# Identification of a membrane-associated element (MAE) in the C-terminal region of SARS-CoV-2 nsp6 that is essential for viral replication

Yuying Han,[1] Zhenghong Yuan,[1] Zhigang Yi[1,2]

**AUTHOR AFFILIATIONS** See affiliation list on p. 16.

**ABSTRACT** The coronavirus disease 2019 (COVID-19) pandemic, caused by the novel coronavirus severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2), has rapidly spread worldwide since its emergence in late 2019. Its ongoing evolution poses challenges for antiviral drug development. Coronavirus nsp6, a multiple-spanning transmembrane protein, participates in the biogenesis of the viral replication complex, which accommodates the viral replication-transcription complex. The roles of its structural domains in viral replication are not well studied. Herein, we predicted the structure of the SARS-CoV-2 nsp6 protein using AlphaFold2 and identified a highly folded C-terminal region (nsp6C) downstream of the transmembrane helices. The enhanced green fluorescent protein (EGFP)-fused nsp6C was found to cluster in the cytoplasm and associate with membranes. Functional mapping identified a minimal membrane-associated element (MAE) as the region from amino acids 237 to 276 (LGV–KLL), which is mainly composed of the α-helix H1 and the α-helix H2; the latter exhibits characteristics of an amphipathic helix (AH). Mutagenesis studies and membrane flotation experiments demonstrate that AH-like H2 is required for MAE-mediated membrane association. This MAE was functionally conserved across MERS-CoV, HCoV-OC43, HCoV-229E, HCoV-HKU1, and HCoV-NL63, all capable of mediating membrane association. In a SARS-CoV-2 replicon system, mutagenesis studies of H2 and replacements of H1 and H2 with their homologous counterparts demonstrated requirements of residues on both sides of the H2 and properly paired H1–H2 for MAE-mediated membrane association and viral replication. Notably, mutations I266A and K274A significantly attenuated viral replication without dramatically affecting membrane association, suggesting a dual role of the MAE in viral replication: mediating membrane association as well as participating in protein-protein interactions.

**IMPORTANCE** Severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2) assembles a double-membrane vesicle (DMV) by the viral non-structural proteins for viral replication. Understanding the mechanisms of the DMV assembly is of paramount importance for antiviral development. Nsp6, a multiple-spanning transmembrane protein, plays an important role in the DMV biogenesis. Herein, we predicted the nsp6 structure of SARS-CoV-2 and other human coronaviruses using AlphaFold2 and identified a putative membrane-associated element (MAE) in the highly conserved C-terminal regions of nsp6. Experimentally, we verified a functionally conserved minimal MAE composed of two α-helices, the H1, and the amphipathic helix-like H2. Mutagenesis studies confirmed the requirement of H2 for MAE-mediated membrane association and viral replication and demonstrated a dual role of the MAE in viral replication, by mediating membrane association and participating in residue-specific interactions. This functionally conserved MAE may serve as a novel anti-viral target.

**KEYWORDS**    coronavirus, SARS-CoV-2, nsp6, α-helix, replication

The coronavirus disease 2019 (COVID-19) pandemic caused by the severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2) has posed significant global health challenges (1–3). Vaccines, therapeutic antibodies, and antiviral drugs are considered the primary approaches for addressing the COVID-19 pandemic (4–6). Nevertheless, as the spike protein continues to undergo mutations, an increasing number of mutated strains, often referred to as variants of concern, are steadily emerging (7, 8). This presents a more formidable challenge to the protective efficacy of vaccines and therapeutic antibodies (9). A comprehensive understanding of the various stages of viral replication is of great significance for the design and development of antiviral drugs.

SARS-CoV-2 is a member of the *Coronaviridae* family (10). Currently, seven pathogenic human coronaviruses have been identified, including SARS-CoV-2, SARS-CoV, MERS-CoV, HCoV-HKU1, HCoV-NL63, HCoV-OC43, and HCoV-229E (11). Of these, SARS-CoV-2, SARS-CoV, and MERS-CoV are associated with more severe respiratory illnesses that can lead to multisystem organ failure and potentially fatal outcomes (12). In contrast, the other four coronaviruses are commonly associated with milder respiratory symptoms, similar to the common cold (11, 13).

SARS-CoV-2 is a positive-sense, single-stranded RNA virus (14). Its genome is approximately 30 kilobases (kb) in length, bearing a 5' cap and a 3' polyadenylated tail (15). The viral genome acts as messenger RNA to translate the open reading frame ORF1a, producing the polyprotein pp1a. The ORF1ab is generated by a frameshift of the ORF1a to produce the polyprotein pp1ab (11). These polyproteins are processed by the viral proteases nsp3 and nsp5 into at least 16 non-structural proteins (16). These non-structural proteins play crucial roles in viral replication as well as in viral assembly (3, 17, 18). Similar to other single-stranded RNA viruses, the non-structural proteins of SARS-CoV-2 assemble the virus replicase on the endoplasmic reticulum (ER), inducing the formation of double-membrane vesicles (DMVs) (19–22). These DMVs are believed to be the sites for virus replication and transcription (23, 24). Overexpression of the viral membrane proteins nsp3 and nsp4 is sufficient to induce the formation of coronavirus DMVs in cells (25, 26). Recently, the nsp3 of coronaviruses was found to assemble a molecular pore spanning the DMVs. This molecular pore is believed to play a critical role in DMV biogenesis and viral RNA export from the DMV (27). Very recently, Chlanda et al. demonstrated that the SARS-CoV-2 nsp3 and nsp4 are the minimal components required for assembling the molecular pore (28).

Coronavirus nsp6 is a multiple-spanning transmembrane protein (29, 30). The role of nsp6 in the formation of DMV remains controversial. In overexpression systems, it has been reported that nsp6 is either required (25) or dispensable for DMV formation (26, 31). A recent study demonstrates that nsp6 of SARS-CoV-2 regulates and organizes the DMV clusters, probably by zippering the ER lumen through homologous oligomerization (31). Natural mutation of nsp6 in the Omicron variant BA.1 attenuates viral replication, highlighting a critical role of nsp6 in viral replication (32, 33). In this study, we identified a minimal membrane-associated element (MAE) in the conserved C-terminal region downstream of the nsp6 transmembrane helices and demonstrated the critical role of the MAE in viral replication.

## RESULTS

### Prediction of the SARS-CoV-2 nsp6 structure

The structure of coronavirus nsp6 has not been resolved. We used AlphaFold2 (34) to predict the structure of SARS-CoV-2 nsp6. The predicted structure shows that it has eight transmembrane helices (TMs) (in black), a short N-terminal region, and a highly folded C-terminal domain (CTD). The CTD predominantly consists of secondary structures, including two β-sheets and two α-helices, namely helix 1 (H1, 242–257aa) and helix 2 (H2, 264–276aa) (highlighted in orange) (Fig. 1A). H2 is laterally positioned near the cytosolic surface of the ER membrane and exhibits characteristics of an amphipathic
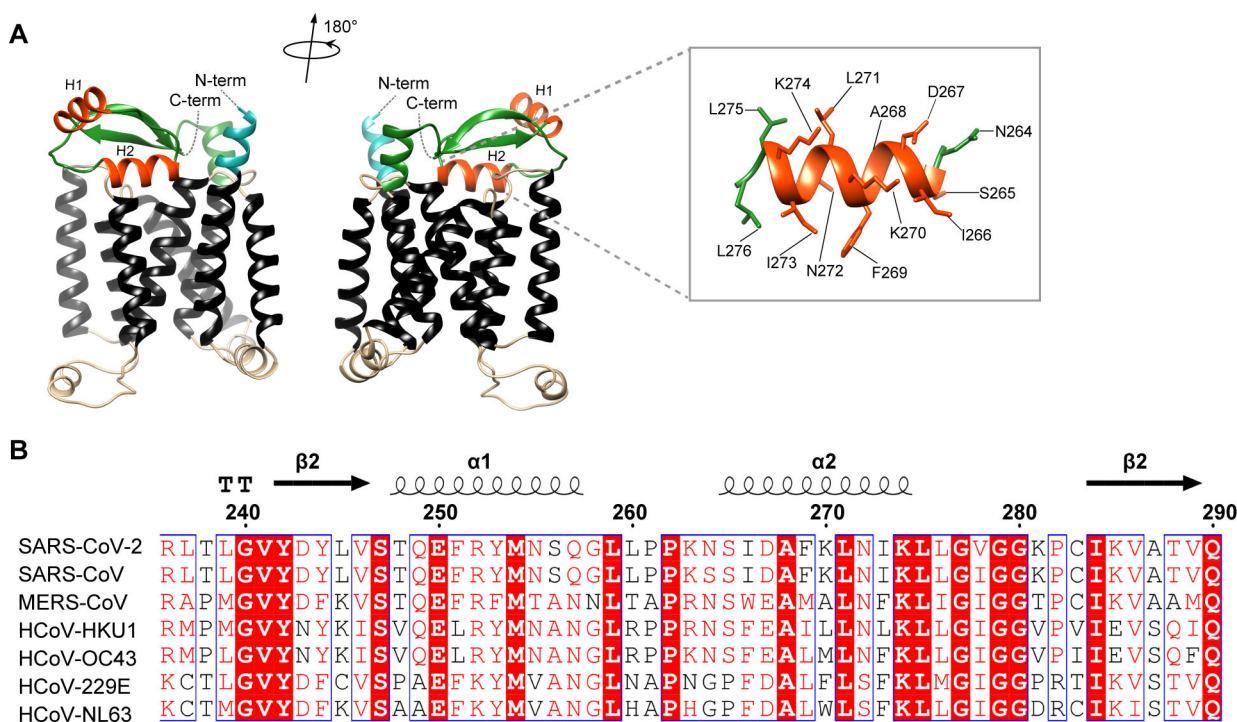
FIG 1 The predicted structure of SARS-CoV-2 nsp6 and alignment of the C-terminal regions of the human coronavirus nsp6. (A) The SARS-CoV-2 nsp6 structure predicted by AlphaFold2. The lumenal loop (beige), N-terminal domain (blue), CTD (green), and TMs (black) are shown. Two α-helices in the CTD (helix 1 and helix 2) are highlighted in orange. The enlarged region (boxed) shows the side chain structure of helix 2 (residues 264–276) in nsp6C. (B) Alignment of the C-terminal regions of the nsp6 of human coronaviruses. The amino acids that are conserved are highlighted in red.

helix (AH), with the hydrophobic residues (I266, F269, I273, L276) facing the ER membrane and the hydrophilic residues (N264, D267, K274) on the opposite side (Fig. 1A). We speculated that H2 might interact with the membrane. The α-helix of TM8 is extended to residue F235 (Fig. 1A). In order to verify if H2 in the CTD mediates membrane association and to exclude the potential effect of TM8, we studied the C-terminal region of nsp6 (236–290aa), which we designate as nsp6C. This nsp6C region is conserved among the human coronaviruses (Fig. 1B) (35).

## Expression and distribution of EGFP-fused nsp6 variants

We in-frame fused the coding sequences of the SARS-CoV-2 nsp6, nsp6[1–235] (1–235aa), and nsp6C (236–290aa) to the C-terminus of EGFP (Fig. 2A). The proteins were transiently expressed in HEK293T cells, which were harvested using an SDS-containing buffer for western blotting. We detected the expression of the EGFP-nsp6C, but failed to detect the EGFP-nsp6 and the EGFP-nsp6[1–235] in the boiled samples (Fig. 2B). We speculated that the boiling condition might have denatured the nsp6 thoroughly, leading to its aggregation mediated by the hydrophobic TMs. Referring to a reported protocol (31), we optimized the protein sampling conditions. The cells were harvested with SDS-containing buffer as previously described, and the samples were analyzed without boiling. Under these conditions, we successfully detected the expression of EGFP-nsp6 and EGFP-nsp6[1–235] with their expected molecular weights. In addition, we detected high-molecular-weight species, which were equivalent to the expected dimers of the EGFP-nsp6 and the EGFP-nsp6[1–235] (Fig. 2B), suggesting that these putative dimers might be stable in SDS-loading buffer.

We then employed confocal microscopy to examine the intracellular distribution of the EGFP-fused nsp6 variants in HEK293T cells. In contrast to EGFP, the EGFP-nsp6 fusion exhibited tubular structures (Fig. 2C), which is similar to previous studies when nsp6 was fused with other fluorescent proteins and tags (31). The EGFP-nsp6[1–235] fusion mainly
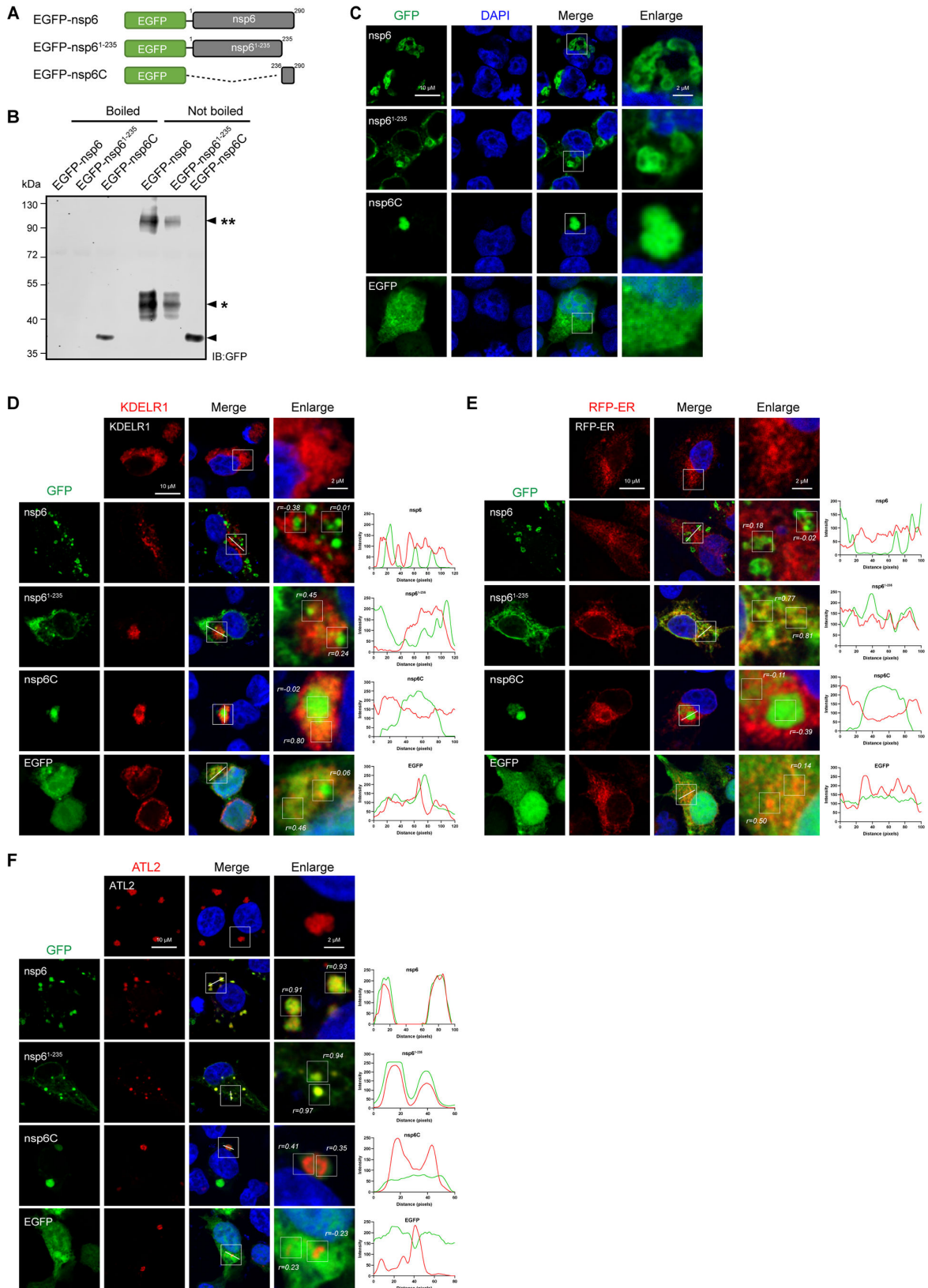
**FIG 2** Evaluating the expression and localization of EGFP-nsp6 fusion proteins. (A) Schematic diagram of the expression constructs. The nsp6 (1–290aa), nsp6$^{1-235}$(1–235aa), and nsp6C (236–290aa) regions were fused to the C-terminus of EGFP. (B) The expression of EGFP-nsp6C variants. HEK293T cells were transfected with EGFP-nsp6, EGFP-nsp6$^{1-235}$, and EGFP-nsp6C, and then harvested at 36 hours post-transfection. The cells were lysed by SDS as described in the Materials and

(Continued on next page)

**FIG 2** (Continued)

Methods. Protein samples were either boiled (Boiled) or not boiled (Not boiled) and loaded onto a 10% SDS-PAGE gel and then subjected to western blotting. The lowermost arrow corresponds to the EGFP-nsp6C. The arrow marked with an asterisk represents the EGFP-nsp6 and EGFP-nsp6$^{1-235}$. The arrow marked with two asterisks indicates the putative dimers of EGFP-nsp6 and EGFP-nsp6$^{1-235}$. (C) The distribution of the EGFP-fused nsp6 variants. The plasmids of EGFP-fused nsp6 variants were transfected into HEK293T cells. After 36 hours, cells were fixed and observed by confocal microscopy. Nuclei were visualized by DAPI staining (blue); the scale bars represent 10 µm. The boxed regions were acquired at higher magnification (Enlarge); the scale bars represent 2 µm. (D–F) Co-localization of EGFP-nsp6 variants with ER proteins. The EGFP-fused nsp6 variants were co-transfected with KDELR1-RFP (D), RFP-ER (E), or RFP-ATL2 (F) into HEK293T cells. After 36 hours, the cells were fixed and observed by confocal microscopy. Nuclei were visualized by DAPI (blue); the scale bar represents 10 µm. The boxed regions were acquired at higher magnification (Enlarge); the scale bars represent 2 µm. The fluorescence intensity profiles of EGFP or RFP along the indicated line were plotted. The Pearson correlation coefficients ($r$) for the boxed regions in the enlarged pictures were quantified and indicated.

exhibited reticular and tubular structures whereas the EGFP-nsp6C was observed to exclusively cluster in the cytoplasm (Fig. 2C). In a previous study, nsp6 was observed to co-localize with KDEL receptor 1 (KDELR1) and Atlastin2 (ATL2) on the ER (31). KDELR1, a receptor for the C-terminal sequence motif K-D-E-L, promotes the recycling of ER-resident proteins from the Golgi (36). ATL2, a member of the atlastin GTPase family, tethers membranes by forming trans-homo-oligomers and facilitates the homotypic fusion of ER membranes (37). To examine the co-localization of the EGFP-fused nsp6 variants with these ER proteins, we co-expressed EGFP-nsp6 variants with KDELR1-RFP, RFP-ATL2, and RFP containing an ER retention sequence (RFP-ER), respectively. EGFP-nsp6C and Nsp6$^{1-235}$ didn't exhibit obvious co-localization with KDELR1-RFP ($r < 0.5$). When co-expressed with KDELR1-RFP, the EGFP-nsp6C exhibited intensely clustered (IC) and less intensely clustered (LIC) patterns and the LIC nsp6C surrounded the IC nsp6C. The IC nsp6C didn't co-localize with KDELR1-RFP ($r = -0.02$), whereas the LIC nsp6C co-localized with KDELR1-RFP ($r = 0.80$) (Fig. 2D). When co-expressed with RFP-ER, nsp6 didn't co-localize with RFP-ER whereas nsp6$^{1-235}$ exhibited co-localization with RFP-ER ($r = 0.79 \pm 0.03$) (Fig. 2E). Neither the IC nsp6C ($r = -0.39$) nor the LIC nsp6C ($r = -0.11$) co-localized with ER-RFP. It was notable that, when co-expressed with RFP-ATL2, nsp6 and nsp6$^{1-235}$ mainly exhibited an IC pattern and co-localized with RFP-ATL2 (Fig. 2F), in contrast to the scenarios when expressed alone (Fig. 2C), indicating that expression of ATL2 may change the distributions of ER proteins. When co-expressed with RFP-ATL2, EGFP-nsp6C exhibited an IC pattern and didn't co-localize with RFP-ATL2 ($r = 0.38 \pm .04$) (Fig. 2F). Thus, EGFP-nsp6C exhibited a distinct distribution pattern compared to nsp6 and nsp6$^{1-235}$, the latter of which lacks the nsp6C region (see Discussion).

## Identification of a minimal membrane-associated element in the SARS-CoV-2 nsp6C

We speculated that the clustering of the EGFP-nsp6C is attributed to its membrane association. We aimed to verify the membrane association of nsp6C and identify the minimal essential region for membrane association. First, we progressively truncated the nsp6C region from its N terminus and C terminus (Fig. 3A) and then employed confocal microscopy to examine the intracellular distribution of the EGFP-fused nsp6C variants. Meanwhile, we examined the membrane association of these variants by subcellular fractionation. The cells were disrupted using the hypotonic buffer and the postnuclear cell lysates were treated with 0.5 M NaCl and then fractionated into the crude membrane fraction (P) and the cytosolic fraction (S). In contrast to EGFP, which is mainly distributed in the cytosolic fraction along with the cytosolic protein HSP70, EGFP-nsp6C largely distributed in the membrane fraction (greater than 90%), as the ER-resident protein Calnexin did (Fig. 3F through K).

Similar to the EGFP-nsp6C, the C-terminally truncated variant EGFP-nsp6C$^{236-279}$ exhibited IC and LIC structures. The IC structures didn't co-localize with KDELR1-RFP ($r = -0.51$) whereas the LIC structures co-localized with KDELR1-RFP ($r = 0.81$) (Fig. 2D and 3B). The cellular fractionation experiment showed that the EGFP-nsp6C$^{236-279}$ largely distributed in the membrane fraction (greater than 90%) (Fig. 3G and K). In contrast, the C-terminally truncated variants EGFP-nsp6C$^{236-268}$, EGFP-nsp6C$^{236-257}$, and EGFP-nsp6C$^{236-}$
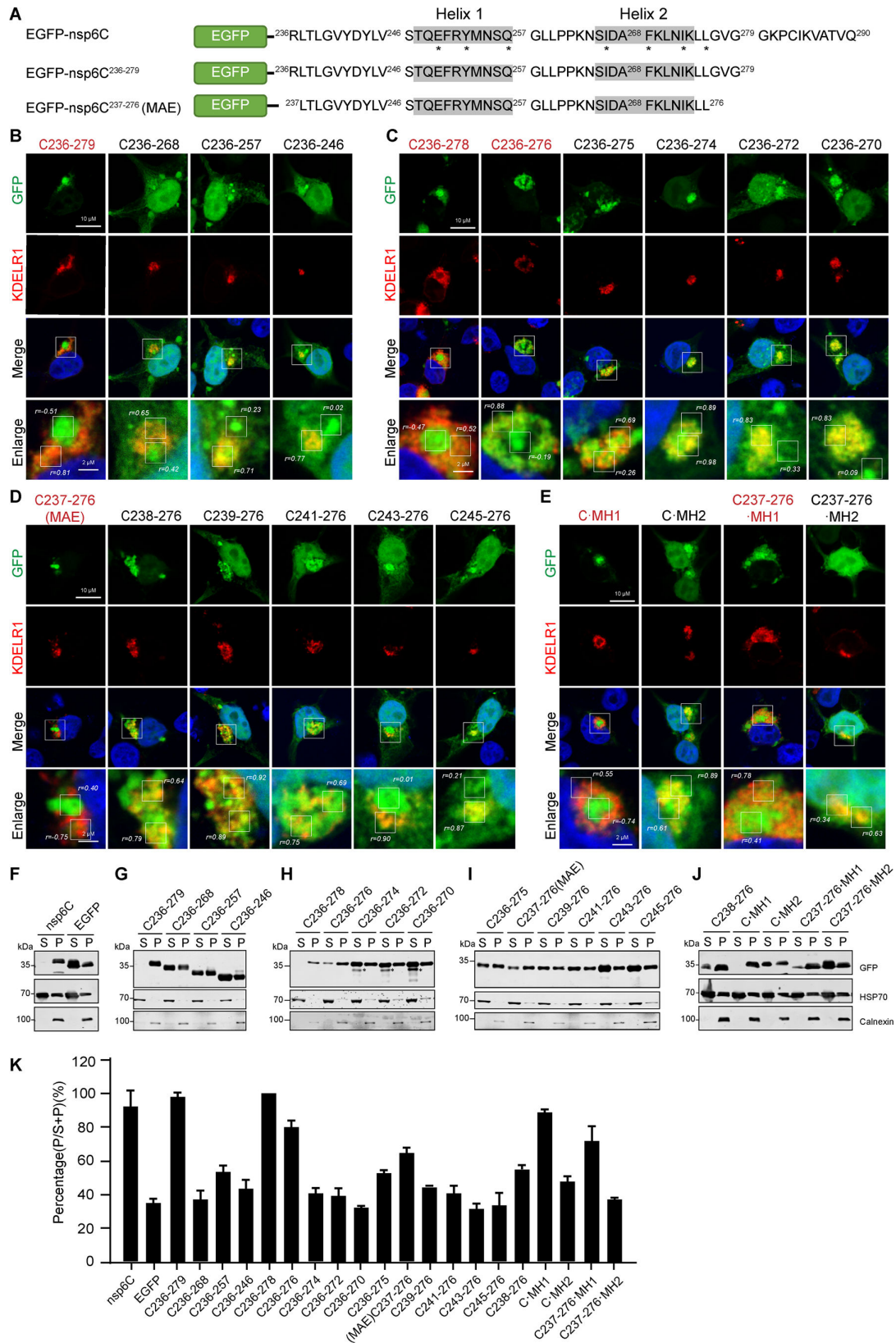
**FIG 3** Evaluating the membrane versus cytosolic distribution of EGFP-nsp6C variants. (A) Schematic diagram of EGFP-nsp6C truncations and mutations. The SARS-CoV-2 nsp6C variants were fused to the C-terminus of EGFP. The nsp6C was truncated from its N-terminus and C-terminus. Representative variants are shown. The amino acids of H1 and H2 are highlighted in gray. The asterisks indicate alanine substitutions at the indicated amino acids. (B–E) The distribution

(Continued on next page)

**FIG 3** (Continued)

pattern of the EGFP-nsp6C variants. HEK293T cells were co-transfected with the plasmids for nsp6C variants and KDELR1-RFP. After 36 hours, the cells were fixed and observed using confocal microscopy. Nuclei were visualized by DAPI staining (blue). MH1 indicates the nsp6 variants bearing the substitutions (E250A, Y253A, and Q257A) in H1. MH2 indicates the nsp6C variants bearing the substitutions (I266A, F269A, I273A, and L276A) in H2; the scale bars represent 10 μM. Images in the boxed regions were acquired at higher magnification (Enlarge); the scale bars represent 2 μm. The Pearson correlation coefficients (*r*) for the boxed regions in the enlarged pictures were quantified and indicated. (F–K) Membrane distribution of the nsp6C variants. HEK293T cells were transfected with the EGFP-nsp6C variants. Cells were disrupted by hypotonic buffer and postnuclear lysates were treated with 0.5 M NaCl. The postnuclear lysates were then separated by centrifugation into the crude membrane fraction (pellet, P) and the cytosolic fraction (supernatant, S). (F–J) The fractions were analyzed with the indicated antibodies by western blotting. The asterisk indicates unclarified bands. (K) Protein abundance was quantified, and the proportion of proteins in the pellets (P) and the supernatants (S) was calculated and plotted. Mean values ± SDs are shown (*n* = 3). Data represent the combined results from three independent experiments.

$^{246}$ were more diffuse in the cell, although minor IC and LIC structures were still present. Cellular fractionation experiments showed that these variants were mainly distributed in the cytosolic fractions (greater than 50%) (Fig. 3G and K). These experiments mapped the MAE within the region from amino acids 236 to 279.

We then further truncated the nsp6C$^{236–279}$, creating a series of variants with additional truncations from the C-terminus or the N-terminus. The C-terminally truncated variants EGFP-nsp6C$^{236–278}$ and EGFP-nsp6C$^{236–276}$ exhibited similar structures as EGFP-nsp6C$^{236–279}$ (Fig. 3C) and were largely distributed in the membrane fractions (greater than 80%) (Fig. 3H and K). In contrast, the variants EGFP-nsp6C$^{236–275}$, EGFP-nsp6C$^{236–274}$, EGFP-nsp6C$^{236–272}$, and EGFP-nsp6C$^{236–270}$ were mainly diffuse (Fig. 3C) and distributed in the cytosolic fractions (greater than 50%) (Fig. 3H and K). Thus, these experiments mapped the MAE within the region from amino acids 236 to 276aa.

Then we truncated the nsp6C$^{236–276}$ from its N-terminus. The N-terminally truncated variant EGFP-nsp6C$^{237–276}$ exhibited a distribution pattern similar to that of EGFP-nsp6C$^{236–276}$, with IC structures that did not co-localize with KDELR1-RFP (*r* < 0.5) and LIC structures that co-localized with KDELR1-RFP (*r* > 0.5) (Fig. 3D). Cellular fractionation experiments showed that EGFP-nsp6C$^{237–276}$ was mainly found in the membrane fraction, comprising over 60% of its distribution (Fig. 3I and K). In contrast, the variants EGFP-nsp6C$^{239–276}$, EGFP-nsp6C$^{241–276}$, EGFP-nsp6C$^{243–276}$, and EGFP-nsp6C$^{245–276}$ displayed IC and diffused distribution patterns (Fig. 3D) and were mainly located in the cytosolic fractions, comprising over 50% of its distribution (Fig. 3I and K). We further truncated the N-terminus to the residue T238. The EGFP-nsp6C$^{238–276}$ was found diffusely in the cytoplasm to some extent (Fig. 3D), and in contrast with the MAE (EGFP-nsp6C$^{237–276}$) (Fig. 3I through K), less than 60% of the EGFP-nsp6C$^{238–276}$ distributed in the membrane fraction. Thus, taking these data together, we unambiguously mapped the minimal MAE as the region from amino acids 237 to 276 (Fig. 3A).

## The AH-like helix in the MAE is essential for MAE-mediated membrane association

AHs mediate membrane association (38). To test if the AH-like H2 in the MAE is essential for MAE-mediated membrane association, we mutated the residues in the hydrophobic face of H2 to alanine to create mutant H2 (MH2: I266A, F269A, I273A, and L276A) in the context of nsp6C and nsp6C$^{237–276}$, respectively. We also mutated the residues in H1 that face the ER membrane to create mutant H1 (MH1: E250A, Y253A, and Q257A) in the context of nsp6C and nsp6C$^{237–276}$, respectively. The EGFP-nsp6C·MH1 (C·MH1) and the EGFP-nsp6C$^{237–276}$·MH1 both exhibited IC and LIC structures (Fig. 3E) and were mainly distributed in the membrane fractions (greater than 60%) (Fig. 3J and K). In contrast, the EGFP-nsp6C·MH2 (C·MH2) and the EGFP-nsp6C$^{237–276}$·MH2 were mainly diffuse in the cytoplasm (Fig. 3E) and distributed in the cytosolic fractions (greater than 60%) (Fig. 3J and K). These results indicate that the AH-like H2 in the MAE is essential for MAE-mediated membrane association.

## The MAE in the SARS-CoV-2 nsp6C is sufficient to direct membrane association of EGFP

To further validate the MAE-mediated membrane association, we performed membrane flotation assays with the EGFP-fused nsp6C and the MAE, as well as their variants that bear the H2 mutations (MH2: I266A, F269A, I273A, and L276A). The EGFP-nsp6C, EGFP-nsp6C$^{236-276}$ (MAE), EGFP-nsp6C·MH2, and EGFP-nsp6C$^{237-276}$ (MAE)·MH2 proteins were expressed in HEK293T cells. The cells were disrupted using the hypotonic buffer and the postnuclear lysates were subjected to membrane flotation assays as described in the Materials and Methods (Fig. 4A). The protein distribution in the flotation gradients was examined by western blotting. Less than 5% ± 7% of the EGFP was found distributed in the float fractions (fractions 1, 2, and 3) (Fig. 4B). The presence of the residual EGFP in the float fractions is probably due to the high abundance of the EGFP. More than 90% of the EGFP-nsp6C distributed in the float fractions, indicating efficient membrane association (Fig. 4B). In contrast, the H2 mutation dramatically reduced the membrane association of the EGFP-nsp6C·MH2, indicating that the H2 is essential for the membrane association of nsp6C. About 42% of the EGFP-nsp6C$^{237-276}$ (MAE) was found to distribute in the float fractions. In contrast, the H2 mutation nearly abolished the membrane association of EGFP-nsp6C$^{237-276}$ (MAE)·MH2. These data indicate that the nsp6C is membrane-associated, that the minimal MAE is sufficient to mediate membrane association, albeit less efficiently than the nsp6C (see Discussion), and that H2 is essential for MAE-mediated membrane association.

## Membrane association of the EGFP-fused MAEs of nsp6 from other human coronaviruses

Given that the nsp6C region is conserved across the human coronaviruses (Fig. 1B), we reasoned that the MAEs of other human coronaviruses have similar function as the MAE of SARS-CoV-2, and are likely capable of mediating membrane association. First, by using AlphaFold2, we predicted the nsp6 structures of SARS-CoV, MERS-CoV, HCoV-HKU1, HCoV-OC43, HCoV-229E, and HCoV-NL63. Overall, the nsp6s of these human coronaviruses adopt similar folds as the SARS-CoV-2 nsp6 does. Secondary protein structures were also found in the nsp6 CTD regions of these viruses, including potential MAEs (highlighted in orange) (Fig. 5A), albeit with sequence heterogeneity (Fig. 5B).
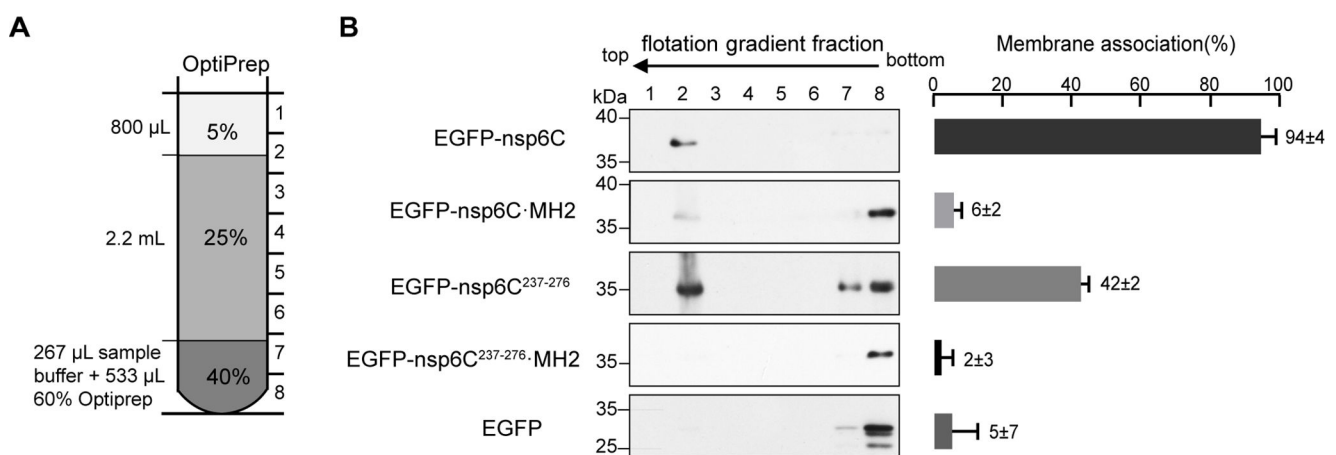


FIG 4  A minimal MAE in the SARS-CoV-2 nsp6C is sufficient to direct membrane association of EGFP. (A) The schematic diagram of the membrane flotation assay. The postnuclear cell lysates were mixed with 60% OptiPrep to make a 40% OptiPrep solution, which was then sequentially overlaid with 25% and 5% OptiPrep. After centrifugation, eight fractions were collected and subjected to western blotting. (B) Analyses of EGFP-nsp6C variants by membrane flotation. The EGFP-nsp6C, EGFP-nsp6C·MH2, EGFP-nsp6C$^{237-276}$, EGFP-nsp6C$^{237-276}$·MH2, and EGFP were expressed in HEK293T cells. Cell lysates were subjected to membrane flotation and the fractions were analyzed by western blotting using an anti-GFP antibody. The amount of protein in each fraction was quantified. The membrane association efficiency was calculated by comparing the abundance of protein in the float fractions (fractions 1, 2, and 3) to those in the total fractions, and is plotted to the right. Representative blots from two independent experiments are shown. Mean values ± SDs are shown (n = 2).
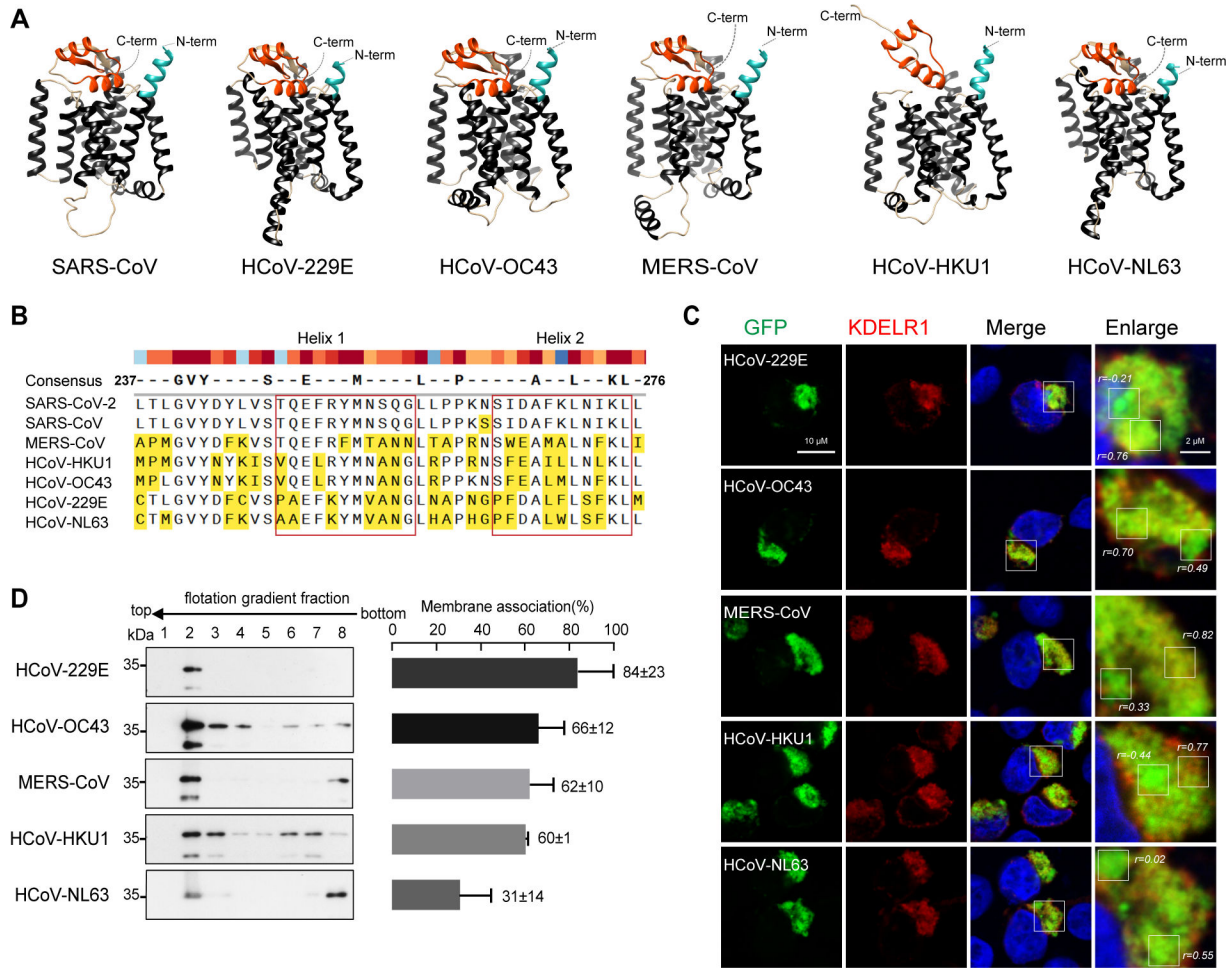
FIG 5 The MAEs of nsp6 from other human coronaviruses are capable of mediating membrane association. (A) The structures of nsp6 from the indicated human coronaviruses were predicted by AlphaFold2. The MAEs are highlighted in orange. (B) Multiple sequence alignment of the MAEs by ClustalW. The conserved residues are depicted at the top. The non-conserved amino acids are highlighted in yellow. H1 and H2 are highlighted with red boxes. (C) Distribution of EGFP-fused MAEs. The plasmids of the EGFP-MAEs and KDELR1-RFP were co-transfected into HEK293T cells. After 36 hours, the cells were fixed and observed by confocal microscopy. The nuclei were stained with DAPI (blue); the scale bar represents 10 µM. Images in the boxed regions were acquired at higher magnification (Enlarge); the scale bars represent 2 µm. The Pearson correlation coefficients (*r*) were analyzed as in Fig. 2. (D) Analyses of EGFP-MAEs by membrane flotation. The EGFP-MAEs were expressed in HEK293T cells and subjected to membrane flotation, and membrane association efficiency was calculated as described in Fig. 4. Representative blots from two independent experiments are shown. Mean values ± SDs are shown (*n* = 2).

To test whether these potential MAEs mediate membrane association, we fused the nsp6C regions that are equivalent to the minimal MAE of SARS-CoV-2 (237–276aa) to the C-terminus of EGFP (Fig. 5B). We omitted the MAE of SARS-CoV as it shares almost the same amino acid sequence as SARS-CoV-2, with only one amino acid difference (Fig. 5B). The EGFP-fused MAEs of MERS-CoV, HCoV-HKU1, HCoV-OC43, HCoV-229E, and HCoV-NL63 were co-expressed with KDELR1-RFP in HEK293T cells and the cells were observed by confocal microscopy. The EGFP-fused MAEs of these human coronaviruses exhibited two structural patterns similar to those observed in the EGFP-fused MAE of SARS-CoV-2. Specifically, the IC structures showed no co-localization with KDELR1 (*r* < 0.5), whereas the LIC patterns surrounding the IC patterns exhibited co-localization with KDELR1 (*r* > 0.5) (Fig. 2C and 5C). Then we performed membrane flotation assay for these EGFP-fused MAEs as described in Fig. 4. Compared with the EGFP-MAE of SARS-CoV-2 (42% float) (Fig. 4B), greater than 80% of the EGFP-HCoV-229E MAE and 60% of the EGFP-MAEs of HCoV-43, MERS-CoV, and HCoV-HKU1 distributed in the float fractions. To a lesser extent (about 30%) of the EGFP-HCoV-NL63 distributed in the float fractions

(Fig. 5D). Taken together, these data suggest that the MAEs of nsp6s from other human coronaviruses are functionally conserved to mediate membrane association.

## Mutagenesis of the SARS-CoV-2 MAE and the effects on viral replication

We sought to perform mutagenesis studies of the MAE and examine the effect on viral replication as well as the MAE-mediated membrane association. First, we used an *in vitro*-ligation SARS-CoV-2 replicon system (39), and introduced alanine substitutions in the H2 as shown in Fig. 6A. We substituted the hydrophobic residues (I266, F269, I273) facing the ER membrane (Fig. 6B) individually or in combination (MH2: I266A, F269A, I273A, L276A). We also substituted the residues (D267, L271, K274) on the opposite side of the H2 (Fig. 6B). These mutations were first introduced into a plasmid containing the nsp6 region and then the mutated fragments were ligated with other replicon fragments. After *in vitro* transcription, the replicon RNAs were co-transfected with N mRNA as reported (39). Viral replication was monitored by measuring the expression of luciferase activity from the replicon reporter gene. The viral replication capacity of each replicon RNA was calculated by the values of the luciferase activity at each time point post-transfection, normalized to the values obtained at 4 hours post-transfection. Mutation of these residues resulted in variable effects on viral replication. The D267A mutation (green oval, Fig. 6B) resulted in a slight reduction of virus replication, with approximately fivefold reduction compared to the wild-type. The I266A mutation (blue oval, Fig. 6B) highly attenuated viral replication, with ~60-fold reduction. The F269A mutation (lower right red oval, Fig. 6B) resulted in a ~300-fold reduction of viral replication. The L271A, I273A, K274A (red ovals, Fig. 6B), and the combined mutation MH2 (I266A, F269A, I273A, L276A) completely abrogated replication, resulting in an approximately 1,000-fold reduction in luciferase measurements (Fig. 6B). These data indicate that both the residues (I273 and F269) toward the membrane face and the residues (K274 and L271) on the opposite face of the AH-like helix H2 are essential for viral replication (see Discussion).

Given that the amino acids (K274A, L271A) on the cytosolic face of the H2 also completely abolished viral replication (Fig. 1, 6A and B), we speculated that the residues on H2 probably also participate in protein-protein interactions other than solely mediating membrane association. We then replaced the SARS-CoV-2 H1 and H2 with the homologous counterparts from the MERS-CoV and HCoV-OC43 in the SARS-CoV-2 replicon system (Fig. 6), expecting that the replacement will not change the MAE-mediated membrane association. Compared to the replication of the SARS-CoV-2 replicon at 48 hours post-transfection, replacement with the MERS-CoV H1 resulted in about a fourfold reduction in viral replication. Strikingly, replacement with HCoV-OC43 H1 completely abolished the viral replication. Similarly, replacement with MERS-CoV H2 or HCoV-OC43 H2 nearly abrogated viral replication to the level of the non-replicating SAA polymerase dead mutant (Fig. 6C). These data suggest that the residues in both the H1 and the AH-like H2 are required for viral replication (see Discussion).

## Mutagenesis of the SARS-CoV-2 MAE and the effects on MAE-mediated membrane association

We then investigated the intracellular distributions and membrane association of the EGFP-nsp6C variants that bear the H2 mutations and H1 and H2 replacements. The EGFP-nsp6C-I266A and EGFP-nsp6C-K274A mutants exhibited a similar distribution pattern as EGFP-nsp6C, exhibiting the IC structures that didn't co-localize with KDELR1-RFP ($r < 0.5$) and the LIC structures that co-localized with KDELR1-RFP ($r > 0.5$) (Fig. 6D). Accordingly, EGFP-nsp6C-I266A and EGFP-nsp6C-K274A exhibited similar membrane association efficacy as EGFP-nsp6C in the membrane flotation experiment (Fig. 6F), indicating that these mutations didn't change the membrane association and distribution of nsp6C. In contrast, EGFP-nsp6C-D267A, EGFP-nsp6C-F269A, EGFP-nsp6C-L271A, and EGFP-nsp6C-I273A mainly exhibited a LIC structure that co-localized with KDELR1-RFP ($r > 0.5$) (Fig. 6D). The membrane flotation experiment showed that these variants
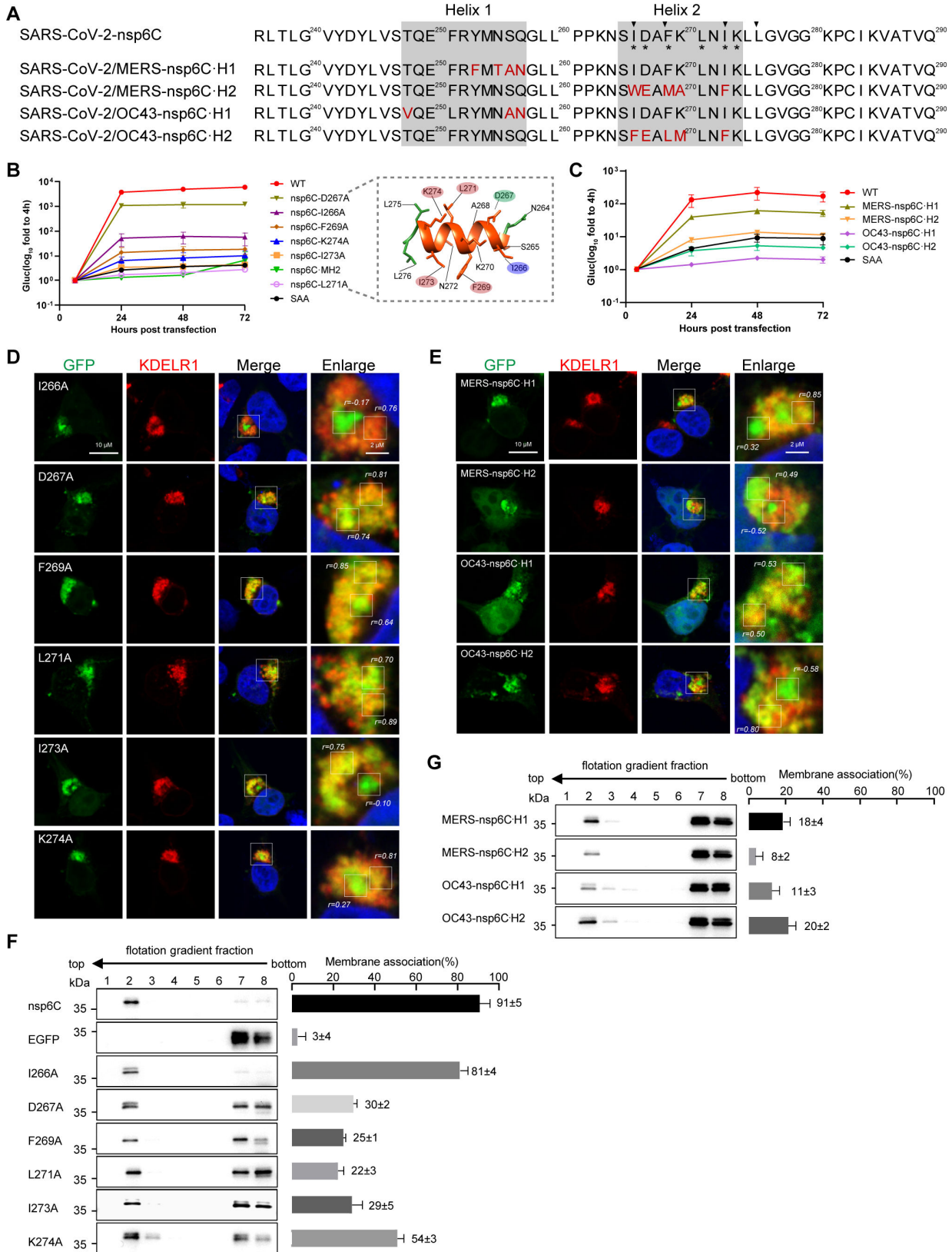
FIG 6 Mutagenesis of SARS-CoV-2 MAE and its effects on viral replication and nsp6C-mediated membrane association. (A) Schematic diagram of the MAE mutagenesis. The single alanine substitutions in I266, D267, F269, L271, I273, or K274 are marked with an asterisk. A four-alanine substitution (I266, F269, I273, and L276) is marked with arrowheads. The helix 1 and helix 2 in the MAE of SARS-CoV-2 nsp6 were replaced with their counterparts from MERS-CoV

(Continued on next page)

**FIG 6** (Continued)

and HCoV-OC43. The mutations were introduced in the plasmid pLC-nCoV-B-BsaI as described in the Materials and Methods. (B) Viral replication kinetics of SARS-CoV-2 replicons with H2 mutations. Huh7 cells were co-transfected with *in vitro*-transcribed SARS-CoV-2 replicon RNAs and an mRNA encoding the SARS-CoV-2 N protein. The medium was changed at 4 hours post-transfection. The luciferase activity in the supernatants was assessed at the indicated intervals. SAA, the nsp12 polymerase active-site mutant. The data were normalized to the values obtained 4 hours after transfection. Mean values ± SDs ($n = 3$) of triplicate wells are shown. Similar data were observed in another independent experiment. The green oval indicates a diminution of viral replication by less than 10-fold. The blue oval indicates a reduction of viral replication by 10–100-fold. The red ovals indicate a reduction of viral replication by over 100-fold. (C) The viral replication kinetics of SARS-CoV-2 replicons with H1 and H2 replacement. Huh7 cells were co-transfected with *in vitro*-transcribed SARS-CoV-2 replicon RNAs and an mRNA encoding the SARS-CoV-2 N protein. The medium was changed at 4 hours post-transfection. The luciferase activity in the supernatants was measured at the indicated intervals. SAA, the nsp12 polymerase active-site mutant. The data were normalized to the values obtained 4 hours post-transfection. Mean values ± SDs ($n = 3$) of triplicate wells are shown. Similar data were observed in the other two independent experiments. (D and E) Distribution of the EGFP-fused nsp6C variants. HEK293T cells were co-transfected with the plasmid for KDELR1-RFP along with the plasmids for the EGFP-fused nsp6C bearing H2 mutations (D) and H1 and H2 replacements (E). After 36 hours, the cells were fixed and observed by confocal microscopy. The nuclei were stained with DAPI (blue); the scale bars represent 10 µM. The boxed regions were acquired at higher magnification (Enlarge); the scale bars represent 2 µm. The Pearson correlation coefficients (*r*) for the boxed regions were quantified and are shown. (F and G) Membrane flotation of EGFP-nsp6C variants with H2 mutations (F) and H1 and H2 replacements (G). The EGFP-nsp6C variants were expressed in HEK293T cells and the cells were subjected to membrane flotation and membrane association efficiency was calculated as described in Fig. 4. Representative blots from two independent experiments are shown. Mean values ± SDs are shown ($n = 2$).

were less efficient at membrane association than EGFP-nsp6C (Fig. 6F). EGFP-nsp6C with the H1 replaced by the MERS-nsp6C.H1 and the H2 replacement by the HCoV-OC43-nsp6C.H2 mainly exhibited LIC structures that co-localized with KDELR1-RFP (*r* > 0.5) (Fig. 6E) and were less efficient at membrane association (Fig. 6G). Strikingly, EGFP-nsp6C with the H2 replaced by the MERS-nsp6C.H2 and the H1 replaced by the HCoV-OC43-nps6C.H1 were diffused in the cytoplasm to some extent (Fig. 6E) and showed a dramatically reduced membrane association (<20%) (Fig. 6G). Taking these data together, residues in H1 and residues on both sides of the AH-like H2 affect the membrane association of nsp6C as well as the viral replication (see Discussion).

## DISCUSSION

The C-terminal region of nsp6 is conserved among the coronaviruses, while its function is unknown. Based on the AlphaFold2-predicted structures of nsp6, we experimentally identified an MAE in this region. The minimal MAE was mapped to the region from amino acids 237 to 276, which contains two helices, H1 and H2. The H2 possesses the characteristics of an AH (Fig. 1C). Mutating the residues on the hydrophobic face abolished MAE-mediated membrane association (Fig. 4), confirming that this AH-like helix H2 is essential for MAE-mediated membrane association. It is noteworthy that, judged by the membrane float efficiency, the MAE-mediated membrane association (42%) was less efficient than the nsp6C-mediated membrane association (94%) (Fig. 4B), suggesting that residues outside the MAE in the nsp6C region also contribute to membrane association. In addition, the EGFP-nsp6C$^{245–276}$ that contains the H1 and H2 primarily distributed in the cytosolic fraction (Fig. 3D and I), suggesting that H1 and H2 alone are not competent for membrane association and that the residues from 237 to 244 in the MAE are required for MAE function, probably by properly folding of the H1 and H2.

It is noteworthy that residues on the cytosolic side of the AH-like H2 also contribute to membrane association, as the L271A mutation dramatically reduced the membrane association of nsp6C (Fig. 6F). Although mutating the H1 residues on the ER side had no effect on MAE-mediated membrane association (Fig. 3E and J), replacement of the H1 with its counterparts from MERS-CoV and HCoV-OC43 dramatically reduced the membrane association of nsp6C (Fig. 6G), indicating a requirement of a species-specific H1 and H2 pair for MAE-mediated membrane association. There might be interactions between the H1 and H2 that are essential for membrane association.

It has been reported that the C-terminal 80 amino acids (C80, 211–290aa) of SARS-CoV-2 nsp6 co-localizes with lipid droplets (LD) and nsp6 mediates the interaction between LDs and replication-organelle-like structures (31). The C80 region (211–290aa) contains the predicted TM8 (208–235aa, Fig. 1A). Here, we didn't observe the

co-localization of nsp6C (236–290aa) with LD (data not shown). This may imply that nsp6C alone is insufficient for co-localization with LD. Expression of EGFP-nsp6 exhibited clustered and tubular structures whereas EGFP-nsp6C was found to exclusively cluster in the cytoplasm (Fig. 2 through F). It has been reported that nsp6 co-localizes with two ER proteins KDELR1 and ATL2 (31). In this study, we didn't observe obvious co-localization of nsp6 with KDELR1-RFP and RFP-ER. We found less intensively clustered EGFP-nsp6C and EGFP-MAE (C237–276) structures when they were co-expressed with RFP-KDELR1 and ER-RFP. These LIC EGFP-nsp6C co-localized with KDELR1-RFP but not with ER-RFP. Notably, the IC EGFP-nsp6C and EGFP-MAE were surrounded by KDELR1-RFP and ER-RFP (Fig. 2E and 3D). According to the results of the cellular fractionation and membrane flotation experiments, the majority of the EGFP-nsp6C (>90%) was distributed in the membrane fractions. These data demonstrate that the IC nsp6C indicates a novel membrane-associated structure that is probably connected to the ER membrane, especially KDELR1-associated structures. We speculate that the MAE may have the capacity to bind to specific lipids or clustered host factors, and in the absence of the transmembrane helices, it mediates nsp6C clustering. Further studies are needed to confirm this hypothesis. EGFP-nsp6 and its truncated mutant EGFP-nsp6$^{1-235}$ that lacks the nsp6C exhibited distinct structures when co-expressed with KDELR1-RFP and ER-RFP (Fig. 2D and E), indicating a role of the nsp6C in nsp6 distribution.

Mutagenesis analysis demonstrated a correlation between MAE-mediated membrane association and viral replication for the nsp6C mutants F269A, L271A, I273A, and the H1 and H2 replacement mutants. These mutations reduced nsp6C-mediated membrane association and viral replication (Fig. 6). Notably, the nsp6C mutant D267A had a similar flotation efficiency as the I273A and F269A mutants, with reduced MAE-mediated membrane association, but didn't dramatically affect viral replication (Fig. 6B and F). These data indicate that the reduction of viral replication by the F269A and I273A mutations may not be solely attributed to the reduction of MAE-mediated membrane association. Accordingly, the I266A and K274A mutations, despite not dramatically affecting MAE-mediated membrane association, highly attenuated viral replication (Fig. 6B and F), suggesting an involvement of these residues in protein-protein interactions other than membrane association.

AHs are widely used by viral proteins of positive-stranded RNA viruses. Hepatitis C virus (HCV) NS5A contains an N-terminal amphipathic helix, which alone is capable of membrane association (40). Similarly, the replication protein 1a of Barley Mosaic Virus (BMV) has an amphipathic helix B that localizes 1a to the ER (41). The HCV NS5A AH has been suggested to mediate membrane anchoring and was recently demonstrated to play a dual role in viral replicase assembly, by properly positioning the viral polyprotein for processing and protein-protein interactions (42). The HCV NS5A AH also participates in NS5A dimerization (43). The function of the AH of BMV 1a can be separated into two categories: one is membrane association and induction of membrane invagination, and the other is recruitment of the viral polymerase (41, 44). The AH-like H2 in the MAE of nsp6 may also play a dual role in the viral replication; one is mediating membrane association, and the other is participating in interactions. The presence of a membrane association element downstream of a bundle of transmembrane helices may provide flexibility for membrane protein movements for precise protein-protein interactions or, on the contrary, attenuate protein mobility by engaging protein-protein and protein-lipid interactions.

Several studies used AlphaFold2 to predict the structure of nsp6 and identify drugs that target nsp6 through computer-aided drug screening and molecular docking simulations (45, 46). Interestingly, the results of molecular docking studies reveal that numerous compounds selected in the screening bind to sites located within the MAE of the nsp6, suggesting the potential of MAE as a novel anti-viral target.

In summary, we identified an MAE in the conserved nsp6C region of coronaviruses. This MAE is essential for viral replication, likely playing a dual role in membrane

association and protein-protein interaction. This novel functional viral element may serve as a target for anti-viral drugs.

## MATERIALS AND METHODS

### Plasmids

Full length and the variants of nsp6C were in-frame fused to the C-terminus of EGFP in the EGFP-C2 plasmid to construct the EGFP-fused plasmids. The sequences of nsp6 were synthesized for each of the five human coronaviruses by BGI (Wuxi, China) and cloned into the EGFP-C2 plasmid as described above. The gene sequences of KDELR1 and ATL2 were synthesized by BGI (Wuxi) and cloned into a homemade RFP plasmid to generate the KDELR1-RFP and RFP-ATL2 fusion plasmids. The RFP-ER plasmid was generated by fusing a signal sequence to the N-terminus of RFP and an ER retention sequence (KDEL) to the C-terminus of RFP.

Mutagenesis of nsp6 in the SARS-CoV-2 replicon system was performed by *in vitro* ligation. Mutations containing single or multiple amino acids at the H1 and H2 at the C-terminus of nsp6 were introduced into the plasmid pLC-nCoV-B-BsaI by fusion-PCR mediated mutagenesis. Subsequently, the fragments released by *BsaI* digestion were ligated with the fragments from the plasmids pLC-nCoV-A-BsaI, pLC-nCoV-C-BsaI, and pnCoV-D-sGluc-BsaI as previously described (39). Similarly, plasmids harboring substitutions of H1 and H2 at the C-terminus of nsp6 from MERS-CoV and HCoV-OC43 were introduced into the plasmid pLC-nCoV-B-BsaI and subjected to the *in vitro* ligation, as described below. All the plasmids were verified by Sanger sequencing and the details are available upon request.

### Cells and antibodies

The human embryonic kidney cell line HEK293T and the human hepatoma cell line Huh7 were purchased from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China, www.cellbank.org.cn) and routinely maintained in Dulbecco's modified medium supplemented with 10% FBS (Gibco) and 25 mM HEPES (Gibco), and non-essential amino acids (Gibco). The anti-GFP antibody (Santa Cruz; sc-9996) was used at 1:2,000 dilution. The anti-Hsp70 (Abclonal; A0284) antibody was used at 1:2,000 dilution. The anti-calnexin antibody (BD; 610523) was used at 1:2,000 dilution. The goat-anti-mouse HRP IgG (Santa Cruz; sc-2004) was used at 1:5,000. The goat-anti-mouse IRDye 800CW secondary antibody (licor; 926-32210) was used at 1:10,000 dilution. The goat-anti-rabbit IRDye 800CW secondary antibody (licor; 926-32211) was used at 1:10,000 dilution.

### Transfection

For plasmid transfection, HEK293T cells were seeded onto poly-L-lysine-coated 6-well plates at a density of $9.0 \times 10^5$ cells per well and then transfected with plasmids using a Lipofectamine 3000 transfection kit (Invitrogen; L3000008) according to the manufacturer's instructions. For RNA transfection, Huh7 cells were seeded onto 48-well plates at a density of $1.875 \times 10^5$ cells per well and then transfected with 0.3 µg *in vitro*-transcribed RNA using a TransIT-mRNA transfection kit (Mirus; MIR2250) according to the manufacturer's instructions.

### Fluorescence and confocal imaging

HEK293T cells were fixed with 4% paraformaldehyde solution in PBS for 10 minutes at room temperature and then washed three times with PBS. DAPI staining was done with a 1:1,000 dilution of 1 mg/mL DAPI solution (ThermoFisher; 62248), and incubated for 15 minutes at room temperature in the dark. The samples were then washed three times with PBS, and mounted with ProLong Glass Antifade Moutant medium (ThermoFisher; P36980). Confocal microscopy imaging was performed with a Leica TCS SP8 confocal laser microscope at an excitation of ~488 nm for GFP and ~594 nm for RFP. Images

were acquired with a 60× oil immersion objective, captured with the LAS software, and processed by ImageJ. The analysis of co-localization was conducted by calculating Pearson's correlation coefficients for the selected regions using the ImageJ plugin, as previously described (47). All immunofluorescence assays were performed at least three times.

## Western blotting immunofluorescence

After washing with PBS, cells were lysed with 2× SDS loading buffer [100 mM Tris-Cl (pH 6.8), 4% SDS, 0.2% bromophenol blue, 20% glycerol, and 10% 2-mercaptoethanol] and boiled for 5 minutes or not boiled. Cells expressing EGFP-nsp6 and EGFP-nsp6$^{1-235}$ were lysed with SDS lysis buffer. Lysates were aspirated 10 times with a 27-gauge needle to achieve a fluid consistency and were loaded without boiling. Proteins were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. The membranes were incubated for 1 hour in a blocking buffer (PBS, 5% milk, 0.05% Tween) and then incubated with the primary antibody, which was diluted in the blocking buffer. After three washes with PBST (PBS, 0.05% Tween), the membranes were incubated with a secondary antibody. After three washes with PBST, the membrane was visualized by the Western Lightning Plus-ECL substrate (PerkinElmer; NEL10500) or by the Odyssey CLx Imaging System. If necessary, the protein bands were quantified by densitometry using ImageJ.

## Cell fractionation

Cell fractionation was performed as described previously (42). In brief, HEK293T cells were seeded in 6-well plates and transfected with EGFP-fused proteins. After 36 hours, the cells were harvested by scraping into 500 µL of hypotonic buffer [5 mM Tris-Cl (pH 7.5), 15 mM KCl, 2.5 mM MgCl$_2$]. After 15 minutes of swelling on ice, the cells were passed 20 times through a 27-gauge needle and centrifuged (900 × $g$ for 5 minutes) to remove nuclei and cell debris. One-tenth volume of 5 M NaCl was added to the postnuclear supernatants followed by incubation on ice for 20 minutes, then the membranes were collected by centrifugation at 15,000 × $g$ for 20 minutes. The pellets (membrane fraction, P) were resuspended in a 30 µL SDS loading buffer. The supernatant (cytosol fraction, S) was concentrated by adding 4 volumes of methanol and centrifuged 12,000 × $g$ for 10 minutes. The protein was resuspended in a 30 µL SDS loading buffer.

## Membrane flotation assay

Membrane flotation assays were performed as described previously with modification (34). HEK293T cells were seeded into a 6 cm$^2$ dish and transfected with the EGFP-nsp6C and variant plasmids. After 48 hours, cells were scraped into 270 µL solution B [0.25 M sucrose, 1 mM, EDTA, Hepes-NaOH (pH 7.4)] and passed through a 27-gauge needle 20 times. After centrifugation at 900 × $g$ for 5 minutes, the supernatants were adjusted to 40% Optiprep (Sigma; D1556) by mixing 267 µL of the sample with 533 µL of 60% Optiprep. An amount of 800 µL of the mixture was placed at the bottom of a Beckman SW60 centrifuge tube and then sequentially overlaid with 1.4 mL of 25% Optiprep in Solution B and 800 µL of 5% Optiprep in Solution B. The gradients were centrifuged at 250,000 × $g$ at 4°C for 3 hours. A total of eight fractions were collected (500 µL/each) and the proteins were concentrated by adding 4 volumes of methanol. After centrifugation at 12,000 × $g$ for 10 minutes, the pellets were resuspended in 30 µL of SDS loading buffer for SDS-PAGE and western blotting analysis.

## *In vitro* ligation and transcription

The pLC-nCoV-B-BsaI plasmid was used as the backbone to introduce point mutations or substitutions in the nsp6C region as described previously (39). The *BsaI*-digested fragments were purified using the Gel Extraction Kit (OMEGA; D2500-1) and ligated using vT4 DNA Ligase (New England Biolabs; M0202L) with the fragments similarly released

from the plasmids pLCnCoV-A-BsaI, pLC-nCoV-C-BsaI, and pLC-nCoV-D-sGluc-BsaI at room temperature for 1 hour. The ligated products were purified through phenol-chloroform extraction and resuspended in nuclease-free water. The purified *in vitro* ligated products were used as templates for the *in vitro* transcription using the mMESSAGE mMACHINE T7 Transcription Kit (Invitrogen; AM1344) according to the manufacturer's protocol. The RNA was purified by RNeasy Mini Kit (QIAGEN; 74106) and then resuspended in nuclease-free water.

## Luciferase activity

Ten microliters of supernatants were mixed with an equal volume of 2× passive lysis buffer (Promega; E2820) prior to analysis of luciferase activity using the Renilla luciferase substrate (Promega; E2820) according to the manufacturer's protocol.

## Statistical analysis

Statistical analysis was conducted using the GraphPad Prism 8 software. The details of the tests are described in the figure legends.

## AUTHOR AFFILIATIONS

[1]Key Laboratory of Medical Molecular Virology (MOE/NHC/CAMS), School of Basic Medical Sciences, and Shanghai Institute of Infectious Disease and Biosecurity, Fudan University, Shanghai, China
[2]Shanghai Public Health Clinical Center, Fudan University, Shanghai, China

## AUTHOR ORCIDs

Yuying Han ⓘ http://orcid.org/0009-0006-1569-6794
Zhigang Yi ⓘ http://orcid.org/0000-0002-4560-4970

## DATA AVAILABILITY

The structure of nsp6 from SARS-CoV-2/human/CHN/SH01/2020 (GenBank: MT121215.1) was predicted by AlphaFold2 in Google coLab (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb).

The protein structures predicted by AlphaFold2 for nsp6 do not meet the Protein Data Bank (PDB) website's upload criteria and therefore lack accession numbers. Researchers may obtain the PDB files for nsp6 through structural prediction on the AlphaFold2 website as required.

## REFERENCES

1. Hu B, Guo H, Zhou P, Shi Z-L. 2021. Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol 19:141–154. https://doi.org/10.1038/s41579-020-00459-7

2. Harrison AG, Lin T, Wang P. 2020. Mechanisms of SARS-CoV-2 transmission and pathogenesis. Trends Immunol 41:1100–1115. https://doi.org/10.1016/j.it.2020.10.004

3. Malone B, Urakova N, Snijder EJ, Campbell EA. 2022. Structures and functions of coronavirus replication-transcription complexes and their relevance for SARS-CoV-2 drug design. Nat Rev Mol Cell Biol 23:21–39. https://doi.org/10.1038/s41580-021-00432-z

4. Kumari M, Lu R-M, Li M-C, Huang J-L, Hsu F-F, Ko S-H, Ke F-Y, Su S-C, Liang K-H, Yuan J-Y, Chiang H-L, Sun C-P, Lee IJ, Li W-S, Hsieh H-P, Tao M-H, Wu H-C. 2022. A critical overview of current progress for COVID-19: development of vaccines, antiviral drugs, and therapeutic antibodies. J Biomed Sci 29:68. https://doi.org/10.1186/s12929-022-00852-9

5. Adamson CS, Chibale K, Goss RJM, Jaspars M, Newman DJ, Dorrington RA. 2021. Antiviral drug discovery: preparing for the next pandemic. Chem Soc Rev 50:3647–3655. https://doi.org/10.1039/d0cs01118e

6. Robinson PC, Liew DFL, Tanner HL, Grainger JR, Dwek RA, Reisler RB, Steinman L, Feldmann M, Ho L-P, Hussell T, Moss P, Richards D, Zitzmann N. 2022. COVID-19 therapeutics: challenges and directions for the future. Proc Natl Acad Sci U S A 119:e2119893119. https://doi.org/10.1073/pnas.2119893119

7. Scovino AM, Dahab EC, Vieira GF, Freire-de-Lima L, Freire-de-Lima CG, Morrot A. 2022. SARS-CoV-2's variants of concern: a brief characterization. Front Immunol 13:834098. https://doi.org/10.3389/fimmu.2022.834098

8. Tian D, Sun Y, Xu H, Ye Q. 2022. The emergence and epidemic characteristics of the highly mutated SARS-CoV-2 Omicron variant. J Med Virol 94:2376–2383. https://doi.org/10.1002/jmv.27643

9. Liu M, Liang Z, Cheng ZJ, Liu L, Liu Q, Mai Y, Chen H, Lei B, Yu S, Chen H, Zheng P, Sun B. 2023. SARS-CoV-2 neutralising antibody therapies: recent advances and future challenges. Rev Med Virol 33:e2464. https://doi.org/10.1002/rmv.2464

10. Yang H, Rao Z. 2021. Structural biology of SARS-CoV-2 and implications for therapeutic development. Nat Rev Microbiol 19:685–700. https://doi.org/10.1038/s41579-021-00630-8

11. Hartenian E, Nandakumar D, Lari A, Ly M, Tucker JM, Glaunsinger BA. 2020. The molecular virology of coronaviruses. J Biol Chem 295:12910–12934. https://doi.org/10.1074/jbc.REV120.013930

12. Abdelrahman Z, Li M, Wang X. 2020. Comparative review of SARS-CoV-2, SARS-CoV, MERS-CoV, and influenza A respiratory viruses. Front Immunol 11:552909. https://doi.org/10.3389/fimmu.2020.552909

13. Gupta R, Charron J, Stenger CL, Painter J, Steward H, Cook TW, Faber W, Frisch A, Lind E, Bauss J, Li X, Sirpilla O, Soehnlen X, Underwood A, Hinds D, Morris M, Lamb N, Carcillo JA, Bupp C, Uhal BD, Rajasekaran S, Prokop JW. 2020. SARS-CoV-2 (COVID-19) structural and evolutionary dynamicome: insights into functional evolution and human genomics. J Biol Chem 295:11742–11753. https://doi.org/10.1074/jbc.RA120.014873

14. Bar-On YM, Flamholz A, Phillips R, Milo R. 2020. SARS-CoV-2 (COVID-19) by the numbers. Elife 9:e57309. https://doi.org/10.7554/eLife.57309

15. Kim D, Lee JY, Yang JS, Kim JW, Kim VN, Chang H. 2020. The architecture of SARS-CoV-2 transcriptome. Cell 181:914–921. https://doi.org/10.1016/j.cell.2020.04.011

16. Rohaim MA, El Naggar RF, Clayton E, Munir M. 2021. Structural and functional insights into non-structural proteins of coronaviruses. Microb Pathog 150:104641. https://doi.org/10.1016/j.micpath.2020.104641

17. Wang Q, Wu J, Wang H, Gao Y, Liu Q, Mu A, Ji W, Yan L, Zhu Y, Zhu C, et al. 2020. Structural basis for RNA replication by the SARS-CoV-2 polymerase. Cell 182:417–428. https://doi.org/10.1016/j.cell.2020.05.034

18. V'kovski P, Kratzel A, Steiner S, Stalder H, Thiel V. 2021. Coronavirus biology and replication: implications for SARS-CoV-2. Nat Rev Microbiol 19:155–170. https://doi.org/10.1038/s41579-020-00468-6

19. Cortese M, Lee J-Y, Cerikan B, Neufeldt CJ, Oorschot VMJ, Köhrer S, Hennies J, Schieber NL, Ronchi P, Mizzon G, et al. 2020. Integrative imaging reveals SARS-CoV-2-induced reshaping of subcellular morphologies. Cell Host Microbe 28:853–866. https://doi.org/10.1016/j.chom.2020.11.003

20. Gosert R, Kanjanahaluethai A, Egger D, Bienz K, Baker SC. 2002. RNA replication of mouse hepatitis virus takes place at double-membrane vesicles. J Virol 76:3697–3708. https://doi.org/10.1128/jvi.76.8.3697-3708.2002

21. Goldsmith CS, Tatti KM, Ksiazek TG, Rollin PE, Comer JA, Lee WW, Rota PA, Bankamp B, Bellini WJ, Zaki SR. 2004. Ultrastructural characterization of SARS coronavirus. Emerg Infect Dis 10:320–326. https://doi.org/10.3201/eid1002.030913

22. Knoops K, Kikkert M, Worm S van den, Zevenhoven-Dobbe JC, van der Meer Y, Koster AJ, Mommaas AM, Snijder EJ. 2008. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. PLoS Biol 6:e226. https://doi.org/10.1371/journal.pbio.0060226

23. Snijder EJ, Limpens R, de Wilde AH, de Jong AWM, Zevenhoven-Dobbe JC, Maier HJ, Faas F, Koster AJ, Bárcena M. 2020. A unifying structural and functional model of the coronavirus replication organelle: tracking down RNA synthesis. PLoS Biol 18:e3000715. https://doi.org/10.1371/journal.pbio.3000715

24. Wolff G, Melia CE, Snijder EJ, Bárcena M. 2020. Double-membrane vesicles as platforms for viral replication. Trends Microbiol 28:1022–1033. https://doi.org/10.1016/j.tim.2020.05.009

25. Angelini MM, Akhlaghpour M, Neuman BW, Buchmeier MJ. 2013. Severe acute respiratory syndrome coronavirus nonstructural proteins 3, 4, and 6 induce double-membrane vesicles. mBio 4:e00524-13. https://doi.org/10.1128/mBio.00524-13

26. Oudshoorn D, Rijs K, Limpens RWAL, Groen K, Koster AJ, Snijder EJ, Kikkert M, Bárcena M. 2017. Expression and cleavage of Middle East respiratory syndrome coronavirus nsp3-4 polyprotein induce the formation of double-membrane vesicles that mimic those associated with coronaviral RNA replication. mBio 8:e01658-17. https://doi.org/10.1128/mBio.01658-17

27. Wolff G, Limpens RWAL, Zevenhoven-Dobbe JC, Laugks U, Zheng S, de Jong AWM, Koning RI, Agard DA, Grünewald K, Koster AJ, Snijder EJ, Bárcena M. 2020. A molecular pore spans the double membrane of the coronavirus replication organelle. Science 369:1395–1398. https://doi.org/10.1126/science.abd3629

28. Zimmermann L, Zhao X, Makroczyova J, Wachsmuth-Melm M, Prasad V, Hensel Z, Bartenschlager R, Chlanda P. 2023. SARS-CoV-2 nsp3 and nsp4 are minimal constituents of a pore spanning replication organelle. Nat Commun 14:7894. https://doi.org/10.1038/s41467-023-43666-5

29. Baliji S, Cammer SA, Sobral B, Baker SC. 2009. Detection of nonstructural protein 6 in murine coronavirus-infected cells and analysis of the transmembrane topology by using bioinformatics and molecular approaches. J Virol 83:6957–6962. https://doi.org/10.1128/JVI.00254-09

30. Bills C, Xie X, Shi P-Y. 2023. The multiple roles of nsp6 in the molecular pathogenesis of SARS-CoV-2. Antiviral Res 213:105590. https://doi.org/10.1016/j.antiviral.2023.105590

31. Ricciardi S, Guarino AM, Giaquinto L, Polishchuk EV, Santoro M, Di Tullio G, Wilson C, Panariello F, Soares VC, Dias SSG, Santos JC, Souza TML, Fusco G, Viscardi M, Brandi S, Bozza PT, Polishchuk RS, Venditti R, De Matteis MA. 2022. The role of NSP6 in the biogenesis of the SARS-CoV-2 replication organelle. Nature 606:761–768. https://doi.org/10.1038/s41586-022-04835-6

32. Chen DY, Chin CV, Kenney D, Tavares AH, Khan N, Conway HL, Liu G, Choudhary MC, Gertje HP, O'Connell AK, Adams S, Kotton DN, Herrmann

A, Ensser A, Connor JH, Bosmann M, Li JZ, Gack MU, Baker SC, Kirchdoerfer RN, Kataria Y, Crossland NA, Douam F, Saeed M. 2023. Spike and nsp6 are key determinants of SARS-CoV-2 Omicron BA.1 attenuation. Nature 615:143–150. https://doi.org/10.1038/s41586-023-05697-2

33. Taha TY, Chen IP, Hayashi JM, Tabata T, Walcott K, Kimmerly GR, Syed AM, Ciling A, Suryawanshi RK, Martin HS, Bach BH, Tsou CL, Montano M, Khalid MM, Sreekumar BK, Renuka Kumar G, Wyman S, Doudna JA, Ott M. 2023. Rapid assembly of SARS-CoV-2 genomes reveals attenuation of the Omicron BA.1 variant through NSP6. Nat Commun 14:2308. https://doi.org/10.1038/s41467-023-37787-0

34. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction with alphafold. Nature 596:583–589. https://doi.org/10.1038/s41586-021-03819-2

35. Feng S, O'Brien A, Chen D-Y, Saeed M, Baker SC. 2023. SARS-CoV-2 nonstructural protein 6 from alpha to Omicron: evolution of a transmembrane protein. mBio 14:e0068823. https://doi.org/10.1128/mbio.00688-23

36. Jia J, Yue X, Zhu L, Jing S, Wang Y, Gim B, Qian Y, Lee I. 2021. KDEL receptor is a cell surface receptor that cycles between the plasma membrane and the golgi via clathrin-mediated transport carriers. Cell Mol Life Sci 78:1085–1100. https://doi.org/10.1007/s00018-020-03570-3

37. Niu L, Ma T, Yang F, Yan B, Tang X, Yin H, Wu Q, Huang Y, Yao Z-P, Wang J, Guo Y, Hu J. 2019. Atlastin-mediated membrane tethering is critical for cargo mobility and exit from the endoplasmic reticulum. Proc Natl Acad Sci U S A 116:14029–14038. https://doi.org/10.1073/pnas.1908409116

38. Giménez-Andrés M, Čopič A, Antonny B. 2018. The many faces of amphipathic helices. Biomolecules 8:45. https://doi.org/10.3390/biom8030045

39. Zhang Y, Song W, Chen S, Yuan Z, Yi Z. 2021. A bacterial artificial chromosome (BAC)-vectored noninfectious replicon of SARS-CoV-2. Antiviral Res 185:104974. https://doi.org/10.1016/j.antiviral.2020.104974

40. Elazar M, Cheong KH, Liu P, Greenberg HB, Rice CM, Glenn JS. 2003. Amphipathic helix-dependent localization of NS5A mediates hepatitis C

virus RNA replication. J Virol 77:6055–6061. https://doi.org/10.1128/jvi.77.10.6055-6061.2003

41. Sathanantham P, Zhao W, He G, Murray A, Fenech E, Diaz A, Schuldiner M, Wang X. 2022. A conserved viral amphipathic helix governs the replication site-specific membrane association. PLoS Pathog 18:e1010752. https://doi.org/10.1371/journal.ppat.1010752

42. Zhang Y, Zhao X, Zou J, Yuan Z, Yi Z. 2019. Dual role of the amphipathic helix of hepatitis C virus NS5A in the viral polyprotein cleavage and replicase assembly. Virology 535:283–296. https://doi.org/10.1016/j.virol.2019.07.017

43. Zhang Y, Chen S, Yuan Z, Yi Z. 2021. Bioorthogonal dissection of the replicase assembly of hepatitis C virus. Cell Chem Biol 28:1366–1378. https://doi.org/10.1016/j.chembiol.2021.03.006

44. Liu L, Westler WM, den Boon JA, Wang X, Diaz A, Steinberg HA, Ahlquist P. 2009. An amphipathic alpha-helix controls multiple roles of brome mosaic virus protein 1A in RNA replication complex assembly and function. PLoS Pathog 5:e1000351. https://doi.org/10.1371/journal.ppat.1000351

45. Sundar S, Thangamani L, Piramanayagam S, Rahul CN, Aiswarya N, Sekar K, Natarajan J. 2021. Screening of FDA-approved compound library identifies potential small-molecule inhibitors of SARS-CoV-2 nonstructural proteins NSP1, NSP4, NSP6 and NSP13: molecular modeling and molecular dynamics studies. J Proteins Proteom 12:161–175. https://doi.org/10.1007/s42485-021-00067-w

46. Abdelkader A, Elzemrany AA, El-Nadi M, Elsabbagh SA, Shehata MA, Eldehna WM, El-Hadidi M, Ibrahim TM. 2022. In-silico targeting of SARS-CoV-2 NSP6 for drug and natural products repurposing. Virology 573:96–110. https://doi.org/10.1016/j.virol.2022.06.008

47. French AP, Mills S, Swarup R, Bennett MJ, Pridmore TP. 2008. Colocalization of fluorescent markers in confocal microscope images of plant cells. Nat Protoc 3:619–628. https://doi.org/10.1038/nprot.2008.31