



HAL
open science

MERS-CoV and SARS-CoV-2 membrane proteins are modified with polylectosamine chains

Dylan Juckel, Lowiese Desmarets, Adeline Danneels, Yves Rouillé, Jean Dubuisson, Sandrine Belouzard

► **To cite this version:**

Dylan Juckel, Lowiese Desmarets, Adeline Danneels, Yves Rouillé, Jean Dubuisson, et al.. MERS-CoV and SARS-CoV-2 membrane proteins are modified with polylectosamine chains. *Journal of General Virology*, 2023, 104 (10), pp.001900. 10.1099/jgv.0.001900 . hal-04234661

HAL Id: hal-04234661

<https://hal.science/hal-04234661>

Submitted on 10 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MERS-CoV and SARS-CoV-2 membrane proteins are modified with polylectosamine chains

Dylan Juckel, Lowiese Desmaretst, Adeline Danneelst, Yves Rouillé, Jean Dubuisson and Sandrine Belouzard*

Abstract

Coronaviruses are positive-stranded RNA enveloped viruses. The helical nucleocapsid is surrounded by a lipid bilayer in which are anchored three viral proteins: the spike (S), membrane (M) and envelope (E) proteins. The M protein is the major component of the viral envelope and is believed to be its building block. The M protein of Middle East respiratory syndrome coronavirus (MERS-CoV) and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) contains a short N-terminal domain with an N-glycosylation site. We investigated their N-glycosylation and show that polylectosamine chains are conjugated to SARS-CoV-2 and MERS-CoV M proteins in transfected and infected cells. Acidic residues present in the first transmembrane segments of the proteins are required for their glycosylation. No specific signal to specify polylectosamine conjugation could be identified and high mannose-conjugated protein was incorporated into virus-like particles.

INTRODUCTION

Coronaviruses are positive-stranded RNA enveloped viruses. The helical ribonucleocapsid is surrounded by a lipid bilayer in which are anchored three viral proteins: the spike (S), the membrane (M) and the envelope (E) proteins. The M protein is the major component of the viral envelope and is believed to be its building block. Indeed, the M protein has been shown to be able to undertake molecular interactions with all the other structural proteins of the viral envelope and has an essential role in virus-like particle (VLP) formation [1]. The M protein ranges from 219 to 262 residues and is composed of a small N-terminal glycosylated domain, three transmembrane segments and a long C-terminal domain. The N-terminal domain of the protein is exposed at the surface of the virion and the C-terminal domain inside the virion (N-exo/C-endo topology). However, a dual topology was demonstrated for the M protein of transmissible gastroenteritis virus (TGEV) with roughly one-third of the protein having its C-terminal domain exposed at the virion surface (N-exo/C-exo) suggesting the presence of a fourth transmembrane segment [2]. TGEV and feline coronavirus (FCoV) M proteins have a signal peptide sequence. The structure of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) M protein was recently solved by cryo-electron microscopy [3, 4]. The M protein forms homodimers and is structurally related to the ORF3a protein suggesting a shared ancestral origin. Depending on the structure, the N-terminal domain is composed of nine or 16 residues and is followed by three transmembrane helices (amino acids 10/17–36, 43–71 and 79–105) spaced by short linkers. The C-terminal domain is highly structured and rich in beta-strands.

The short N-terminal domain of M proteins can be either N- or O-glycosylated depending on the virus. Mouse hepatitis virus (MHV) and other betacoronaviruses such as HCoV-OC43 M proteins are O-glycosylated whereas M proteins of Middle East respiratory syndrome coronavirus (MERS-CoV), SARS-CoV or SARS-CoV-2 are N-glycosylated. It is well established that M protein glycosylation is dispensable for virus assembly but a role in viral pathogenesis has been suggested. Notably, the glycosylation of M proteins of MHV and TGEV has been linked to its interferogenic capacity and, for MHV-A59, abolition of M

Received 13 July 2023; Accepted 22 September 2023; Published 06 October 2023

Author affiliations: ¹Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 9017 - CIL - Center for Infection and Immunity of Lille, F-59000 Lille, France.

***Correspondence:** Sandrine Belouzard, sandrine.belouzard@ibl.cnrs.fr

Keywords: N-linked glycosylation; polylectosamines; coronavirus; membrane protein.

Abbreviations: DSL, *Datura stramonium* lectin; E, envelope; ER, endoplasmic reticulum; FCoV, feline coronavirus; HCV, hepatitis C virus; IBV, infectious bronchitis virus; LacNac, N-acetylglucosamine; LEL, *Lycopersicon esculentum* lectin; M, membrane; MERS-CoV, Middle East respiratory syndrome coronavirus; MHV, mouse hepatitis virus; OST, oligosaccharyltransferase; S, spike; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; TGEV, transmissible gastroenteritis virus; TGN, *Trans*-Golgi network; VLP, virus-like particle.

†These authors contributed equally to this work

001900 © 2023 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

glycosylation decreased the replication of the virus in the liver whereas replication in the brain was not affected [5, 6]. Furthermore, the glycosylation of infectious bronchitis virus (IBV) M protein induced endoplasmic reticulum (ER) stress and apoptosis [7].

In this study, we investigated the N-glycosylation of M proteins of the highly pathogenic viruses MERS-CoV and SARS-CoV-2. Our results demonstrated that M proteins are modified by poly-lactosamine chain(s) in transfected and infected cells. Moreover, we found that not only the N-linked glycosylation site but also acidic residues in the first transmembrane segment are important for the glycosylation of the protein.

METHODS

Cell culture

Huh-7 cells were grown at 37°C with 5% CO₂ in Dulbecco's modified Eagle medium (DMEM; Life Technologies) supplemented with 10% FBS and 2 mM glutamax (Life Technologies).

Plasmids

cDNA was obtained from total RNA extracted from SARS-CoV-2 (strain LillehCoV-19/France/HDF-IPL/2020, GenBank MW575140) infected cells. Then the coding sequence corresponding to the M protein gene was amplified with the forward primer: 5'-CGgatccgagattccaacggtactatta-3' and the following reverse primer: 5'-CAgaattctactgtacaagcaagcaata-3'. The PCR product was then inserted between *Bam*HI and *Eco*RI restriction sites of the pcDNA3.1 vector containing a coding sequence for the V5 tag. Mutations were introduced by PCR-directed mutagenesis. For pcDNA3.1-V5-N5Q-M, the coding sequence of the M protein was amplified by using the following primers: Fw-1 5'-CGgatccgagattccCAAggtactattaccgtt-3' and Rv-1 5'-CAgaattctactgtacaagcaagcaata-3'.

For mutation of three residues of the N-terminus into alanine, the coding sequence were amplified in two fragments to insert the mutations. Then, for SARS-CoV-2 M mutants, the two fragments were fused by using the same Fw-2 (5'-ATTGACGTCAATGGGAGTTT-3') and Rv-2 (5'-AGCGGCAGCagtagcttggatctcgcatccCG-3') primers. The following primers were used for the first round of PCR for SARS-CoV-2: ₈ITV₁₀-M: Fw-2 with Rv (5'-AGCGGCAGCagtagcttggatctcgcatccCG-3') and Rv-2 with (5'-ctGCTGCCGCTgaagagcttaaaaagctcc-3'), for ₁₁EEL₁₃-M: Fw-2 with Rv (5'-AGCCGCTGCaaggttaagtagcttgg-3') and Rv-2 with Fw (5'-tGCAGCGCTTaaaagctcttgaacaatgg-3'), for ₁₄KKL₁₆-M: Fw-2 with Rv (5'-AGCCGCTGCaagctctcaacgtaatagtagc-3') and Rv-2 with Fw (5'-tGCAGCGCTTcttgaacaatggaaactagtagc-3'), for ₁₇LEQ₁₉-M: Fw-2 with Rv (5'-TGCTGCAGCgagcttttaagctctcaac-3') and Rv-2 with Fw (5'-cGCTGCAGCAtggaactagtagtattcc-3') and for E11,12,18A-M: Fw-2 with Rv (5'-attgCGCaaggagcttttaagCGCTGCaac-3') and Rv-2 with Fw (5'-aaaagctcttGCGcaatgg-3'). Similarly, two mutated MERS-CoV M fragments were generated by PCR using pcDNA3.1-V5-MERS-M as a template. The following primers were used for ₆QLT₈-M: Fw2 with Rv (5'-tcAGCGGCTGcgcgtcatattggatccCGTAGAATC-3') and Rv-4 (5'-aaaagctcttGCGcaatgg-3') with Fw (5'-CAGCCGCTgagcgcgagattattgcatattaaa-3'), for ₉EAQ₁₁-M: Fw2 with Rv (5'-atCGCcgCGCagtgagtgctcatattggatcc-3') and Rv-4 with Fw (5'-GCGcgcGCGattattgcatattaaaagctgga-3'), for ₁₂IIA₁₄-M: Fw2 with Rv (5'-aatggcTGCAGCctgagcctcagtgagtgctcatatt-3') and Rv-4 with Fw (5'-gcagGCTGCAGcattattaaaagctggaactttgcatgg-3'), for ₁₅IIKD₁₈-M: Fw2 with Rv (5'-GCTGCAGCAGCggcaataatctgagcctcagtg-3') and Rv-4 with Fw (5'-CAgaattcctgaagctgcaagcaatgca-3'), for E9A-M: Fw2 with Rv (5'-tctgcgCGCagtgagtgctcatattgg-3') and Rv-4 with Fw (5'-tGCGcgcgagattattgcatattaaa-3'), for E9D-M Fw2 with Rv (5'-cgcATCagtgagtgctcata-3') and Rv-4 with Fw (5'-cactGATgcgagattattgcatattaaa-3'), for E9D18A-M Fw2 with Rv (5'-caaagttc-caCGTtaataatgcaataatctgagcCGCagtg-3') and Rv-4 with Fw (5'-cactGCGcgcgagattattgcatattaaaGCGtgaactttg-3'). Then the two PCR products were fused by using the primers Fw-3 (5'-CGgatccaatgatgacgcaactcact-3') and Rv-4.

To generate the M-linker construct (pcDNA3.1-MERS-M-linker), two PCR products were generated by using the primers Fw-3 with Rv (5'-CTTAATAGTTGTATTTAATTTATAGTTTCCAATACGGTAGCGGTTGTATGCGTAtagatcaattggataaacg-3') and Rv-4 with Fw (5'-TAAATTAATAACAATTAAGGACCTGGACTACAAGGACGACGATGACAAGCCAAAAGAGgctcc-cagataatctc-3'). Then the two PCR products were assembled by a second PCR with Fw-3 and Rv-4 primers. Then the pcDNA3.1-MERS-M-linker was used as a template to introduce the N3Q and N88Q mutations. For the N3Q mutations, only one round of PCR was used to insert the mutation. The M-linker coding sequence was amplified with Fw-4 (5'-CGgatccCAAatgacgcaactcactga-3') and Rv-4. For the N88Q mutations, the two first PCRs were performed with Fw-3 and Rv (5'-CCTTAATAGTTGTTTGTAAATTTATAGTTTCC-3') and with Fw (5'-GGAAACTATAAATTACAAACAATTAAGG-3') and Rv-4. The two PCR products were assembled by using Fw-3 and Rv-4 for the single mutation or with Fw-4 and Rv-4 for the double mutation (N3Q,N88Q). The E9,D18-A mutation were introduced as done for MERS-M without linker and by using pcDNA3.1-V5-M-linker or pcDNA3.1-V5-N3Q-M-linker as a template for the first round of PCR.

For the E1-M chimaera, an E1 coding sequence was added into the forward primer used for PCR. The SARS-CoV-2 M protein coding sequence was amplified with 5'-ACgatccGCCACcatgGCTCAGGTGAAGAATAGCAGTGGCCTCTACgttgaagagct-taaaagct-3' and Rv-1.

The ORF3a coding sequence was amplified from infected cell cDNA by using 5'-actctgagccagtgtctcaag-3' and 5'-gaacgaatgag-tacataagt-3' primers. Then a second PCR was performed to introduce the restriction site with the following primers: TCggatcc GCCACCATGGATTTGTTTATGAGAA-3' and 5'-ccgaattcCAAAGGCACGCTAGTAGTCG-3'.

To produce the Nt-M/ORF3a chimaera, the M sequence was amplified with Fw-2 and Rv (5'-GCCAACAATAAGCCA TCCggaatagtagctgtggaatc-3') and ORF3a with Fw (5'-gattccaacgtactattaccGGATGGCTTATGTTGGC-3') and Rv-5 (5'-ccgaattcCAAAGGCACGCTAGTAGTCG-3'). For the TM1-M/ORF3a chimaera, the M protein sequence was amplified with Fw-2 and Rv (5'-CTCTTTTGTAGGGTTATGATAaattgtagaagacaaatcc-3') and ORF3a with Fw (5'-ggattgtctctacaatt ATCATAACCCTCAAAAAGAG-3') and Rv-5. For both chimaeras, the PCR products were assembled by amplification with Fw-2 and Rv-5.

Finally, for all of the constructs, the final PCR products were inserted between the *Bam*HI and *Eco*RI restriction sites of the vector and the constructions were verified by DNA sequencing.

To generate the knockout (KO) cell lines, targeted DNA sequences were inserted into the pX330 vector as previously described [8]. Briefly, 10 μ M of oligonucleotides were annealed and phosphorylated with T4 polynucleotide kinase (NEB). Then the oligonucleotides and the vector were digested with *Bbs*I and ligated in a single step. Residual linearized DNA was removed by exonuclease treatment and the ligation product transformed in Top10 *Escherichia coli* competent cells. The following sequences were used for β 3GNT2: 5'-CACCGACTACTGCGAACCTGACCTG-3' and 5'-AAACCAGGTCAGGTTTCGCAGTAGTC-3', for STT3A: 5'-CACCGTACTCACGGATCAAACCTCA-3' and 5'-AAACTGAGTTTGATCCGTGAGTAC-3' and for STT3B: 5'-CACCGTACAGCAAAGAGTCTACAT-3' and 5'-AAACATGTAGACTCTTTTGCTGTA-3'.

KO cell line

Huh-7 cells were seeded in six-well plates and co-transfected with 1 μ g of pX330 vector coding for small guide RNA targeting β 3GNT2, STT3A, STT3B or pX330 empty vector and 50 ng of plasmid containing a puromycin resistance gene. Transfection was performed with TransIT-LT1 transfection reagent (Mirus Bio) according to the manufacturer's instruction. Twenty-four hours after transfection, 5 μ g ml⁻¹ of puromycin was added to the cell culture medium. Selection was maintained for 4 days. For STT3A and STT3B, cell populations were used for expression of M proteins. STT3A and STT3B expression was analysed in an immunoblot with polyclonal anti-STT3A (Fisher Scientific) and STT3B (Sigma)

For β 3GNT2-KO cells, clones were isolated by serial dilutions of the cell population obtained after puromycin selection. After sequencing, a β 3GNT2-KO clone was selected showing a truncated protein.

Glycosidase treatment

Huh-7 cells plated into six well-plates were transfected with 1 μ g of pcDNA3.1-V5-MERS-M, pcDNA3.1-V5-N3Q-MERS-M, pcDNA3.1-V5-SARS-2-M or pcDNA3.1-V5-N5Q-SARS-2-M by using TransIT-LT1 transfection reagent (Mirus Bio) according to the manufacturer's instruction. Twenty-four hours later, cells were lysed on ice with B1 buffer (50 mM Tris/HCl pH 7.5, 100 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS) containing protease inhibitors. Cell lysates were cleared by centrifugation at 14000 r.p.m. for 5 min. Then, lysates were mock-treated or treated with Endoglycosidase H or PNGase F according to the manufacturer's recommendations. Proteins were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Membrane proteins were detected by using an anti-V5 antibody (Invitrogen; R96025) and HRP-conjugated anti-mouse secondary antibody (Jackson ImmunoResearch). Detection was carried out by chemiluminescence (Pierce).

For quantification of the unglycosylated form of the protein, cell lysates were either mock-treated or PNGase F-treated and then samples were analysed in a Western blot. The band corresponding to the unglycosylated form was quantified in both samples using the band quantification function of Image J. The PNGase F-treated samples allow for the quantification of the total amount of M protein in the sample.

Lectin pulldown

Huh-7 cells plated into six well-plates were transfected with 1 μ g of pcDNA3.1-V5-MERS-M, pcDNA3.1-V5-N3Q-MERS-M, pcDNA3.1-V5-SARS-2-M or pcDNA3.1-V5-N5Q-SARS-2-M by using TransIT-LT1 transfection reagent (Mirus Bio) according to the manufacturer's instruction. Twenty-four hours later, the cells were lysed on ice with B1 buffer containing protease inhibitors. Cell lysates were cleared by centrifugation at 14000 r.p.m. for 5 min. An aliquot of lysates was kept for Western blot analysis, while the rest was incubated with *Datura stramonium* lectin (DSL)- or *Lycopersicon esculentum* lectin (LEL)-conjugated agarose beads (Eurobio) for 2 h at 4 °C on a wheel. Then, beads were washed three times with B1 buffer and resuspended in 2 \times Laemmli loading buffer. Precipitated proteins were detected by immunoblot. Proteins were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Membrane proteins were detected by using an anti-V5 antibody (Invitrogen; R96025) and HRP-conjugated anti-mouse secondary antibody (Jackson ImmunoResearch). Detection was carried out by chemiluminescence (Pierce).

Infection

β 3GNT2-KO and control cells were infected at an m.o.i. of 0.1 with MERS-CoV or SARS-CoV-2. For SARS-CoV-2 infection, cells were previously transduced with a lentivirus expressing ACE2 48 h before infection with SARS-CoV-2. Sixteen hours after infection, cells were lysed in B1 buffer and centrifuged at 14 000 r.p.m. Cell lysates were then heated at 95 °C for 20 min before removing from the Biosafety Level (BSL)3 facility. Proteins were separated on SDS-PAGE and transferred on a nitrocellulose membrane. Membrane proteins were detected with anti-M antibodies (proteogenix for MERS-CoV M and Abnova PAB31758 for SARS-CoV-2 M) and HRP-conjugated anti-rabbit secondary antibodies (Jackson ImmunoResearch). Detection was carried out by chemiluminescence (Pierce).

Immunofluorescence and selective permeabilization assay

Huh-7 cells plated on coverslips were transfected with vector expressing the different membrane protein mutants. The next day, cells were fixed with 3% paraformaldehyde, then permeabilized with 0.1% Triton X-100 or by a freeze–thaw cycle for selective permeabilization assay. The cells were frozen at –80 °C in PBS then allowed to thaw at room temperature. Then, cells were blocked in 5% horse serum in PBS for double stainings or 5% goat serum for simple labelling. Next, cells were incubated with anti-V5 mAbs (Invitrogen; mAb 2F11F7) or anti-E1 (A6 mAb) and a sheep anti-TGN46 pAbs (Bio-Rad) or a rabbit anti-calnexin (Abcam) diluted in blocking solution. Cells were rinsed three times with PBS and incubated with Cy3-conjugated donkey anti-mouse (Jackson immunoresearch) and A488-conjugated donkey anti-sheep secondaries antibodies for V5 and TGN-46 co-staining and Cy3-conjugated donkey anti-mouse and A488-conjugated donkey anti-rabbit secondaries antibodies (Jackson immunoresearch) for V5 and calnexin co-staining. Anti-E1 was detected with A488-conjugated donkey anti-human (Jackson immunoresearch) in combination with Cy3-conjugated donkey anti-rabbit (Jackson immunoresearch) or Cy3-conjugated donkey anti-sheep for calnexin or TGN46 co-staining. Nuclei were stained with 1 μ g ml⁻¹ 4,6-diamidino-2-phenylindole.

Images were acquired using a laser-scanning confocal microscope (LSM 880; Zeiss) using a 63 \times oil immersion objective.

RESULTS

SARS-CoV-2 and MERS-CoV M proteins are N-glycosylated

In a previous study, we noticed that the MERS-CoV M protein migrated as three bands in an immunoblot, corresponding to three M species differing in their glycosylation state [9]. When analysing the migration profile of the SARS-CoV-2 M protein in an immunoblot after transient expression in Huh-7 cells, a similar migration pattern was observed (Fig. 1a). The first band around 25 kDa corresponded to the non-glycosylated protein as shown by the migration profile of the M proteins with a mutated N-glycosylation site (N5Q-M for SARS-CoV-2 and N3Q-M for MERS-CoV) or the M proteins treated with PNGase F. A second band between 25 and 35 kDa was sensitive to Endo-H treatment and hence corresponded to a high-mannose form of the glycans linked to the M protein (obtained in ER/*cis*-Golgi) that has not reached the medial Golgi for further modification. The third species of M protein migrated as a diffuse band (35–40 kDa) that was insensitive to Endo-H but sensitive to PNGase F, confirming that this band corresponded to M proteins modified with complex N-glycans. For SARS-CoV-2, in the sample treated with PNGase or in cells expressing the N5Q-M mutant, a band around 50 kDa was observed that probably represents M-dimers.

Then, we transfected the cells to express the SARS-CoV-2 and MERS-CoV M proteins and treated them with cycloheximide to block protein synthesis for different periods of time. Starting at 2 h of cycloheximide treatment, the band corresponding to high-mannose-conjugated M disappeared showing that the high-mannose-conjugated M proteins are then processed into complex glycan-conjugated M (Fig. 1b). The band corresponding to the unglycosylated protein remains stable over time suggesting that some proportion of the protein does not undergo glycosylation and therefore is not a precursor of the glycosylated M.

SARS-CoV-2 and MERS-CoV M proteins are modified with polylectosamines

We further investigated the nature of these complex N-glycans. It has been suggested that the IBV M protein is modified by polylectosamine chains [10]. Polylectosamine chains are constituted by repeats of *N*-acetylglucosamine (LacNAc). Each LacNAc is composed of galactose linked to *N*-acetylglucosamine through a β 1–4 bond (Gal β 1–4GlcNAc). In polylectosamine chains, LacNAc repeats are linked by β 1–3 bonds. Polylectosamine chains can occur on multi-antennary glycans, and the chains can be linear or may also become branched by the addition of β 1–6GlcNAc to internal galactose [11].

To test if SARS-CoV-2 and MERS-CoV M proteins could harbour such modifications, we performed a protein pulldown with beads conjugated with lectins known for their affinity to polylectosamine motifs (Fig. 1c, d). We used two different lectins, LEL and DSL, to precipitate proteins from cells transfected with vectors expressing either the wild-type or the non-glycosylated M proteins of MERS-CoV and SARS-CoV-2. As shown in Fig. 1(c), only LEL-conjugated beads precipitated the M protein modified with complex N-glycans. The lack of recognition by DSL could be due to the fact that this lectin preferentially binds polyLacNAc on branched glycans. Furthermore, its affinity increases with higher branching and DSL binding is inhibited by α 2–6 sialic acid terminal modification whereas LEL tolerates terminal fucose and sialic acid [12]. To confirm that M proteins are modified with

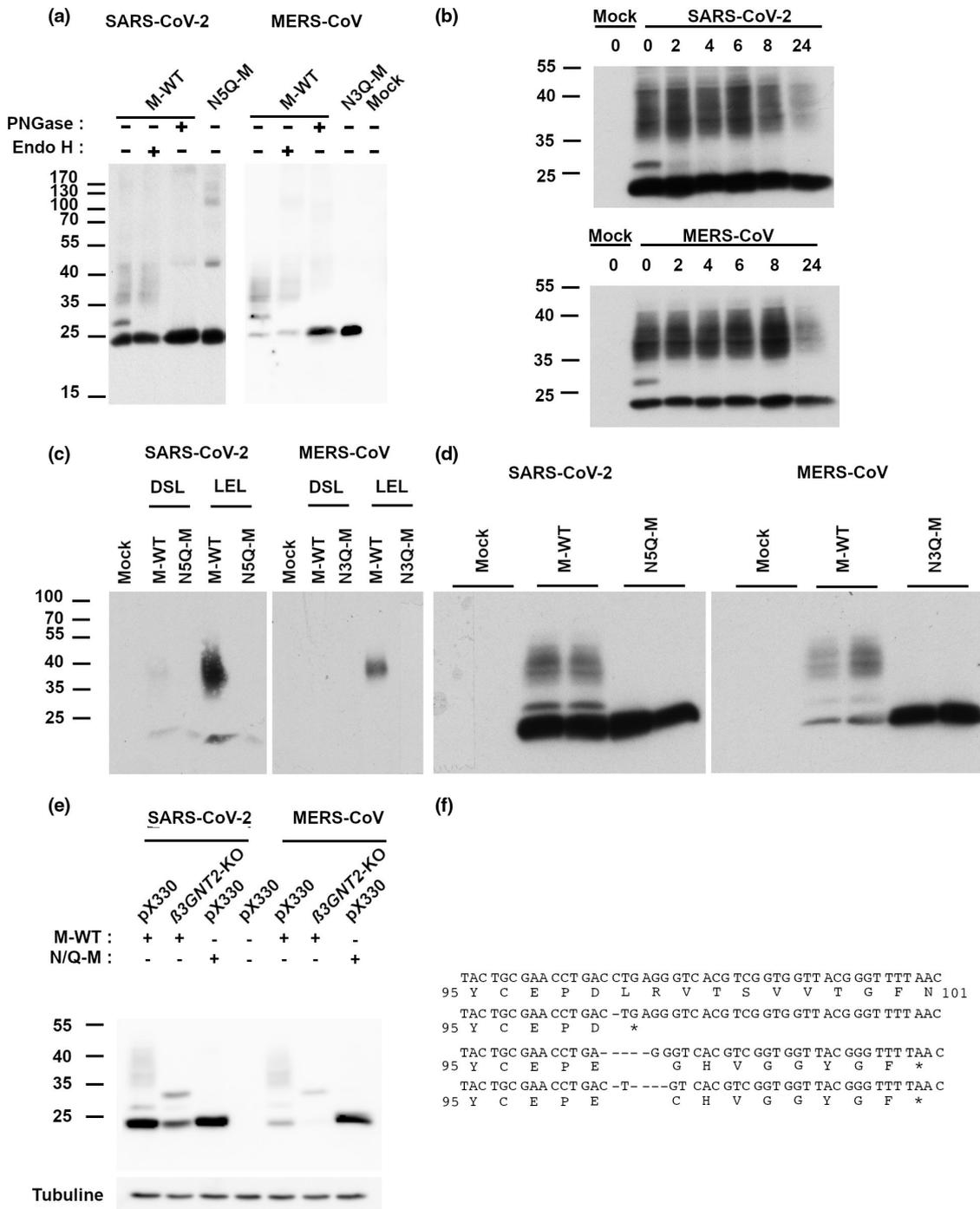


Fig. 1. N-glycosylation of MERS-CoV and SARS-CoV-2 M proteins. (a) Lysates of cells transfected with plasmids encoding the V5-tagged SARS-CoV-2 or MERS-CoV M proteins were treated with Endo-H or PNGase F. M proteins were detected with an anti-V5 antibody. Lysates of cells expressing an N-glycosylation defective mutant (N3Q-M for MERS-CoV or N5Q-M for SARS-CoV-2) were loaded as controls. (b) Cells were transfected with vector expressing the SARS-CoV-2 or MERS-CoV WT-M. Then, cells were treated with 100 μ M cycloheximide to block protein synthesis for different periods of time (2, 4, 6, 8 or 24 h) before lysis. M proteins were detected with an anti-V5 antibody. (c) Lysates of cells transfected with vectors expressing the membrane proteins were precipitated with *Datura stramonium* lectin (DSL)- or *Lycopersicon esculentum* lectin (LEL)-conjugated beads and M proteins were detected in Western blots (precipitation, left panels). (d) M protein expression in cell lysates (=input) was also analysed by Western blotting. (e) Control (pX330) and β 3GnT2-KO cells were transfected with vectors expressing the wild-type or the N-glycosylation deficient form of the MERS-CoV and SARS-CoV-2 M proteins. Cells were lysed and M expression was detected in Western blots. (f) Sequence of the β 3GnT2 protein corresponding to the edited region. Three different editions were detected in similar proportion, all of them resulting in a truncated protein.

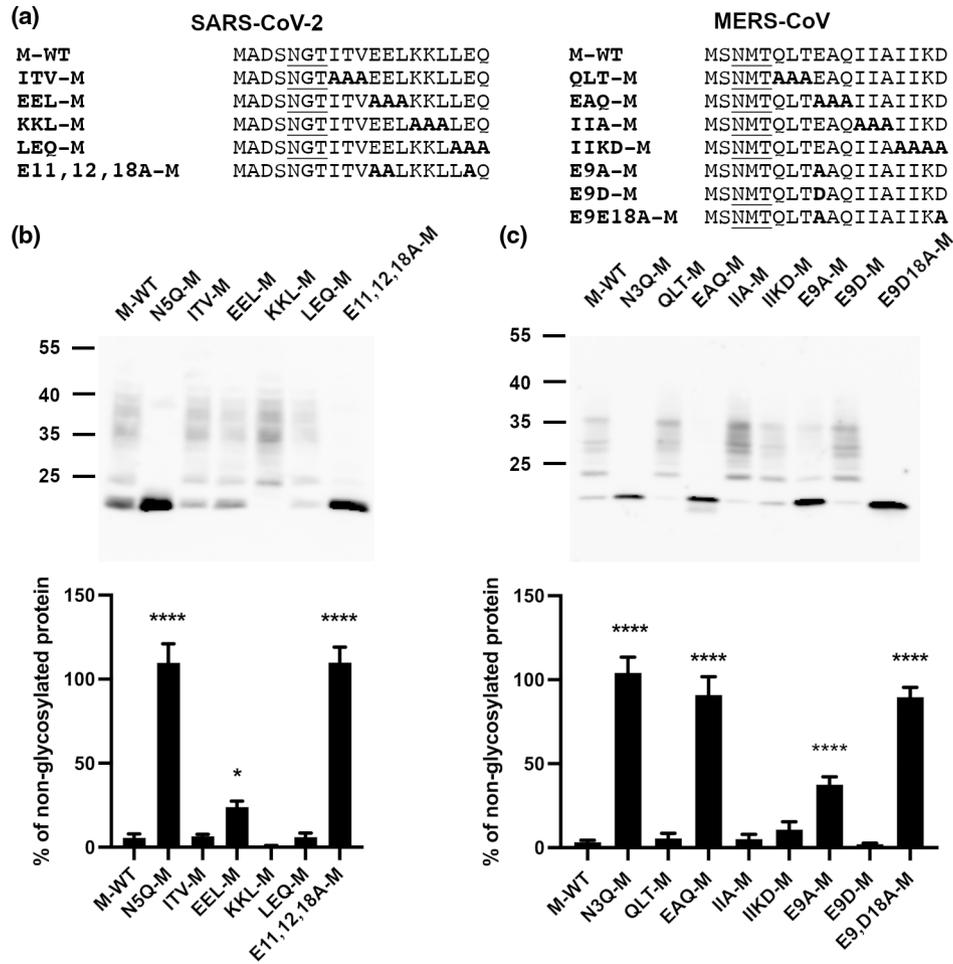


Fig. 2. Mutation of acidic residues of the SARS-CoV-2 and MERS-CoV M proteins abrogates N-glycosylation. (a) Amino acid sequences of the different mutants that were generated. (b) Western blot analysis of the expression of the different mutants in transfected cells. (c) Quantification of the non-glycosylated M protein for each mutant. Cells were transfected, then lysates were either mock- or PNGase F-treated before detection in Western blots. Then the intensity of the band corresponding to the non-glycosylated form of the protein was quantified and normalized with the PNGase F-treated sample reflecting the total amount of protein. Results are presented as the percentage of non-glycosylated proteins and are expressed as the mean of at least four independent experiments. Error bars represent the standard error of the mean (SEM). Results were analysed by using an ANOVA test (* $P < 0.1$; **** $P < 0.0001$).

polylactosamine glycans, we generated a β 1,3-*N*-acetylglucosaminyltransferase 2 (β 3GNT2) knockout Huh-7 cell line by using CRISPR-Cas 9-mediated gene inactivation and expressed the M proteins in these β 3GNT2-KO cells, since β 3GnT2 is a major polylactosamine synthase [13]. In the absence of an antibody recognizing β 3GnT2 in Western blots, we isolated clones and sequenced the β 3GNT2 gene. We choose to use a clone with a stop codon in the protein at position 109 for our experiment. As shown in Fig. 1(e), the third M protein species appeared as a more defined band migrating slightly slower than the high-mannose-conjugated M proteins in this cell line.

These results suggest that MERS-CoV and SARS-CoV-2 M proteins are modified with polylactosamine chain(s).

Nearby acidic residues in the first transmembrane segment are important for M protein glycosylation

The N-terminal domain of M proteins is very short, so we investigated whether amino acid residues in this region were important for the protein to become modified with polylactosamine chains. As we started this study before the release of the SARS-CoV-2 M protein structure, we predicted that the N-terminal domain would be around 20 residues long. Therefore, several residues located further downstream of the N-glycosylation site were mutated three by three into alanine, generating ITV-M, EEL-M, KKL-M and LEQ-M mutants for SARS-CoV-2 and QLT-M, EAQ-M, IIA-M and IIKD-M mutants for MERS-CoV (Fig. 2a). All the mutants that were glycosylated were modified with polylactosamine chains (Fig. 2b, c, top panels). However, an accumulation of non-glycosylated proteins was noted when the residues $_{11}$ EEL $_{13}$ in the SARS-CoV-2 M protein and $_{9}$ EAQ $_{11}$ in the M protein

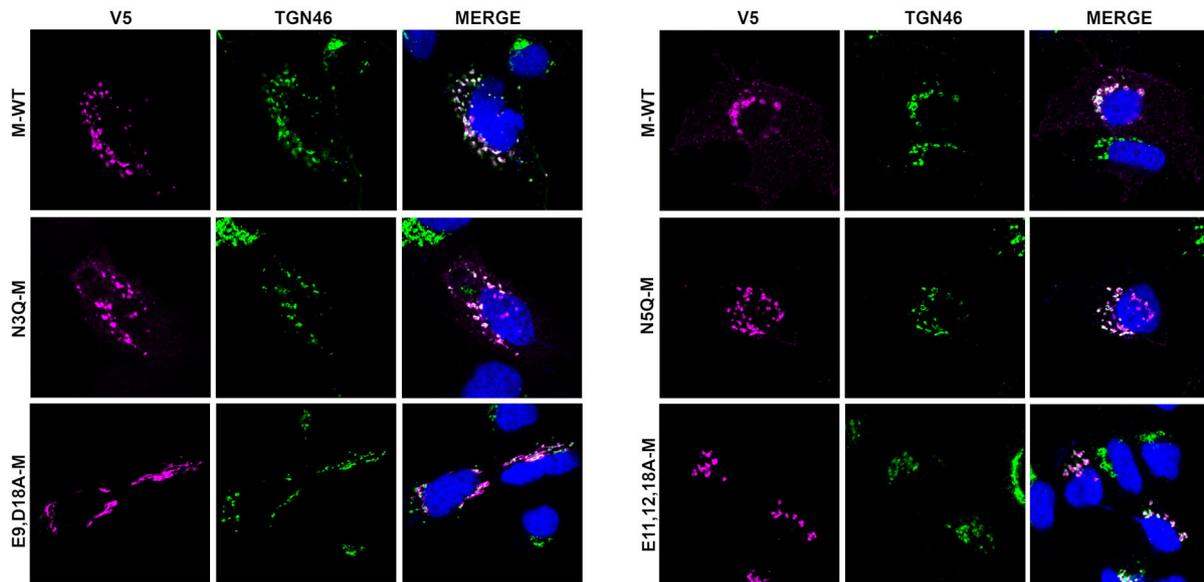


Fig. 3. Mutants are located in the *trans*-Golgi network. Cells were transfected with vector expressing the MERS-CoV or SARS-CoV-2 wild-type M protein or non-glycosylated proteins (MERS-CoV N3Q-M and E9D18A-M or SARS-CoV-2 N5Q-M and E11,12,18A-M), fixed and processed for immunofluorescence. Cells were stained with anti-V5 (magenta) and anti-TGN46 (green) antibodies.

of MERS-CoV were mutated. Because of its heterogeneity, the complex glycan-conjugated M protein can be difficult to detect in immunoblots, so to better compare the glycosylation state of the mutants with the wild-type protein, the non-glycosylated species were quantified for each mutant. The total amount of M protein was quantified after deglycosylation of the sample by PNGase F treatment (Fig. 2c). We found that the majority of wild-type proteins became glycosylated as only 3.32 ± 1.13 and $5.73 \pm 2.18\%$ of MERS-CoV and SARS-CoV-2 M proteins were found to be unglycosylated, respectively. As expected, MERS-CoV N3Q-M and SARS-CoV-2 N5Q-M remained completely unglycosylated. These quantifications also confirmed that the SARS-CoV-2 EEL-M mutant showed a slight defect in glycosylation, whereas the MERS-CoV EAQ-M mutant remained largely unglycosylated ($90.9 \pm 10.82\%$). As both mutants involved glutamic acid residues, the role of such acidic residues was further investigated with additional mutant proteins, including E11,12,18A-M for SARS-CoV-2 and E9A, E9D and E9A,D18A-M for MERS-CoV. Mutation of E11,E12 and E18 in the SARS-CoV-2 M protein completely abolished the glycosylation of the protein (Fig. 2b). For MERS-CoV, the E9A,D18A-M mutant showed a strong defect of its glycosylation whereas the single mutant E9A had a partial effect and E9D mutation had no impact (Fig. 2c). As the above-mentioned „EAQ₁₁“ mutation had a stronger effect than the E9A mutation alone, we also constructed a Q11A-M mutant, but its glycosylation state was similar to the wild-type protein (data not shown). Nonetheless, these results suggest that apart from the intrinsic N-glycosylation site, acidic residues in the N-terminal part of SARS-CoV-2 and MERS-CoV M proteins are important for glycosylation of the protein.

To make sure that the absence of glycosylation was not due to a misfolding of the mutated protein, we investigated the intracellular location of the E9A,D18A-MERS-M and E11,12,18A-SARS-2-M, expecting that if the mutants were misfolded, then they would no longer localize in the *trans*-Golgi network but would be retained in the ER. We did not detect any change of the localization of the mutants, all of them colocalizing with TGN46, a marker of the *trans*-Golgi network (Fig. 3). These results suggest that the E9A,D18A-MERS-M and E11,12,18A-SARS-2-M are correctly folded.

The M protein is a triple membrane-spanning protein with its N-terminus in the lumen of intracellular compartments and the C-terminus exposed in the cytosol. This topology with the N-terminus facing the lumen of the intracellular compartments is very important for the glycosylation of the protein, so we tested if the N-terminus of our mutants is accessible to antibodies when performing selective permeabilization of the cells. To specifically permeabilize the plasma membrane of the cells, we performed freeze/thaw cycles of the cells after fixation. As observed for the wild-type protein, the N-terminal V5-tagged extremity of the wild-type or the mutated M proteins was not accessible to antibodies with selective permeabilization but was detected upon permeabilization with 0.1% Triton X-100 (Fig. 4). In contrast, the protein was detected with an antibody directed against the C-terminus of the protein with both permeabilization methods. These results suggest that the topology of the protein is not affected when the acidic residues were mutated into alanine.

Finally, to confirm these findings we also inserted an N-glycosylation site in a linker that we inserted between the second and third transmembrane segment of the MERS-CoV M protein as previously reported for the SARS-CoV M protein [14] (Fig. 5a).

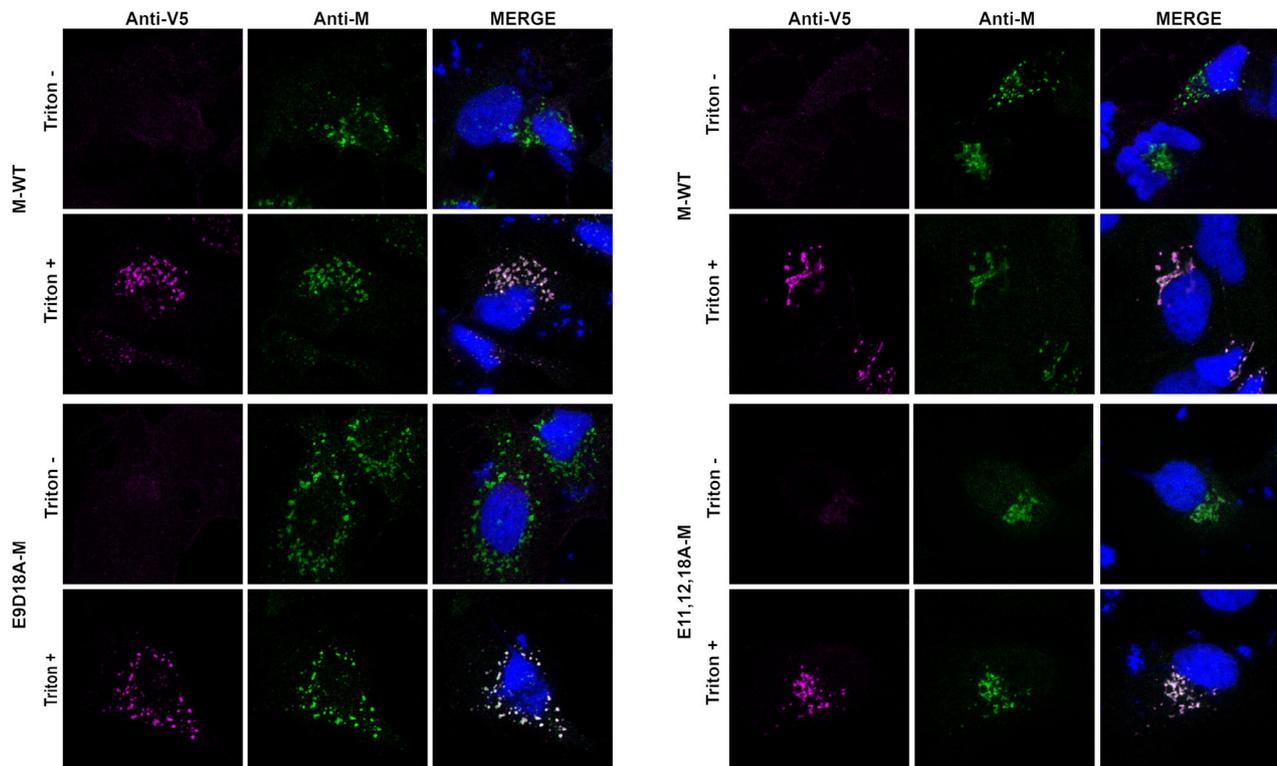


Fig. 4. The N-terminal extremity of mutant proteins is exposed in the luminal side of intracellular compartments. Cells expressing the N-terminal V5 tagged MERS-CoV wild-type or E9,D18A-M or the SARS-CoV-2 wild-type or E11,12,18A-M proteins were fixed. Cells were either permeabilized by freeze/thaw (Triton-) or with 0.1% Triton X-100 (Triton+) and stained with anti-V5 (magenta) or with an anti-MERS-CoV M or anti-SARS-CoV-2 M that are both directed against the C-terminal domain of the proteins (green).

To have a better understanding of the different migration profiles of our construct with the linker, we also constructed the M-linker protein with mutations of the two glycosylation sites (N3Q and/or N88Q). As shown by the different migration profiles of N3Q-M, N88Q-M and N3Q,N88Q-M, both sites were glycosylated and could be modified by polylectosamine chains (Fig. 5b). The E9,D18 mutation abolished only the glycosylation of the N3 sites. Indeed, the E9D18A-M-linker had a similar migration profile as the N3Q-M-linker or N88Q-M-linker mutants and the E9,D18A-N88Q-M-linker was similar to the double mutant N3Q,N88Q-M-linker.

First, these results confirm the correct topology of the protein as N88 in the linker is accessible to glycosyl-transferases. Moreover, these results also show that the acidic residues act locally on the N-glycosylation site of the protein as their mutation did not affect glycosylation of the N-glycosylation site inserted into the linker.

HCV E1-M and SARS-CoV-2 M-ORF3A chimeric proteins confirm the importance for correct subcellular trafficking and the role of the first transmembrane segment of M for its glycosylation.

To further assess whether a motif in the N-terminal extremity of the M protein could be involved in modification of the protein by polylectosamine chains, we constructed a chimaera of the SARS-CoV-2 M protein in which the first nine residues (MADSN~~GT~~IT) were replaced by the 11 N-terminal residues of hepatitis C virus (HCV) E1 protein (MAQVKNS~~SG~~LY). HCV envelope proteins are highly glycosylated but it has been shown that they do not contain complex N-glycans as they mainly localize in the ER. The E1 envelope protein also contains an N-glycosylation site in its N-terminals extremity [15]. When the N-terminus of E1 was transferred on the SARS-CoV-2 M protein (E1-M), it was also modified by polylectosamine chains as shown by the migration profile in Western blots (Fig. 6a). This result suggests that trafficking of the protein through the Golgi is important for the modification of the protein with polylectosamine chains. We then analysed the intracellular distribution of E1-M. E1 was mainly distributed in the ER, but E1-M had the same localization as M-WT and co-localized with TGN46 (Fig. 6b).

The structure of the SARS-CoV-2 M protein is very similar to that of ORF3a [3, 4]. The N-terminal domain of ORF3a is composed of 42 residues, being much longer than the N-terminal domain of SARS-CoV-2 M. According to the M structure solved by Zhang and colleagues, the glutamic residues important for SARS-CoV-2 M protein glycosylation are located within the first transmembrane segment. As the first transmembrane segment of ORF3a is shorter and does not contain any acidic residues,

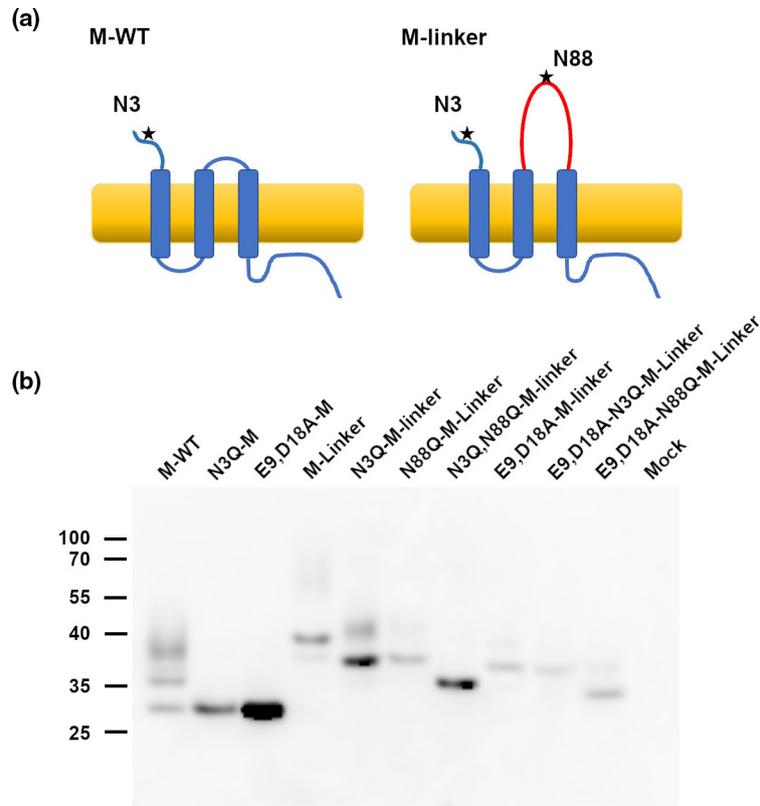


Fig. 5. N-glycosylation of a linker inserted between the second and third transmembrane segment of the MERS-CoV M protein is not dependent on the acidic residues of the N-terminal extremity. (a) Schematic representation of the wild-type M protein and the protein with a linker (red) inserted between the second and third transmembrane segment (M-linker). The positions of the glycosylation sites are depicted with stars. (b) Cells were transfected with vectors expressing constructs containing the linker in combination with different mutations (E9,D18A and/or N3Q and/or N88Q). Proteins were analysed in immunoblots with an anti-V5 antibody.

we replaced its N-terminus by that of SARS-CoV-2 M protein (Nt-M/ORF3a, Fig. 7a). This chimaera remained unglycosylated (Fig. 7b). However, when we also replaced the first transmembrane segment of ORF3a with that of SARS-CoV-2 M (TM1-M/ORF3a), then the protein was modified with polylactosamine chains as shown by its migration profile in Western blots. ORF3a has been shown to localize in the endocytic pathway and at the cell surface [16]. As expected, we did not detect any co-localization of ORF3a with TGN46 in immunofluorescence and observed a punctuated staining consistent with endocytic compartments (Fig. 7c). Both M-ORF3a chimaeras had similar patterns as ORF3a staining and did not colocalize with TGN46. These results confirm the important role played by the first transmembrane segment of the SARS-CoV-2 M protein for its glycosylation.

OST-B is responsible for glycosylation of the M protein

Next, we investigated the machinery involved in glycosylation of the M protein. Two oligosaccharyltransferases (OSTs) can mediate the transfer of the preassembled oligosaccharide onto the asparagine of the N-glycosylation sequon. OST-A is associated with the protein translocon channel and acts co-translationally. OST-B is not associated with the sec61 complex but localizes to the rough ER and is considered as a proofreading machinery. The main difference between OST-A and B is in the catalytic subunit STT3. OST-A contains STT3A whereas OST-B contains STT3B. Therefore, we invalidated the expression of STT3A or STT3B by CRISPR/Cas9 (Fig. 8a) and expressed the M protein and the spike protein that is highly N-glycosylated in these KO cell populations [17]. SARS-CoV-2 M protein was no longer glycosylated in the STT3B-KO cells but was fully glycosylated in the STT3A-KO cells (Fig. 8b). For MERS-CoV-M protein, we also detected an accumulation of non-glycosylated protein in STT3B-KO cells. A weak N-glycosylation could be detected that is probably due to residual expression of STT3B. In contrast, the S protein of both MERS-CoV and SARS-CoV-2 were glycosylated in STT3A- and STT3B-KO cells (Fig. 8c), but in the double knock-out cells, expression of the protein was severely impaired suggesting a misfolding of the protein caused by the lack of glycosylation. For SARS-CoV-2 S protein, a smear could be observed suggesting a modification of its glycosylation. These results suggest that M proteins are glycosylated by the OST-B complex.

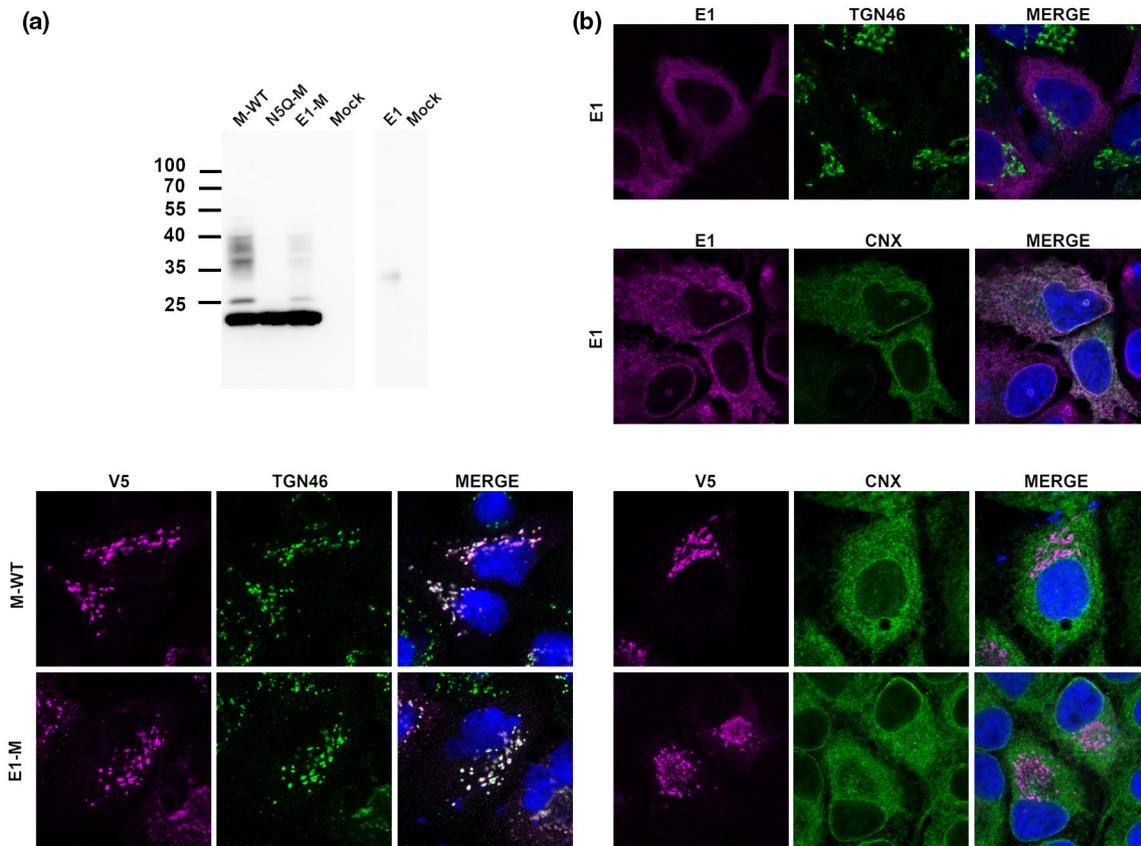


Fig. 6. HCV E1 N-terminal extremity is conjugated to complex glycans when transferred on SARS-CoV-2 M protein. (a) Western blot analysis of SARS-CoV-2 M and the E1-M chimaera with an anti-V5; E1 was detected with an anti-E1 (A6 mAb). (b) Cells expressing V5-M, V5-E1-M and E1 were fixed and co-stained with either anti-V5 and anti-TGN46 or anti-calnexin (CNX) antibodies or anti-E1 and anti-TGN46 or anti-calnexin antibodies.

Glycosylated proteins in infected cells and viral assembly

When expressed alone, the M protein is located in the *trans*-Golgi network, while in infected cells, the M protein is spread throughout the biosynthetic pathway, including the ERGIC (ER–Golgi intermediate compartment) where the new virions are formed. This difference in intracellular distribution of the protein may affect its glycosylation as glycosidases and glycosyltransferases implicated in complex glycan formation are located in the *trans*-Golgi. Because of the heterogeneity of the complex-glycan conjugated to the M proteins, these forms are difficult to observe in Western blots. To confirm that the M protein is also modified with a poly-lactosamine chain in infected cells, we took advantage of our $\beta 3GNT2$ -KO cell line. Therefore, we infected this cell line with MERS-CoV or with SARS-CoV-2. However, as Huh-7 cells do not express the receptor of SARS-CoV-2, they were transduced with a lentivirus expressing ACE2 prior to SARS-CoV-2 infection. Cells were infected and lysed 24 h later and the migration profile of membrane proteins was analysed in immunoblots (Fig. 9a). In control cells, two main bands could be observed for SARS-CoV-2 and MERS-CoV M protein. These bands correspond to the non-glycosylated protein and the high-mannose glycosylated protein. With longer exposure of the Western blot, in the lanes corresponding to the control cells, a smear could be observed that is not present in $\beta 3GNT2$ -KO cells. In these cells, an additional band migrating slower than the high-mannose glycosylated protein was present. This result shows that M protein is also modified by poly-lactosamine chains in the context of infected cells.

As new virions bud in the ERGIC, it is likely that poly-lactosamines are not linked to the M proteins before budding. After budding, virions transit through the Golgi and further modifications of the proteins may occur. To identify the glycosylated forms of the protein that are incorporated during viral assembly, we used a VLP production assay. Huh-7 cells were co-transfected with a vector expressing the M protein alone or in combination with E protein expression vector. VLPs released in the extracellular medium were concentrated by ultracentrifugation on sucrose cushions. For both SARS-CoV-2 and MERS-CoV, the three M glycoforms were detected in the cell lysates (Fig. 9b). In contrast, the VLP fraction showed a clear enrichment of the high-mannose form of the M protein for both viruses (Fig. 9c). When using both glycosylated and N3Q/N5Q unglycosylated M mutant proteins, VLP formation was visible, confirming previous findings that coronaviruses do not use the N-glycans on the M proteins for assembly.

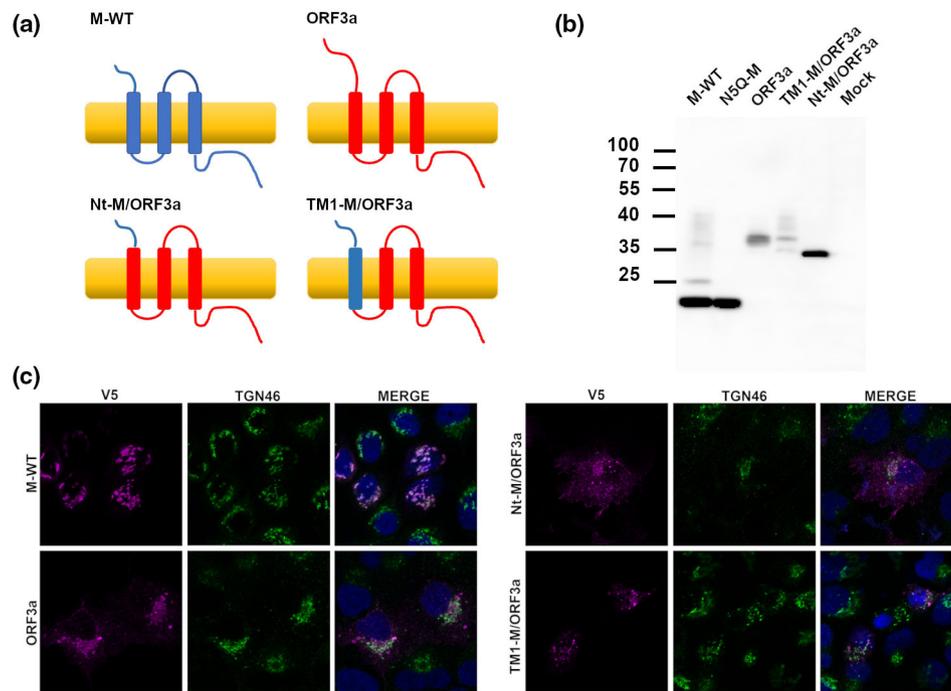


Fig. 7. The first transmembrane segment of SARS-CoV-2 M is required for a chimaera between M and ORF3a to be glycosylated. (a) Drawing of the different chimaeras that were constructed between SARS-CoV-2 M (blue) and ORF3a (green). (b) Western blot analysis of the different chimaeras. (c) Cells expressing the V5-tagged SARS-CoV-2 M, ORF3a or the different chimaeras were fixed and stained with anti-V5 (magenta) and anti-TGN46 (green).

Interestingly, VLP formation was more efficient for SARS-CoV-2 than MERS-CoV under the same experimental conditions. Taken together, these results seem to be in accordance with virus budding occurring in the ERGIC compartment, but also indicate that the M protein is not greatly modified after formation of the new VLPs.

DISCUSSION

N-glycosylation occurs on an N-X-T/S sequon with X being any amino acid residue except proline, but it is not known how the protein environment of the sequon influences its glycosylation efficiency. Here, we show that nearby acidic residues in MERS-CoV and SARS-CoV-2 M proteins are important for their glycosylation, their mutations completely abolishing the glycosylation of the proteins. The structure of SARS-CoV-2 M protein has been recently solved by two different groups [3, 4]. In the structure solved by Dolan *et al.*, the N-terminal domain structure of the protein comprising residues 1–16 was not solved suggesting that this domain was disordered and the first transmembrane segment was found to start at residue 17. In the structure of Zhang *et al.*, the first transmembrane segment was found to start at residue 10, and only the structure of the first nine residues was not solved. Therefore, in this second structure, the acidic residues that are important for M glycosylation are included in the first transmembrane segment. If the acidic residues are not part of the transmembrane segment, then we can hypothesize that because of their charge, these residues may push the N-terminus away from the surface of the ER membrane and facilitates recognition of the sequon by the OST. Membrane electrostatics depend largely on phosphatidylserine (PS), a negatively charged lipid. This lipid only accounts for a few per cent of the ER membrane composition, but PS is mostly present on the luminal side of the ER membrane [18]. On the other hand, these residues may also promote a direct interaction of OST with the M protein. Two different machineries are responsible for the initial transfer of the oligosaccharide precursor to the nascent protein. OST-A is directly associated with the translocon channel whereas OST-B is not and is considered as a proofreading machinery [19]. We found that OST-B is probably responsible for SARS-CoV-2 and MERS-CoV M N-glycosylation. Indeed, OST-B has been shown to be responsible for N-glycosylation of sites located near the N-terminus of the protein such as the N-glycosylation site found in SARS-CoV-2 and MERS-CoV M proteins. Acidic residues are relatively rare in transmembrane helices and their flanking region, but it has been suggested that these negatively charged residues are more likely to be depleted from the cytosolic side of the transmembrane region and enriched on the ‘outside’ face [20]. It has also been suggested that changing the flanking charges of a transmembrane segment can reverse the topology of the transmembrane [21], so we tested if the N-terminus of the M protein in which charged residues were mutated into alanine was still directed to the luminal side of intracellular compartments. We

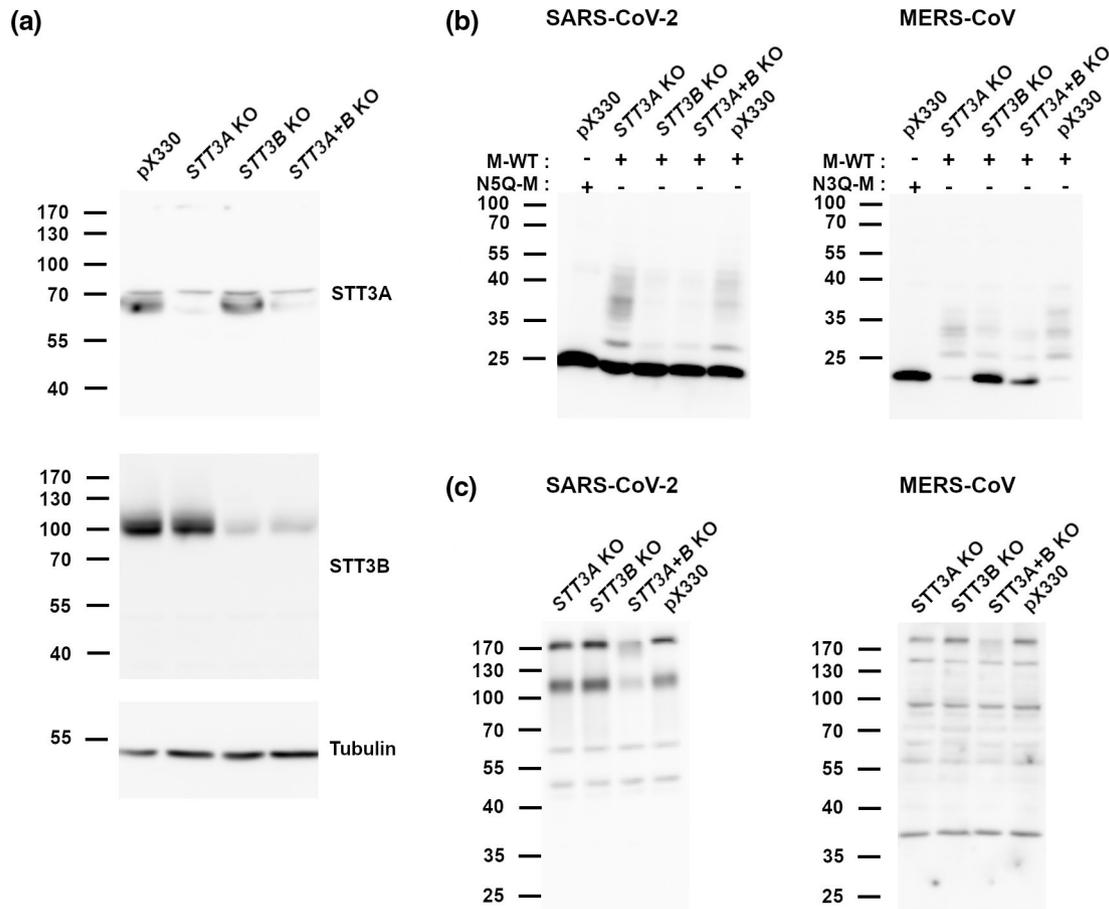


Fig. 8. OST-B mediates M protein glycosylation. (a) Expression level of STT3A and STT3B in a KO cell population. (b) STT3A-KO, STT3B-KO or control cells (pX330) were transfected with expression vector for M-WT or unglycosylated-M mutants (N3Q-M for MERS-CoV and N5Q-M for SARS-CoV-2). (c) STT3A-KO, STT3B-KO or control cells (pX330) were transfected with expression vector for SARS-CoV-2 or MERS-CoV S protein. SARS-CoV-2 S protein was detected with an anti-S1 mAb and the MERS-CoV S protein with a polyclonal anti-spikes.

did not see any change of protein topology either with selective permeabilization and staining with an antibody recognizing the N-terminus of the protein or by glycosylation of a linker inserted between the second and third transmembrane segment of the MERS-CoV M protein. Moreover, mutation of the acidic residues in the first transmembrane segment did not have any effect on the glycosylation of the linker inserted between the second and third transmembrane segment of the protein, suggesting that the acidic residues act locally on the structure of the M protein for its accessibility to OST.

Our results suggest that M proteins are modified with poly-lactosamine chain(s), but the number of arms that are modified with poly-lactosamines or the number of lactosamines linked on one arm remain to be determined. In $\beta 3GNT2$ -KO cells, each arm would be modified with only one LacNac. The difference of migration profile with the control cells suggests that either one arm is modified with a long chain of poly-lactosamine or two or more arms are modified with shorter chains of poly-lactosamines.

We first thought that the M protein must contain a signal that specifies the modification of the protein with poly-lactosamines, and thus we performed serial mutations of the N-terminal domain. We did not find any signal responsible for poly-lactosamine modification but residues that are important for glycosylation of the protein. We confirmed this absence of signal by switching the N-terminal extremity of the M protein with the one of the E1 envelope protein of HCV. Glycosylation of E1 has been extensively studied [15, 22]. E1 contains five potential N-glycosylation sites but only four of them are glycosylated. Of particular interest, E1 contains an N-glycosylation site (E1N1) in its N-terminal extremity that is branched with high-mannose glycans. Our E1-M chimaera containing the E1N1 site is glycosylated with complex glycans. This result suggests that trafficking through the Golgi is probably more important for the modification with poly-lactosamines than a specific signal. Therefore, we constructed a chimaera between the M protein and ORF3a. The ORF3a structure is very similar to that of SARS-CoV-2 M protein, but its N-terminal extremity with 42 residues is longer than that of SARS-CoV-2 M. The first transmembrane segment is 28 residues long for SARS-CoV-2 M and 20 residues long for ORF3a; moreover, ORF3a TM1 does not contain any acidic residues. The chimaera

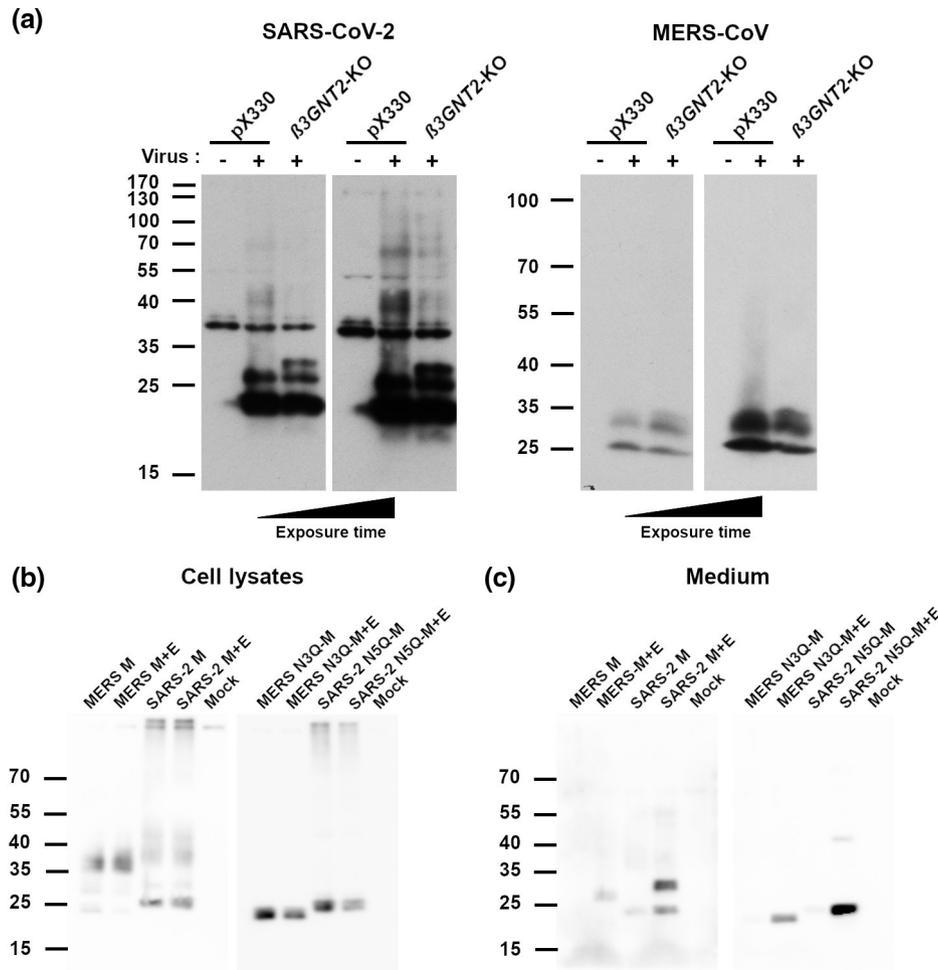


Fig. 9. M proteins are modified with poly-lactosamine chains in infected cells but high-mannose-conjugated M proteins are incorporated in VLPs. (a) β 3GNT2 cells were infected with SARS-CoV-2 (left panel) or MERS-CoV (right panel) and lysed 24 h later. Cell lysates were analysed in Western blots. M proteins were detected by using an anti-M antibody. (b) Cells were transfected with plasmids encoding either the V5-tagged wild-type M proteins of SARS-CoV-2 and MERS-CoV (left panel) or their non-glycosylated forms (right panel), in the absence or presence of an E-encoding plasmid. Cell lysates were loaded onto SDS-PAGE gel and M proteins were detected in an immunoblot with an anti-V5 antibody. (c) Supernatants were collected and ultracentrifuged on a 20% sucrose cushion to concentrate the VLPs secreted in the medium. Then VLPs were resuspended and analysed in Western blots. The M protein was detected with an anti-V5 antibody.

containing only the N-terminal domain of M was not glycosylated, whereas addition of the SARS-CoV-2 M TM1 rescues the glycosylation of the chimaera that is modified with poly-lactosaminoglycans. It has been reported that ORF3a colocalizes with markers of the plasma membrane, the endocytic pathway and the Golgi. Therefore the protein is synthesized in the ER and then crosses the Golgi before being sorted to the plasma membrane or endosomes [16]. Similar to ORF3a, our chimaeras did not colocalize with TGN46 suggesting that these chimaeras have a similar trafficking pattern as ORF3a. Taken together, our results suggest that no signal specifies the modification with poly-lactosaminoglycans but trafficking through the Golgi that allows for encounters of the proper glycosylases and glycosyl-transferases is important. LAMP-2 is a lysosomal protein that contains 18 N-glycans some of which are modified with poly-lactosaminoglycans [23]. It has been shown that the Golgi residence time of the protein is associated with the extent of poly-lactosamine conjugation [24]. It is interesting to note for LAMP-2 that it is not known why some N-glycosylation sites are modified with poly-lactosamine glycans and the other ones with high-mannose. It has been suggested that only membrane-anchored proteins can be modified with poly-lactosamines and even if the proximity to the membrane has been evoked as a possible requirement for poly-lactosamine conjugation, this has not been demonstrated [25].

The M proteins of MERS-CoV and SARS-CoV-2 are N-glycosylated and are processed from a high-mannose form to a heterogeneous form of higher molecular weight. Because of their heterogeneity, these forms can be difficult to visualize in immunoblots and leading to the conclusion that the majority of the M protein is not glycosylated. However, when we quantified the non-glycosylated form of the protein in cell lysates, only around 5% of the SARS-CoV-2 M and 3% of the MERS-CoV M protein were

not glycosylated. Our results show that these two proteins are modified by poly lactosamine chains. Other viral envelope proteins have been shown to harbour similar complex glycans. Indeed, the respiratory syncytial virus small hydrophobic (SH) protein has been shown to be modified with poly lactosamines [26]. The function of this protein remains unclear, but it has been shown to be incorporated into viral particles [27]. Influenza B virus encodes a unique NB protein that has also been shown to harbour poly lactosamine modifications and to be incorporated into the virion [28, 29].

N-glycans have a wide variety of functions; they can affect the conformation, solubility, activity or antigenicity of glycosylated proteins. The functions of poly lactosamine chains remain unclear. β 3GnT2 is a key enzyme for the synthesis of poly lactosamine chains. β 3GnT2 KO mice present a hyperactivation of lymphocytes, showing that poly lactosamine chains are important for immunological biofunctions. Poly lactosamine chains are also involved in 'I' and 'T' blood group antigens. Moreover, they can be recognized by glycan-binding proteins. For example, poly lactosamines conjugated to LAMP-2 can bind galectin-9. This binding is important for lysosome homeostasis [30]. The role of N-glycosylation of the M protein of coronaviruses remains unclear. It is well-established that glycosylation of the M protein is not necessary for IBV or MHV virus assembly, and we obtained similar results with MERS-CoV and SARS-CoV-2 M protein, as we observed VLP formation with N-glycosylation defective M proteins. However, we observed that the high-mannose glycoconjugated form of the M protein was mainly incorporated in VLPs. For SARS-CoV-2, a small proportion of non-glycosylated protein was also observed, suggesting that the non-glycosylated M protein could dimerize with N-glycosylated M. As budding of coronaviruses occurs in the ERGIC, it is not surprising that high-mannose-conjugated M proteins are incorporated into nascent particles. After assembly of the virus, it is not clear how new virions are released from the cells, but our results suggest that the M protein glycan does not undergo any further maturation after its incorporation into the viral particle. Indeed, it has been suggested that β -coronaviruses may egress by lysosomal exocytosis [31], but it is not clear where virions are sorted toward the lysosomal pathway, if it is directly from the ERGIC or after crossing the Golgi.

The role of the glycosylation of coronavirus M proteins is not clear. It has been shown that M proteins of IBV, TGEV or MHV have interferogenic activity. The glycosylation status of the protein is probably important for this activity. MHV M protein is O-glycosylated, and recombinant viruses with an M protein either O-glycosylated, N-glycosylated or non-glycosylated have been generated by De Haan *et al.* [6]. The recombinant virus with an N-glycosylated protein was a stronger inducer of type I IFN than the recombinant virus with an O-glycosylated protein in *in vitro* assays. The recombinant virus with a non-glycosylated protein was the poorer inducer. *In vivo*, replication of the virus in the liver but not in the brain was also dependent on the glycosylation status of the protein. Viral replication was highest for the virus with an N-glycosylated protein. The pathogenicity of the different strains of MHV varies. Whereas MHV-A59 causes an acute hepatitis, an acute encephalitis and chronic demyelination in infected mice, MHV-2 produces only hepatitis and mild meningitis without encephalitis or demyelination. Interestingly, it has been shown that the M protein of MHV-2 is N-glycosylated [32]. Although coronavirus cell tropism depends mainly on the S protein, we cannot exclude the possible implication of M through its glycosylation status in modulating cell tropism. Poly lactosamines can also be linked to O-glycans, and we hypothesize that MHV M could also be modified with poly lactosamines. Glycosylation of the M protein is conserved suggesting that it has a functional role. It is probably involved in regulating host-pathogen interactions but further investigations are necessary to characterize this interaction and the role played by M protein poly lactosaminoglycans.

Funding information

L.D. is supported by a fellowship from the Region Hauts-de-France, and D.J. is supported by a fellowship from the fondation pour la recherche médicale (ECO202006011744). This work was supported by the Agence Nationale de la Recherche (ANR-22-0E15-0030-01).

Acknowledgements

We thank Dr Philip Meuleman (Ghent University) for kindly providing the A6 mAb. Immunofluorescence analyses were performed with the help of the imaging core facility of the Bioluminescence Centre Lille Nord-de-France. We thank Robin Prath and Nicolas Vandenabeele for their technical help in the BSL-3 facility

Author contributions

S.B. and J.D. conceived and designed the experiments. D.J., L.D., A.D. and Y.R. performed the experiments. S.B., D.J., L.D., Y.R., J.D. analyzed the data. S.B. wrote the manuscript. All authors read and approved the final manuscript.

Conflicts of interest

The authors declare that they have no conflicts of interest with the contents of this article

References

1. Masters PS. The molecular biology of coronaviruses. *Adv Virus Res* 2006;66:193–292.
2. Escors D, Camafeite E, Ortego J, Laude H, Enjuanes L. Organization of two transmissible gastroenteritis coronavirus membrane protein topologies within the virion and core. *J Virol* 2001;75:12228–12240.
3. Zhang Z, Nomura N, Muramoto Y, Ekimoto T, Uemura T, *et al.* Structure of SARS-CoV-2 membrane protein essential for virus assembly. *Nat Commun* 2022;13:4399.
4. Dolan KA, Dutta M, Kern DM, Kotecha A, Voth GA, *et al.* Structure of SARS-CoV-2 M protein in lipid nanodiscs. *Elife* 2022;11:e81702.
5. Charley B, Lavenant L, Delmas B. Glycosylation is required for coronavirus TGEV to induce an efficient production of IFN alpha by blood mononuclear cells. *Scand J Immunol* 1991;33:435–440.
6. de Haan CAM, de Wit M, Kuo L, Montalto-Morrison C, Haagmans BL, *et al.* The glycosylation status of the murine hepatitis coronavirus M protein affects the interferogenic capacity of the virus in vitro

- and its ability to replicate in the liver but not the brain. *Virology* 2003;312:395–406.
7. Liang JQ, Fang S, Yuan Q, Huang M, Chen RA, et al. N-Linked glycosylation of the membrane protein ectodomain regulates infectious bronchitis virus-induced ER stress response, apoptosis and pathogenesis. *Virology* 2019;531:48–56.
 8. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 2013;8:2281–2308.
 9. Perrier A, Bonnin A, Desmarests L, Danneels A, Goffard A, et al. The C-terminal domain of the MERS coronavirus M protein contains a trans-Golgi network localization signal. *J Biol Chem* 2019;294:14406–14421.
 10. Machamer CE, Mentone SA, Rose JK, Farquhar MG. The E1 glycoprotein of an avian coronavirus is targeted to the cis Golgi complex. *Proc Natl Acad Sci U S A* 1990;87:6944–6948.
 11. Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, (eds). *Essentials of Glycobiology*. 4th edn. <http://www.ncbi.nlm.nih.gov/books/NBK579918/> (Accessed 7 July 2023). Cold Spring Harbor Laboratory Press, 2022.
 12. Bojar D, Meche L, Meng G, Eng W, Smith DF, et al. A useful guide to lectin binding: machine-learning directed annotation of 57 unique lectin specificities. *ACS Chem Biol* 2022;17:2993–3012.
 13. Togayachi A, Kozono Y, Kuno A, Ohkura T, Sato T, et al. A major polylysosamine synthase: analysis of B3gnt2-deficient Mice. In: *Methods in Enzymology*. Elsevier, . pp. 185–204.
 14. Voss D, Kern A, Tragglia E, Eickmann M, Stadler K, et al. Characterization of severe acute respiratory syndrome coronavirus membrane protein. *FEBS Lett* 2006;580:968–973.
 15. Dubuisson J, Hsu HH, Cheung RC, Greenberg HB, Russell DG, et al. Formation and intracellular localization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia and Sindbis viruses. *J Virol* 1994;68:6147–6160.
 16. Miller AN, Houlihan PR, Matamala E, Cabezas-Bratesco D, Lee GY, et al. The SARS-CoV-2 accessory protein Orf3a is not an ion channel, but does interact with trafficking proteins. *Elife* 2023;12:e84477.
 17. Shrimal S, Gilmore R, Cherepanova NA. Cotranslational and post-translational N-glycosylation of proteins in the endoplasmic reticulum. *Semin Cell Dev Biol* 2015;41:71–78.
 18. Bigay J, Antony B. Curvature, lipid packing, and electrostatics of membrane organelles: defining cellular territories in determining specificity. *Dev Cell* 2012;23:886–895.
 19. Shrimal S, Gilmore R, Cherepanova NA. Cotranslational and post-translational N-glycosylation of proteins in the endoplasmic reticulum. *Semin Cell Dev Biol* 2015;41:71–78.
 20. Baker JA, Wong W-C, Eisenhaber B, Warwicker J, Eisenhaber F. Charged residues next to transmembrane regions revisited: “Positive-inside rule” is complemented by the “negative inside depletion/ outside enrichment rule.” *BMC Biol* 2017;15:72.
 21. Martínez-Gil L, Saurí A, Martí-Renom MA, Mingarro I. Membrane protein integration into the endoplasmic reticulum: membrane protein integration. *FEBS J* 2011;278:3846–3858.
 22. Goffard A, Callens N, Bartosch B, Wychowski C, Cosset F-L, et al. Role of N-linked glycans in the functions of hepatitis C virus envelope glycoproteins. *J Virol* 2005;79:8400–8409.
 23. Carlsson SR, Fukuda M. The polylysosaminoglycans of human lysosomal membrane glycoproteins lamp-1 and lamp-2. Localization on the peptide backbones. *J Biol Chem* 1990;265:20488–20495.
 24. Nabi IR, Dennis JW. The extent of polylysosamine glycosylation of MDCK LAMP-2 is determined by its Golgi residence time. *Glycobiology* 1998;8:947–953.
 25. Fukuda M, Guan JL, Rose JK. A membrane-anchored form but not the secretory form of human chorionic gonadotropin-alpha chain acquires polylysosaminoglycan. *J Biol Chem* 1988;263:5314–5318.
 26. Anderson K, King AMQ, Lerch RA, Wertz GW. Polylysosaminoglycan modification of the respiratory syncytial virus small hydrophobic (SH) protein: a conserved feature among human and bovine respiratory syncytial viruses. *Virology* 1992;191:417–430.
 27. Huong TN, Ravi Iyer L, Lui J, Wang DY, Tan BH, et al. The respiratory syncytial virus SH protein is incorporated into infectious virus particles that form on virus-infected cells. *Virology* 2023;580:28–40.
 28. Elderfield RA, Koutsakos M, Frise R, Bradley K, Ashcroft J, et al. NB protein does not affect influenza B virus replication in vitro and is not required for replication in or transmission between ferrets. *J Gen Virol* 2016;97:593–601.
 29. Williams MA, Lamb RA. Polylysosaminoglycan modification of a small integral membrane glycoprotein, influenza B virus NB. *Mol Cell Biol* 1988;8:1186–1196.
 30. Sudhakar JN, Lu H-H, Chiang H-Y, Suen C-S, Hwang M-J, et al. Luminal Galectin-9-Lamp2 interaction regulates lysosome and autophagy to prevent pathogenesis in the intestine and pancreas. *Nat Commun* 2020;11:4286.
 31. Ghosh S, Dellibovi-Ragheb TA, Kerviel A, Pak E, Qiu Q, et al. β -coronaviruses use lysosomes for egress instead of the biosynthetic secretory pathway. *Cell* 2020;183:1520–1535.
 32. Yamada YK, Yabe M, Ohtsuki T, Taguchi F. Unique N-linked glycosylation of murine coronavirus MHV-2 membrane protein at the conserved O-linked glycosylation site. *Virus Res* 2000;66:149–154.

Five reasons to publish your next article with a Microbiology Society journal

1. When you submit to our journals, you are supporting Society activities for your community.
2. Experience a fair, transparent process and critical, constructive review.
3. If you are at a Publish and Read institution, you'll enjoy the benefits of Open Access across our journal portfolio.
4. Author feedback says our Editors are 'thorough and fair' and 'patient and caring'.
5. Increase your reach and impact and share your research more widely.

Find out more and submit your article at microbiologyresearch.org.