofectamine 3000 Transfection Reagent, 6µl P3000 Reagent, 1.5µg of pBAC-$OC43^{FL}$ or other modified pBAC DNA, and 500ng of pcDNA(OC-E) or empty pcDNA plasmid per well. The plates were incubated at 37°C for 8 hours and then medium replaced with MEM-α supplemented with 10% (vol/vol) FBS and 0.01% (vol/vol) gentamycin and incubated for 3 days.

The cells from 6-well culture plates were harvested either to recover total RNA or total protein while the supernatant (P0) was recovered by aspiration after centrifugation at 500 × *g* for 7 min and then clarified at 1000 × *g* for 10 min. The supernatant (P0) served to inoculate HRT-18 cells in order to amplify the viral stocks. The supernatant from this first-round amplification (P1) served for a second round of viral amplification on HRT-18 cells from which supernatant was recovered (P2) and in some cases was repeated again for a third round of amplification (P3). The production of infectious viral particles corresponding to the different pBAC-OC43 cDNA clones was titrated by an immunoperoxidase assay (IPA) prior to each amplification step in order that titers could be normalized to the lowest detectable titer and replication rates be compared.

Infection of human cell lines and of primary mouse CNS cultures. The HRT-18 and LA-N-5 cells as well as mixed primary cultures of mouse CNS cells were infected at a MOI equivalent to the lowest detectable titer of the series of recombinant virus stock used during each experiment or mock-infected and then incubated at 33°C (HRT-18) or 37°C (LA-N-5 cell line and primary CNS cultures), for 2 h (for virus adsorption), and incubated at 33°C with fresh MEM-α supplemented with 1% (vol/vol) FBS (for HRT-18 cells), at 37°C with fresh RPMI medium supplemented with 2.5% (vol/vol) FBS (for LA-N-5 cells) or at 37°C with fresh Neurobasal Medium with B27-GlutaMAX-I (for primary murine CNS cell cultures) for different periods of time before fixing cells for immunofluorescence detection or harvesting the cell-associated and/or cell-free medium fractions for infectious virus titer determination by IPA.

Titration of infectious virus using an immunoperoxidase assay (IPA). The IPA was performed on HRT-18 cells, as previously described (Lambert *et al*., 2008). Briefly, the primary antibody used was mAb 4.3E4 (hybridoma supernatant; ½ dilution) directed against the S protein of HCoV-OC43. The secondary antibody was horseradish peroxidase-conjugated goat anti-mouse immunoglobulin (KPL; 1/500). Immune complexes were detected by incubation with 0.025% (w/v) 3,3'-diaminobenzidine tetrahydrochloride (Bio-Rad) and 0.01% (vol/vol) hydrogen peroxide in PBS and infectious virus titers were calculated by the Karber method, as previously described (Lambert *et al*., 2008).

Immunofluorescence for semi-quantification of transfection efficiency or virus propagation. BHK-21 and LAN-5 cells as well as mixed primary cultures of mouse CNS were fixed onto glass coverslips with 4% (wt/vol) paraformaldahyde for 30 min at room temperature and permeabilized for 5 minutes with 100% methanol at -20°C. For LA-N-5 and BHK-21 cells, to detect HCoV-OC43 spike (S) protein, one hour incubations of primary 4.3.E.4 (hybridoma supernatant; ½ dilution;) followed by secondary antibody AlexaFluor 488 donkey anti-mouse IgG (H + L) (1/1000; Life Technologies-Molecular probes) were conducted with three PBS washes between

steps. For primary mouse CNS cultures, after blocking with a PBS-BSA 2% (wt/vol) solution for one hour at room temperature, primary antibody polyclonal rabbit anti-S protein (dilution 1/1000) and mouse monoclonal antibody against the neuron-specific MAP2 protein (1/1000; BD Pharmagen, catalog no. 556320) were diluted in PBS + 0.1% Triton X-100 and incubated on cells for one hour at room temperature followed by three PBS washes. Cells were then incubated one hour at room temperature with anti-rabbit Alexa Fluor 568- and anti-mouse Alexa Fluor 488-conjugated secondary antibodies (1/1000; Life Technologies-Molecular probes) in PBS. For all cell types, nucleus detection was accomplished by a 5 min incubation with 4',6-diamidino-2-phenylindole (DAPI; 1µg/ml; Life Technologies). Triplicate samples were mounted on glass slides with Immuno-Mount medium (Fisher Scientific). Immunofluorescent staining was observed under a Nikon Eclipse E800 microscope with a QImaging Retiga-EXi Fast 1394 digital camera using Procapture system software.

RNA extraction, cDNA synthesis and gene amplification. After transfection of BHK-21 cells or infection of HRT-18 cells, cells were scraped from wells or plates, centrifuged at 500 × *g* for 7 min at 4°C, medium was removed and cell pellet resuspended with 1.5ml ice-cold PBS and centrifuged at 1000 × *g* for 10 min at 4°C. PBS was aspirated and the dry pellet stored at -80°C until use. Total RNA was extracted using the RNeasy Mini Kit (QIAGEN) with QIAshredder spin columns (QIAGEN) to lyse cells according to manufacturer's instructions. RNA quality was verified using the Agilent 2100 Bioanalyzer using the Agilent RNA 6000 Nano Assay protocol according to manufacturer's instructions and concentration measured using a ND1000 spectrophotometer (Nanodrop). To produce cDNA, 5µg of total extracted RNA was reverse transcribed using the SuperScript III First-Strand Synthesis Supermix Kit using oligo(dT) primer (Invitrogen) according to manufacturer's instructions.

PCR was conducted using Accuprime *Pfx* Supermix (Life Technologies) with one cycle at 95°C for 1 min, followed by 40 cycles at 95°C for 35 sec, 50°C for 45 sec and 68°Cfor 2 min, followed by one cycle at 68°C for 7 min using two sets of primers (Table

4.1, Section B) to amplify the E gene (forward primer, OC-ns5-116-E-For; reverse primers, OC-M127-E-Rev or OC43-M241-E-Rev) and GAPDH gene as control (forward primer, GAPDH-For; reverse primer, GADPH-Rev).

Protein extraction and Western blot analysis. To confirm E protein production after transfection of pcDNA(OC-E) in BHK-21 cells, proteins from whole cell lysates were extracted. Harvested cells were pipetted up and down into RIPA buffer (150 mM NaCl, 50 mM Tris, pH 7.4, 1% (v/v) NP-40, 0.25% (wt/vol) sodium deoxycholate, 1 mM EDTA) supplemented with protease cocktail inhibitor (Sigma). Lysates were incubated on ice for 20 min and centrifuged at 17,000 × *g* for 10 min at 4°C. Supernatants were harvested, aliquoted and stored at −80°C until further analyzed.

Protein concentrations were determined using a bicinchoninic acid (BCA) protein assay kit (Novagen) according to the manufacturer's instructions. Ten µg of protein was loaded on a Tris-Glycine 4–15% gradient gel, transferred to PVDF membrane with a semi-dry trans-blot apparatus (Bio-Rad). Membranes were blocked overnight at 4 °C with TBS buffer containing 1% (vol/vol) Tween (TBS-T) and 5% (wt/vol) non-fat milk. The following day all steps were conducted at room temperature, with, or solutions diluted in, TBS-T and milk. A primary rabbit polyclonal antibody was used to detect either the HCoV-OC43 E protein (1/5000) or GAPDH (1/10 000) for 1 hour, followed by three 10 min washes. Anti-rabbit IgG horseradish peroxidase linked whole antibody (from donkey) (GE Healthcare) was the secondary antibody used, followed by three 10 min washes. Detection was performed using 1/1 solution of Clarity Western ECL Substrate (Bio-Rad) for one minute followed by membrane exposure on CL-X-Posure Film (Thermo Scientific).

4.5 Results

Deletion of HCoV-OC43 E protein prevents infectious virion detection but can be rescued by transient complementation with wild-type protein. In order to evaluate the importance of the HCoV-OC43 E protein in infectious virion production, a stop codon was introduced at the beginning of the E gene of our cDNA infectious clone, pBAC-OC43^{FL} (St-Jean *et al.*, 2006), preventing corresponding full-length E protein production in the resultant recombinant virus (Figure 4.1A). Transfection of BHK-21 cells with the pBAC-OC43^{FL} led to the detection of reference HCoV-OC43 recombinant infectious virus (rOC/ATCC) whereas transfection with the pBAC-OC43-E-Stop mutant did not lead to any detectable infectious virus (rOC/E-Stop) (Figure 4.1B).

To confirm that the inability to detect infectious viral particles was due to the lack of E protein expression, we wished to verify whether viral production could be rescued with wild-type E protein. Transfection of a plasmid containing the reference HCoV-OC43 E gene, pcDNA(OC-E), in BHK-21 cells clearly showed via Western blot assay (WB) that the E protein was produced compared to an empty plasmid condition (Figure 4.1C). Subsequently, a transient co-transfection was conducted in the same cells with pBAC-E-Stop and pcDNA(OC-E) and infectious particles production was rescued to detectable levels in a dose-dependent manner (Figure 4.1D). Making use of a monoclonal antibody against the S protein of HCoV-OC43, we confirmed that the viral S protein was produced at equivalent levels in cells transfected with pBAC-OC43^{FL} alone or pBAC-E-Stop with pcDNA(OC-E) or empty plasmid (Figure 4.1E). Viral RNA was harvested and cDNA sequenced to confirm that the infectious particles detected after transfection corresponded to rOC/ATCC and rOC/E-Stop mutant (data not shown).

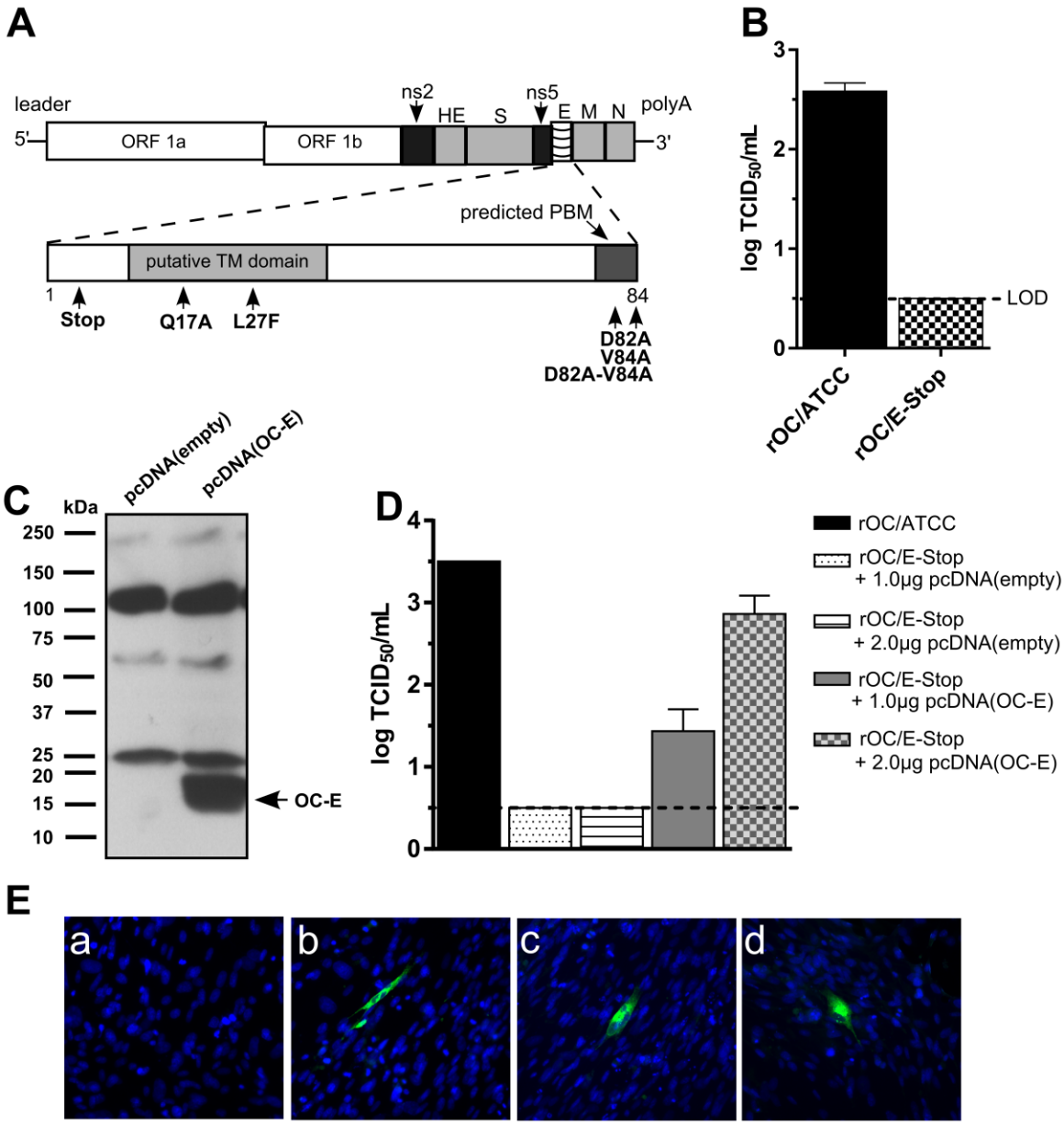
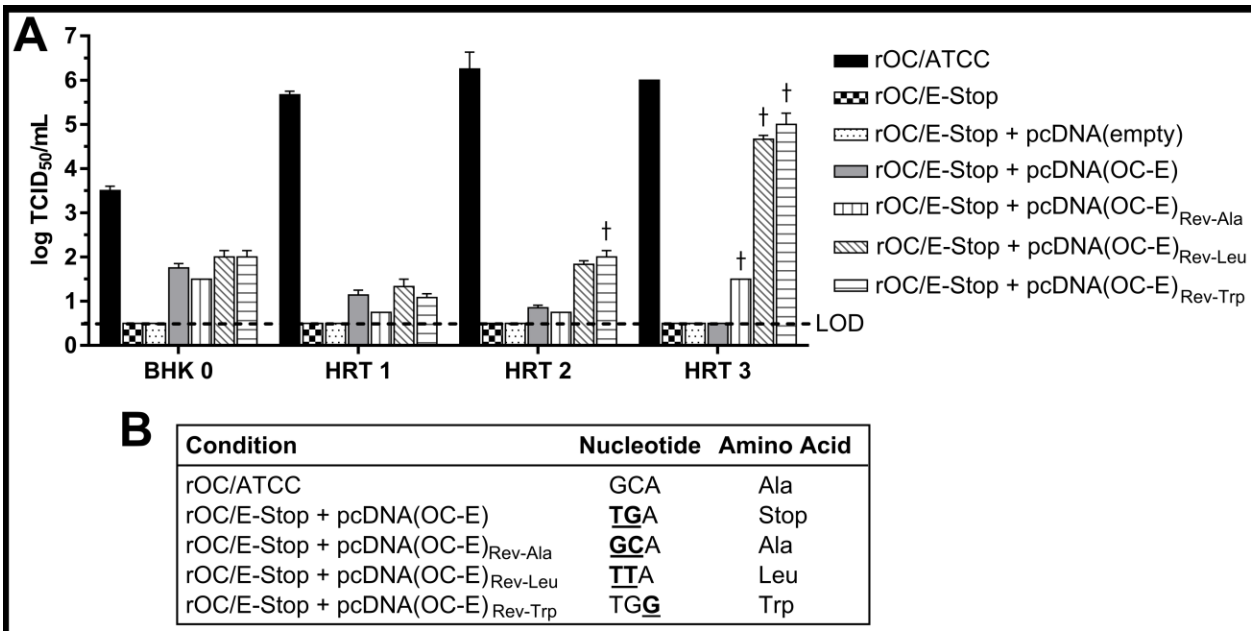


Figure 4.1 **The HCoV-OC43 E protein is critical for infectious particle production.** (A) Representation of the full-length HCoV-OC43 genome found within the pBAC- OC43^{FL} infectious clone (top) with a schematic representation of the HCoV-OC43 E gene to be modified at various amino acid positions indicated at their relative positions within the protein (bottom). TM, transmembrane; PBM, PDZ-binding motif. (B) Evaluation of infectious virus production corresponding to pBAC-E-Stop transfection of BHK-21 cells compared to pBAC-OC43^{FL}. (C) Insertion of wild-type E gene into pcDNA3.1(+) expression vector, pcDNA(OC-E), yielded corresponding HCoV-OC43 E protein expression compared to empty vector. (D) Transient co-transfection of pBAC-E-Stop and 1or 2µg pcDNA(OC-E) in BHK-21 cells rescued detectable infectious virus in a dose-dependent manner. (E) Visualization of transfection efficiency on BHK-21 cells of various conditions by immunofluorescence detection of HCoV-OC43 S protein (green), nucleus staining with DAPI (blue). Subpanels: (a) mock, (b) pBAC-OC43^{FL} (rOC/ATCC), (c) pBAC-E-Stop (rOC/E-Stop + 2µg pcDNA(OC-E)), (d) pBAC-E-Stop (rOC/E-Stop + 2µg pcDNA(empty)). LOD, limit of detection.

HCoV-OC43 lacking E protein can infect epithelial cells but has reduced efficiency in progeny production and introduces a strong selection pressure for reversion. As we were previously able to rescue infectious particles production through transient complementation, we wondered whether this resultant virus, still lacking the E gene, could be amplified further in subsequent passages. To this end, we amplified the viral stocks of all transfected plasmids three times on HRT-18 epithelial cells, each time normalizing viral titers to the lowest detectable titer in order to compare replication rates (Figure 4.2A). Throughout the amplification process we were consistently unable to detect infectious viral particles issuing from viral stocks of rOC/E-Stop. Amplifications of initially complemented viral stocks of rOC/E-Stop led to detectable titers which decreased over the course of each subsequent amplification compared to rOC/ATCC. Sequencing of viral RNA confirmed that the E gene in the viral rOC/E-Stop stocks contained the introduced stop codon at each amplification step (data not shown). These results demonstrate that production of progeny infectious HCoV-OC43 virions is still possible in the absence E protein, however the efficiency of the process is dramatically diminished.

Interestingly, when conducting independent experiments following the same experimental approach, the titers of initially complemented rOC/E-Stop sometimes increased substantially after two or three amplifications, approaching reference virus titer levels after three rounds of amplification on HRT-18 cells (Figure 4.2A). Sequence analysis of the E gene of the corresponding viral stocks revealed that a reversion of sequence appeared at the position where the stop codon had been initially introduced; representing reversion to wild-type or new amino acids (Figure 4.2B). Taken together, these data demonstrate that the HCoV-OC43 E protein is critical for efficient infectious virion production in epithelial cells.



B

Condition	Nucleotide	Amino Acid
rOC/ATCC	GCA	Ala
rOC/E-Stop + pcDNA(OC-E)	**TG**A	Stop
rOC/E-Stop + pcDNA(OC-E)$_{Rev\text{-}Ala}$	**GC**A	Ala
rOC/E-Stop + pcDNA(OC-E)$_{Rev\text{-}Leu}$	**TT**A	Leu
rOC/E-Stop + pcDNA(OC-E)$_{Rev\text{-}Trp}$	TG**G**	Trp

Figure 4.2 **Transient rescue of HCoV-OC43 E protein deletion mutant yields infectious virions but introduces strong selection pressure for reversion.** (A) Evaluation of infectious recombinant virus production after transient co-transfection of pBAC-E-Stop and 1µg of pcDNA(OC-E) in BHK-21 cells (BHK 0) and following subsequent amplifications on HRT-18 epithelial cells (HRT 1-3). Infectious viral titer differences observed between experiments, revealed, by sequencing (B), the appearance of reversions at the position in the E gene where a stop codon was introduced are indicated by bold and underline. LOD, limit of detection. ┼ (cross) indicates appearance of reversion(s) in the HCoV-OC43 E gene in viral stocks as detected by sequencing.

Neuronal cells are susceptible to infection with HCoV-OC43 lacking E protein but progeny virus production is severely inhibited. HCoV-OC43 is neuroinvasive (Arbour *et al.*, 2000) and neurotropic, with the neuron being the main target of infection in the CNS (Jacomy *et al.*, 2006, Jacomy *et al.*, 2003). Therefore, we sought to investigate whether the absence of the E protein would modify these neurotropic capacities by infecting a susceptible differentiated human neuronal cell line (LA-N-5) or mixed primary cultures of murine CNS cells. Initially complemented rOC/E-Stop, previously recovered from transfection on BHK-21 cells (P0), was used for infection and infectious viral titers determined over a period of 72 hours post-infection (hpi). This revealed an important decrease of infectious virus production for human cells (Figure 4.3A), which was exacerbated in primary murine cells, where virus titers were under the limit of detection (Figure 4.3B). However, in these primary cultures, low levels of infected cells were visualized by immunofluorescence (IFA), suggesting that infection was possible even for the complemented rOC/E-Stop virus but that propagation was severely inhibited (Figure 4.3C, Figure 4.6 panels a-i).

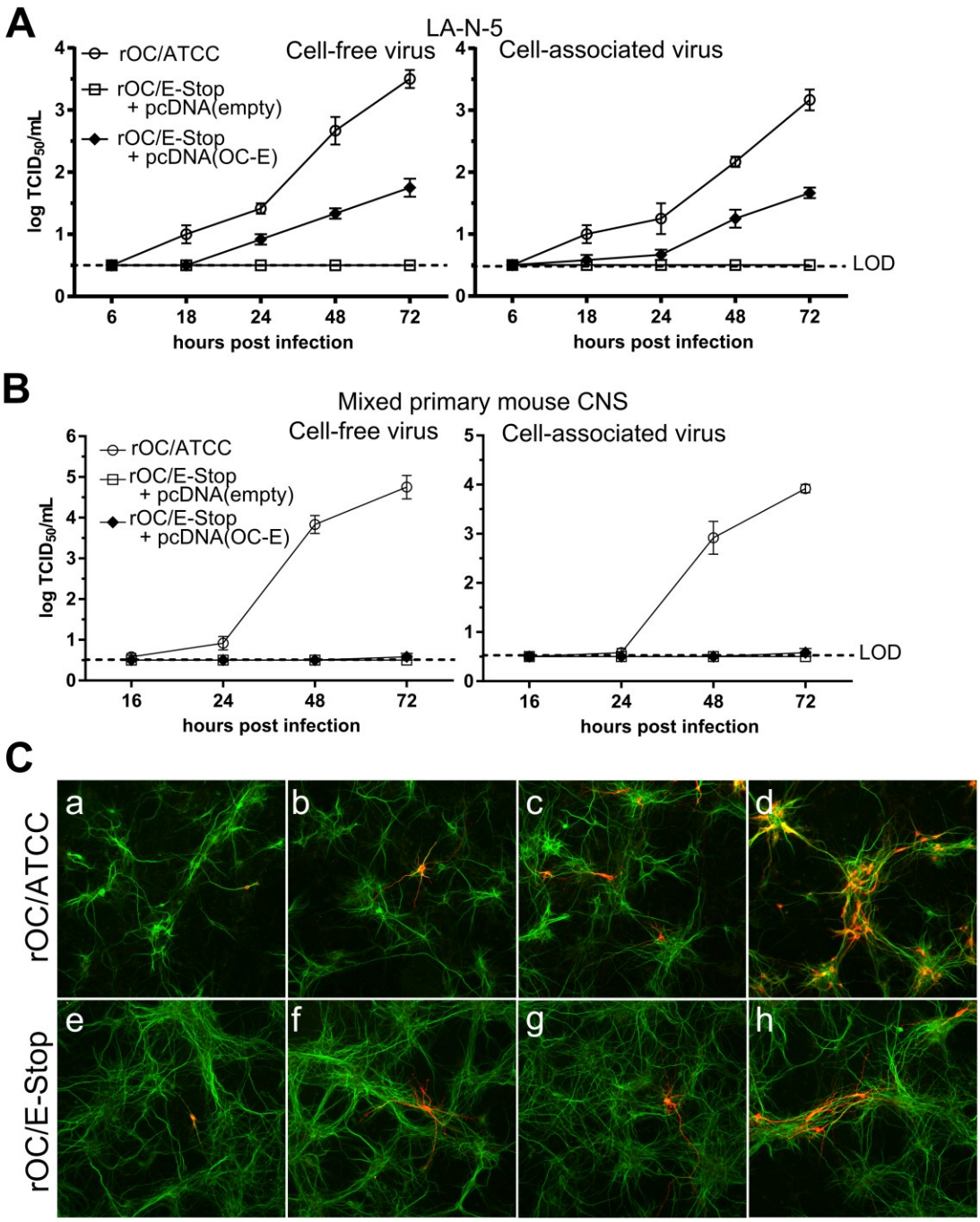


Figure 4.3 **HCoV-OC43 lacking E protein can infect neuronal cells but replication and propagation are severely impaired.** LA-N-5 (A) and mixed primary mouse CNS cells (B) were infected with virus lacking E protein, rOC/E-Stop + pcDNA(empty), or initially complemented virus, rOC/E-Stop + 2µg pcDNA(OC-E), and cell-free and cell-associated virus infectious titers were determined over 72 hours post-infection. (C) Immunofluorescence assay on mixed primary murine CNS cells showed restricted propagation of initially complemented rOC/E-Stop compared to reference virus at 48 hours post-infection. Subpanels represent 8 (a, e), 16 (b,f), 24 (c, g), and 48 (d, h) hours post-infection. Green represents the microtubule associated protein 2 (MAP2) staining in neurons; red represents viral antigens.

HCoV-OC43 E protein putative transmembrane domain integrity is important for efficient infectious virion production and efficient infection of neuronal cells. The transmembrane domain of some coronavirus E proteins is known to homo-oligomerize in membranes and appears to modulate infectious virus production (Nieto-Torres *et al.*, 2014). In order to determine the effect of HCoV-OC43 E protein transmembrane domain on virus production in cell culture, pBAC-OC43^{FL} was modified at key amino acid positions previously identified in other coronaviruses to be critical for the stability of this specific domain (Nieto-Torres *et al.*, 2014, Ruch *et al.*, 2012b) and compared against wild-type virus during infection of cells. The large, polar glutamine at position 17 of the HCoV-OC43 E protein putative transmembrane domain was modified into a smaller, non-polar alanine (pBAC-E-TM-Q17A) in order to diminish any possible ion channel selectivity conveyed by this amino acid (Pervushin *et al.*, 2009) at the opening of the putative ion channel. Using the same methods, a second amino acid, the small leucine at position 27, was substituted with a much larger and bulkier tryptophan (pBAC-E-TM-L27F) in order to destabilize and/or block a region of the TM domain that has previously been determined to be the monomer-monomer interface of the ion channel for SARS-CoV E protein (Pervushin *et al.*, 2009).

Transfection of transmembrane mutants in BHK-21 cells yielded detectable virus titers of rOC/E-TM$_{Q17A}$, while rOC/E-TM$_{L27F}$ remained under the limit of detection (Figure 4.4A). In both cases, viral RNA encoding the E gene was detected, indicating that the transfection was successful with active transcription even if no infectious virus was detected for the latter (Fig. 4B). Recovered rOC/E-TM$_{Q17A}$ was further amplified on HRT-18 cells (P1) and replicated significantly less compared to reference virus (Figure 4.4A). Infection of human LA-N-5 cells and mixed primary cultures of mouse CNS cells showed a similar deficiency in virus production over a period of 72 hpi in the cell-free fractions (Figure 4.4C and D, left panel). However, in the cell-associated fractions, the amount of recovered infectious virus particles was almost identical to those of the reference virus (Figure 4.4C and D, right panel), suggesting a possible defect in virus release. These findings are supported by immunofluorescence assay where there was a

delay in propagation of rOC/E-TM$_{Q17A}$ (Figure 4.6 panels j-l) over 48 hpi compared to reference virus (Figure 4.6 panels a-c).

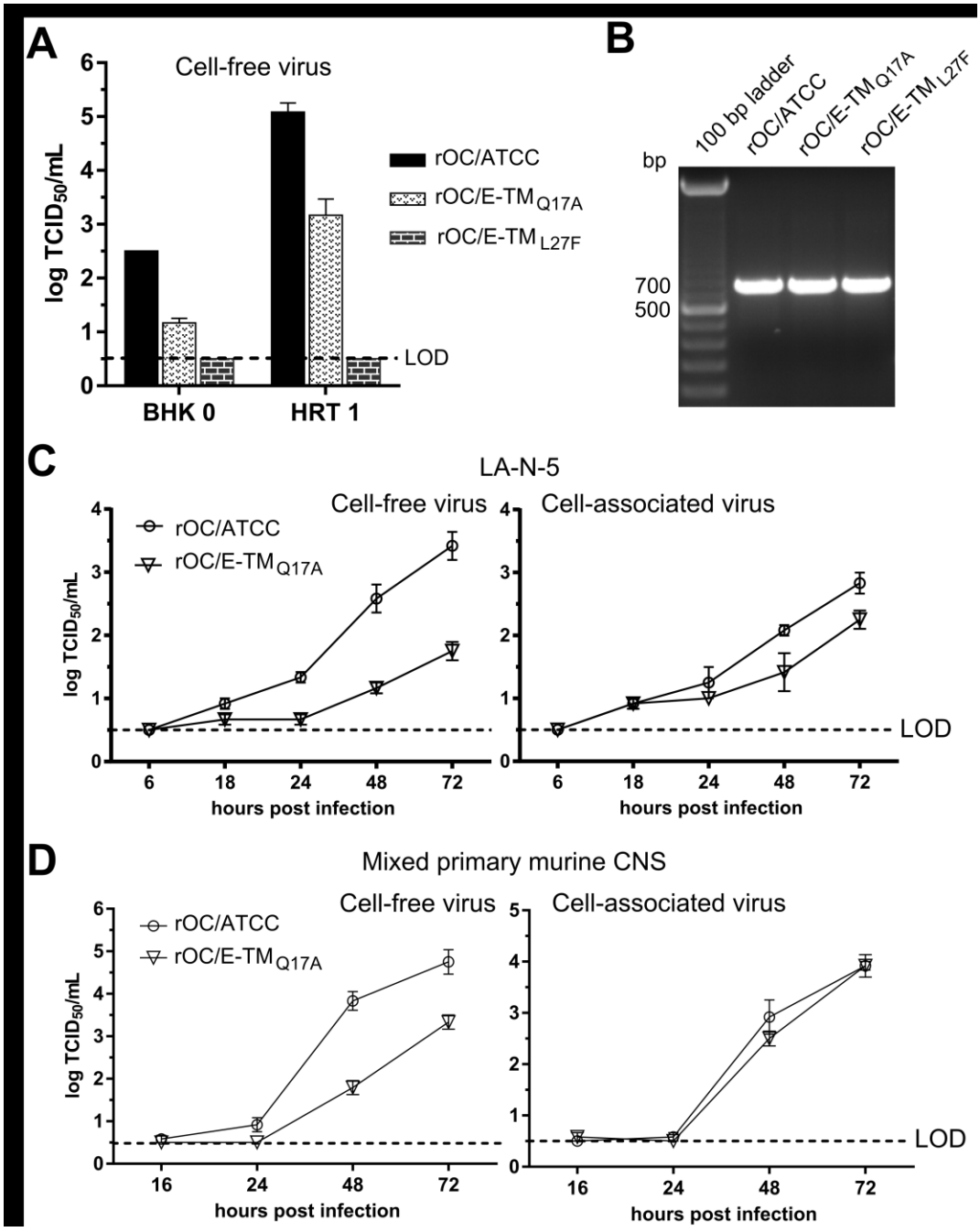


Figure 4.4 **The putative HCoV-OC43 transmembrane domain plays an important role in infectious virion production and replication in neuronal cells.** (A) Production of infectious virus after transfection of E protein transmembrane mutants in BHK-21 cells (BHK 0) and amplification on HRT-18 cells (HRT 1). (B) HCoV-OC43 E gene was detected by RT-PCR after transfection of BHK-21 cells. (C) LA-N-5 human neuronal cells and (D) mixed primary mouse CNS cells were infected with rOC/E-TM$_{Q17A}$ and cell-free and cell-associated virus fractions were recovered over 72 hours. LOD, limit of detection.

HCoV-OC43 E protein putative C-terminal protein-protein interaction motif is critical for efficient infectious virion production and dissemination. Bioinformatics analysis and modeling suggest that several coronavirus species possess a PDZ-domain binding motif (PBM) at the extreme C-terminus of their E proteins that could interact with cellular and viral proteins (Jimenez-Guardeno *et al.*, 2014). We sought to investigate whether the putative four-amino acid E protein PBM of HCoV-OC43 modulates production of infectious particles and infection of susceptible cells. To this end, we modified our cDNA infectious clone to change the two key amino acids of the putative PBM motif, at the -0 and -2 positions from the C-terminal end respectively, into inert alanines, and thereby abrogated putative motif recognition by potential interaction partner(s). Single amino acid mutants (pBAC-OC-E-PBM-D82A and pBAC-OC-E-PBM-V84A) or double mutant (pBAC-OC-E-PBM-D82A-V84A) were transfected in BHK-21 cells and amplified on HRT-18 cells at the same multiplicity of infection and compared to reference virus (Figure 4.5A). Amplification on HRT-18 cells demonstrated that viral titers of the double mutant were significantly decreased compared to other viruses.

To investigate whether the ability to infect susceptible cells, replicate and disseminate is affected by the putative C-terminal PBM in the context of the CNS, LA-N-5 or mixed primary cultures of murine CNS cells were infected with single or double mutant PBM viruses and viral titers and propagation were analyzed. In LA-N-5 cells, after 18 hpi, the titers of rOC/E-PBM$_{D82A-V84A}$ were significantly decreased in the cell-free and cell-associated virus fraction compared to single PBM mutants or reference viruses and total infectious virus titers of the double mutant was severely altered over 72 hours (Figure 4.5B). This trend was exacerbated in primary mixed murine CNS cultures, which showed no detection of infectious rOC/E-PBM$_{D82A-V84A}$ compared to single mutant PBM and reference viruses (Figure 4.5C). Immunofluorescence analysis detected a significant difference in propagation for both LA-N-5 cells (Figure 4.5D) and primary mixed murine CNS cultures (Figure 4.6). The reference virus and single mutant PBM viruses (Figure 4.5D, panels b-d, g-i; Figure 4.6 panels a-c, m-r) were able to disseminate similarly through the cell culture after 48 hpi, while the propagation of the

double PBM mutant rOC/E-PBM$_{D82A-V84A}$ (Figure 4.5D, panels e, j; Figure 4.6, panels s-u) was significantly reduced at the same time point.

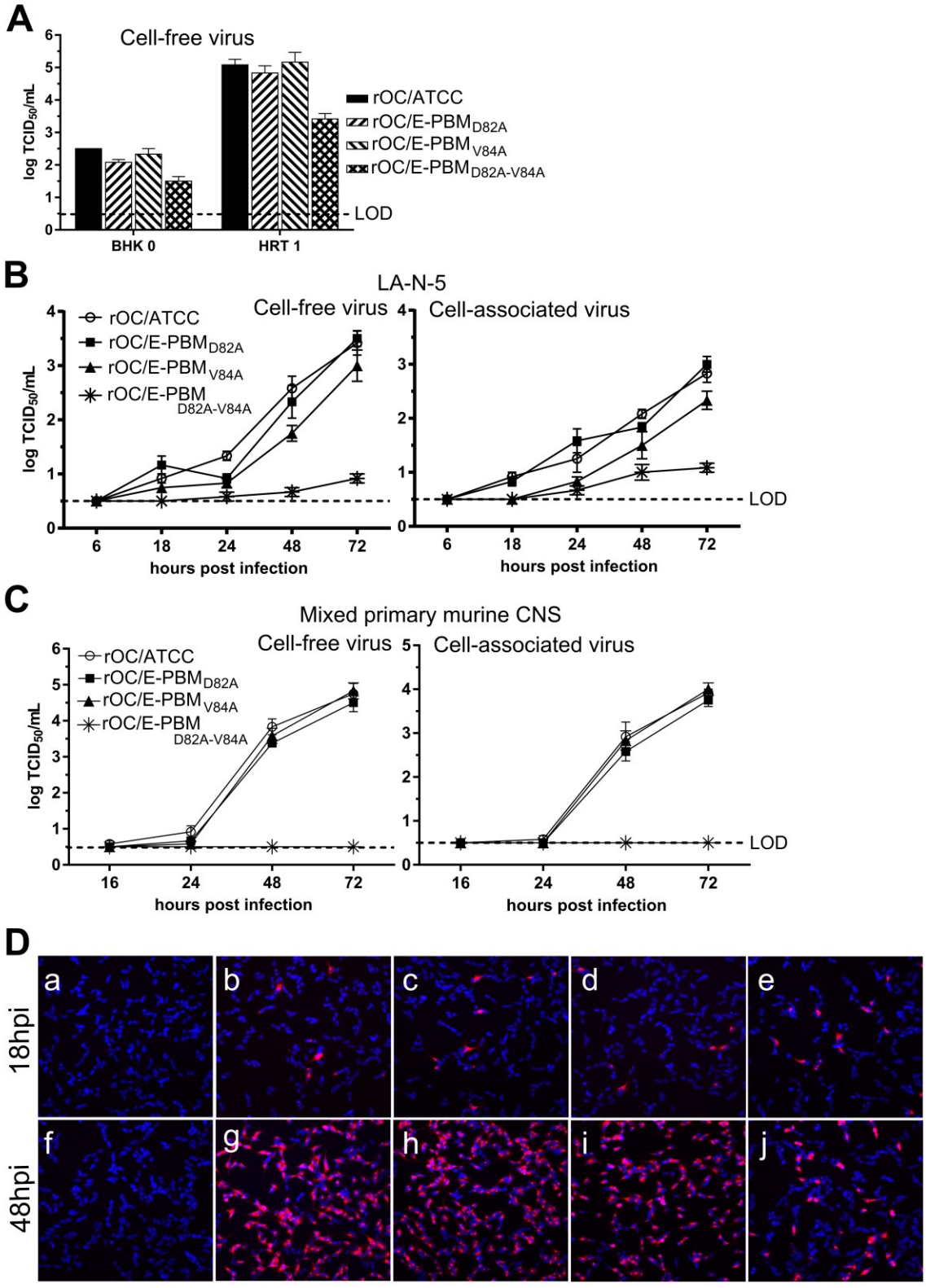


Figure 4.5 **A functional PDZ-domain binding motif at the C-terminal end of the HCoV-OC43 E protein is critical for efficient virus production and spread in neuronal cells.** (A) Production of infectious virus after transfection of various E protein mutants with fully or partially abrogated predicted PBM after transfection in BHK-21 cells (BHK 0) and amplification on HRT-18 cells (HRT 1). The results show a representative experiment. Cell-free and cell-associated infectious virus titers were determined at indicated timepoints after infection at an MOI of 0.05 on (B) LA-N-5 cells and (C) mixed primary cultures of mouse CNS cells. (D) Viral propagation trends in LA-N-5 cells evaluated at 18 and 48 hours post-infection using immunofluorescence assay. Virus detected using an antibody against the HCoV-OC43 S protein (red) and nuclei detected using DAPI (blue). Subpanels: (a, f) mock, (b, g) rOC/ATCC, (c, h) rOC/E-PBM$_{D82A}$, (d, i) rOC/E-PBM$_{V84A}$, (e, j) rOC/E-PBM$_{D82A-V84A}$. LOD, limit of detection.

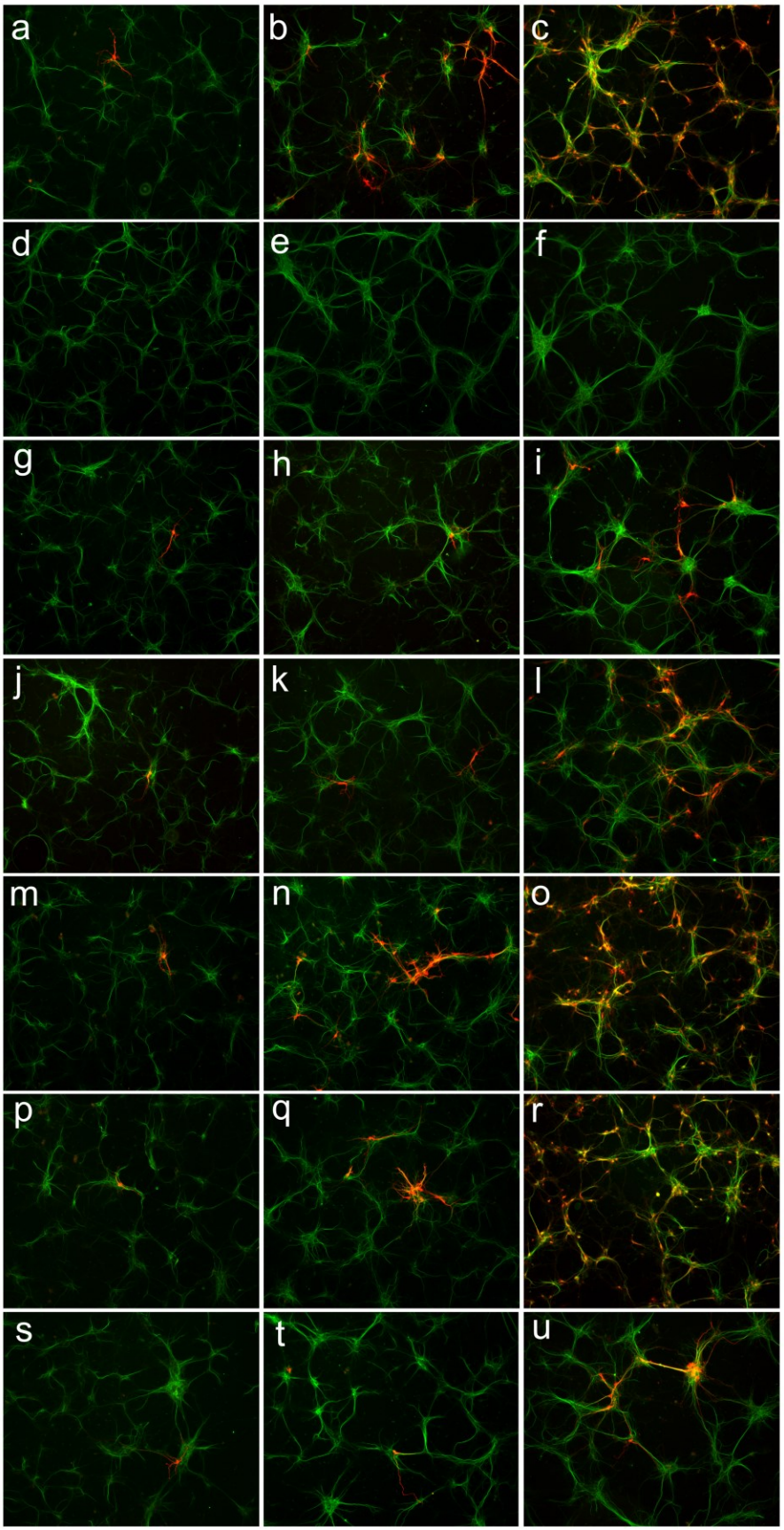


Figure 4.6 Deletion of HCoV-OC43 E protein or key positions of specific domains found therein affect virus spread in mixed primary mouse CNS cells. IFA on infected mixed primary murine CNS cells over 48h. Green represents the microtubule associated protein 2 (MAP2) staining in neurons; red represents viral antigens.

First, second and third column represent 16, 24 and 48 hpi, respectively.

Subpanels:

(a-c) rOC/ATCC,

(d-f) rOC/E-Stop initially co-transfected with 2µg pcDNA(empty),

(g-i) initially complemented rOC/E-Stop (with 2µg pcDNA(OC-E)),

(j-l) rOC/E-TM$_{Q17A}$,

(m-o) rOC4/E-PBM$_{D82A}$,

(p-r) rOC4/E-PBM$_{V84A}$,

(s-u) rOC4/E-PBM$_{D82A-V84A}$.

4.6 Discussion

In this study, by modifying a full-length cDNA infectious clone of the human HCoV-OC43 virus, we demonstrate that its fully functional E protein was critical for the production of infectious virions and spreading in epithelial and neuronal cell culture. Furthermore, transient complementation with reference E protein rescued infectious viral production and a strong selection pressure to revert to a functional E protein was observed, underlining the importance of the protein for efficient infection.

Deletion of the E protein leads to varying degrees of defects for coronaviruses. Indeed, whereas murine hepatitis virus (MHV) and SARS-CoV, are attenuated, showing a reduced ability to produce infectious virus without E protein, in a cell-type specific manner (DeDiego *et al.*, 2007, DeDiego *et al.*, 2008, Jimenez-Guardeno *et al.*, 2015, Kuo *et al.*, 2003), transmissible gastroenteritis virus (TGEV) (Ortego *et al.*, 2002) and MERS-CoV (Almazan *et al.*, 2013) are replication competent, but completely propagation defective, with no detectable infectious virus production when the E protein is deleted. Similarly, we were able to rescue infectious rOC/E-Stop production by providing wild-type E protein in trans, as the recovery of initially complemented rOC/E-Stop through complementation, and amplification on epithelial cells yielded detectable infectious virus. As suggested for MERS-CoV (Almazan *et al.*, 2013), the apparent low titer detected after the first passage on HRT-18 cells (especially at P1), could be due to a transfer of detached cells transfected with the initially complemented pBAC-OC43-E-Stop. Furthermore, the production of infectious particles (for initially complemented mutant lacking the E protein) was low and decreased with subsequent amplification attempts. On the other hand, these results may also suggest that production of infectious virions without E protein is possible but with severely affected efficiency, underlining the requirement of a fully functional E protein. This concept is emphasized by the appearance of recombinant HCoV-OC43 E protein revertants at a very low passage number (P2 or P3) on HRT-18 cells. Indeed, the appearance of revertants with different E sequence after only a few rounds of amplification on HRT-18 cells indicates that viral particles must have been produced earlier during the process as we already

observed previously for HE-deleted recombinant HCoV-OC43 (Desforges *et al.*, 2013). We observed strong selective pressure at the position where we introduced a stop codon by reverse genetics where one or two nucleotide changes led to reversion to reference E sequence or to another amino acid residue (tryptophan or leucine). It was previously described that SARS-CoV (Jimenez-Guardeno *et al.*, 2015) and MHV (Kuo *et al.*, 2010) E protein deletion mutants underwent compensatory mutations after a few passages in culture to utilize a partially duplicated version of the adjacent M protein to recover partial virus production. A second type of reversion was observed for SARS-CoV E protein deletion mutant after an intranasal infection of susceptible mice as the small transmembrane ion channel forming 8a was modified to incorporate a potential PBM associated with increased infectious virus production compared to E protein deletion mutant (Jimenez-Guardeno *et al.*, 2015). Our results rather support the hypothesis that there is selective pressure to specifically restore the E protein functionality itself. Production of infectious virus was reduced after infection of neuronal cell cultures with initially complemented rOC/E-Stop compared to reference virus. An assay for viral proteins confirmed these latter results showing no defect in entry for the initially complemented rOC/E-Stop virus compared to reference virus at 16 hpi, while observations at 72 hours indicated an important defect of viral propagation. A decrease in virus spread was also reported for MHV (Kuo *et al.*, 2007, Kuo *et al.*, 2003) and SARS-CoV (DeDiego *et al.*, 2007) ΔE mutants which formed smaller and less numerous plaques.

Recently, the coronavirus E protein has emerged as a virulence factor (reviewed extensively in (DeDiego *et al.*, 2014)) and there have been extensive efforts to characterize the different domains of this relatively small transmembrane protein and the possibility that it acts as a viroporin, with ion channel activity. The transmembrane domain of several coronavirus E proteins was predicted (Torres *et al.*, 2005) and shown to form ion channels permeable to small cations in artificial membranes for species such as IBV, MHV, HCoV-229E (Wilson *et al.*, 2006), MERS-CoV (Surya *et al.*, 2015) and SARS-CoV (Wilson *et al.*, 2004). Furthermore, it was demonstrated that the IBV E protein can exist in a penta-oligomeric state (Westerbeck *et al.*, 2015), and that in

MERS-CoV (Surya *et al.*, 2015) and SARS-CoV (Pervushin *et al.*, 2009, Torres *et al.*, 2006, Verdia-Baguena *et al.*, 2013, Verdia-Baguena *et al.*, 2012) it forms pentameric channels in lipid membranes. Chemical inhibition of ion channel activity (Wilson *et al.*, 2006), destruction of the integrity of (Almazan *et al.*, 2014, Regla-Nava *et al.*, 2015, Y. Ye *et al.*, 2007) or replacement of the TM domain with those of other viral species (Ruch *et al.*, 2011) all led to reduced viral titers for other coronaviruses species. The E protein of coronaviruses is largely localized within the secretory pathway (Cohen *et al.*, 2011, Nieto-Torres *et al.*, 2011, Venkatagopalan *et al.*, 2015). Furthermore, it has recently been described that the E protein of IBV exists in two different pools; one of monomeric E proteins that disrupt the secretory pathway and a second pool in an oligomeric state, likely serving to facilitate the assembly of progeny virions (Westerbeck *et al.*, 2015). Our results show that by altering key amino acids in the predicted transmembrane domain of HCoV-OC43 E protein, infectious virus production is severely altered (Figure 4.4). Indeed, mutation of position L27 in the middle of the putative transmembrane domain of HCoV-OC43 E protein prevented the formation of any detectable infectious virus. Cells expressing IBV E with mutations at homologous position A26 were unable to produce virus-like particles (VLPs), suggested to be due to the inability of these mutants to homo-oligomerize (Westerbeck *et al.*, 2015) while nuclear magnetic resonance (NMR) study of the SARS-CoV E pentameric channel suggests this position represents the most constricted region of the channel (Pervushin *et al.*, 2009). Residue L27 of HCoV-OC43 E protein could represent a critical amino acid during infection for homo-oligomerization and/or ion channel activity and could explain why we were unable to recover infectious virus.

Mutation at the Q17 position of the HCoV-OC43 E protein, also in the TM domain, significantly reduced infectious viral titers in epithelial and neuronal cells. Interestingly, after infection of both human cells and murine primary CNS cultures by rOC/E-TMQ17A, cell-associated virus fractions showed equivalent levels compared to reference infectious virus but the amount of cell-free virus was significantly reduced, suggesting a defect in virus release, as what was observed for IBV harboring an E protein without ion conductance activity associated with the TM domain (Ruch *et al.*,

2011). Furthermore, mutation at homologous position T16 of IBV did not have an effect on VLP formation but rather was required for secretory pathway disruption (Ruch *et al.*, 2012b, Westerbeck *et al.*, 2015). These results suggest that mutation Q17A (homologous to IBV T16A) in the HCoV-OC43 E protein putative transmembrane domain plays a role in modulating infectious virus release in the cell culture medium. Equivalent amino acid positions N15 and V25 in SARS-CoV were shown to be important in ion channel activity (Verdia-Baguena *et al.*, 2012) and viral fitness but mutations at these positions did not have a significant effect in viral titers compared to wild-type virus (Nieto-Torres *et al.*, 2014). However, smaller plaques were observed in mutant conditions, suggested by the authors to be due to a lack of ion channel activity which could confer cytopathic effects as visualized by plaque assay (Nieto-Torres *et al.*, 2014).

A four-amino acid C-terminal PBM protein-protein interaction motif has been predicted for HCoV-OC43 (Jimenez-Guardeno *et al.*, 2014). Through the replacement of the key amino acids of this motif by alanines, we demonstrated its importance in infectious virion production in the epithelial and neuronal cells tested (Figure 4.5). Defective propagation without an effect on the ability to infect cells was observed with the recombinant virus with an abrogated putative PBM. Deletion of the PBM in SARS-CoV E protein led to slight decreases in viral titers in some cell types (Jimenez-Guardeno *et al.*, 2014) while in others viral titers remained unaffected (Regla-Nava *et al.*, 2015); for HCoV-OC43, we observed significant infectious virus production defects in an epithelial cell line which was further accentuated in neurons. Among coronavirus E proteins, only the SARS-CoV E protein has been shown to per se possess such a functional motif, which interacts with PALS1 to disturb secretory pathway membranes to alter tight junction formation (Teoh *et al.*, 2010) and syntenin to play a role in the exacerbated inflammatory response typical of infection via p38 MAPK activation (Jimenez-Guardeno *et al.*, 2014). For SARS-CoV E protein, the PBM is suggested to be important in two independent functions: virus stability and virulence/pathogenesis rather than virus production (Jimenez-Guardeno *et al.*, 2014, Jimenez-Guardeno *et al.*, 2015, Regla-Nava *et al.*, 2015). It is interesting to note that alanine substitutions in the MHV E

protein C-terminus, effectively abrogating the PBM, did not yield recoverable virus, and is suggested by the authors to be a lethal mutation (Fischer *et al.*, 1998); however there have not been further investigations to help explain these results. Given that the functions of PBM is dependent on their sequence and surrounding sequence context (F. Ye *et al.*, 2013), further study of the HCoV-OC43 E protein neuronal interactome is warranted and is currently underway to help explain our results. Prevention of a functional interaction between HCoV-OC43 E protein PBM and its cellular or viral PDZ domain-containing ligand could conceivably function in the same vein as seen in case of neurotropic encephalitic rabies virus, for which differences in disease phenotype, rapid versus attenuated spread of virus infection, was recently attributed to differences in PBM sequences on the C-terminal of the rabies envelope glycoprotein G leading to different cellular interaction partners to mediate either neuronal cell survival or death (Prehaud *et al.*, 2010). With the putative C-terminal HCoV-OC43 E protein PBM abrogated, critical cellular and/or viral interactions for infectious virion assembly could be lacking, making infectious virus production excessively inefficient, and could potentially be linked to neuropathologies and glutamate excitoxicity, that we have previously observed in mice (Brison *et al.*, 2011, Jacomy *et al.*, 2006, Jacomy *et al.*, 2003), possibly by interfering with PDZ-domain containing proteins found in neuronal cells (Feng *et al.*, 2009).

In summary, the current study demonstrates the critical importance of a fully functional HCoV-OC43 E protein to infectious virus production and efficient spread in both epithelial and neuronal cell culture. Modifications to key amino acids in different functionally important domains the transmembrane domain (potentially associated with a viroporin function) modulated infectious virion production and delayed virus spread in human and murine neuronal cell culture and points towards the presence of a true transmembrane domain and possible a role in the secretory pathway, as seen for other coronaviruses. Furthermore, our results demonstrate for the first time, that the HCoV-OC43 E protein putative C-terminal PDZ-domain binding motif plays a significant role in infectious virion production and efficient virus spread, especially in neuronal cell cultures. Functional viroporins and viral PBM are associated with viral pathogenesis for

a growing number of viruses and their studies to better understand virus-host interaction represent an emerging field (Javier *et al.*, 2011, Nieva *et al.*, 2012, Scott *et al.*, 2015). Considering that the HCoV-OC43 E protein seems to possess both functions, future studies regarding its impact and function in HCoV-OC43 neuropathogenesis already described (Brison *et al.*, 2011, Jacomy *et al.*, 2006, Le Coupanec *et al.*, 2015) is necessary.

4.7 Acknowledgements

We thank Alain Le Coupanec and Mathieu Meessen-Pinard for their preparation of mixed primary murine CNS cells. J.K.S. wishes to thank Alain Le Coupanec, Mathieu Meessen-Pinard, Guillaume Dubois, Mathieu Dubé, and especially M.D. and P.J.T for their advice and support throughout the project. Additionally, J.K.S. wishes to thank M.D. and P.J.T. for guidance in the preparation of this manuscript.

4.8 Funding Information

This study was supported by Discovery grant 42619-2009 from the National Sciences and Engineering Research Council of Canada and operating grant MT-9203 from the Institute of Infection and Immunity of the Canadian Institutes for Health Research to P.J.T. who is the holder of the Tier-1 (Senior) Canada Research Chair in Neuroimmunovirology award. J.K.S. gratefully acknowledges a masters studentship from the *Fondation Universitaire Armand-Frappier de l'INRS*. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

LA PROTÉINE D'ENVELOPPE (E) DU CORONAVIRUS RESPIRATOIRE HUMAIN HCOV-OC43 EST NÉCESSAIRE POUR LA FORMATION DE VIRIONS INFECTIEUX DANS LES CELLULES ÉPITHÉLIALES ET NEURONALES

III – DISCUSSION & CONCLUSION

5 DISCUSSION & CONCLUSION

This thesis presents the first study of the human coronavirus OC43 E protein with the findings outlined clearly in the article manuscript submission presented in Chapter 4. Several characteristics of this small transmembrane protein were studied, the first being its overall importance for infectious virus production. By modifying a full-length cDNA infectious clone to prevent E protein production, we demonstrated that the E protein is critical to infectious virus production and spread. Prevention of E protein production and subsequent complementation studies with wild-type HCoV-OC43 E protein led to a similar phenotype seen with studies of TGEV (Ortego *et al.*, 2002) and MERS-CoV (Almazan *et al.*, 2013) which are replication competent but propagation defective in the absence of E protein. This is vastly different from coronavirus species such as MHV and SARS-CoV which can still produce significant levels of infectious virus after several passages (although compared to wild-type virus, may be decreased anywhere from 20- to 1000-fold) (DeDiego *et al.*, 2007, DeDiego *et al.*, 2008, Jimenez-Guardeno *et al.*, 2015, Kuo *et al.*, 2003).

HCoV-OC43 appears to most closely resemble the MERS-CoV E phenotype in terms of E protein importance; it has been demonstrated that low infectious virus titers can be detected after one passage of an initially-complemented E-deleted MERS-CoV (Almazan *et al.*, 2013). However, the HCoV-OC43 virus appears to be marginally more tolerant to E protein deletion compared to MERS-CoV as we were often able to recover detectable viral stocks lacking the E protein for up to three passages in epithelial HRT-18 cells. Nonetheless, this low efficiency in infectious virus propagation increases substantially the apperance of revertants (likely quasi-species that arise at each replication cycle) after two or three passages, which underscores the critical need of the E protein in the virus replication cycle. In our studies of CNS cells (differentiated human LA-N-5 neurons and primary cultures from murine CNS) known to be susceptible to HCoV-OC43 infection (Favreau *et al.*, 2009, Le Coupanec *et al.*, 2015), we saw a similar phenotype of reduced and undetectable levels of infectious virus production and

substantially decreased spread in the primary cultures. This study points to the importance of the E protein in the CNS, an area of great interest for our laboratory.

A logical next step in these studies would be to evaluate the effect of E protein deletion in *in vivo* studies. It would be interesting to see if intracranially inoculated mice would be able to produce detectable progeny virions. Of interest would also be to determine the capacity for spreading if initially-complemented rOC/E-Stop were used and if the resultant virions would contain rOC/E-Stop virus or various revertant viruses would arise as seen in our *in vitro* studies after several passages. Intranasal infection of HCoV-OC43 leading to infection in the brain has been previously demonstrated by our research group (Le Coupanec *et al.*, 2015, St-Jean *et al.*, 2004), however in a pilot study of the nature proposed above, it would be prudent to compare intranasal infections with direct inoculations into the brain to evaluate the relative importance of E protein in the mechanism(s) of HCoV-OC43 neuroinvasion (Desforges *et al.*, 2014). It is interesting to note that intranasal inoculation of ΔE SARS-CoV (which can produce infectious virus *in vitro*, reduced 20- to 200-fold compared to wild-type depending on the cell type) in hACE2 transgenic mice gave titers reduced 100-fold in lung while there was no detectable virus in the brain, even after intracranial inoculation (DeDiego *et al.*, 2008). The recent development of a transgenic hDPP4 mouse suceptible to wild-type MERS-CoV infection has demonstrated infection of the brain 6 days post intranasal inoculation (K. Li *et al.*, 2015), which could help guide the relevance of future in-depth *in vivo* studies and further supports the neuroinvasive capacity of coronaviruses in general.

The coronavirus E protein represents an emerging virulence factor (DeDiego *et al.*, 2014) in addition to the S protein that traditionally holds this role for many coronaviruses (Gallagher *et al.*, 2001, Millet *et al.*, 2015). The coronavirus E protein of SARS-CoV has been associated with the induction of the cell stress response and apoptosis (DeDiego *et al.*, 2011, Nieto-Torres *et al.*, 2014). There have been a number of studies across coronavirus species to determine which domains of the E protein

could be implicated during infection, with two, the transmembrane domain and its potential homo-oligomerization capacity and the C-terminal domain PDZ-binding motif, being intensely investigated in recent years. The investigation of E protein viroporins and C-terminal PBMs (and E protein in general) hopes to yield therapeutic targets/solutions that would be pan-coronavirus effective (Jimenez-Guardeno *et al.*, 2015, Torres *et al.*, 2015).

Although the work presented in this thesis was not designed specifically to confim or refute the presence of an E protein viroporin or the *per se* function of this domain or of the C-terminal PBM, our results suggest their possible involvement during infection of susceptible cells. Deletion of consensus sequence amino acids Q17 and L27 in the TM domain, previously determined to be important for SARS-CoV fitness and pathogenesis, modulated infectious virus production for HCoV-OC43. In the case of L27 we were unable to detect infectious virus after the transfection of the modified cDNA infectious clone on susceptible cells (BHK-21) whereas RT-PCR results suggested that the transfection was successful. However, we did not investigate the production of non-infectious particles nor conduct in-depth genomic analyses using qPCR to fully conclude that no particles were being made or that only particle release was affected. For our Q17 HCoV-OC43 E mutant, we were able to recover infectious virus titers after transfection, however there was a marked decrease in production compared to wild-type virus in epithelial cells. Both LA-N-5 and primary murine CNS cells were susceptible to infection, however infectious titers were decreased compared to wild-type in cell-free virus fractions in both cell types and a delay in propagation was observed in the primary murine cultures, suggesting a defect in virus release. The mutations introduced in the putative TM domain of the HCoV-OC43 E protein show a larger effect on infectious virus production than was demonstrated previously for SARS-CoV (Nieto-Torres *et al.*, 2014); one could speculate therefore that these positions play a role in modulating infection, perhaps due to the mutants inability to homo-oligomerize or inability to function as a viroporin, further studies in artificial membranes would need to be conducted to reach this conclusion.

Many RNA and DNA viruses contain viroporins which can affect viral genome replication and assembly, virus particle entry and release from infected cells and modify general cellular homeostasis (Nieva *et al.*, 2012). Pertubation of membrane gradients and leaking of ionic species from defined organelles can have far-ranging effects processes on the cell, such as trafficking, signalling and induction of cell death (Scott *et al.*, 2015). The rotavirus NSp4 protein is able to homo-oligomerize in endoplasmic reticulum membranes to release Ca^{2+} ions in order to promote the formation of viroplasms and expedite virus release (Hyser *et al.*, 2010, Hyser *et al.*, 2013). Similarly, the 2B protein of the *picornaviridae* family (including poliovirus, coxsackie virus among others) forms tetramers within mammalian cells and and readily permeabilizes vesicles (Agirre *et al.*, 2002, Agirre *et al.*, 2008, Sanchez-Martinez *et al.*, 2008). When localized to the Golgi, 2B expression leads to elevated cytosolic Ca^{2+} levels, which alters vesicle trafficking, induces apoptosis and directly lyses cells as protein levels accumulate (Campanella *et al.*, 2004, de Jong *et al.*, 2004, Sandoval *et al.*, 1997). Interestingly, 2B proteins appear to cause inflammasome activation (Ito *et al.*, 2012) which is similar to recent findings for SARS-CoV E protein ion channel activity induction to NLRP3 inflammasome activation (Nieto-Torres *et al.*, 2015). In the latter case, the E protein was demonstrated to form protein-lipid channels permeable to calcium; this ion channel activity boosted the activation of NLRP3 inflammasome leading to IL-1β overproduction, suggesting a link between ionic disturbances to immunopathological consequences leading to disease worsening in the organism (Nieto-Torres *et al.*, 2015).

Further investigation of the HCoV-OC43 E protein transmembrane domain, in addition to the studies mentioned above, should consist of expressing either the full-length or synthetic peptide of the transmembrane domain of the E protein in artificial membranes and determining their permeability to ions. Furthermore, infections of different cell types (epithelial and neuronal) could be conducted with wild-type HCoV-OC43 and chemical inhibition studies conducted to verify viroporin activity as demonstrated previously for various coronaviruses (Wilson *et al.*, 2006). A panel of

chemical inhibitors could be compiled and a comprehensive study performed using the most common viroporin inhibitors used in the literature such as : hexamethylene amiloride (HMA) and amiloride which block sodium channels and used in HCoV-229E, MHV and SARS-CoV studies (Pervushin *et al.*, 2009, Wilson *et al.*, 2006); amantadine which blocks Influenza A M2 channels and HCV and Dengue replication (Griffin *et al.*, 2003, OuYang *et al.*, 2014); and DIDS [4,40-disothiocyano-2,20-stilbenedisulphonic acid], a chloride channel blocker used recently in a feline coronavirus study (Takano *et al.*, 2015). Another interesting addition to this panel would be memantine, an NDMA receptor blocker, recently demonstrated to alleviate neuropathological symptoms due to glutamate exicitotoxicity after intracranial HCoV-OC43 inoculation in mice (Brison *et al.*, 2014). If the potential HCoV-OC43 E protein viroporin activity can be blocked by memantine, it could suggest a direct role of the E protein in the development of neuropathological symptoms due to glutamate exocitotoxicity after inoculation and an ideal therapeutic target to be further investigated.

The hepatitis C virus (HCV) p7 protein forms hexameric viroporins through oligomerization within the membrane of the endoplasmic reticulum and mitochondria (Carrere-Kremer *et al.*, 2002, Clarke *et al.*, 2006, Griffin *et al.*, 2003). p7 has been studied as a potential therapeutic target due to being demonstrated as essential to HCV propagation *in vivo* in chimpanzee models (Sakai *et al.*, 2003). Of interest is the fact that in addition to its channel activity after oligomerization, the p7 protein is known to perform multiple functions within infected cells through distinct protein-protein interactions (Scott *et al.*, 2015). Specific interactions with viral proteins NS2, core and envelope glycoproteins have been demonstrated (reviewed in (Scott *et al.*, 2015)) and are suggested to be interactions that underpin recently described roles for p7 during capsid assembly and the envelopment of HCV particles (Gentzsch *et al.*, 2013, Scott *et al.*, 2015). These roles for viroporin proteins, independent of ion channel conductance capacity, further emphasize the multiple roles of these small transmembrane proteins within the cell.

In our studies, we investigated the effect of full or partial abrogation of the C-terminal PBM of the HCoV-OC43 E protein, previously identified by the Enjuanes laboratory (Jimenez-Guardeno *et al.*, 2014), on infectious virus production and spread in an initial investigation of true function during infection. Our studies show that when the predicted motif is fully abrogated, no effect on ability to infect cells is seen, however infectious virus production is decreased in epithelial cells and further significantly diminished in neuronal cells compared to wild-type virus. Similarly to modifications to the transmembrane domain, SARS-CoV E protein infectious virus production was only slightly decreased or not at all affected when its C-terminal PBM was abrogated or deleted (Jimenez-Guardeno *et al.*, 2014, Regla-Nava *et al.*, 2015). Generally speaking it appears that a fully functional E protein has a larger overall effect on HCoV-OC43 production and spread than on SARS-CoV, a logical conclusion given the first results described regarding the effects of full E protein deletion (**Figure 4.1**).

Additional studies are needed to determine if the same effect in infectious virus production after E protein PBM abrogation can be replicated *in vivo*, using our mouse model. These results can bring us closer to concluding the PBM is functional, however, when considering PBMs, the determination of actual protein interaction(s) that during HCoV-OC43 infection of susceptible cells is the only source of conclusive proof of functionality. Any potential partners elucidated from such a study could help clarify mechanisms of pathogenesis and/or help explain the difference in importance of the E protein between species such as HCoV-OC43 and SARS-CoV. To date, the SARS-CoV E protein PBM has been found to interact with two different proteins to disrupt the development of cell polarity in the lungs (Teoh *et al.*, 2010) and participate in the characteristic infection immunopathology in the same organ (Jimenez-Guardeno *et al.*, 2014). The targeting of viral PBMs to disrupt cellular PDZ domains and corresponding cellular processes is an emerging field and until recently mostly studied within the field of oncogenic viruses (James *et al.*, 2016, Javier *et al.*, 2011). However, neurons and their post-synaptic densities have been shown to be home to a plethora of PDZ domain containing proteins (Feng *et al.*, 2009), but surprisingly the literature regarding viruses "taking advantage" of these domains is still in its infancy. The rabies virus G protein

PBM is the one of the only viral protein PBM extensively studied from a CNS-attaining virus (Caillet-Saguy *et al.*, 2015). Given this information and ability of HCoV-OC43 to invade the CNS, and lack of this type of information for any coronavirus, it would be interesting to study any potential HCoV-OC43 E protein PBM-cellular protein/viral protein interactions in the context of the CNS.

Influenza A virus infects the respiratory tracts of birds and mammals and encodes the NS1 protein which is an important virulence determinant and whose general function is to counteract antiviral mechanisms of the innate immune system (Javier *et al.*, 2011). Influenza A NS1 has been determined to encode distinct C-terminal PBM sequences depending on whether the strain infects humans or avian species (Obenauer *et al.*, 2006). Infections with the two types of strains demonstrated more severe disease in mice infected with the avian strain compared to the human strain (Jackson *et al.*, 2008, Soubies *et al.*, 2010) suggesting a difference in protein-protein interactions mediated by the different PBM. Indeed, a GST pull-down assay showed that the avian PBM sequence, but not the human strain equivalent, could specifically associate *in vitro* with at least five cellular PDZ proteins (H. Liu *et al.*, 2010, Thomas *et al.*, 2011). Two of these, proteins Scribble and Dlg, were found to disrupt tight junction formation and potentially contribute to viral pathogenesis (Golebiewski *et al.*, 2011). Depending on which cellular/viral proteins are found to interact with the HCoV-OC43 (assuming the PBM is functional), the results could help explain why replication of HCoV-OC43 is significantly more affected for this species compared to SARS-CoV and/or why the pathology after infection of the lungs is much milder in the case of HCoV-OC43 compared to SARS-CoV and MERS-CoV, for example.

In conclusion, this study demonstrated for the first time the critical importance of the HCoV-OC43 E protein for infectious virus production and propagation in both epithelial and neuronal cells, showing the closest deletion phenotype to MERS-CoV E protein. Studies of the putative transmembrane domain region of this protein revealed key amino acids implicated in modulating infectious virus production and spread in the

same cell types and opened the door to further investigations to determine if the HCoV-OC43 E protein could act as a true viroporin. Furthermore, by modifying a predicted C-terminal PDZ binding motif (PBM), we were able to demonstrate the importance of this motif in infectious virus production and spread in epithelial and neuronal cells. Future studies should focus on studying the effect of infection of mice with the various mutants produced herein to give an initial understanding of the HCoV-OC43 E protein and its specific domains' role(s) in neuropathogenesis. Concurrently, studies determining interaction partners of the HCoV-OC43 E protein PBM in the context of the CNS (beginning *in vitro*) should be conducted in order to determine a neural-interactome that could potentially solidify the connection between HCoV-OC43 infection and the development of neuropathologies seen in mice.

6 REFERENCES

Agirre A, Barco A, Carrasco L & Nieva JL (2002) Viroporin-mediated membrane permeabilization. Pore formation by nonstructural poliovirus 2B protein. *The Journal of biological chemistry* 277(43):40434-40441.

Agirre A, Lorizate M, Nir S & Nieva JL (2008) Poliovirus 2b insertion into lipid monolayers and pore formation in vesicles modulated by anionic phospholipids. *Biochimica et biophysica acta* 1778(11):2621-2626.

Aldabe R, Barco A & Carrasco L (1996) Membrane permeabilization by poliovirus proteins 2B and 2BC. *The Journal of biological chemistry* 271(38):23134-23137.

Aldabe R, Irurzun A & Carrasco L (1997) Poliovirus protein 2BC increases cytosolic free calcium concentrations. *Journal of virology* 71(8):6214-6217.

Almazan F, DeDiego ML, Sola I, Zuniga S, Nieto-Torres JL, Marquez-Jurado S, Andres G & Enjuanes L (2013) Engineering a replication-competent, propagation-defective Middle East respiratory syndrome coronavirus as a vaccine candidate. *MBio* 4(5):e00650-00613.

Almazan F, Sola I, Zuniga S, Marquez-Jurado S, Morales L, Becares M & Enjuanes L (2014) Coronavirus reverse genetic systems: infectious clones and replicons. *Virus research* 189:262-270.

Alvarez E, DeDiego ML, Nieto-Torres JL, Jimenez-Guardeno JM, Marcos-Villar L & Enjuanes L (2010) The envelope protein of severe acute respiratory syndrome coronavirus interacts with the non-structural protein 3 and is ubiquitinated. *Virology* 402(2):281-291.

Arabi YM, Harthi A, Hussein J, Bouchama A, Johani S, Hajeer AH, Saeed BT, Wahbi A, Saedy A, AlDabbagh T, Okaili R, Sadat M & Balkhy H (2015) Severe neurologic syndrome associated with Middle East respiratory syndrome corona virus (MERS-CoV). *Infection* 43(4):495-501.

Arbely E, Khattari Z, Brotons G, Akkawi M, Salditt T & Arkin IT (2004) A highly unusual palindromic transmembrane helical hairpin formed by SARS coronavirus E protein. *Journal of molecular biology* 341(3):769-779.

Arbour N, Day R, Newcombe J & Talbot PJ (2000) Neuroinvasion by human respiratory coronaviruses. *Journal of virology* 74(19):8913-8921.

Bastien N, Anderson K, Hart L, Van Caeseele P, Brandt K, Milley D, Hatchette T, Weiss EC & Li Y (2005) Human coronavirus NL63 infection in Canada. *J Infect Dis* 191(4):503-506.

Baudoux P, Carrat C, Besnardeau L, Charley B & Laude H (1998) Coronavirus pseudoparticles formed with recombinant M and E proteins induce alpha interferon synthesis by leukocytes. *Journal of virology* 72(11):8636-8643.

Belouzard S, Millet JK, Licitra BN & Whittaker GR (2012) Mechanisms of coronavirus cell entry mediated by the viral spike protein. *Viruses* 4(6):1011-1033.

Bonavia A, Arbour N, Yong VW & Talbot PJ (1997) Infection of primary cultures of human neural cells by human coronaviruses 229E and OC43. *Journal of virology* 71(1):800-806.

Boscarino JA, Logan HL, Lacny JJ & Gallagher TM (2008) Envelope protein palmitoylations are crucial for murine coronavirus assembly. *Journal of virology* 82(6):2989-2999.

Brison E, Jacomy H, Desforges M & Talbot PJ (2011) Glutamate excitotoxicity is involved in the induction of paralysis in mice after infection by a human coronavirus with a single point mutation in its spike protein. *Journal of virology* 85(23):12464-12473.

Brison E, Jacomy H, Desforges M & Talbot PJ (2014) Novel treatment with neuroprotective and antiviral properties against a neuroinvasive human respiratory virus. *Journal of virology* 88(3):1548-1563.

Brohm C, Steinmann E, Friesland M, Lorenz IC, Patel A, Penin F, Bartenschlager R & Pietschmann T (2009) Characterization of determinants important for hepatitis C

virus p7 function in morphogenesis by using trans-complementation. *Journal of virology* 83(22):11682-11693.

Caillet-Saguy C, Maisonneuve P, Delhommel F, Terrien E, Babault N, Lafon M, Cordier F & Wolff N (2015) Strategies to interfere with PDZ-mediated interactions in neurons: What we can learn from the rabies virus. *Prog Biophys Mol Biol* 10.1016/j.pbiomolbio.2015.02.007.

Campanella M, de Jong AS, Lanke KW, Melchers WJ, Willems PH, Pinton P, Rizzuto R & van Kuppeveld FJ (2004) The coxsackievirus 2B protein suppresses apoptotic host cell responses by manipulating intracellular Ca2+ homeostasis. *The Journal of biological chemistry* 279(18):18440-18450.

Carrasco L (1978) Membrane leakiness after viral infection and a new approach to the development of antiviral agents. *Nature* 272(5655):694-699.

Carrasco L (1995) Modification of membrane permeability by animal viruses. *Adv Virus Res* 45:61-112.

Carrasco L & Smith AE (1976) Sodium ions and the shut-off of host cell protein synthesis by picornaviruses. *Nature* 264(5588):807-809.

Carrere-Kremer S, Montpellier-Pala C, Cocquerel L, Wychowski C, Penin F & Dubuisson J (2002) Subcellular localization and topology of the p7 polypeptide of hepatitis C virus. *Journal of virology* 76(8):3720-3730.

Chami M, Ferrari D, Nicotera P, Paterlini-Brechot P & Rizzuto R (2003) Caspase-dependent alterations of Ca2+ signaling in the induction of apoptosis by hepatitis B virus X protein. *The Journal of biological chemistry* 278(34):31745-31755.

Chen CC, Kruger J, Sramala I, Hsu HJ, Henklein P, Chen YM & Fischer WB (2011) ORF8a of SARS-CoV forms an ion channel: experiments and molecular dynamics simulations. *Biochimica et biophysica acta* 1808(2):572-579.

Chen SC, Lo SY, Ma HC & Li HC (2009) Expression and membrane integration of SARS-CoV E protein and its interaction with M protein. *Virus genes* 38(3):365-371.

Cho KO, Hunt CA & Kennedy MB (1992) The rat brain postsynaptic density fraction contains a homolog of the Drosophila discs-large tumor suppressor protein. *Neuron* 9(5):929-942.

Clarke D, Griffin S, Beales L, Gelais CS, Burgess S, Harris M & Rowlands D (2006) Evidence for the formation of a heptameric ion channel complex by the hepatitis C virus p7 protein in vitro. *The Journal of biological chemistry* 281(48):37057-37068.

Cohen JR, Lin LD & Machamer CE (2011) Identification of a Golgi complex-targeting signal in the cytoplasmic tail of the severe acute respiratory syndrome coronavirus envelope protein. *Journal of virology* 85(12):5794-5803.

Corse E & Machamer CE (2000) Infectious bronchitis virus E protein is targeted to the Golgi complex and directs release of virus-like particles. *Journal of virology* 74(9):4319-4326.

Corse E & Machamer CE (2002) The cytoplasmic tail of infectious bronchitis virus E protein directs Golgi targeting. *Journal of virology* 76(3):1273-1284.

Corse E & Machamer CE (2003) The cytoplasmic tails of infectious bronchitis virus E and M proteins mediate their interaction. *Virology* 312(1):25-34.

Curtis KM, Yount B & Baric RS (2002) Heterologous gene expression from transmissible gastroenteritis virus replicon particles. *Journal of virology* 76(3):1422-1434.

David-Ferreira JF & Manaker RA (1965) An Electron Microscope Study of the Development of a Mouse Hepatitis Virus in Tissue Culture Cells. *J Cell Biol* 24:57-78.

de Groot RJ (2006) Structure, function and evolution of the hemagglutinin-esterase proteins of corona- and toroviruses. *Glycoconjugate journal* 23(1-2):59-72.

de Groot RJ, Baker SC, Baric RS, Enjuanes L, Gorbalenya AE, Holmes KV, Perlman S, Poon L, Rottier PJM, Talbot PJ, Woo PCY & Ziebuhr J (2012) Coronaviridae. *Virus Taxonomy: ninth report of the International Committee on Taxonomy of Viruses*, King AMQ, Adams MJ, Carstens EB & Lefkowitz EJ (Édit.) Academic

Press, Inc, New York, NYhttp://dx.doi.org/10.1016/B978-0-12-384684-6.00068-9. p 806-828.

de Jong AS, Melchers WJ, Glaudemans DH, Willems PH & van Kuppeveld FJ (2004) Mutational analysis of different regions in the coxsackievirus 2B protein: requirements for homo-multimerization, membrane permeabilization, subcellular localization, and virus replication. *The Journal of biological chemistry* 279(19):19924-19935.

DeDiego ML, Alvarez E, Almazan F, Rejas MT, Lamirande E, Roberts A, Shieh WJ, Zaki SR, Subbarao K & Enjuanes L (2007) A severe acute respiratory syndrome coronavirus that lacks the E gene is attenuated in vitro and in vivo. *Journal of virology* 81(4):1701-1713.

DeDiego ML, Nieto-Torres JL, Jimenez-Guardeno JM, Regla-Nava JA, Alvarez E, Oliveros JC, Zhao J, Fett C, Perlman S & Enjuanes L (2011) Severe acute respiratory syndrome coronavirus envelope protein regulates cell stress response and apoptosis. *PLoS pathogens* 7(10):e1002315.

DeDiego ML, Nieto-Torres JL, Jimenez-Guardeno JM, Regla-Nava JA, Castano-Rodriguez C, Fernandez-Delgado R, Usera F & Enjuanes L (2014) Coronavirus virulence genes with main focus on SARS-CoV envelope gene. *Virus research* 194:124-137.

DeDiego ML, Pewe L, Alvarez E, Rejas MT, Perlman S & Enjuanes L (2008) Pathogenicity of severe acute respiratory coronavirus deletion mutants in hACE-2 transgenic mice. *Virology* 376(2):379-389.

Desforges M, Desjardins J, Zhang C & Talbot PJ (2013) The acetyl-esterase activity of the hemagglutinin-esterase protein of human coronavirus OC43 strongly enhances the production of infectious virus. *Journal of virology* 87(6):3097-3107.

Desforges M, Le Coupanec A, Stodola JK, Meessen-Pinard M & Talbot PJ (2014) Human coronaviruses: viral and cellular factors involved in neuroinvasiveness and neuropathogenesis. *Virus research* 194:145-158.

Dijkman R, Jebbink MF, Gaunt E, Rossen JW, Templeton KE, Kuijpers TW & van der Hoek L (2012) The dominance of human coronavirus OC43 and NL63 infections in infants. *J Clin Virol* 53(2):135-139.

Doedens JR & Kirkegaard K (1995) Inhibition of cellular protein secretion by poliovirus proteins 2B and 3A. *EMBO J* 14(5):894-907.

Drosten C, Gunther S, Preiser W, van der Werf S, Brodt HR, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RA, Berger A, Burguiere AM, Cinatl J, Eickmann M, Escriou N, Grywna K, Kramme S, Manuguerra JC, Muller S, Rickerts V, Sturmer M, Vieth S, Klenk HD, Osterhaus AD, Schmitz H & Doerr HW (2003) Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* 348(20):1967-1976.

Ewart GD, Sutherland T, Gage PW & Cox GB (1996) The Vpu protein of human immunodeficiency virus type 1 forms cation-selective ion channels. *Journal of virology* 70(10):7108-7115.

Facciuto F, Cavatorta AL, Valdano MB, Marziali F & Gardiol D (2012) Differential expression of PDZ domain-containing proteins in human diseases - challenging topics and novel issues. *FEBS J* 279(19):3538-3548.

Favreau DJ, Desforges M, St-Jean JR & Talbot PJ (2009) A human coronavirus OC43 variant harboring persistence-associated mutations in the S glycoprotein differentially induces the unfolded protein response in human neurons as compared to wild-type virus. *Virology* 395(2):255-267.

Favreau DJ, Meessen-Pinard M, Desforges M & Talbot PJ (2012) Human coronavirus-induced neuronal programmed cell death is cyclophilin d dependent and potentially caspase dispensable. *Journal of virology* 86(1):81-93.

Feng W & Zhang M (2009) Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nat Rev Neurosci* 10(2):87-99.

Fischer F, Stegen CF, Masters PS & Samsonoff WA (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(10):7885-7894.

Fouchier RA, Hartwig NG, Bestebroer TM, Niemeyer B, de Jong JC, Simon JH & Osterhaus AD (2004) A previously undescribed coronavirus associated with respiratory disease in humans. *Proceedings of the National Academy of Sciences of the United States of America* 101(16):6212-6216.

Gabanyi MJ, Adams PD, Arnold K, Bordoli L, Carter LG, Flippen-Andersen J, Gifford L, Haas J, Kouranov A, McLaughlin WA, Micallef DI, Minor W, Shah R, Schwede T, Tao YP, Westbrook JD, Zimmerman M & Berman HM (2011) The Structural Biology Knowledgebase: a portal to protein structures, sequences, functions, and methods. *J Struct Funct Genomics* 12(2):45-54.

Gallagher TM & Buchmeier MJ (2001) Coronavirus spike proteins in viral entry and pathogenesis. *Virology* 279(2):371-374.

Gaunt ER, Hardie A, Claas EC, Simmonds P & Templeton KE (2010) Epidemiology and clinical presentations of the four human coronaviruses 229E, HKU1, NL63, and OC43 detected over 3 years using a novel multiplex real-time PCR method. *J Clin Microbiol* 48(8):2940-2947.

Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, Mazet JK, Hu B, Zhang W, Peng C, Zhang YJ, Luo CM, Tan B, Wang N, Zhu Y, Crameri G, Zhang SY, Wang LF, Daszak P & Shi ZL (2013) Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 503(7477):535-538.

Gentzsch J, Brohm C, Steinmann E, Friesland M, Menzel N, Vieyres G, Perin PM, Frentzen A, Kaderali L & Pietschmann T (2013) hepatitis c Virus p7 is critical for capsid assembly and envelopment. *PLoS pathogens* 9(5):e1003355.

Gerna G, Campanini G, Rovida F, Percivalle E, Sarasini A, Marchi A & Baldanti F (2006) Genetic variability of human coronavirus OC43-, 229E-, and NL63-like strains and their association with lower respiratory tract infections of hospitalized infants and immunocompromised patients. *J Med Virol* 78(7):938-949.

Godet M, L'Haridon R, Vautherot JF & Laude H (1992) TGEV corona virus ORF4 encodes a membrane protein that is incorporated into virions. *Virology* 188(2):666-675.

Golebiewski L, Liu H, Javier RT & Rice AP (2011) The avian influenza virus NS1 ESEV PDZ binding motif associates with Dlg1 and Scribble to disrupt cellular tight junctions. *Journal of virology* 85(20):10639-10648.

Gonzalez ME & Carrasco L (1998) The human immunodeficiency virus type 1 Vpu protein enhances membrane permeability. *Biochemistry* 37(39):13710-13719.

Gonzalez ME & Carrasco L (2001) Human immunodeficiency virus type 1 VPU protein affects Sindbis virus glycoprotein processing and enhances membrane permeabilization. *Virology* 279(1):201-209.

Gonzalez ME & Carrasco L (2003) Viroporins. *FEBS letters* 552(1):28-34.

Graham RL, Donaldson EF & Baric RS (2013) A decade after SARS: strategies for controlling emerging coronaviruses. *Nat Rev Microbiol* 11(12):836-848.

Griffin SD, Beales LP, Clarke DS, Worsfold O, Evans SD, Jaeger J, Harris MP & Rowlands DJ (2003) The p7 protein of hepatitis C virus forms an ion channel that is blocked by the antiviral drug, Amantadine. *FEBS letters* 535(1-3):34-38.

Gu J, Gong E, Zhang B, Zheng J, Gao Z, Zhong Y, Zou W, Zhan J, Wang S, Xie Z, Zhuang H, Wu B, Zhong H, Shao H, Fang W, Gao D, Pei F, Li X, He Z, Xu D, Shi X, Anderson VM & Leong AS (2005) Multiple organ infection and the pathogenesis of SARS. *J Exp Med* 202(3):415-424.

Hamre D & Procknow JJ (1966) A new virus isolated from the human respiratory tract. *Proc Soc Exp Biol Med* 121(1):190-193.

Hara M & Takao S (2015) Coronavirus Infections in Pediatric Outpatients with Febrile Respiratory Tract Infections in Hiroshima, Japan, over a 3-Year Period. *Jpn J Infect Dis* 10.7883/yoken.JJID.2014.591.

Hill DP & Robertson KA (1998) Differentiation of LA-N-5 neuroblastoma cells into cholinergic neurons: methods for differentiation, immunohistochemistry and reporter gene introduction. *Brain research. Brain research protocols* 2(3):183-190.

Hsieh P-K, Chang SC, Huang C-C, Lee T-T, Hsiao C-W, Kou Y-H, Chen IY, Chang C-K, Huang T-H & Chang M-F (2005) Assembly of Severe Acute Respiratory Syndrome Coronavirus RNA Packaging Signal into Virus-Like Particles Is Nucleocapsid Dependent. *Journal of virology* 79(22):13848-13855.

Hsieh YC, Li HC, Chen SC & Lo SY (2008) Interactions between M protein and other structural proteins of severe, acute respiratory syndrome-associated coronavirus. *J Biomed Sci* 15(6):707-717.

Huang X, Dong W, Milewska A, Golda A, Qi Y, Zhu QK, Marasco WA, Baric RS, Sims AC, Pyrc K, Li W & Sui J (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. *Journal of virology* 89(14):7202-7213.

Huang Y, Yang Z-y, Kong W-p & Nabel GJ (2004) Generation of Synthetic Severe Acute Respiratory Syndrome Coronavirus Pseudoparticles: Implications for Assembly and Vaccine Production. *Journal of virology* 78(22):12557-12565.

Humbert P, Russell S & Richardson H (2003) Dlg, Scribble and Lgl in cell polarity, cell proliferation and cancer. *Bioessays* 25(6):542-553.

Hyser JM, Collinson-Pautz MR, Utama B & Estes MK (2010) Rotavirus disrupts calcium homeostasis by NSP4 viroporin activity. *MBio* 1(5).

Hyser JM, Utama B, Crawford SE, Broughman JR & Estes MK (2013) Activation of the endoplasmic reticulum calcium sensor STIM1 and store-operated calcium entry by rotavirus requires NSP4 viroporin activity. *Journal of virology* 87(24):13579-13588.

Ito M, Yanagi Y & Ichinohe T (2012) Encephalomyocarditis virus viroporin 2B activates NLRP3 inflammasome. *PLoS pathogens* 8(8):e1002857.

Ivarsson Y, Arnold R, McLaughlin M, Nim S, Joshi R, Ray D, Liu B, Teyra J, Pawson T, Moffat J, Li SS, Sidhu SS & Kim PM (2014) Large-scale interaction profiling of PDZ domains through proteomic peptide-phage display using human and viral

phage peptidomes. *Proceedings of the National Academy of Sciences of the United States of America* 111(7):2542-2547.

Jackson D, Hossain MJ, Hickman D, Perez DR & Lamb RA (2008) A new influenza virus virulence determinant: the NS1 protein four C-terminal residues modulate pathogenicity. *Proceedings of the National Academy of Sciences of the United States of America* 105(11):4381-4386.

Jacomy H, Fragoso G, Almazan G, Mushynski WE & Talbot PJ (2006) Human coronavirus OC43 infection induces chronic encephalitis leading to disabilities in BALB/C mice. *Virology* 349(2):335-346.

Jacomy H, St-Jean JR, Brison E, Marceau G, Desforges M & Talbot PJ (2010) Mutations in the spike glycoprotein of human coronavirus OC43 modulate disease in BALB/c mice from encephalitis to flaccid paralysis and demyelination. *J Neurovirol* 16(4):279-293.

Jacomy H & Talbot PJ (2003) Vacuolating encephalitis in mice infected by human coronavirus OC43. *Virology* 315(1):20-33.

Jacotot E, Ravagnan L, Loeffler M, Ferri KF, Vieira HL, Zamzami N, Costantini P, Druillennec S, Hoebeke J, Briand JP, Irinopoulou T, Daugas E, Susin SA, Cointe D, Xie ZH, Reed JC, Roques BP & Kroemer G (2000) The HIV-1 viral protein R induces apoptosis via a direct effect on the mitochondrial permeability transition pore. *J Exp Med* 191(1):33-46.

James CD & Roberts S (2016) Viral Interactions with PDZ Domain-Containing Proteins-An Oncogenic Trait? *Pathogens* 5(1).

Javier RT & Rice AP (2011) Emerging theme: cellular PDZ proteins as common targets of pathogenic viruses. *Journal of virology* 85(22):11544-11556.

Jimenez-Guardeno JM, Nieto-Torres JL, DeDiego ML, Regla-Nava JA, Fernandez-Delgado R, Castano-Rodriguez C & Enjuanes L (2014) The PDZ-binding motif of severe acute respiratory syndrome coronavirus envelope protein is a determinant of viral pathogenesis. *PLoS pathogens* 10(8):e1004320.

Jimenez-Guardeno JM, Regla-Nava JA, Nieto-Torres JL, DeDiego ML, Castano-Rodriguez C, Fernandez-Delgado R, Perlman S & Enjuanes L (2015) Identification of the Mechanisms Causing Reversion to Virulence in an Attenuated SARS-CoV for the Design of a Genetically Stable Vaccine. *PLoS pathogens* 11(10):e1005215.

Jones CT, Murray CL, Eastman DK, Tassello J & Rice CM (2007) Hepatitis C virus p7 and NS2 proteins are essential for production of infectious virus. *Journal of virology* 81(16):8374-8383.

Keng CT, Akerstrom S, Leung CS, Poon LL, Peiris JS, Mirazimi A & Tan YJ (2011) SARS coronavirus 8b reduces viral replication by down-regulating E via an ubiquitin-independent proteasome pathway. *Microbes Infect* 13(2):179-188.

Kennedy MB (1995) Origin of PDZ (DHR, GLGF) domains. *Trends Biochem Sci* 20(9):350.

Kim E & Sheng M (2004) PDZ domain proteins of synapses. *Nat Rev Neurosci* 5(10):771-781.

Kiyono T, Hiraiwa A, Fujita M, Hayashi Y, Akiyama T & Ishibashi M (1997) Binding of high-risk human papillomavirus E6 oncoproteins to the human homologue of the Drosophila discs large tumor suppressor protein. *Proceedings of the National Academy of Sciences of the United States of America* 94(21):11612-11616.

Klumperman J, Locker JK, Meijer A, Horzinek MC, Geuze HJ & Rottier PJ (1994) Coronavirus M proteins accumulate in the Golgi complex beyond the site of virion budding. *Journal of virology* 68(10):6523-6534.

Koetzner CA, Kuo L, Goebel SJ, Dean AB, Parker MM & Masters PS (2010) Accessory protein 5a is a major antagonist of the antiviral action of interferon against murine coronavirus. *Journal of virology* 84(16):8262-8274.

Kranjec C & Banks L (2011) A systematic analysis of human papillomavirus (HPV) E6 PDZ substrates identifies MAGI-1 as a major target of HPV type 16 (HPV-16) and HPV-18 whose loss accompanies disruption of tight junctions. *Journal of virology* 85(4):1757-1764.

Krijnse-Locker J, Ericsson M, Rottier PJ & Griffiths G (1994) Characterization of the budding compartment of mouse hepatitis virus: evidence that transport from the RER to the Golgi complex requires only one vesicular transport step. *J Cell Biol* 124(1-2):55-70.

Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, Tong S, Urbani C, Comer JA, Lim W, Rollin PE, Dowell SF, Ling AE, Humphrey CD, Shieh WJ, Guarner J, Paddock CD, Rota P, Fields B, DeRisi J, Yang JY, Cox N, Hughes JM, LeDuc JW, Bellini WJ, Anderson LJ & Group SW (2003) A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* 348(20):1953-1966.

Kuo L, Hurst KR & Masters PS (2007) Exceptional flexibility in the sequence requirements for coronavirus small envelope protein function. *Journal of virology* 81(5):2249-2262.

Kuo L & Masters PS (2003) The small envelope protein E is not essential for murine coronavirus replication. *Journal of virology* 77(8):4597-4608.

Kuo L & Masters PS (2010) Evolved variants of the membrane protein can partially replace the envelope protein in murine coronavirus assembly. *Journal of virology* 84(24):12872-12885.

Lama J & Carrasco L (1992) Expression of poliovirus nonstructural proteins in Escherichia coli cells. Modification of membrane permeability induced by 2B and 3A. *The Journal of biological chemistry* 267(22):15932-15937.

Lambert F, Jacomy H, Marceau G & Talbot PJ (2008) Titration of human coronaviruses, HcoV-229E and HCoV-OC43, by an indirect immunoperoxidase assay. *Methods Mol Biol* 454:93-102.

Lau SK, Woo PC, Yip CC, Tse H, Tsoi HW, Cheng VC, Lee P, Tang BS, Cheung CH, Lee RA, So LY, Lau YL, Chan KH & Yuen KY (2006) Coronavirus HKU1 and other coronavirus infections in Hong Kong. *J Clin Microbiol* 44(6):2063-2071.

Lavi E, Wang Q, Weiss SR & Gonatas NK (1996) Syncytia formation induced by coronavirus infection is associated with fragmentation and rearrangement of the Golgi apparatus. *Virology* 221(2):325-334.

Le Coupanec A, Desforges M, Meessen-Pinard M, Dube M, Day R, Seidah NG & Talbot PJ (2015) Cleavage of a Neuroinvasive Human Respiratory Virus Spike Glycoprotein by Proprotein Convertases Modulates Neurovirulence and Virus Spread within the Central Nervous System. *PLoS pathogens* 11(11):e1005261.

Lee SS, Weiss RS & Javier RT (1997) Binding of human virus oncoproteins to hDlg/SAP97, a mammalian homolog of the Drosophila discs large tumor suppressor protein. *Proceedings of the National Academy of Sciences of the United States of America* 94(13):6670-6675.

Lepiller Q, Barth H, Lefebvre F, Herbrecht R, Lutz P, Kessler R, Fafi-Kremer S & Stoll-Keller F (2013) High incidence but low burden of coronaviruses and preferential associations between respiratory viruses. *J Clin Microbiol* 51(9):3039-3046.

Letunic I, Doerks T & Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* 40(Database issue):D302-305.

Li K, Wohlford-Lenane C, Perlman S, Zhao J, Jewell AK, Reznikov LR, Gibson-Corley KN, Meyerholz DK & McCray PB, Jr. (2015) Middle East Respiratory Syndrome Coronavirus Causes Multiple Organ Damage and Lethal Disease in Mice Transgenic for Human Dipeptidyl Peptidase 4. *J Infect Dis* 10.1093/infdis/jiv499.

Li W, Shi Z, Yu M, Ren W, Smith C, Epstein JH, Wang H, Crameri G, Hu Z, Zhang H, Zhang J, McEachern J, Field H, Daszak P, Eaton BT, Zhang S & Wang LF (2005) Bats are natural reservoirs of SARS-like coronaviruses. *Science* 310(5748):676-679.

Li Y, Surya W, Claudine S & Torres J (2014) Structure of a conserved Golgi complex-targeting signal in coronavirus envelope proteins. *The Journal of biological chemistry* 289(18):12535-12549.

Liao Y, Lescar J, Tam JP & Liu DX (2004) Expression of SARS-coronavirus envelope protein in Escherichia coli cells alters membrane permeability. *Biochem Biophys Res Commun* 325(1):374-380.

Liao Y, Yuan Q, Torres J, Tam JP & Liu DX (2006) Biochemical and functional characterization of the membrane association and membrane permeabilizing activity of the severe acute respiratory syndrome coronavirus envelope protein. *Virology* 349(2):264-275.

Lim KP & Liu DX (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 biological chemistry* 276(20):17515-17523.

Lissenberg A, Vrolijk MM, van Vliet AL, Langereis MA, de Groot-Mijnes JD, Rottier PJ & de Groot RJ (2005) Luxury at a cost? Recombinant mouse hepatitis viruses expressing the accessory hemagglutinin esterase protein display reduced fitness in vitro. *Journal of virology* 79(24):15054-15063.

Liu DX, Fung TS, Chong KK, Shukla A & Hilgenfeld R (2014) Accessory proteins of SARS-CoV and other coronaviruses. *Antiviral Res* 109:97-109.

Liu DX & Inglis SC (1991) Association of the infectious bronchitis virus 3c protein with the virion envelope. *Virology* 185(2):911-917.

Liu H, Golebiewski L, Dow EC, Krug RM, Javier RT & Rice AP (2010) The ESEV PDZ-binding motif of the avian influenza A virus NS1 protein protects infected cells from apoptosis by directly targeting Scribble. *Journal of virology* 84(21):11164-11174.

Lopez LA, Jones A, Arndt WD & Hogue BG (2006) Subcellular localization of SARS-CoV structural proteins. *Adv Exp Med Biol* 581:297-300.

Lopez LA, Riffle AJ, Pike SL, Gardner D & Hogue BG (2008) Importance of conserved cysteine residues in the coronavirus envelope protein. *Journal of virology* 82(6):3000-3010.

Lu W, Zheng BJ, Xu K, Schwarz W, Du L, Wong CK, Chen J, Duan S, Deubel V & Sun B (2006) Severe acute respiratory syndrome-associated coronavirus 3a protein forms an ion channel and modulates virus release. *Proceedings of the National Academy of Sciences of the United States of America* 103(33):12540-12545.

Luck K, Charbonnier S & Trave G (2012) The emerging contribution of sequence context to the specificity of protein interactions mediated by PDZ domains. *FEBS letters* 586(17):2648-2661.

Ludford-Menting MJ, Oliaro J, Sacirbegovic F, Cheah ET, Pedersen N, Thomas SJ, Pasam A, Iazzolino R, Dow LE, Waterhouse NJ, Murphy A, Ellis S, Smyth MJ, Kershaw MH, Darcy PK, Humbert PO & Russell SM (2005) A network of PDZ-containing proteins regulates T cell polarity and morphology during migration and immunological synapse formation. *Immunity* 22(6):737-748.

Madan V, Garcia Mde J, Sanz MA & Carrasco L (2005) Viroporin activity of murine hepatitis virus E protein. *FEBS letters* 579(17):3607-3612.

Maeda J, Repass JF, Maeda A & Makino S (2001) Membrane topology of coronavirus E protein. *Virology* 281(2):163-169.

McBride R, van Zyl M & Fielding BC (2014) The coronavirus nucleocapsid is a multifunctional protein. *Viruses* 6(8):2991-3018.

McIntosh K, Becker WB & Chanock RM (1967a) Growth in suckling-mouse brain of "IBV-like" viruses from patients with upper respiratory tract disease. *Proceedings of the National Academy of Sciences of the United States of America* 58(6):2268-2273.

McIntosh K, Dees JH, Becker WB, Kapikian AZ & Chanock RM (1967b) Recovery in tracheal organ cultures of novel viruses from patients with respiratory disease. *Proceedings of the National Academy of Sciences of the United States of America* 57(4):933-940.

Mehnert T, Routh A, Judge PJ, Lam YH, Fischer D, Watts A & Fischer WB (2008) Biophysical characterization of Vpu from HIV-1 suggests a channel-pore dualism. *Proteins* 70(4):1488-1497.

Millet JK & Whittaker GR (2015) Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus research* 202:120-134.

Morfopoulou S, Brown JR, Davies EG, Anderson G, Virasami A, Qasim W, Chong WK, Hubank M, Plagnol V, Desforges M, Jacques TS, Talbot PJ & Breuer J (2016) Coronavirus HCoV-OC43 associated with fatal encephalitis. *N Engl J Med*.

Nal B, Chan C, Kien F, Siu L, Tse J, Chu K, Kam J, Staropoli I, Crescenzo-Chaigne B, Escriou N, van der Werf S, Yuen KY & Altmeyer R (2005) Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. *The Journal of general virology* 86(Pt 5):1423-1434.

Nieto-Torres JL, Dediego ML, Alvarez E, Jimenez-Guardeno JM, Regla-Nava JA, Llorente M, Kremer L, Shuo S & Enjuanes L (2011) Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein. *Virology* 415(2):69-82.

Nieto-Torres JL, DeDiego ML, Verdia-Baguena C, Jimenez-Guardeno JM, Regla-Nava JA, Fernandez-Delgado R, Castano-Rodriguez C, Alcaraz A, Torres J, Aguilella VM & Enjuanes L (2014) Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis. *PLoS pathogens* 10(5):e1004077.

Nieto-Torres JL, Verdia-Baguena C, Jimenez-Guardeno JM, Regla-Nava JA, Castano-Rodriguez C, Fernandez-Delgado R, Torres J, Aguilella VM & Enjuanes L (2015) Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome. *Virology* 485:330-339.

Nieva JL, Madan V & Carrasco L (2012) Viroporins: structure and biological functions. *Nat Rev Microbiol* 10(8):563-574.

Obenauer JC, Denson J, Mehta PK, Su X, Mukatira S, Finkelstein DB, Xu X, Wang J, Ma J, Fan Y, Rakestraw KM, Webster RG, Hoffmann E, Krauss S, Zheng J, Zhang Z & Naeve CW (2006) Large-scale sequence analysis of avian influenza isolates. *Science* 311(5767):1576-1580.

Ortego J, Ceriani JE, Patino C, Plana J & Enjuanes L (2007) Absence of E protein arrests transmissible gastroenteritis coronavirus maturation in the secretory pathway. *Virology* 368(2):296-308.

Ortego J, Escors D, Laude H & Enjuanes L (2002) Generation of a replication-competent, propagation-deficient virus vector based on the transmissible gastroenteritis coronavirus genome. *Journal of virology* 76(22):11518-11529.

OuYang B & Chou JJ (2014) The minimalist architectures of viroporins and their therapeutic implications. *Biochimica et biophysica acta* 1838(4):1058-1067.

Parthasarathy K, Lu H, Surya W, Vararattanavech A, Pervushin K & Torres J (2012) Expression and purification of coronavirus envelope proteins using a modified beta-barrel construct. *Protein Expr Purif* 85(1):133-141.

Parthasarathy K, Ng L, Lin X, Liu DX, Pervushin K, Gong X & Torres J (2008) Structural flexibility of the pentameric SARS coronavirus envelope protein ion channel. *Biophysical journal* 95(6):L39-41.

Pavlovic D, Neville DC, Argaud O, Blumberg B, Dwek RA, Fischer WB & Zitzmann N (2003) The hepatitis C virus p7 protein forms an ion channel that is inhibited by long-alkyl-chain iminosugar derivatives. *Proceedings of the National Academy of Sciences of the United States of America* 100(10):6104-6108.

Pedersen SW, Pedersen SB, Anker L, Hultqvist G, Kristensen AS, Jemth P & Stromgaard K (2014) Probing backbone hydrogen bonding in PDZ/ligand interactions by protein amide-to-ester mutations. *Nat Commun* 5:3215.

Peiris JS, Guan Y & Yuen KY (2004) Severe acute respiratory syndrome. *Nat Med* 10(12 Suppl):S88-97.

Pene F, Merlat A, Vabret A, Rozenberg F, Buzyn A, Dreyfus F, Cariou A, Freymuth F & Lebon P (2003) Coronavirus 229E-related pneumonia in immunocompromised patients. *Clin Infect Dis* 37(7):929-932.

Penkert RR, DiVittorio HM & Prehoda KE (2004) Internal recognition through PDZ domain plasticity in the Par-6-Pals1 complex. *Nat Struct Mol Biol* 11(11):1122-1127.

Perez M, Garcia-Barreno B, Melero JA, Carrasco L & Guinea R (1997) Membrane permeability changes induced in Escherichia coli by the SH protein of human respiratory syncytial virus. *Virology* 235(2):342-351.

Pervushin K, Tan E, Parthasarathy K, Lin X, Jiang FL, Yu D, Vararattanavech A, Soong TW, Liu DX & Torres J (2009) Structure and inhibition of the SARS coronavirus envelope protein ion channel. *PLoS pathogens* 5(7):e1000511.

Pielak RM & Chou JJ (2011) Influenza M2 proton channels. *Biochimica et biophysica acta* 1808(2):522-529.

Pinto LH, Dieckmann GR, Gandhi CS, Papworth CG, Braman J, Shaughnessy MA, Lear JD, Lamb RA & DeGrado WF (1997) A functionally defined model for the M2 proton channel of influenza A virus suggests a mechanism for its ion selectivity. *Proceedings of the National Academy of Sciences of the United States of America* 94(21):11301-11306.

Pinto LH, Holsinger LJ & Lamb RA (1992) Influenza virus M2 protein has ion channel activity. *Cell* 69(3):517-528.

Pinto LH & Lamb RA (2006) The M2 proton channels of influenza A and B viruses. *The Journal of biological chemistry* 281(14):8997-9000.

Prehaud C, Wolff N, Terrien E, Lafage M, Megret F, Babault N, Cordier F, Tan GS, Maitrepierre E, Menager P, Chopy D, Hoos S, England P, Delepierre M, Schnell MJ, Buc H & Lafon M (2010) Attenuation of rabies virulence: takeover by the cytoplasmic domain of its envelope protein. *Sci Signal* 3(105):ra5.

Pyrc K, Berkhout B & van der Hoek L (2007) The novel human coronaviruses NL63 and HKU1. *Journal of virology* 81(7):3051-3057.

Raamsman MJ, Locker JK, de Hooge A, de Vries AA, Griffiths G, Vennema H & Rottier PJ (2000) Characterization of the coronavirus mouse hepatitis virus strain A59 small membrane protein E. *Journal of virology* 74(5):2333-2342.

Regla-Nava JA, Nieto-Torres JL, Jimenez-Guardeno JM, Fernandez-Delgado R, Fett C, Castano-Rodriguez C, Perlman S, Enjuanes L & DeDiego ML (2015) Severe acute respiratory syndrome coronaviruses with mutations in the E protein are

attenuated and promising vaccine candidates. *Journal of virology* 89(7):3870-3887.

Rousset R, Fabre S, Desbois C, Bantignies F & Jalinot P (1998) The C-terminus of the HTLV-1 Tax oncoprotein mediates interaction with the PDZ domain of cellular proteins. *Oncogene* 16(5):643-654.

Ruch TR & Machamer CE (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(2):675-685.

Ruch TR & Machamer CE (2012a) The coronavirus E protein: assembly and beyond. *Viruses* 4(3):363-382.

Ruch TR & Machamer CE (2012b) A single polar residue and distinct membrane topologies impact the function of the infectious bronchitis coronavirus E protein. *PLoS pathogens* 8(5):e1002674.

Ruebner BH, Hirano T & Slusser RJ (1967) Electron microscopy of the hepatocellular and Kupffer-cell lesions of mouse hepatitis, with particular reference to the effect of cortisone. *Am J Pathol* 51(2):163-189.

Sakaguchi T, Leser GP & Lamb RA (1996) The ion channel activity of the influenza virus M2 protein affects transport through the Golgi apparatus. *J Cell Biol* 133(4):733-747.

Sakai A, Claire MS, Faulk K, Govindarajan S, Emerson SU, Purcell RH & Bukh J (2003) The p7 polypeptide of hepatitis C virus is critical for infectivity and contains functionally important genotype-specific sequences. *Proceedings of the National Academy of Sciences of the United States of America* 100(20):11646-11651.

Sanchez-Martinez S, Huarte N, Maeso R, Madan V, Carrasco L & Nieva JL (2008) Functional and structural characterization of 2B viroporin membranolytic domains. *Biochemistry* 47(40):10731-10739.

Sandoval IV & Carrasco L (1997) Poliovirus infection and expression of the poliovirus protein 2B provoke the disassembly of the Golgi complex, the organelle target for the antipoliovirus drug Ro-090179. *Journal of virology* 71(6):4679-4693.

Sanz MA, Perez L & Carrasco L (1994) Semliki Forest virus 6K protein modifies membrane permeability after inducible expression in Escherichia coli cells. *The Journal of biological chemistry* 269(16):12106-12110.

Sawicki SG, Sawicki DL & Siddell SG (2007) A contemporary view of coronavirus transcription. *Journal of virology* 81(1):20-29.

Schubert U, Ferrer-Montiel AV, Oblatt-Montal M, Henklein P, Strebel K & Montal M (1996) Identification of an ion channel activity of the Vpu transmembrane domain and its involvement in the regulation of virus release from HIV-1-infected cells. *FEBS letters* 398(1):12-18.

Scott C & Griffin S (2015) Viroporins: structure, function and potential as antiviral targets. *The Journal of general virology* 96(8):2000-2027.

Siu YL, Teoh KT, Lo J, Chan CM, Kien F, Escriou N, Tsao SW, Nicholls JM, Altmeyer R, Peiris JSM, Bruzzone R & Nal B (2008) 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. *Journal of virology* 82(22):11318-11330.

Smith EC & Denison MR (2012) Implications of altered replication fidelity on the evolution and pathogenesis of coronaviruses. *Curr Opin Virol* 2(5):519-524.

Snijder EJ, van der Meer Y, Zevenhoven-Dobbe J, Onderwater JJ, van der Meulen J, Koerten HK & Mommaas AM (2006) Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome coronavirus replication complex. *Journal of virology* 80(12):5927-5940.

Songyang Z, Fanning AS, Fu C, Xu J, Marfatia SM, Chishti AH, Crompton A, Chan AC, Anderson JM & Cantley LC (1997) Recognition of unique carboxyl-terminal motifs by distinct PDZ domains. *Science* 275(5296):73-77.

Soubies SM, Volmer C, Croville G, Loupias J, Peralta B, Costes P, Lacroux C, Guerin JL & Volmer R (2010) Species-specific contribution of the four C-terminal amino acids of influenza A virus NS1 protein to virulence. *Journal of virology* 84(13):6733-6747.

Spaller MR (2006) Act globally, think locally: systems biology addresses the PDZ domain. *ACS Chem Biol* 1(4):207-210.

St-Jean JR, Desforges M, Almazan F, Jacomy H, Enjuanes L & Talbot PJ (2006) Recovery of a neurovirulent human coronavirus OC43 from an infectious cDNA clone. *Journal of virology* 80(7):3670-3674.

St-Jean JR, Jacomy H, Desforges M, Vabret A, Freymuth F & Talbot PJ (2004) Human respiratory coronavirus OC43: genetic stability and neuroinvasion. *Journal of virology* 78(16):8824-8834.

Stadler K, Masignani V, Eickmann M, Becker S, Abrignani S, Klenk HD & Rappuoli R (2003) SARS--beginning to understand a new virus. *Nat Rev Microbiol* 1(3):209-218.

Steinhauer DA, Domingo E & Holland JJ (1992) Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene* 122(2):281-288.

Steinmann E, Penin F, Kallis S, Patel AH, Bartenschlager R & Pietschmann T (2007) Hepatitis C virus p7 protein is crucial for assembly and release of infectious virions. *PLoS pathogens* 3(7):e103.

Stewart JN, Mounir S & Talbot PJ (1992) Human coronavirus gene expression in the brains of multiple sclerosis patients. *Virology* 191(1):502-505.

Stiffler MA, Chen JR, Grantcharova VP, Lei Y, Fuchs D, Allen JE, Zaslavskaia LA & MacBeath G (2007) PDZ domain binding selectivity is optimized across the mouse proteome. *Science* 317(5836):364-369.

Stricker NL, Christopherson KS, Yi BA, Schatz PJ, Raab RW, Dawes G, Bassett DE, Jr., Bredt DS & Li M (1997) PDZ domain of neuronal nitric oxide synthase recognizes novel C-terminal peptide sequences. *Nat Biotechnol* 15(4):336-342.

Suhy DA, Giddings TH, Jr. & Kirkegaard K (2000) Remodeling the endoplasmic reticulum by poliovirus infection and by individual viral proteins: an autophagy-like origin for virus-induced vesicles. *Journal of virology* 74(19):8953-8965.

Surya W, Li Y, Verdia-Baguena C, Aguilella VM & Torres J (2015) MERS coronavirus envelope protein has a single transmembrane domain that forms pentameric ion channels. *Virus research* 201:61-66.

Svoboda D, Nielson A, Werber A & Higginson J (1962) An electron microscopic study of viral hepatitis in mice. *Am J Pathol* 41:205-224.

Takano T, Nakano K, Doki T & Hohdatsu T (2015) Differential effects of viroporin inhibitors against feline infectious peritonitis virus serotypes I and II. *Arch Virol* 160(5):1163-1170.

Teoh KT, Siu YL, Chan WL, Schluter MA, Liu CJ, Peiris JS, Bruzzone R, Margolis B & Nal B (2010) The SARS coronavirus E protein interacts with PALS1 and alters tight junction formation and epithelial morphogenesis. *Mol Biol Cell* 21(22):3838-3852.

Thomas M, Kranjec C, Nagasaka K, Matlashewski G & Banks L (2011) Analysis of the PDZ binding specificities of Influenza A virus NS1 proteins. *Virology journal* 8:25.

Tonikian R, Zhang Y, Sazinsky SL, Currell B, Yeh JH, Reva B, Held HA, Appleton BA, Evangelista M, Wu Y, Xin X, Chan AC, Seshagiri S, Lasky LA, Sander C, Boone C, Bader GD & Sidhu SS (2008) A specificity map for the PDZ domain family. *PLoS Biol* 6(9):e239.

Tooze J, Tooze S & 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(2):281-293.

Tooze J & Tooze SA (1985) Infection of AtT20 murine pituitary tumour cells by mouse hepatitis virus strain A59: virus budding is restricted to the Golgi region. *Eur J Cell Biol* 37:203-212.

Torres J, Maheswari U, Parthasarathy K, Ng L, Liu DX & Gong X (2007) Conductance and amantadine binding of a pore formed by a lysine-flanked transmembrane domain of SARS coronavirus envelope protein. *Protein Sci* 16(9):2065-2071.

Torres J, Parthasarathy K, Lin X, Saravanan R, Kukol A & Liu DX (2006) Model of a putative pore: the pentameric alpha-helical bundle of SARS coronavirus E protein in lipid bilayers. *Biophysical journal* 91(3):938-947.

Torres J, Surya W, Li Y & Liu DX (2015) Protein-Protein Interactions of Viroporins in Coronaviruses and Paramyxoviruses: New Targets for Antivirals? *Viruses* 7(6):2858-2883.

Torres J, Wang J, Parthasarathy K & Liu DX (2005) The transmembrane oligomers of coronavirus protein E. *Biophysical journal* 88(2):1283-1290.

Toyooka K, Iritani S, Makifuchi T, Shirakawa O, Kitamura N, Maeda K, Nakamura R, Niizato K, Watanabe M, Kakita A, Takahashi H, Someya T & Nawa H (2002) Selective reduction of a PDZ protein, SAP-97, in the prefrontal cortex of patients with chronic schizophrenia. *J Neurochem* 83(4):797-806.

Tseng YT, Wang SM, Huang KJ & Wang CT (2014) SARS-CoV envelope protein palmitoylation or nucleocapid association is not required for promoting virus-like particle production. *J Biomed Sci* 21:34.

Tyrrell DA & Bynoe ML (1965) Cultivation of a Novel Type of Common-Cold Virus in Organ Cultures. *Br Med J* 1(5448):1467-1470.

Ujike M & Taguchi F (2015) Incorporation of spike and membrane glycoproteins into coronavirus virions. *Viruses* 7(4):1700-1725.

Vabret A, Dina J, Brison E, Brouard J & Freymuth F (2009) [Human coronaviruses]. *Pathol Biol (Paris)* 57(2):149-160.

Vabret A, Dina J, Gouarin S, Petitjean J, Corbet S & Freymuth F (2006) Detection of the new human coronavirus HKU1: a report of 6 cases. *Clin Infect Dis* 42(5):634-639.

Vabret A, Dina J, Gouarin S, Petitjean J, Tripey V, Brouard J & Freymuth F (2008) Human (non-severe acute respiratory syndrome) coronavirus infections in hospitalised children in France. *J Paediatr Child Health* 44(4):176-181.

van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJ, Wolthers KC, Wertheim-van Dillen PM, Kaandorp J, Spaargaren J & Berkhout B (2004) Identification of a new human coronavirus. *Nat Med* 10(4):368-373.

van der Hoek L, Sure K, Ihorst G, Stang A, Pyrc K, Jebbink MF, Petersen G, Forster J, Berkhout B & Uberla K (2005) Croup is associated with the novel coronavirus NL63. *PLoS Med* 2(8):e240.

van Kuppeveld FJ, Hoenderop JG, Smeets RL, Willems PH, Dijkman HB, Galama JM & Melchers WJ (1997) Coxsackievirus protein 2B modifies endoplasmic reticulum membrane and plasma membrane permeability and facilitates virus release. *EMBO J* 16(12):3519-3532.

Venkatagopalan P, Daskalova SM, Lopez LA, Dolezal KA & Hogue BG (2015) Coronavirus envelope (E) protein remains at the site of assembly. *Virology* 478:75-85.

Vennema H, Godeke GJ, Rossen JW, Voorhout WF, Horzinek MC, Opstelten DJ & Rottier PJ (1996) Nucleocapsid-independent assembly of coronavirus-like particles by co-expression of viral envelope protein genes. *The EMBO Journal* 15(8):2020-2028.

Verdia-Baguena C, Nieto-Torres JL, Alcaraz A, Dediego ML, Enjuanes L & Aguilella VM (2013) Analysis of SARS-CoV E protein ion channel activity by tuning the protein and lipid charge. *Biochimica et biophysica acta* 1828(9):2026-2031.

Verdia-Baguena C, Nieto-Torres JL, Alcaraz A, DeDiego ML, Torres J, Aguilella VM & Enjuanes L (2012) Coronavirus E protein forms ion channels with functionally and structurally-involved membrane lipids. *Virology* 432(2):485-494.

Westerbeck JW & Machamer CE (2015) A Coronavirus E Protein is Present in Two Distinct Pools with Different Effects on Assembly and the Secretory Pathway. *Journal of virology* 10.1128/JVI.01237-15.

Willott E, Balda MS, Fanning AS, Jameson B, Van Itallie C & Anderson JM (1993) The tight junction protein ZO-1 is homologous to the Drosophila discs-large tumor

suppressor protein of septate junctions. *Proceedings of the National Academy of Sciences of the United States of America* 90(16):7834-7838.

Wilson L, Gage P & Ewart G (2006) Hexamethylene amiloride blocks E protein ion channels and inhibits coronavirus replication. *Virology* 353(2):294-306.

Wilson L, McKinlay C, Gage P & Ewart G (2004) SARS coronavirus E protein forms cation-selective ion channels. *Virology* 330(1):322-331.

Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, Huang Y, Wong BH, Poon RW, Cai JJ, Luk WK, Poon LL, Wong SS, Guan Y, Peiris JS & Yuen KY (2005) Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *Journal of virology* 79(2):884-895.

Woo PC, Lau SK, Huang Y & Yuen KY (2009) Coronavirus diversity, phylogeny and interspecies jumping. *Exp Biol Med (Maywood)* 234(10):1117-1127.

Woods DF & Bryant PJ (1993) ZO-1, DlgA and PSD-95/SAP90: homologous proteins in tight, septate and synaptic cell junctions. *Mech Dev* 44(2-3):85-89.

Wozniak AL, Griffin S, Rowlands D, Harris M, Yi M, Lemon SM & Weinman SA (2010) Intracellular proton conductance of the hepatitis C virus p7 protein and its contribution to infectious virus production. *PLoS pathogens* 6(9):e1001087.

Wu PS, Chang LY, Berkhout B, van der Hoek L, Lu CY, Kao CL, Lee PI, Shao PL, Lee CY, Huang FY & Huang LM (2008) Clinical manifestations of human coronavirus NL63 infection in children in Taiwan. *Eur J Pediatr* 167(1):75-80.

Xu J, Zhong S, Liu J, Li L, Li Y, Wu X, Li Z, Deng P, Zhang J, Zhong N, Ding Y & Jiang Y (2005) Detection of severe acute respiratory syndrome coronavirus in the brain: potential role of the chemokine mig in pathogenesis. *Clin Infect Dis* 41(8):1089-1096.

Ye F & Zhang M (2013) Structures and target recognition modes of PDZ domains: recurring themes and emerging pictures. *Biochem J* 455(1):1-14.

Ye Y & Hogue BG (2007) Role of the coronavirus E viroporin protein transmembrane domain in virus assembly. *Journal of virology* 81(7):3597-3607.

Yeh EA, Collins A, Cohen ME, Duffner PK & Faden H (2004) Detection of coronavirus in the central nervous system of a child with acute disseminated encephalomyelitis. *Pediatrics* 113(1 Pt 1):e73-76.

Yokomori K, Banner LR & Lai MM (1991) Heterogeneity of gene expression of the hemagglutinin-esterase (HE) protein of murine coronaviruses. *Virology* 183(2):647-657.

Yu X, Bi W, Weiss SR & Leibowitz JL (1994) Mouse hepatitis virus gene 5b protein is a new virion envelope protein. *Virology* 202(2):1018-1023.

Yuan Q, Liao Y, Torres J, Tam JP & Liu DX (2006) Biochemical evidence for the presence of mixed membrane topologies of the severe acute respiratory syndrome coronavirus envelope protein expressed in mammalian cells. *FEBS letters* 580(13):3192-3200.

Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD & Fouchier RA (2012) Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med* 367(19):1814-1820.

Zhang R, Wang K, Lv W, Yu W, Xie S, Xu K, Schwarz W, Xiong S & Sun B (2013) The ORF4a protein of human coronavirus 229E functions as a viroporin that regulates viral production. *Biochimica et biophysica acta* 10.1016/j.bbamem.2013.07.025.

Zhang R, Wang K, Ping X, Yu W, Qian Z, Xiong S & Sun B (2015) The ns12.9 accessory protein of human coronavirus OC43 is a viroporin involved in virion morphogenesis and pathogenesis. *Journal of virology* 10.1128/JVI.01986-15.

Zhang Y, Yeh S, Appleton BA, Held HA, Kausalya PJ, Phua DC, Wong WL, Lasky LA, Wiesmann C, Hunziker W & Sidhu SS (2006) Convergent and divergent ligand specificity among PDZ domains of the LAP and zonula occludens (ZO) families. *The Journal of biological chemistry* 281(31):22299-22311.

Zhao L, Jha BK, Wu A, Elliott R, Ziebuhr J, Gorbalenya AE, Silverman RH & Weiss SR (2012) Antagonism of the interferon-induced OAS-RNase L pathway by murine

coronavirus ns2 protein is required for virus replication and liver pathology. *Cell Host Microbe* 11(6):607-616.

Zumla A, Hui DS & Perlman S (2015) Middle East respiratory syndrome. *Lancet* 10.1016/S0140-6736(15)60454-8.

Zuniga S, Cruz JL, Sola I, Mateos-Gomez PA, Palacio L & Enjuanes L (2010) Coronavirus nucleocapsid protein facilitates template switching and is required for efficient transcription. *Journal of virology* 84(4):2169-2175.

Zuniga S, Sola I, Moreno JL, Sabella P, Plana-Duran J & Enjuanes L (2007) Coronavirus nucleocapsid protein is an RNA chaperone. *Virology* 357(2):215-227.

7 ANNEX I – FRENCH TRANSLATION OF TITLES AND LEGENDS OF TABLES

Tableau 3.1 **Énumération de viroporines classés selon leur appartenance au virus ARN ou ADN et taille.** Adapté de (Gonzalez *et al.*, 2003) avec informations supplémentaires de (Lu *et al.*, 2006, Nieva *et al.*, 2012, R. Zhang *et al.*, 2013) et (R. Zhang *et al.*, 2015).

Tableau 3.2 **Motifs canoniques connus pour intéragir avec les domaines PDZ retrouvés dans la region C-terminale des protéines.**

Légende : X = n'importe quel acide aminé; Φ = acide aminé hydrophobe

Tableau 3.3 **Les protéines cellulaires PDZ connues comme étant des cibles des protéines virales, *circa* 2011.**

Légende : Human Ad, adénovirus humain; HBV; virus de l'hépatite B; DV, virus de la dengue; HPV, virus du papillome humain; RhPV; virus du papillome du singe rhésus; CRPV, virus du papillome du Shope; HIV, virus de l'immunodéficience humaine. Prise de (Javier *et al.*, 2011) et informations plus récentes ajoutées de (Jimenez-Guardeno *et al.*, 2014).

Tableau 4.1 **Amorces utlisées pour introduire des substitutions nucléotidiques dans le gène E du pBAC-OC43^{FL} (St-Jean *et al.*, 2006) pour la production des virus recombinants mutés dans des régions spécifiques dans la protéine E (Partie A). Amorces utilisées pour vérifier la séquence du gène E complet et une partie de la protéine M des virus recombinants (Partie B).** Les parties soulignées en gras représentent de nouvelles substitutions nucléotidiques. Les sections soulignées pour le mutant E-PBM 82-84 représentent des nucléotides préalablement introduits dans la séquence.

8 ANNEX II – FRENCH TRANSLATION OF TITLES AND LEGENDS OF FIGURES

Figure 1.1 **Analyse phylogénique des ARN polymérases dépendentes à l'ARN des 10 coronavirus pour lesquels il y a le génome complet disponible avant la découverte du SARS-CoV en 2002 (A) et la même analyse pour ceux découverts et disponibles à la fin de 2008 (avec le génome du MERS-CoV découvert en 2012 dans le regroupement du coronavirus de chauve-souris HKU4/U5 ajouté pour référence).** Les différents genres de coronavirus sont repertoriés avec différentes couleurs et le genre Deltacoronavirus est un genre en attente de classement. L'échelle représente le nombre estimé de substitutions par 10 acides aminés. CoV, coronavirus; SARS, syndrôme respiratoire aïgu sevère; MERS, syndrôme respiratoire sévère du moyen-orient; HCoV, coronavirus humain; BCoV, coronavirus bovin; MHV, virus des hépatites murines; IBV, virus de la bronchite infectieuse; PEDV, virus de la diarrhée épidémique porcine; TGEV; virus de la gastroentérite porcine; FcoV, coronavirus félin; PRCoV, coronavirus porcin respiratoire; bat-CoV, coronavirus de la chauve-souris; PHEV, virus de l'encéphalomyélite hémagglutinante du porc; ECoV, coronavirus équin; TCoV, coronavirus de la dinde; SW1, coronavirus du beluga; BulbCoV, coronavirus du bulbul; ThCoV, coronavirus du muguet; MuCoV, coronavirus du munia. Modifié de (Woo *et al.*, 2009) avec informations supplémentaires de (de Groot *et al.*, 2012) et (Graham *et al.*, 2013).

Figure 2.1 **L'organisation génomique des coronavirus basée sur la lignée A du betacoronavirus HCoV-OC43.** Les gènes de la replicase ORF1a/b sont représentés en gris pâle, les gènes structurales en gris foncée et les protéines accessoires en carrées pointillés. UTR, région non-traduite; ORF, cadre de lecture ouvert; ns, non-structurale; HE, hémagglutinine-estérase; S, spicule; E, enveloppe; M, membrane; N, nucléocapside. Modifié de (St-Jean *et al.*, 2006).

Figure 2.2 **Cycle de replication des coronavirus.** Au lieu de s'assembler à la membrane plasmique, le cycle viral des coronavirus se dinstingue des autres virus enveloppés en s'assemblant et en bourgeonnant dans le lumen de la région intermédiaire du réticulum endoplasmique (ERGIC), suivi par le relâchement par exocytose. Image pris de (Stadler *et al.*, 2003).

Figure 2.3 **Morphologie virale et détails sur la structure des particules virales du coronavirus.** (A) Image prise au microscope électronique de la morphologie de HCoV-OC43 ce qui est typique de tous les coronavirus (image non-publié, Alain LeCoupanec, Laboratoire du P.J. Talbot). (B) Représentation schématique des virions du genre betacoronavirus. S, protéine du spicule; HE, protéine hémaglutinine-estérase; M, protéine membranaire; N, protéine de la nucléocapside; E, protéine d'enveloppe. Image prise de (de Groot *et al.*, 2012).

Figure 3.1 **Alignement des séquences protéiques de la protéine E des coronavirus représentatifs.** Les acides aminés cystéines qui pourraient être palmitoylés sont encadrés en bleu. Alignement séquentiel produit avec le programme ClustalW2. Modifié avec informations présentes en (Ruch *et al.*, 2012a) et (Ruch *et al.*, 2012b).

Figure 3.2 **Structure du canal ionique produit par l'homo-oligomérisation pentamérique de la protéine E du SARS-CoV indentifiée par une étude RMN.** (A) Structure générale du pore homo-oligomérique formée de monomères de la protéine E (en cylindres rouge) avec les phospholipides adjacents (en bleu) (DeDiego *et al.*, 2014). (B) Une représentation en deux dimensions du canal ionique formé par la protéine transmembranaire α-hélicale E du SARS-CoV (en cylindres rouge) et phospholipides adjacents (en bleu). Les acides aminés importants pour la fonction du canal ionique du SARS-CoV sont indiqués. N15 et V25 correspondent au positions Q17 et L27 chez le HCoV-OC43, respectivement (Nieto-Torres *et al.*, 2014). Étude initiale RMN a été effectué par Pervushin et collègues (Pervushin *et al.*, 2009).

Figure 3.3 **Alignement des séquences protéiques de la protéine E de plusieurs coronavirus avec les différents acides aminés impliqués dans la fonction des canaux ioniques encadrés en rouge.** Alignement séquentiel produit avec le programme ClustalW2. Modifié avec informations présentes en (Ruch *et al.*, 2012a) et (Ruch *et al.*, 2012b).

Figure 3.4 **Interaction représentative entre une protéine ayant un domaine PDZ et une protéine cible avec un motif d'attachement avec un domain PDZ (PBM) dans sa région C-terminale.** Représentation en rubans avec les hélices-α visualisées en spirales et feuillets-β représentés avec des flèches (A) et un diagramme topologique simplifié en deux dimensions avec les hélices-α représentées par des tubes et les feuillets-β par des flèches. Les protéines avec un motif PDZ sont en bleu et la région C-terminale des ligands (PBM) en vert. Panneau A pris du PSI Structural Biology Knowledgebase (Gabanyi *et al.*, 2011) et le panneau B est inspiré de (F. Ye *et al.*, 2013).

Figure 3.5 **Représentation schématique de plusieurs protéines contenant un motif PDZ.** Les domaines PDZ sont souvent retrouvés en tandem ou avec d'autres motifs et/ou domaines qui interagissent avec d'autres protéines. Les domaines PDZ sont indiqués en élipses violet. Protéines : PDS-95, post-synaptic density 95; Dlg, discs large; ZO-1, zona occludens protein 1 ; PALS1, membrane-associated palmitoylated protein 5; Syntenin, syndecan binding protein; GRIP1, glutamate-receptor-interacting protein; nNOS, neuronal nitric oxide synthase; CASK/LIN2, Calcium/Calmodulin-dependent serine protein kinase; Par, partitioning defective homologue. Domaines: SH3, Src homology 3 domain; GK, guanylate kinase domain; BP1, Phox and Bem1p domain; CRIB, Cdc42/Rac-interactive binding domain; NO, nitric oxide; FABD, FAD-binding domain; NADB, NAD-binding domain ; Zu5, domain present in ZO-1; CaM kinase, calmodulin-dependent kinase. Adapté de (Kim *et al.*, 2004) avec informations à propos la structure du Par-6 et Scribble venant du (F. Ye *et al.*, 2013) et (Penkert *et al.*, 2004), respectivement.

Figure 3.6 **Identification bioinformatique du motif PBM potentiel dans la région C-terminale de plusieurs coronavirus.** L'encadrement rouge indique l'endroit des PBM potentiels. Modifié de (Jimenez-Guardeno *et al.*, 2014).

Figure 4.1 **La protéine E est critique pour la production des particules infectieuses.** (A) Représentation schématique du genome complet du HCoV-OC43 retrouvé dans le clone infectieux pBAC-OC43^{FL} (en haut) avec une représentation du gène E d'HCoV-OC43 où les positions des acides aminés mutés sont indiqués (en bas). TM, région transmembranaire; PBM, motif d'interaction PDZ. (B) Evaluation de la production des virions infectieux suite à la transfection de cellules avec pBAC-E-Stop comparé à une transfection avec pBAC-OC43^{FL}. (C) Production de la protéine E d'HCoV-OC43 suite à son insertion dans le vecteur d'expression pcDNA3.1(+), nommé pcDNA(OC-E). pcDNA(empty) représente le plasmide vide, sans la séquence de la protéine E. (D) La co-transfection transitoire sur les cellules BHK-21 de pBAC-E-Stop avec 1 ou 2µg pcDNA(OC-E) peut ramener la production de virus infecteux à un niveau détectable de façon dose-dépendante. (E) La détection semi-quantative de la protéine S d'HCoV-OC43 (vert) et du noyeau avec DAPI (bleu) démontrent l'efficacité de transfection de plusieurs clones infectieux sur les cellules BHK-21. Sous-panneaux: (a) mock, (b) pBAC-OC43^{FL} (rOC/ATCC), (c) pBAC-E-Stop (rOC/E-Stop + 2µg pcDNA(OC-E)), (d) pBAC-E-Stop (rOC/E-Stop + 2µg pcDNA(vide)). LOD, limite de détection.

Figure 4.2 **Quand la protéine E est supprimée, la production de particules infectieuses d'HCoV-OC43 peut être rétablie avec la transfection transitoire mais avec une forte sélection pour les révertants.** (A) Evaluation de la production de particules virales infectieuses après co-transfection transitoire de pBAC-E-Stop et de 1µg de pcDNA(OC-E) sur les cellules BHK-21 (BHK 0), ainsi que des amplifications subséquentes sur les cellules HRT-18 épithéliales (HRT 1-3). Les différents titres viraux obtenues de différents essais ont dévoilés par séquençage qu'il y a des réversions de séquence dans les stocks viraux à la position du codon stop introduite. LOD, limite de détection. ┼ (croix) indique la détection par séquençage de réversions dans le gène E d'HCoV-OC43 dans les stocks viraux.

Figure 4.3 **HCoV-OC43 sans E est capable d'infecter les cellules neuronales mais la réplication et la propagation sont fortement compromises.** Les cellules LA-N-5 (A) et les cultures mixtes du SNC murin (B) ont été infectés avec du virus manquant la protéine E, rOC/E-Stop + pcDNA(vide), ou le virus initialement complémenté, rOC/E-Stop + 2µg pcDNA(OC-E), et les titres des particules infectieuses associées et non-associées à la cellule ont été récoltés pendant 72 heures. (C) À 48 heures post-infection, l'immunofluorescence sur les cultures mixtes du SNC murin ont démontré une propagation restreinte du virus rOC/E-Stop initialement complémenté comparé au virus sauvage. Les sous panneaux représentent 8 (a, e), 16 (b,f), 24 (c, g), et 48 (d, h) heures post-infection. La protéine associée aux microtubules 2 (MAP2) des neurones est en vert et l'antigène viral est en rouge.

Figure 4.4 **La potentielle région transmembranaire d'HCoV-OC43 joue un rôle important dans la production de virions infectieux et la réplication dans les cellules neuronales.** (A) Production de particules virales infectieuses après la transfection des mutants transmembranaires de la protéine E sur les cellules BHK-21 (BHK 0) et amplification sur les cellules HRT-18 (HRT 1). (B) Le gène E d'HCoV-OC43 a été détecté par RT-PCR après une transfection sur les cellules BHK-21. (C) Les cellules neuronales humaines LA-N-5 et (D) les cultures mixtes du SNC murin ont été infectées avec rOC/E- TM$_{Q17A}$ et les fractions de virus associé et non-associé aux cellules ont été récoltées et titrées pendant 72 heures. LOD, limite de detection.

Figure 4.5 **Un motif d'interaction des domaines PDZ dans la région C-terminale de la protéine E d'HCoV-OC43 est critique pour la production virale et propagation efficace dans les cellules neuronales**. (A) Production de virus infectieux après la transfection de plusieurs mutants abrogeant partiellement ou complètement le motif PBM de la protéine E sur des cellules BHK-21 (BHK 0) et l'amplification subséquente sur des cellules HRT-18 (HRT 1). Les résultats montrés viennent d'une manipulation représentative. Les titres viraux associés et non-associé à la cellule ont été déterminé pour les temps indiqués sur les graphiques et les infections faites à une multiplicité d'infection de 0.05 sur les cellules LA-N-5 (B) et les cultures mixtes du SNC murin (C). (D) Les tendances de propagation virale sur les cellules LA-N-5 ont été évaluées par immunofluoresecence à 18 et 48 heures post-infection. Le virus a été détecté avec un anticorps contre la protéine S d'HCoV-OC43 (rouge) et le noyau avec le DAPI (bleu). Sous panneaux: (a, f) mock, (b, g) rOC/ATCC, (c, h) rOC/E-PBM$_{D82A}$, (d, i) rOC/E-PBM$_{V84A}$, (e, j) rOC/E-PBM$_{D82A-V84A}$. LOD, limite de détection.

Figure 4.6 **La délétion de la protéine E dans HCov-OC43, ou modification des positions clés de ses domaines spécifiques, affectent la propagation virale dans les cultures mixtes du SNC murin.** Immunofluorescence sur les cultures mixtes du SNC murin pendant 48 heures. Vert représente la protéine associé au microtubules 2 (MAP2) des neurones; rouge représente les antigènes viraux. La première, deuxième et troisième colonne représentent respectivement 16, 24 et 48 heures post-infection. Sous-panneaux: (a-c) rOC/ATCC, (d-f) rOC/E-Stop initialement co-transfecté avec 2µg pcDNA(vide), (g-i) rOC/E-Stop initialement complémenté avec 2µg pcDNA(OC-E), (j-l) rOC/E-TM$_{Q17A}$, (m-o) rOC4/E-PBM$_{D82A}$, (p-r) rOC4/E-PBM$_{V84A}$, (s-u) rOC4/E-PBM$_{D82A-V84A}$.