SARS-CoV-2 accessory protein ORF8 is secreted extracellularly as a glycoprotein homodimer

Kazuhiro Matsuoka[1] ◉, Nobuhiko Imahashi[2] ◉, Miki Ohno[1], Hirotaka Ode[1], Yoshihiro Nakata[1,3], Mai Kubota[1], Atsuko Sugimoto[1], Mayumi Imahashi[1], Yoshiyuki Yokomaku[1], and Yasumasa Iwatani[1,3,]* ◉

From the [1]Department of Infectious Diseases and Immunology, Clinical Research Center, and [2]Department of Hematology, National Hospital Organization Nagoya Medical Center, Nagoya, Aichi, Japan; [3]Department of AIDS Research, Nagoya University Graduate School of Medicine, Nagoya, Aichi, Japan

Edited by Karin Musier-Forsyth

ORF8 is an accessory protein encoded by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Consensus regarding the biological functions of ORF8 is lacking, largely because the fundamental characteristics of this protein in cells have not been determined. To clarify these features, we herein established an ORF8 expression system in 293T cells. Using this system, approximately 41% of the ORF8 expressed in 293T cells were secreted extracellularly as a glycoprotein homodimer with inter/intramolecular disulfide bonds. Intracellular ORF8 was sensitive to the glycosidase Endo H, whereas the secreted portion was Endo-H-resistant, suggesting that secretion occurs *via* a conventional pathway. Additionally, immunoblotting analysis showed that the total amounts of the major histocompatibility complex class I (MHC-I), angiotensin-converting enzyme 2 (ACE2), and SARS-CoV-2 spike (CoV-2 S) proteins coexpressed in cells were not changed by the increased ORF8 expression, although FACS analysis revealed that the expression of the cell surface MHC-I protein, but not that of ACE2 and CoV-2 S proteins, was reduced by ORF8 expression. Finally, we demonstrate by RNA-seq analysis that ORF8 had no significant stimulatory effects in human primary monocyte-derived macrophages (MDMs). Taken together, our results provide fundamental evidence that the ORF8 glycoprotein acts as a secreted homodimer, and its functions are likely associated with the intracellular transport and/or extracellular signaling in SARS-CoV-2 infection.

Despite extensive studies on severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the functions of viral gene products have not been characterized, hampering our understanding of the biology of coronavirus disease 2019 (COVID-19). To further understand the functions of its viral proteins, it is necessary to investigate its biochemical properties and functions in cells using appropriate systems. Except for that of the SARS-CoV-2 spike protein (CoV-2 S), the functions of viral accessory proteins have not yet been widely studied.

Although the viral genes of severe acute respiratory syndrome coronavirus (referred to as SARS-CoV-1 in this study) and SARS-CoV-2 are relatively conserved, many genetic mutations and deletions/insertions are observed in the locus near the *open reading frame 8* (*orf8*) genes among betacoronaviruses (1–10). Therefore, the ORF8 protein is regarded as one of the most rapidly evolving betacoronavirus proteins (4, 6–8). In addition, the ORF8 amino acid sequences tend to be mutated in epidemics and during adaptation to novel hosts, suggesting that they play an important role in pathogenicity (2, 4, 5, 11). In contrast, the expression of the ORF8 protein does not appear to be strictly essential for the *in vitro* replication of SARS-CoV-1 and SARS-CoV-2 (12, 13). In the case of SARS-CoV-1, a 29-nucleotide deletion that occurred early in the human-to-human transmission of the virus in 2002 split *orf8* into *orf8a* and *orf8b*, which correlated with attenuation of the disease (3, 6, 14–16). Furthermore, during the SARS-CoV-1 epidemic, a strain of SARS-CoV-1 with an 82-nucleotide deletion and a large-scale deletion in the *orf8* gene was also detected (3). Moreover, a deletion of 382 bases (Δ382) in SARS-CoV-2, detected in Singapore, was found to correlate with milder disease and lower incidence of hypoxia (11). However, the variant of concern (VOC), the alpha strain (B.1.1.7) that carries a nonsense mutation in *orf8* gene (Q27*) was first detected in the United Kingdom in late 2020 and caused a worldwide pandemic that included many severe cases (17–19). The Q27* of the alpha strain was first thought to be a mutation that splits or inactivates the ORF8 protein (11, 12, 20, 21).

SARS-CoV-2 ORF8 is a 121 amino acid protein that presumably has an N-terminal signal sequence and a glycosylation motif, asparagine-X-serine/threonine (X is any amino acid except proline), as determined by *in silico* analysis (8, 22). Previous studies showed that in cells, the ORF8 proteins of both SARS-CoV-1 and SARS-CoV-2 localize to the endoplasmic reticulum (ER) and that the SARS-CoV-2 ORF8 protein interacts with a variety of host proteins in the ER lumen, including the factors involved in ER-associated degradation and immunity (23–25). Because anti-ORF8 antibodies are produced in patients infected with SARS-CoV-2, the ORF8 protein may be secreted beyond the ER (22). Several other functions of the ORF8 protein have also been reported thus

* For correspondence: Yasumasa Iwatani, iwatani.yasumasa.cp@mail.hosp.go.jp.

far. The SARS-CoV-2 ORF8 protein has been reported to induce ER stress (26, 27), antagonize the interferon (IFN) signaling pathway (26, 28, 29), downregulate major histocompatibility complex class I (MHC-I) molecules in cells (30), and interact with transforming growth factor-beta 2 (TGF-β2) to participate in the TGF-β1 signaling cascade (25). In addition, the ORF8 protein was found to promote the expression of inflammatory factors by interacting with host interleukin(IL)-17 receptor A (IL17RA) (31). However, consensus on these phenomena is limited yet, and the fundamental functions of the SARS-CoV-2 ORF8 protein must be elucidated (9).

One of the most critical points for elucidating its functions is determining its biochemical characteristics in cell using appropriate systems. For example, the molecular size of the ORF8 protein expressed in cells differs among previous reports (22, 23, 25, 27, 28, 30, 31). This prompted us to assess ORF8 protein expression and determine its characteristics in cells. First, when expressed with the ORF8 expression plasmid encoding viral *orf8* cDNA, the mRNA expressed in cells was unexpectedly spliced at cryptic splice sites, which resulted in the production of an ORF8 protein smaller than the theoretical size. Optimization of codons and removal of cryptic splice sites enabled us to build a new ORF8 expression system that allowed the production of mRNA and protein at the appropriate sizes. Using this system, we found that approximately 41% of SARS-CoV-2 ORF8 was secreted out of the cell as a homodimeric glycoprotein with intramolecular disulfide bonds. In addition, we observed ORF8-dependent downregulation of MHC-I. In contrast, stimulatory effects of the secreted ORF8 protein were not detected in primary human monocyte-derived macrophages (MDMs). These studies provide a basic system to further identify the fundamental functions of the ORF8 protein in SARS-CoV-2 infection.

Results

Cryptic splice sites in the ORF8 coding region are used upon *orf8* cDNA expression

SARS-CoV-2 ORF8 is a putative secretory protein comprised of 121 amino acids that contains an N-terminal signal sequence and a glycosylation site at the 78th (N78) position (Fig. 1*A*). Based on the crystal structure obtained by oxidative refolding of the recombinant ORF8 protein from *E. coli*, ORF8 forms a homodimer with three pairs of intramolecular disulfide bonds per protomer and one pair of intermolecular disulfide bonds *via* cysteine 20 (C20) (Fig. 1*A*) (32). To characterize the biological features of ORF8 in cells, we first cloned SARS-CoV-2 ORF8 cDNA directly from viral RNA (vRNA) and constructed an expression plasmid for the ORF8 protein with a mycFLAG tag at the C-terminus (ORF8-mycFLAG). ORF8-mycFLAG-derived cDNA was overexpressed in Human embryonic kidney 293T (293T) cells and analyzed by Western blot. As shown in Figure 1*B*, the band of ORF8-mycFLAG migrated at 14.1 kDa, slightly faster than that of the theoretically sized protein (>15.8 kDa as a glycoprotein of the peptide 16–121), suggesting abnormal ORF8 expression through this system (Fig. 1*B*). Because SARS-CoV-2 is a

positive-sense single-stranded RNA virus, both vRNA replication and viral protein synthesis take place in the cytoplasmic environment. In contrast, in our overexpression system mediated by a mammalian expression vector encoding cDNA, ORF8 mRNA transcribed in the nucleus is transferred to the cytoplasm, where protein synthesis occurs. Therefore, we hypothesized that the abnormal ORF8 expression was due to unexpected splicing as cryptic splice site(s) in ORF8 RNA. To assess this possibility, we isolated total RNA from transfected 293T cells, amplified *orf8* gene products by RT-PCR, and analyzed the sequences by the Sanger method. One major spliced product SF1, but not a full-length (FL) product, was detected. SF1 had a 126-nucleotide (nt) in-frame deletion between the GU-AG splice motif. Inactivation of the splice donor site at the 230th position resulted in production of both a FL product and an alternative spliced fragment SF2 that had a 221-nt out-of-frame deletion (SF1, Fig. 1*C* and Table S1). These results indicate that upon ORF8 cDNA expression, the following cryptic splice sites in the coding region are predominantly used: two active splice donor sites (135 nt and 230 nt) and one splice acceptor site (355 nt) (Fig. 1*D*). When the other plasmid, pQCXIP, was used, FL RNA products were increased (Fig. 1*C*). Therefore, we rebuilt three ORF8 expression plasmids by removing the cryptic splice sites (CO2 and CO3), optimizing the *orf8* codons (CO1, CO2, and CO3), and inserting the genes into the retroviral vector pQCXIP. The sequences of the *orf8* genes designed herein are listed in Table S1. Then, 293T cells were transfected with each plasmid, and the ORF8 proteins were analyzed. The major bands of CO2 and CO3 were found at approximately 19.9 kDa (Fig. 1*B*). In addition, only FL RT-PCR products were detected in CO2 and CO3 (Fig. 1*C*). Therefore, pQCXIP ORF8-mycFLAG (CO3) was used in subsequent experiments for ORF8 protein expression.

The ORF8 protein is secreted extracellularly

Based on our *in silico* analyses, the ORF8 sequences had a signal sequence cleavage site at the 15th peptide bond and no predicted ER-retention motifs (8, 22, 33), suggesting that ORF8 is a secretory protein. To confirm this possibility, we overexpressed ORF8-mycFLAG CO3 in 293T cells, and the proteins in cells and supernatants (sups) were analyzed by an immunoprecipitation assay using anti-Flag magnetic beads. As expected, ORF8-mycFLAG CO3 was detected in both cells and the culture sup, whereas SF1 ORF8 was found only in cells (Fig. 1*E*). Based on the quantification of the intensity of each band in the immunoblots, approximately 41% of the total ORF8 protein was retained in the sup fraction (Fig. S1). These results suggest that ORF8 is efficiently secreted extracellularly, although ORF8 SF1 accumulated inside of the cells and was not secreted extracellularly. We further examined the secreted ORF8 protein in the sup under reducing conditions with 2-mercaptoethanol (2-ME). In the presence of 2-ME, a specific band of the ORF8-mycFLAG protein was detected at approximately 19 kDa, which roughly corresponds to its monomer size (Fig. 1*F*). In contrast, in the absence of 2-ME,
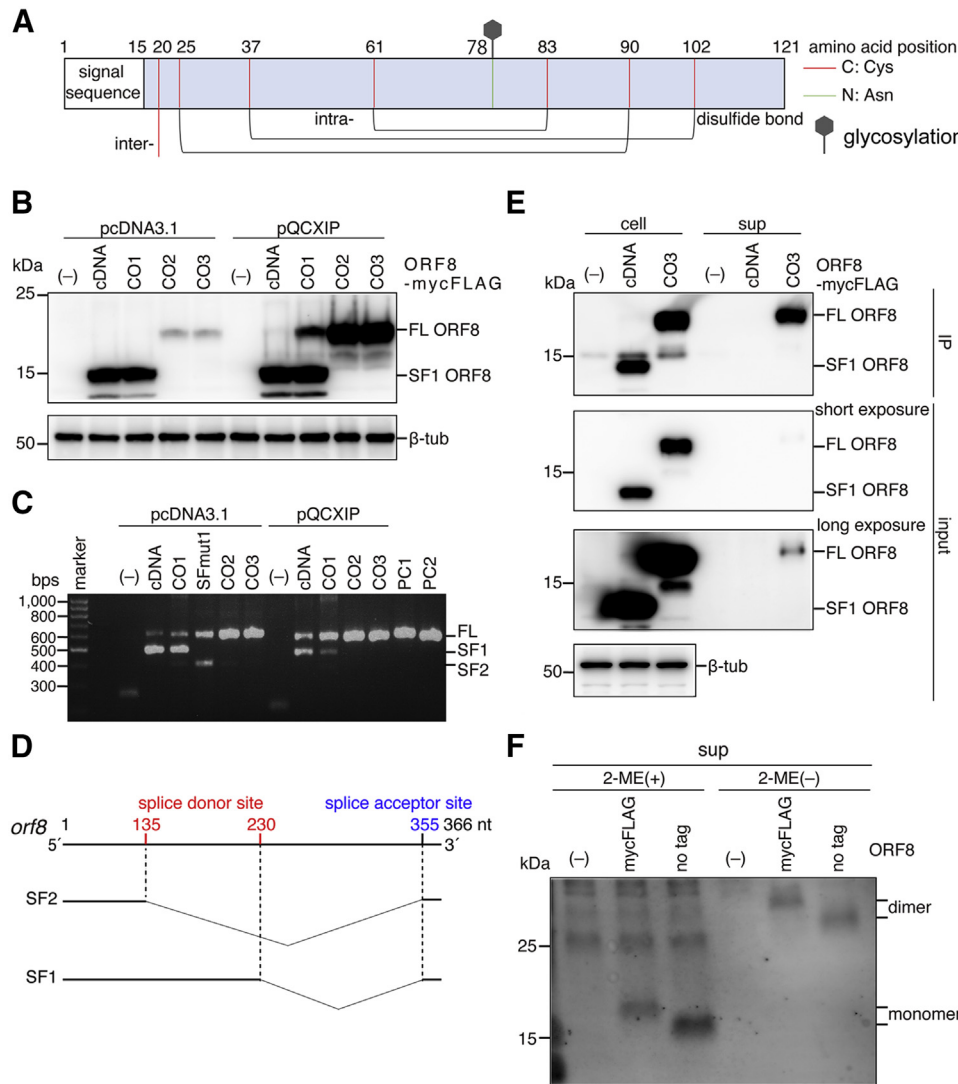
Figure 1. Expression of the SARS-CoV-2 ORF8 protein in 293T cells. *A*, schematic representation of the SARS-CoV-2 ORF8 protein. ORF8 is presumably a secretory protein with an N-terminal signal peptide and a glycosylation site at N78 (*green line*). Cysteine residues involved in the formation of intermolecular (inter) and intramolecular (intra) disulfide bonds are indicated by the *red line*. *B*, Western blot analysis. Briefly, 293T cells were transfected with C-terminally myc-FLAG-tagged ORF8 expression plasmids containing no ORF8 gene (–), ORF8 cDNA (cDNA), codon optimized ORF8 gene 1 (CO1), or the additionally optimized ORF8 gene 2 (CO2) or 3 (CO3). The pcDNA 3.1 or pQCXIP expression vector was used for transient expression. The proteins were detected by Western blotting using anti-FLAG tag mAb and anti-β-tub antibody. The bands corresponding to "FL" ORF8-mycFLAG (FL ORF8) and "spliced form 1" ORF8-mycFLAG (SF1 ORF8) are indicated. *C*, analysis of mRNA in the transfected 293T cells. DNA fragments amplified by RT-PCR were separated by electrophoresis and stained with ethidium bromide. The fragments amplified from pcDNA 3.1 ORF8-mycFLAG (PC1: 678 bps) and pQCXIP 3.1 ORF8-mycFLAG (PC2: 604 bps) were used as positive controls. *D*, schematic representation of the cryptic splice sites in the ORF8 cDNA. Based on the fragment sequences, two major sliced forms were found: spliced form 1 (SF1) and spliced form 2 (SF2). The "SFmut1" is the ORF8 cDNA of which the second cryptic donor site (230 nt) was mutated. The major cryptic splice sites in the open reading frame are indicated. *E*, analysis by pull-down assay. ORF8-mycFLAG proteins were expressed in 293T cells using ORF8-mycFLAG expression plasmids, cDNA-derived ORF8 (cDNA) and optimized ORF8-mycFLAG (CO3). ORF8 proteins in cells (cell) and supernatants (sup) were immunoprecipitated with anti-FLAG mAb beads. The immunoprecipitated (IP) and input fractions were analyzed by Western blot using anti-FLAG mAb and β-tub antibody. *F*, analysis of ORF8 with or without 2-ME. The gene-optimized ORF8 (CO3) with or without mycFLAG was transiently expressed in 293T cells. The samples treated with and without 2-ME were separated by SDS-PAGE and analyzed by Western blot with an anti-SARS-CoV-2 ORF8 rabbit polyclonal antibody. 2-ME, 2-mercaptoethanol; ORF8, open reading frame 8; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

it displayed at 33 kDa, which was approximately twice the size of the monomer. These data indicate that ORF8 is secreted extracellularly as a dimer that is covalently linked *via* a disulfide bond(s). In addition, as expected, because the molecular size of the detected monomer ORF8-mycFLAG was 3.6 kDa larger than the theoretical size of the protein, the results suggest that the secreted ORF8 is likely glycosylated. Of note, these results were obtained even when using the nontagged ORF8 version (Fig. 1*F*).

ORF8 is modified with an endoglycosidase H (Endo H)-resistant glycan chain at the N78 residue in cells

Next, we analyzed the intracellular expression of wild-type (WT) ORF8 and two ORF8 mutants: N78D with an N-linked glycosylation site deficiency and the ORF8 ΔSP mutant with a 15-amino acid signal sequence deletion. The ΔSP mutant, which was expected to be 14.4 kDa and to localize in cytoplasm, was used as control. As shown in Figure 2*A*, the N78D band was detected at approximately 16 kDa, which was

equivalent to that of ΔSP (Fig. 2A). These results demonstrated that the ORF8 protein is translocated into ER, followed by cleavage of the signal sequence and glycosylation at residue N78. To clarify the glycosylation type (high mannose, hybrid, or complex-type), we expressed ORF8-mycFLAG proteins in 293T cells, immunoprecipitated them using anti-Flag magnetic beads, and assessed their sensitivity to two types of glycosidases: peptide-N-glycosidase F (PNGase F) and Endo H. As shown in Figure 2B, the ORF8 WT protein collected from cells

was sensitive to PNGase F and Endo H. In contrast, the ORF8 WT immunoprecipitated from sup was sensitive to PNGase F but resistant to Endo H (Fig. 2B). These results suggest that the majority of intracellular ORF8 proteins have an N-linked high mannose chain, whereas the secreted portion has a complex-type (or hybrid-type) sugar chain. Furthermore, this result suggests that ORF8 is secreted *via* the Golgi-mediated glycoprotein transport and secretion pathway.

To further understand the role of glycosylation in ORF8, we compared the levels of N78D with the WT and C20A mutant in soluble and insoluble cell fractions. The C20A mutant that is deficient for intermolecular disulfide bonds was herein used as control because glycosylation of ORF8 appeared to be independent of stabilization *via* intermolecular disulfide bonding based on information of the ORF8 crystal structure. The results showed that the levels of WT and the C20A mutant were significantly low in the insoluble fraction, although the N78D mutant was highly detected in the insoluble fraction (Fig. 2C). These data demonstrated that glycosylation of ORF8 at the N78 residue increases its solubility in cells.

Glycosylation of ORF8 is required for its efficient secretion, although intermolecular disulfide bonding of the homodimer is dispensable for its secretion

Next, to further characterize intracellular modifications of the ORF8 protein, we investigated the cellular secretions of the ORF8 N78D mutant and the C20A mutant. We expressed either of the ORF8 mutants in 293T cells, pulled down ORF8-mycFLAG proteins in cells (cell) and sups with anti-FLAG-tagged beads, and analyzed them by immunoblotting. As expected, the N78D bands in the lysate and sup migrated 3.6 kDa faster than the WT or C20A mutant in the presence of 2-ME (Fig. 3A), indicating that the secreted ORF8 was glycosylated at the N78 position and that the C20A mutation did not disrupt ORF8 glycosylation. In the sups, the level of ORF8 N78D was dramatically reduced compared with that of the WT and C20A (Fig. 3A). These results indicate that ORF8 glycosylation is critical for its efficient secretion into the cell sup. These trends were also observed in the immunoprecipitated (IP) fractions (Fig. 3B).

When the proteins in the lysate were separated in the absence of 2-ME, the ORF8 bands displayed multiple ladders (Fig. 3C, left). In contrast, the WT and C20A bands were dominantly the dimer size (∼34 kDa) and monomer size (∼18 kDa), respectively (Fig. 3C, right). ORF8 has six cysteine residues at amino acids 25, 37, 61, 83, 90, and 102 in addition to the cysteine at amino acid 20, and the crystal structure shows that each monomer has three pairs of intramolecular disulfide bonds (C25–C90, C37–C102, and C61–C83) (Fig. 1A) (32). The data indicate that the C20 residue is important for stable homodimer formation through an intermolecular disulfide bond, consistent with the ORF8 structure. In contrast, the other multimers retained intracellularly formed intermolecular disulfide bonds because multimers detected in the C20A mutant display a similar ladder pattern. In addition, the IP data show that the ORF8 dimer is
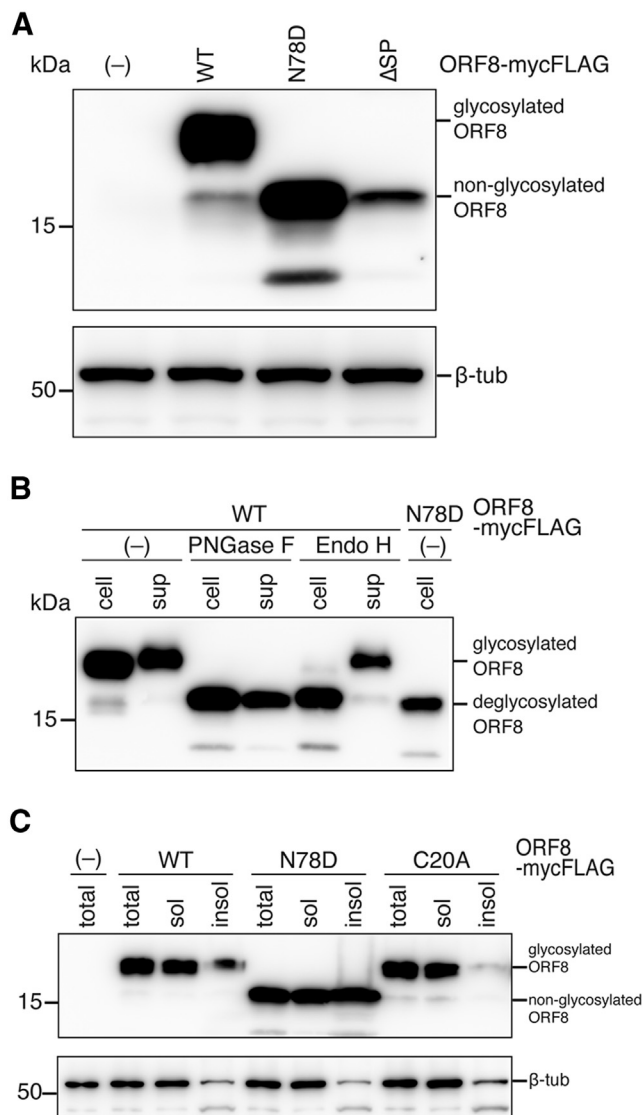
Figure 2. Characterization of the N-linked glycosylation of the SARS-CoV-2 ORF8 protein. *A,* briefly, 293T cells were transfected with the wild-type (WT) or mutant ORF8-mycFLAG expression plasmid. Cell lysates were denatured under reducing conditions with 2-ME and then visualized by immunoblotting with an anti-FLAG mAb and anti-β-tub antibody. The ORF8 mutant deficient for the signal sequence (ΔSP) was used as a negative control. *B,* the Endo H and PNGase F sensitivities of the ORF8 proteins were assessed. The WT and N78D mutant were pulled down with anti-FLAG mAb beads from the cell lysate (cell) and culture supernatant (sup). The IP factions were untreated (–), or incubated with either PNGase F or Endo H glycosidase. The digested samples were subjected to immunoblot analysis using an anti-FLAG mAb. *C,* the total fraction (total), soluble fraction (sol), and insoluble fraction (insol) were analyzed by Western blot using anti-FLAG mAb and anti-β-tub antibody. 2-ME, 2-mercaptoethanol; ORF8, open reading frame 8; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.
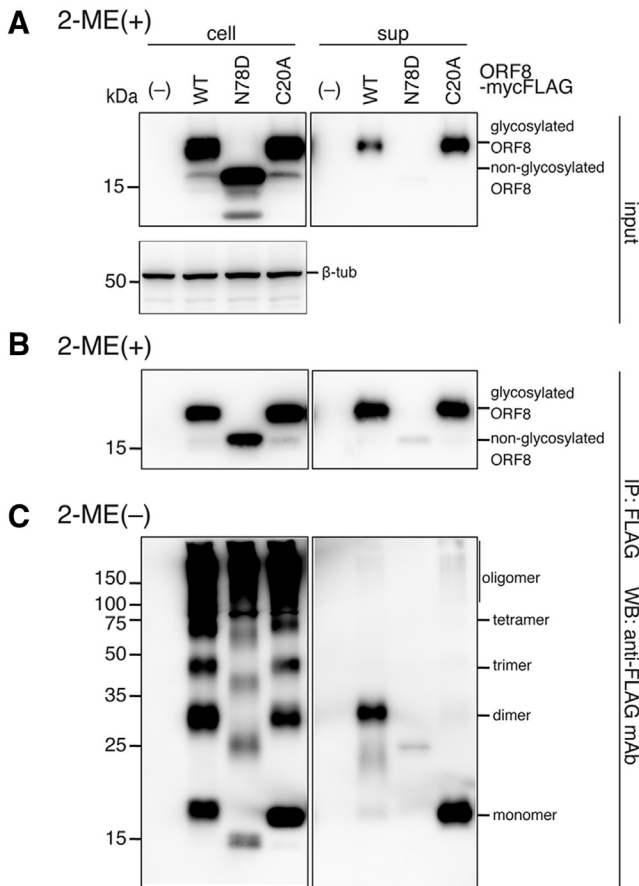
Figure 3. Effects of SARS-CoV-2 ORF8 mutations at glycosylation and cysteine residues on secretion. *A*, wild-type (WT) and mutant ORF8-mycFLAG were expressed in 293T cells and separated by SDS-PAGE. Total fractions (input) treated with 2-ME were analyzed by immunoblotting using anti-FLAG mAb and anti-β-tub antibody. *B* and *C*, the immunoprecipitated (IP) fractions were prepared using anti-FLAG mAb beads and treated with 2-ME (*B*) or without 2-ME (*C*). The IP fractions were analyzed by Western blot using an anti-FLAG mAb. The approximate sizes (kDa) of the molecular weight standards for SDS-PAGE are indicated on the *left*. 2-ME, 2-mercaptoethanol; ORF8, open reading frame 8; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2.

preferentially exported out of cells, although ORF8 dimerization is dispensable for its secretion from cells.

The ORF8 protein preferentially downregulates cell surface MHC-I molecules

Some viruses are known to downregulate antigen-presenting proteins and receptors required for entry expressed on the cell surface upon infection to evade host immunity and to enhance viral replication (34–37). A previous report showed that HA-tagged ORF8 (ORF8-HA), which is dominantly retained in the cytoplasm, directly interacts with MHC-I molecules and leads to their downregulation (30). Because the ORF8-HA was expressed from the ORF8 cDNA, similar to our SF1 ORF8 (Fig. 1), and was ~15 kDa, which is smaller than the FL glycosylated ORF8 (~19 kDa), we needed to reevaluate the effect of FL ORF8 on MHC-I downregulation in our ORF8 expression system. Human leukocyte antigen (HLA-A0201 or HLA-A2) was used as the MHC-I molecule.

HLA-A2, ACE2, or CoV-2 S was coexpressed with ORF8-mycFLAG in 293T cells, and the total amount of each protein in cells was analyzed by Western blotting. The total amounts of the intracellular HLA-A2 protein were similar in the cells with different expression levels of ORF8 (Fig. 4*A*). Elevated expression of the ORF8 protein did not change the total amounts of ACE2 and CoV-2 S in cells (Fig. 4, *B* and *C*). In contrast, FACS analysis of cell surface levels demonstrated that the level of HLA-A2 was significantly reduced (71.9% ORF8 WT *versus* control) in the presence of ORF8 WT (Fig. 4*D*), although ACE2 and CoV-2 S were not affected by ORF8 WT expression (Fig. 4, *E* and *F*). The levels of HLA-A2, ACE2, and CoV-2 S on the cell surface were similar between the control and the ΔSP (Fig. 4, *D–F*). The levels of ACE2 and SARS-CoV-2 S were not downregulated in an ORF8 protein-dependent manner as determined by FACS (Fig. 4, *D–F*). These results suggest that ORF8 protein expression induces the downregulation of MHC-I on the cell surface through a nonproteolytic degradation pathway.

No cytokine-like effects of the ORF8 protein were observed in MDMs

Finally, we assessed whether ORF8 has a stimulatory effect similar to that of cytokines in MDMs because ORF8 is efficiently secreted into the extracellular space as a homodimeric glycoprotein (Fig. 1*E*). In this assay, we used human primary MDM cells that were sensitive to various cytokines. Cell sups with or without ORF8 expression were collected from 293T cell cultures that had been transfected with pQCXIP ORF8-mycFLAG (CO3) or pQCXIP empty vector. The MDM cell culture medium was replaced with the ORF8-containing sup, and the cells were incubated for 2 h. Total RNA was prepared from ORF8-treated cells and subjected to RNA-seq. RNA-seq reads were comparatively analyzed between mock-treated and ORF8-treated MDMs. The results showed that eight genes (C2, U2AF1, NELFE, CD99, PSMC4, TAF15, NDUFA3, and PRKACA) were reduced and six genes (RABGGTA, ITPK1, MGAT4B, ABR, TNNI2, and TAP1) were significantly increased in the ORF8-treated MDMs compared with mock cells (false discovery rate (FDR) < 0.05; Fig. 5, *A* and *B*). Generally, cytokines such as type-I IFNs and IL-6 activate other downstream signals *via* their receptors and induce the expression of a wide variety of genes through master transcription factors, such as members of the signal transducer and activator of transcription (STAT) and/or nuclear factor-kappa B (NF-κB). In contrast, the genes predominantly changed by the addition of secreted ORF8 were not functionally linked to each other. Based on this point, we concluded that there were no cytokine-like effects of the secreted ORF8 protein on MDMs in our assay system.

Discussion

The goal of this study was to characterize the ORF8 protein to thereby elucidate its biological features in cells because several different biological functions of the molecule have been reported (22, 23, 25–31). At the beginning of this study, we
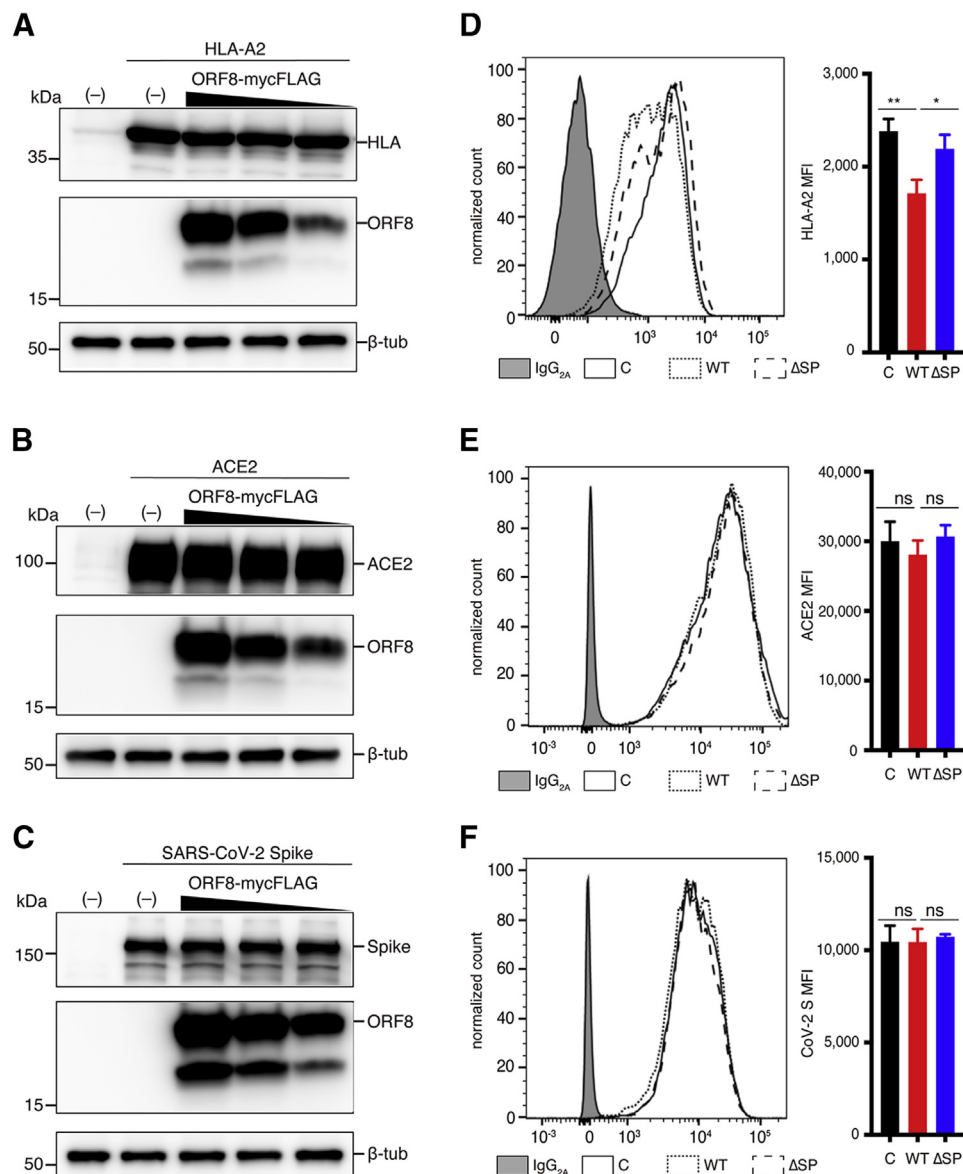
Figure 4. Effects of SARS-CoV-2 ORF8 on expression and downregulation of HLA-A2, ACE2, and CoV-2 S proteins. *A–C*, Western blot analyses of total ORF8 in cells. Briefly, 293T cells were cotransfected with each target expression plasmid (500 ng), (*A*) HLA-A0201-HIS, (*B*)ACE2, or (*C*) CoV-2 S, and different amounts of the expression plasmid encoding ORF8-mycFLAG (CO3) (0, 1500, 1000, and 500 ng). The total plasmid amount for each transfection was normalized with the empty vector. Next, 293T cells were cotransfected with each target plasmid (125 ng), GFP expression plasmid pAcGFP1-N1 (125 ng), plus pQCXIP empty vector control (*C*), ORF8-mycFLAG (WT), or ORF8 ΔSP-mycFLAG (ΔSP)-expressing plasmid (750 ng). *D–F*, FACS analyses of cell surface proteins. *D*, endogenous HLA-2 surface expression. 293T cells were cotransfected with pAcGFP1-N1 (125 ng) and the ORF8 expression plasmid (875 ng) for analysis. *E*, ACE2 and (*F*) CoV-2 S expression on the cell surface. Cells were transfected with each expression plasmid and collected at 40 h after transfection for flow cytometry analysis. The mean fluorescence intensities (MFI) of HLA-, ACE2-, and CoV-2 S-positive cells (gated on GFP-positive cells) were shown (n = 3). The bar graphs on the *right* represent the mean ± SD (error bars). Student's *t* test and one-way ANOVA were used. NS, not significant; *$p < 0.05$, **$p < 0.005$. ORF8, open reading frame 8; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; WT, wild-type.

thought that these different results could be obtained, largely because basic properties of this protein in cells have not yet been fully determined. Therefore, we first set out to establish a protein expression system in cultured cells to prevent the unintended splicing of ORF8 mRNA upon transfection of its expression plasmids. Indeed, the ORF8 cDNA derived from the viral genomic sequence resulted in the expression of a protein with a smaller molecular weight than the theoretical molecular weight, which was caused by unintended splicing at three cryptic splice sites (Fig. 1). Our elaboration of the ORF8 expression plasmids by a combination of codon optimization,

mutagenesis, and expression vectors enabled us to express the ORF8 protein at the theoretical molecular size in cells. Interestingly, further analysis of the protein using our system demonstrated that the SARS-CoV-2 ORF8 protein was secreted out of the cell as a homodimeric glycoprotein (Figs. 2 and 3). The 2-ME-sensitive ORF8 multimers partly accumulate in cells. The N-linked glycosylation of ORF8 at N78 was shown to be required for efficient secretion, although intermolecular disulfide bonding at C20 was dispensable for export out of cells. These results suggest that the ORF8 protein plays important roles in the intracellular transport pathway or
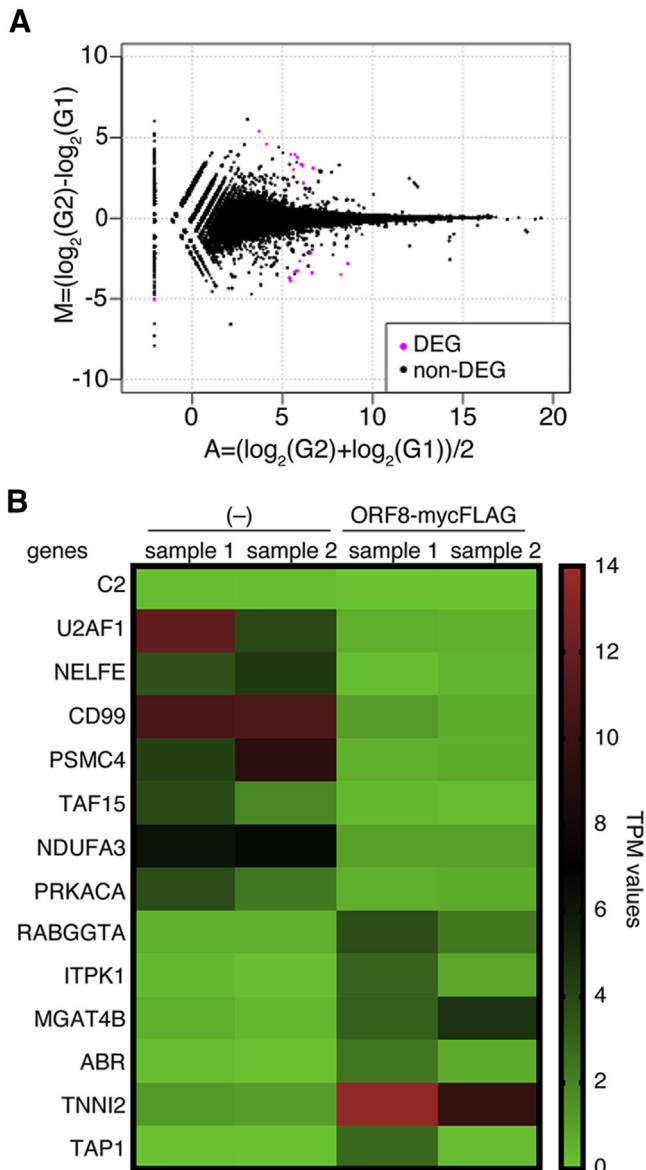
Figure 5. Expression profiles of mRNAs in ORF8-treated and control MDMs as determined by RNA-Seq. *A*, MA plot of human MDMs treated with the control and ORF8. The differentially expressed genes (DEGs) and nondifferentially expressed genes (non-DEGs) are indicated by *pink* and *black dots*, respectively. A false discovery rate (FDR) <0.05 indicated a significant difference. *B*, heatmap representing the differentially expressed transcripts of nontreated control MDMs (−) and ORF8-treated MDMs. Significantly enriched gene sets with an FDR <0.05 were used. The expression levels, defined as transcripts per million (TPM), are shown. MDM, monocyte-derived macrophage; ORF8, open reading frame 8.

extracellularly in SARS-CoV-2 infection. In our attempt to identify the biological function of the ORF8 protein, we found that ORF8 induced the downregulation of MHC-I but not of the ACE2 or CoV-2 S proteins on the cell surface (Fig. 4). In terms of secretory proteins, we examined the role of the secreted ORF8 protein as a cytokine-like agonist, but the induction of relevant gene expression in MDMs was not found in our system (Fig. 5).

Previous studies have shown that the ORF8 protein interacts with a variety of host proteins, including IL17RA and factors involved in ER-related degradation and vesicle trafficking

(23–25). In addition, the SARS-CoV-1 ORF8 protein has been reported to antagonize the IFN signaling pathway (26, 28, 29) and ER stress (26, 27). However, consensus in these reports is limited, and progress on SARS-CoV-2 ORF8 protein research remains limited. We hypothesized that some problems would complicate the overexpression system when using cultured cells for the analysis of ORF8 biological functions. Indeed, the molecular weights of the ORF8 proteins in studies reported thus far are not consistent and varied in a range between 14 and 20 kDa (22, 23, 25–31). Therefore, we first constructed an expression plasmid with a C-terminal mycFLAG tag attached to the cDNA-derived ORF8. Expression of cDNA-derived ORF8 in 293T cells showed that the molecular weight of the expressed ORF8-mycFLAG was slightly smaller than its theoretical value of 15.8 kDa as a glycoprotein (Fig. 1*B*). Our analysis of the mRNA in the transfection revealed that cDNA-derived ORF8 mRNA underwent splicing in the overexpression system (see SF1 ORF8, Fig. 1*B*), which should not naturally occur in the cytoplasm, where SARS-CoV-2 predominantly replicates (Fig. 1*C*). Therefore, we removed cryptic splice sites and optimized the codons and expression vector (retroviral vector) in combination (Fig. 1*B*) and successfully expressed an ORF8 protein of the correct size (approximately 20 kDa including a sugar chain and mycFLAG tag). These results suggest that cryptic splicing events should be considered when cDNA is used to investigate ORF8 functions. Two previous studies showed that a lentiviral vector is efficient for expression of the ORF8 protein at the correct size of 19 kDa (23, 25). This may be partly attributed to the fact that lentiviral vectors are capable of disabling splicing (38). Similarly, in this study, the FL ORF8 was expressed more efficiently with the retroviral vector pQCXIP than with pcDNA3.1 (Fig. 1, *B* and *C*), largely because the pQCXIP has a 5′ long terminal repeat of the murine leukemia virus proviral DNA that potentially enhances the transport of unspliced RNA from nucleus to cytoplasm (39, 40).

The present study showed that ORF8 was efficiently secreted out of the cell as a covalent homodimer with a pair of intermolecular disulfide bonds formed by C20 of each monomer (Fig. 3). Interestingly, the C20A mutation did not affect the glycosylation or extracellular secretion of ORF8 (Fig. 3). This result supports the two previous studies (22, 32). Based on the crystal structure, the intermolecular bonds in the ORF8 dimer are not only intermolecular disulfide bonds, but also residues in the loops connecting β1 and β8 and β3 and β4 that contribute to dimer formation (32). This suggests that the ORF8 protein may feasibly form homodimers, although the C20A mutant does not form intermolecular disulfide bonds. In terms of ORF8 glycosylation modification, our mutational analyses performed with the glycosylase digestion assay revealed that intracellularly accumulated ORF8 had a high mannose-type glycan/hybrid-type glycan, while the secreted ORF8 was conjugated with a complex-type glycan (Fig. 2*B*). Generally, secreted proteins undergo translation and glycosylation (high mannose type) in the ER and are transported to the Golgi apparatus, wherein their glycan structures are mature (41–46). Intracellular ORF8 with a high mannose-type

glycan/hybrid-type glycan was oligomerized (Figs. 2, *B* and *C* and 3*B*). This result suggests that the molecules formed as homodimeric ORF8 may have been transported to the Golgi apparatus, where further maturation of the glycan structure occurred. It is possible that the intracellularly oligomerized ORF8 observed in this study was an unintended byproduct of the virus. Often, when secretory or membrane proteins are expressed, a lack of molecular chaperones or other components in the ER can lead to abnormal protein folding (41–46). As a consequence, such stacking of abnormal proteins, including oligomeric ORF8, induces ER stress. Currently, which biological functions are associated with the complex glycosylation and dimerization of the ORF8 protein remains unknown, and further studies should be carried out.

Viruses are known to downregulate antigen-presenting proteins and the receptors required for entry expressed on infected cells upon infection of host cells (34–37). Similar to a previous study on the downregulation of MHC-I cell surface expression by SARS-CoV-2 ORF8 (30), we also observed a slight effect of ORF8 on the downregulation of an MHC-I protein, but not ACE2 and CoV-2 S proteins, in 293T cells (Fig. 4). MHC-I forms a heterodimer with β2 microglobulin (β2m) (47), whereas the ACE2 and CoV-2 S proteins are dimeric and homotrimeric, respectively (48, 49). Because intracellular ORF8 forms disulfide-mediated oligomers in the ER, ORF8 might trap MHC-I molecules by preferentially binding to them or by selectively binding to heterodimeric proteins during oligomer formation. Although Zhang *et al.* (30) reported that MHC-I is degraded *via* autophagy, we could not find an ORF8-dependent MHC-I reduction in cells in our assay system (Fig. 4*A*). While the factor underlying this differential result is unknown, it may be due to the different ORF8 protein contents (SF1 and FL ORF8s) since the molecular weight of the expressed ORF8 protein was dominantly 15 kDa in their report (30). The detailed molecular mechanism of MHC-I downregulation by ORF8 needs to be further investigated.

Some of the secreted viral proteins function as cytokine-like proteins (50, 51). In this study, we also analyzed the stimulatory effects of the secreted ORF8 protein on MDMs as a model, because MDMs are primary cells with high cytokine sensitivity. However, no significant induction of gene expression by the ORF8 protein was observed in MDMs. These results were potentially attributed to the ORF8 protein functioning as an antagonist and not an agonist and/or the inappropriate using of MDMs to identify stimulatory effects. Because SARS-CoV-2 replicates in the lungs and nasal cavity (52, 53), immune cells in mucous membranes and/or tissues should be targeted rather than immune cells in the blood. The followings are limitations of our study regarding the validation of the ORF8 protein we used: (1) the mycFLAG added to the C-terminus may have affected the bioactivity, and (2) the induction time may have been insufficient.

The ORF8-defective SARS-CoV-2 strain (Δ382) found in Singapore has been shown to have a milder pathology, reduced hypoxia, and decreased release of proinflammatory cytokines (11). In contrast, the VOC alpha strain (B.1.1.7) that has an ORF8 truncation induced by a nonsense mutation (Q27*) and presumably has a nonfunctional ORF8 (20, 21), became prevalent in the winter of 2020 (17, 18, 21). Despite epidemiological evidence that the alpha strain has spread around the world, many severe cases in COVID-19 have been immutably identified worldwide. More recent studies revealed that the VOC omicron variant (B.1.1.529 and the BA lineage) that encodes an intact *orf8* gene displays less pathogenicity (54) and causes mild or even asymptomatic symptoms in infected people (55). This clinical evidence indicates that the *orf8* gene in SARS-CoV-2 does not fully correlate with viral pathogenicity and disease severity. Interestingly, Lin *et al.* have reported that the unglycosylated ORF8 secreted *via* an unconventional pathway is responsible for the COVID-19 cytokine storm through IL-17 receptor-based signaling (31, 56). However, their proposed mechanism seems contradictory to the recent clinical evidence. In addition, the small ORF8 (<14 kDa) was used in the study to assess the effects of the ORF8 protein (31), which corresponds to the size of our observed "SF1 ORF8" (Fig. 1*B*). Therefore, further analyses are required to clarify the ORF8 function using appropriate ORF8 expression systems.

ORF8 is hypothesized to be advantageous for functions within an individual (*e.g.*, viral replication and immune evasion) but neutral or even disadvantageous for transmission (20). Our results revealed that downregulation of MHC-I by ORF8 and the unidentified function of the secreted ORF8 protein might be advantageous, while the intracellular ORF8 protein might be disadvantageous because it induces ER accumulation and ER stress (26, 27) *via* oligomerization. Therefore, the ORF8 protein is considered to be both advantageous and disadvantageous for viruses. Therefore, it is possible that ORF8 is evolving to an optimal state in the individual while balancing these properties.

In conclusion, the SARS-CoV-2 ORF8 protein was found to be secreted extracellularly as a homodimer with a sugar chain of the complex-type glycan. On the other hand, a certain amount of ORF8 is intracellularly accumulated. Therefore, ORF8 may play important roles in the intracellular transport pathway and/or extracellular signaling in SARS-CoV-2 infection. Our study revealed that C20 and N78 play an important role in maintaining the molecular structure of the ORF8 protein in cells. In fact, these two residues of ORF8 in the epidemic strains have been highly conserved according to the Global Initiative on Sharing All Influenza Data (GISAID) (99.9% of 2,156,221 sequences of the FL SARS-CoV-2 genome, registered by December 2021). Therefore, these results might support that these two residues of ORF8 play essential roles in the bioactivity of SARS-CoV-2 infection *in vivo*. Finally, these findings provide an important basis for understanding a novel functional mechanism of the ORF8 protein and for exploring the factors that contribute to human transmission, considering that the *orf8* gene is one of the most rapidly evolving in the betacoronavirus family.

Experimental procedures

Cell culture

293T cells and Vero E6 cells (American Type Culture Collection) were maintained in Dulbecco's modified Eagle

medium (Sigma-Aldrich) supplemented with 10% fetal bovine serum (FBS), penicillin (100 U/ml) and streptomycin (100 μg/ml) (Thermo Fisher Scientific) in 5% CO_2 at 37 °C.

Plasmid construction

To construct the pcDNA 3.1-based ORF8 expression plasmids, we first extracted viral RNA from SARS-CoV-2-infected Vero E6 cells and amplified a DNA fragment of the SARS-CoV-2 *orf8* by nested RT-PCR using the PrimeScript II High Fidelity One Step RT-PCR kit (Takara Bio). The primers for the first PCR (orf8_5′(27841-64)_1(+) and orf8_3′(28327-52)_1(−)) and the second PCR (orf8_ATG(NotI) and orf8_myc-FLAG(HindIII)) were used for amplification, and the primer sequences are listed in Table S2. Each amplified DNA fragment containing the sequence for a tandem tag of myc and FLAG peptides (mycFLAG) was inserted between the NotI and HindIII sites of the pcDNA3.1 (−) empty vector (Thermo Fisher Scientific). For pQCXIP ORF8 expression plasmids, a DNA fragment of SARS-CoV-2 *orf8* with a mycFLAG-coding sequence at the 3′ end was generated from pcDNA3.1 ORF8 using PrimeSTAR Max DNA Polymerase (Takara Bio) and a primer set, orf8_ATG(NotI) and Rev_BamHI-FLAG. An amplified fragment was inserted between the NotI and BamHI sites of the pQCXIP empty vector (Takara Bio). The codon-optimized ORF8 expression plasmids were constructed by inserting the synthetic DNA fragment, SARS-CoV-2 *orf8* (purchased from IDT), into pcDNA3.1 (−) and pQCXIP. The synthetic *orf8* fragment was codon-optimized for mammalian cell expression, in which cryptic splice sites were mutated. For the pQCXIP codon-optimized ORF8 ΔSP expression plasmids or pQCXIP codon-optimized ORF8 (no tag) expression plasmids, a DNA fragment was amplified from pQCXIP codon-optimized ORF8 (CO3) by PCR using the primer sets NotI_ATG_F16__o8 and Rev_BamHI-FLAG for ORF8 ΔSP or orf8_ATG(NotI) and BamHI_stop_o8 for ORF8 (no tag) and was then replaced back into the pQCXIP vector. The C-terminal HIS-tagged HLA expression plasmid of pcDNA3.1-HLA-A0201-HIS was constructed by reinserting the HLA cDNA fragment that was amplified by PCR from the pcDNA3.1 HLA-A*02:01 plasmid (57) using the primer set T7 Promoter and HLA-A0201-C-HIS_XbaI(−). Mutants of the glycosylation site (N78D) and intermolecular disulfide site (C20A) were generated using site-directed mutagenesis with a PrimeSTAR mutagenesis basal kit (Takara Bio). All the sequences of both the insert and the boundary regions of the expression plasmids were verified by conventional DNA sequencing methods. Detailed information about the primers used here is shown in Table S2.

Transfection

For analysis of the intracellular or extracellular ORF8 protein levels, the ORF8 expression plasmid (2 μg) was transfected into 293T cells in 12-well plates using FuGENE HD (Promega). To analyze the effects of ORF8 on the total amounts of surface proteins, a CoV-2 S expression plasmid (pC-SARS2-S WT) (58), ACE2 expression plasmid (pC-ACE2 WT) (58), and HLA expression plasmid (pcDNA3.1-HLA-A0201-HIS)

were used. Each surface protein expression plasmid (500 ng) was cotransfected with different amounts of pQCXIP ORF8(CO3)-mycFLAG WT (0, 1500, 1000, and 500 ng; the plasmid amount was normalized by empty vector) or 1500 ng of the pQCXIP (−) vector control. At 36 h posttransfection, the cell lysates and the culture sup were prepared with Laemmli buffer (Bio-Rad) in the presence or absence of 2.5% (vol/vol) 2-ME. The sup was passed through a 0.45-μm-pore-size filter to remove cellular debris for direct detection by immunoblotting.

For FACS analysis, either pC-SARS2-S WT or pC-ACE2 WT (125 ng), pAcGFP1-N1 (Takara Bio) (125 ng), and 750 ng of pQCXIP ORF8(CO3)-mycFLAG (WT or ΔSP) or pQCXIP (−) were used for cotransfection. To assess the effects of ORF8 expression on endogenous HLA levels on the 293T cell surface, only the pAcGFP1-N1 (125 ng) and ORF8 expression plasmids (875 ng) were transfected. At 40 h posttransfection, the cells were subjected to FACS analyses.

Western blot analysis

We fractionated 293T cell lysates by SDS-PAGE (12% acrylamide gel) and transferred them onto Immobilon-P membranes (Merck-Millipore). The membranes were first probed with the appropriate primary antibodies. The C-terminal mycFLAG-tagged ORF8 and SARS-CoV-2 ORF8 (no Tag) were detected with an anti-DYKDDDDK (FLAG)-tagged mouse monoclonal antibody (mAb) (1:2000; Fujifilm Wako Pure Chemical Co) and an anti-SARS-CoV-2 ORF8 rabbit polyclonal antibody (1:1000; ABclonal; A20235), respectively. For membrane protein detection, an anti-HIS tag mAb (1:1000; MBL; D291-3) for HIS-tagged HLA, an anti-SARS-CoV/SARS-CoV-2 spike mAb (1A9) (1:1000; GeneTex, Inc; GTX632604) for CoV-2 S, an anti-ACE2 rabbit monoclonal antibody (1:1000; Abcam; ab108252) for ACE2, and an anti-β-tubulin rabbit polyclonal antibody (1:2000; Abcam; ab6046) for β-tubulin (β-tub) were used. The immunoblotted membranes were subsequently incubated with horseradish-peroxidase-conjugated secondary antibodies. A goat anti-mouse IgG antibody (1:20,000; Thermo Fisher Scientific) and a goat anti-rabbit IgG antibody (1:20,000; Thermo Fisher Scientific) were used for FLAG, ORF8, HIS, CoV-2 S, ACE2, and β-tub, respectively. The proteins were visualized by enhanced chemiluminescence using SuperSignal West Dura (Thermo Fisher Scientific) and quantified using an ImageQuant LAS 4000 (GE Healthcare Life Sciences). The molecular size of bands on the immunoblot images were estimated by Image-Quant TL 8.2 (Cytiva) with a molecular weight standard, ECL DualVue Western blotting markers (Cytiva).

RNA extraction, RT-PCR, and sequencing

Total RNA was isolated from transfected 293T cells using ISOGEN-LS (Nippon Gene) according to the manufacturer's protocol and treated with RNase-free recombinant DNase I (Takara Bio). The RNA was then reverse-transcribed and amplified by RT-PCR with the PrimeScript One Step RT-PCR Kit Ver. 2 (Takara Bio) using either of two primer sets: T7 Promoter and BGH Reverse (for the pcDNA-based expression) or pQCXIP sequence primer 5′ and pQCXIP sequence primer

≋ASBMB

J. Biol. Chem. (2022) 298(3) 101724 **9**

3′ (for the pQCXIP-based expression) (Table S2). The products were analyzed on a 2.5% agarose gel and visualized by ethidium bromide staining. The sizes of FL ORF8 and SF1 ORF8 fragments amplified from RNA samples of the pcDNA-based expression were 678 and 552 bps, respectively, while those of the pQCXIP-based expression were 604 and 478 bps, respectively. For sequence determination, the RT-PCR products were purified using a DNA gel extraction kit (Qiagen) and subjected to Sanger sequencing using the primer sets listed in Table S2.

Immunoprecipitation assay

To enrich the ORF8 proteins in the culture sup, we performed an immunoprecipitation assay as previously described (59, 60). Briefly, 293T cells were transfected with an ORF8 plasmid containing a C-terminal mycFLAG tag. At 36 h posttransfection, the cells and the culture sup were harvested and then lysed in lysis buffer (150 mM NaCl, 1 mM EDTA, 1% Triton X-100, and 50 mM Tris-HCl pH 7.4). For the sup, cell debris was removed using a 0.45 μm filter. The ORF8 protein was immunoprecipitated with anti-FLAG M2 magnetic beads (Sigma-Aldrich) at 4 °C for 1 h. After the beads were washed with lysis buffer, the IP fractions were eluted competitively with 150 μg/ml of the 3× FLAG peptide (Sigma-Aldrich) and then analyzed by Western blotting. To estimate ratio (%) of the secreted ORF8 protein in total, the band intensity of the cell and sup fractions on one blotted membrane was quantified by ImageQuant TL 8.2 (Cytiva). The ratio (%) of the intensity of sup relative to that of sum (cell + sup) and the average value for five independent experiments was calculated.

Endoglycosidase analysis of ORF8 proteins

For Endo H digestion, eluted fractions of the immunoprecipitants from the cell lysate or culture sup were first diluted with 1× glycoprotein denaturing buffer (0.5% SDS and 40 mM dithiothreitol [DTT]) and boiled at 98 °C for 10 min. The denatured samples were divided into two halves. Half of the sample was diluted with 1× GlycoBuffer 3 (50 mM sodium acetate pH 6.0) (New England Biolabs) and then treated with Endo H (New England Biolabs) at 37 °C for 4.5 h. For PNGase F digestion, the other half of the samples were diluted with 1× GlycoBuffer 2 (50 mM sodium phosphate pH 6.0) and 1% NP-40 and then treated with PNGase F (New England Biolabs) at 37 °C for 4.5 h. Enzyme-treated samples were separated by 12% SDS-PAGE under reducing conditions and analyzed by Western blotting using the anti-FLAG tag mAb.

FACS analysis

For the analysis of cell surface proteins, cells were collected in phosphate-buffered saline (PBS) containing 1 mM EDTA and washed with PBS containing 2% (wt/vol) FBS. Surface proteins were stained for 30 min with monoclonal antibodies in 2% FBS-PBS on ice. After the cells were washed, the proteins were stained with a mouse F(ab)2 IgG (H + L) APC-conjugated

antibody (10 μl/10^6 cells; R&D Systems; F0101B) and LIVE/DEAD Fixable Viability Dye (0.25 μl/10^6 cells; Thermo Fisher Scientific; L34955) in 2% FBS-PBS on ice for 30 min. The levels of cell surface proteins were analyzed by flow cytometry using a FACS Canto II (BD Biosciences). The data were analyzed with FlowJo software (BD Biosciences). The following antibodies were used for FACS: anti-HLA-A2 mAb (BB7.2) (1 μg/10^6 cells; BioLegend; 343302), anti-SARS-CoV/SARS-CoV-2 spike mAb (1A9) (1 μg/10^6 cells), and anti-ACE2 mAb (1 μg/10^6 cells).

Preparation of human MDMs and treatment of ORF8

Human peripheral blood mononuclear cells (PBMCs), obtained by venipuncture from healthy donors, were separated by density centrifugation on Lymphoprep (Veritas) at 800g for 30 min. Monocytes were isolated from PBMCs using a Classical Monocyte Isolation Kit (Miltenyi Biotec) or EasySep Human Monocyte Isolation Kit (Veritas). CD14$^+$ cells were plated at 5×10^5 cells per well in 6-well plates in RPMI 1640 (Sigma-Aldrich) supplemented with penicillin and streptomycin for 3 h, followed by the addition of 10% FBS (Cytiva) and 10 ng/ml macrophage colony stimulating factor (PeproTech). Adherent cells were cultured for 7 days to enable differentiation into MDMs.

To produce cell sups containing secreted ORF8, at 16 h posttransfection, the culture sup was gently washed three times and replaced with fresh RPMI 1640 containing penicillin, streptomycin, and 10% FBS, followed by an additional 24 h incubation. The sups were harvested and clarified by centrifugation at 500g for 10 min. The culture sup was passed through a 0.45-μm-pore-size filter to remove cellular debris. The cell sup containing the secreted ORF8 protein was added to MDMs and incubated at 37 °C for 2 h.

RNA-seq

With 300 ng of total RNA from MDMs, an RNA-seq library was constructed using the MGIEasy RNA Directional Library Prep Set and MGIEasy Circularization Kit (MGI) according to the manufacturer's instructions. The library was loaded on a DNBSEQ-G400RS sequencing flow cell with a DNBSEQ-G400RS high-throughput sequencing kit (MGI) and run on DNBSEQ-G400 (2 × 100 bp). Sequencing data from DNBSEQ-G400 were analyzed using standard protocols. Briefly, raw reads were preprocessed using cutadapt 1.9.1 (61) and sickle 1.33 (https://github.com/najoshi/sickle) and then mapped to the human genome (GRCh38) using HISAT2 version 2.2.0 (62) and Samtools version v.1.10 (63). The gene levels of the obtained read counts were estimated with featureCounts version 2.2.0 (64).

Data availability

RNA seq data are archived at the BioProject of the DNA Data Bank of Japan (DDBJ) (https://www.ddbj.nig.ac.jp/index-e.html) under accession number PRJDB12786 (Dataset title, the effects of the SARS-CoV-2 ORF8 secretory

glycoprotein on human primary monocyte-derived macrophages (MDMs)).

Abbreviations—The abbreviations used are: 2-ME, 2-mercaptoethanol; ACE2, angiotensin-converting enzyme 2; CoV-2 S, SARS-CoV-2 spike; ER, endoplasmic reticulum; FL, full-length; IL17RA, interleukin(IL)-17 receptor A; MDM, monocyte-derived macrophage; MHC-I, major histocompatibility complex class I; ORF8, open reading frame 8; PBS, phosphate-buffered saline; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; sup, supernatant; WT, wild-type.

References

1. Boni, M. F., Lemey, P., Jiang, X., Lam, T. T., Perry, B. W., Castoe, T. A., Rambaut, A., and Robertson, D. L. (2020) Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nat. Microbiol.* **5**, 1408–1417
2. Chen, S., Zheng, X., Zhu, J., Ding, R., Jin, Y., Zhang, W., Yang, H., Zheng, Y., Li, X., and Duan, G. (2020) Extended ORF8 gene region is valuable in the epidemiological investigation of severe acute respiratory syndrome-similar coronavirus. *J. Infect. Dis.* **222**, 223–233
3. Chinese SARS Molecular Epidemiology Consortium (2004) Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. *Science* **303**, 1666–1669
4. Cui, J., Li, F., and Shi, Z. L. (2019) Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **17**, 181–192
5. Laha, S., Chakraborty, J., Das, S., Manna, S. K., Biswas, S., and Chatterjee, R. (2020) Characterizations of SARS-CoV-2 mutational profile, spike protein stability and viral transmission. *Infect. Genet. Evol.* **85**, 104445
6. Lau, S. K., Feng, Y., Chen, H., Luk, H. K., Yang, W. H., Li, K. S., Zhang, Y. Z., Huang, Y., Song, Z. Z., Chow, W. N., Fan, R. Y., Ahmed, S. S., Yeung, H. C., Lam, C. S., Cai, J. P., *et al.* (2015) Severe acute respiratory syndrome (SARS) coronavirus ORF8 protein is acquired from SARS-related coronavirus from Greater Horseshoe bats through recombination. *J. Virol.* **89**, 10532–10547
7. Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., Bi, Y., Ma, X., Zhan, F., Wang, L., Hu, T., *et al.* (2020) Genomic characterisation and epidemiology of 2019 novel coronavirus: Implications for virus origins and receptor binding. *Lancet* **395**, 565–574
8. Mohammad, S., Bouchama, A., Mohammad Alharbi, B., Rashid, M., Saleem Khatlani, T., Gaber, N. S., and Malik, S. S. (2020) SARS-CoV-2 ORF8 and SARS-CoV ORF8ab: Genomic divergence and functional convergence. *Pathogens* **9**, 677
9. Redondo, N., Zaldívar-López, S., Garrido, J. J., and Montoya, M. (2021) SARS-CoV-2 accessory proteins in viral pathogenesis: Knowns and unknowns. *Front. Immunol.* **12**, 708264
10. Wu, Z., Yang, L., Ren, X., Zhang, J., Yang, F., Zhang, S., and Jin, Q. (2016) ORF8-related genetic evidence for Chinese Horseshoe bats as the source of human severe acute respiratory syndrome coronavirus. *J. Infect. Dis.* **213**, 579–583
11. Young, B. E., Fong, S. W., Chan, Y. H., Mak, T. M., Ang, L. W., Anderson, D. E., Lee, C. Y., Amrun, S. N., Lee, B., Goh, Y. S., Su, Y. C. F., Wei, W. E., Kalimuddin, S., Chai, L. Y. A., Pada, S., *et al.* (2020) Effects of a major deletion in the SARS-CoV-2 genome on the severity of infection and the inflammatory response: An observational cohort study. *Lancet* **396**, 603–611
12. Su, Y. C. F., Anderson, D. E., Young, B. E., Linster, M., Zhu, F., Jayakumar, J., Zhuang, Y., Kalimuddin, S., Low, J. G. H., Tan, C. W., Chia, W. N., Mak, T. M., Octavia, S., Chavatte, J. M., Lee, R. T. C., *et al.* (2020) Discovery and genomic characterization of a 382-nucleotide deletion in ORF7b and ORF8 during the early evolution of SARS-CoV-2. *mBio* **11**, e01610-20
13. Gamage, A. M., Tan, K. S., Chan, W. O. Y., Liu, J., Tan, C. W., Ong, Y. K., Thong, M., Andiappan, A. K., Anderson, D. E., Wang, Y., and Wang, L. F. (2020) Infection of human Nasal Epithelial Cells with SARS-CoV-2 and a 382-nt deletion isolate lacking ORF8 reveals similar viral kinetics and host transcriptional profiles. *PLoS Pathog.* **16**, e1009130
14. Guan, Y., Zheng, B. J., He, Y. Q., Liu, X. L., Zhuang, Z. X., Cheung, C. L., Luo, S. W., Li, P. H., Zhang, L. J., Guan, Y. J., Butt, K. M., Wong, K. L., Chan, K. W., Lim, W., Shortridge, K. F., *et al.* (2003) Isolation and characterization of viruses related to the SARS coronavirus from animals in Southern China. *Science* **302**, 276–278
15. Oostra, M., de Haan, C. A., and Rottier, P. J. (2007) The 29-nucleotide deletion present in human but not in animal severe acute respiratory syndrome coronaviruses disrupts the functional expression of open reading frame 8. *J. Virol.* **81**, 13876–13888
16. Muth, D., Corman, V. M., Roth, H., Binger, T., Dijkman, R., Gottula, L. T., Gloza-Rausch, F., Balboni, A., Battilani, M., Rihtarič, D., Toplak, I., Ameneiros, R. S., Pfeifer, A., Thiel, V., Drexler, J. F., *et al.* (2018) Attenuation of replication by a 29 nucleotide deletion in SARS-coronavirus acquired during the early stages of human-to-human transmission. *Sci. Rep.* **8**, 15177
17. Davies, N. G., Abbott, S., Barnard, R. C., Jarvis, C. I., Kucharski, A. J., Munday, J. D., Pearson, C. A. B., Russell, T. W., Tully, D. C., Washburne, A. D., Wenseleers, T., Gimma, A., Waites, W., Wong, K. L. M., van Zandvoort, K., *et al.* (2021) Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* **372**, eabg3055
18. Davies, N. G., Jarvis, C. I., CMMID COVID-19 Working Group, Edmunds, W. J., Jewell, N. P., Diaz-Ordaz, K., and Keogh, R. H. (2021) Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7. *Nature* **593**, 270–274
19. Vöhringer, H. S., Sanderson, T., Sinnott, M., De Maio, N., Nguyen, T., Goater, R., Schwach, F., Harrison, I., Hellewell, J., Ariani, C. V., Gonçalves, S., Jackson, D. K., Johnston, I., Jung, A. W., Saint, C., *et al.* (2021) Genomic reconstruction of the SARS-CoV-2 epidemic in England. *Nature* **600**, 506–511
20. Jungreis, I., Sealfon, R., and Kellis, M. (2021) SARS-CoV-2 gene content and COVID-19 mutation impact by comparing 44 Sarbecovirus genomes. *Nat. Commun.* **12**, 2642
21. Pereira, F. (2021) SARS-CoV-2 variants lacking ORF8 occurred in farmed mink and pangolin. *Gene* **784**, 145596
22. Wang, X., Lam, J. Y., Wong, W. M., Yuen, C. K., Cai, J. P., Au, S. W., Chan, J. F., To, K. K. W., Kok, K. H., and Yuen, K. Y. (2020) Accurate diagnosis of COVID-19 by a novel immunogenic secreted SARS-CoV-2 orf8 protein. *mBio* **11**, e02431-20

23. Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., O'Meara, M. J., Rezelj, V. V., Guo, J. Z., Swaney, D. L., Tummino, T. A., Huttenhain, R., Kaake, R. M., Richards, A. L., Tutuncuoglu, B., *et al.* (2020) A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468

24. Gordon, D. E., Hiatt, J., Bouhaddou, M., Rezelj, V. V., Ulferts, S., Braberg, H., Jureka, A. S., Obernier, K., Guo, J. Z., Batra, J., Kaake, R. M., Weckstein, A. R., Owens, T. W., Gupta, M., Pourmal, S., *et al.* (2020) Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* **370**, eabe9403

25. Stukalov, A., Girault, V., Grass, V., Karayel, O., Bergant, V., Urban, C., Haas, D. A., Huang, Y., Oubraham, L., Wang, A., Hamad, M. S., Piras, A., Hansen, F. M., Tanzer, M. C., Paron, I., *et al.* (2021) Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV. *Nature* **594**, 246–252

26. Rashid, F., Dzakah, E. E., Wang, H., and Tang, S. (2021) The ORF8 protein of SARS-CoV-2 induced endoplasmic reticulum stress and mediated immune evasion by antagonizing production of interferon beta. *Virus Res.* **296**, 198350

27. Echavarria-Consuegra, L., Cook, G. M., Busnadiego, I., Lefevre, C., Keep, S., Brown, K., Doyle, N., Dowgier, G., Franaszek, K., Moore, N. A., Siddell, S. G., Bickerton, E., Hale, B. G., Firth, A. E., Brierley, I., *et al.* (2021) Manipulation of the unfolded protein response: A pharmacological strategy against coronavirus infection. *PLoS Pathog.* **17**, e1009644

28. Li, J. Y., Liao, C. H., Wang, Q., Tan, Y. J., Luo, R., Qiu, Y., and Ge, X. Y. (2020) The ORF6, ORF8 and nucleocapsid proteins of SARS-CoV-2 inhibit type I interferon signaling pathway. *Virus Res.* **286**, 198074

29. Geng, H., Subramanian, S., Wu, L., Bu, H. F., Wang, X., Du, C., De Plaen, I. G., and Tan, X. D. (2021) SARS-CoV-2 ORF8 forms intracellular aggregates and inhibits IFNgamma-induced antiviral gene expression in human lung epithelial cells. *Front. Immunol.* **12**, 679482

30. Zhang, Y., Chen, Y., Li, Y., Huang, F., Luo, B., Yuan, Y., Xia, B., Ma, X., Yang, T., Yu, F., Liu, J., Liu, B., Song, Z., Chen, J., Yan, S., *et al.* (2021) The ORF8 protein of SARS-CoV-2 mediates immune evasion through down-regulating MHC-I. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2024202118

31. Lin, X., Fu, B., Yin, S., Li, Z., Liu, H., Zhang, H., Xing, N., Wang, Y., Xue, W., Xiong, Y., Zhang, S., Zhao, Q., Xu, S., Zhang, J., Wang, P., *et al.* (2021) ORF8 contributes to cytokine storm during SARS-CoV-2 infection by activating IL-17 pathway. *iScience* **24**, 102293

32. Flower, T. G., Buffalo, C. Z., Hooy, R. M., Allaire, M., Ren, X., and Hurley, J. H. (2021) Structure of SARS-CoV-2 ORF8, a rapidly evolving immune evasion protein. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2021785118

33. Michelsen, K., Yuan, H., and Schwappach, B. (2005) Hide and run. Arginine-based endoplasmic-reticulum-sorting motifs in the assembly of heteromultimeric membrane proteins. *EMBO Rep.* **6**, 717–722

34. Schwartz, O., Maréchal, V., Le Gall, S., Lemonnier, F., and Heard, J. M. (1996) Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. *Nat. Med.* **2**, 338–342

35. Coscoy, L., and Ganem, D. (2000) Kaposi's sarcoma-associated herpesvirus encodes two proteins that block cell surface display of MHC class I chains by enhancing their endocytosis. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8051–8056

36. Collins, K. L., Chen, B. K., Kalams, S. A., Walker, B. D., and Baltimore, D. (1998) HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. *Nature* **391**, 397–401

37. Piguet, V., Schwartz, O., Le Gall, S., and Trono, D. (1999) The down-regulation of CD4 and MHC-I by primate lentiviruses: A paradigm for the modulation of cell surface receptors. *Immunol. Rev.* **168**, 51–63

38. Malim, M. H., Hauber, J., Fenrick, R., and Cullen, B. R. (1988) Immunodeficiency virus rev trans-activator modulates the expression of the viral regulatory genes. *Nature* **335**, 181–183

39. Basyuk, E., Boulon, S., Skou Pedersen, F., Bertrand, E., and Vestergaard Rasmussen, S. (2005) The packaging signal of MLV is an integrated module that mediates intracellular transport of genomic RNAs. *J. Mol. Biol.* **354**, 330–339

40. Smagulova, F., Maurel, S., Morichaud, Z., Devaux, C., Mougel, M., and Houzet, L. (2005) The highly structured encapsidation signal of MuLV RNA is involved in the nuclear export of its unspliced RNA. *J. Mol. Biol.* **354**, 1118–1128

41. Dorner, A. J., Wasley, L. C., and Kaufman, R. J. (1992) Overexpression of GRP78 mitigates stress induction of glucose regulated proteins and blocks secretion of selective proteins in Chinese hamster ovary cells. *EMBO J.* **11**, 1563–1571

42. Melnick, J., Aviel, S., and Argon, Y. (1992) The endoplasmic reticulum stress protein GRP94, in addition to BiP, associates with unassembled immunoglobulin chains. *J. Biol. Chem.* **267**, 21303–21306

43. Li, W. W., Hsiung, Y., Zhou, Y., Roy, B., and Lee, A. S. (1997) Induction of the mammalian GRP78/BiP gene by Ca2+ depletion and formation of aberrant proteins: Activation of the conserved stress-inducible grp core promoter element by the human nuclear factor YY1. *Mol. Cell. Biol.* **17**, 54–60

44. McLaughlin, M., and Vandenbroeck, K. (2011) The endoplasmic reticulum protein folding factory and its chaperones: New targets for drug discovery? *Br. J. Pharmacol.* **162**, 328–345

45. Halperin, L., Jung, J., and Michalak, M. (2014) The many functions of the endoplasmic reticulum chaperones and folding enzymes. *IUBMB Life* **66**, 318–326

46. Ninagawa, S., George, G., and Mori, K. (2021) Mechanisms of productive folding and endoplasmic reticulum-associated degradation of glycoproteins and non-glycoproteins. *Biochim. Biophys. Acta Gen. Subj.* **1865**, 129812

47. Peterson, P. A., Rask, L., and Lindblom, J. B. (1974) Highly purified papain-solubilized HL-A antigens contain beta2-microglobulin. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 35–39

48. Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O., Graham, B. S., and McLellan, J. S. (2020) Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263

49. Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., and Zhou, Q. (2020) Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444–1448

50. Yao, Z., Fanslow, W. C., Seldin, M. F., Rousseau, A. M., Painter, S. L., Comeau, M. R., Cohen, J. I., and Spriggs, M. K. (1995) Herpesvirus Saimiri encodes a new cytokine, IL-17, which binds to a novel cytokine receptor. *Immunity* **3**, 811–821

51. Lucas, A., and McFadden, G. (2004) Secreted immunomodulatory viral proteins as novel biotherapeutics. *J. Immunol.* **173**, 4765–4774

52. Sungnak, W., Huang, N., Bécavin, C., Berg, M., Queen, R., Litvinukova, M., Talavera-López, C., Maatz, H., Reichart, D., Sampaziotis, F., Worlock, K. B., Yoshida, M., Barnes, J. L., and HCA Lung Biological Network (2020) SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat. Med.* **26**, 681–687

53. Chen, M., Shen, W., Rowan, N. R., Kulaga, H., Hillel, A., Ramanathan, M., Jr., and Lane, A. P. (2020) Elevated ACE-2 expression in the olfactory neuroepithelium: Implications for anosmia and upper respiratory SARS-CoV-2 entry and replication. *Eur. Respir. J.* **56**, 2001948

54. Shuai, H., Chan, J. F., Hu, B., Chai, Y., Yuen, T. T., Yin, F., Huang, X., Yoon, C., Hu, J. C., Liu, H., Shi, J., Liu, Y., Zhu, T., Zhang, J., Hou, Y., *et al.* (2022) Attenuated replication and pathogenicity of SARS-CoV-2 B.1.1. 529 Omicron. *Nature*. https://doi.org/10.1038/s41586-022-04442-5

55. Wolter, N., Jassat, W., Walaza, S., Welch, R., Moultrie, H., Groome, M., Amoako, D. G., Everatt, J., Bhiman, J. N., Scheepers, C., Tebeila, N., Chiwandire, N., du Plessis, M., Govender, N., Ismail, A., *et al.* (2022) Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: A data linkage study. *Lancet* **399**, 437–446

56. [preprint] Lin, X., Fu, B., Xiong, Y., Xing, N., Xue, W., Guo, D., Zaky, M. Y., Pavani, K. C., Kunec, D., Trimpert, J., and Wu, H. (2021) Unconventional secretion of unglycosylated ORF8 is critical for the cytokine storm during SARS-CoV-2 infection. *bioRxiv*. https://doi.org/10.1101/2021.12.03.471057

57. Mwimanzi, F., Toyoda, M., Mahiti, M., Mann, J. K., Martin, J. N., Bangsberg, D., Brockman, M. A., Goulder, P., Kirchhoff, F., Brumme, Z. L., Ndung'u, T., and Ueno, T. (2018) Resistance of major histocompatibility complex class B (MHC-B) to Nef-mediated downregulation relative to that of MHC-A is conserved among primate lentiviruses and influences antiviral T cell responses in HIV-1-infected individuals. *J. Virol.* **92**, e01409-17

58. Ozono, S., Zhang, Y., Ode, H., Sano, K., Tan, T. S., Imai, K., Miyoshi, K., Kishigami, S., Ueno, T., Iwatani, Y., Suzuki, T., and Tokunaga, K. (2021) SARS-CoV-2 D614G spike mutation increases entry efficiency with enhanced ACE2-binding affinity. *Nat. Commun.* **12**, 848

59. Kitamura, S., Ode, H., Nakashima, M., Imahashi, M., Naganawa, Y., Kurosawa, T., Yokomaku, Y., Yamane, T., Watanabe, N., Suzuki, A., Sugiura, W., and Iwatani, Y. (2012) The APOBEC3C crystal structure and the interface for HIV-1 Vif binding. *Nat. Struct. Mol. Biol.* **19**, 1005–1010

60. Nakashima, M., Ode, H., Kawamura, T., Kitamura, S., Naganawa, Y., Awazu, H., Tsuzuki, S., Matsuoka, K., Nemoto, M., Hachiya, A., Sugiura, W., Yokomaku, Y., Watanabe, N., and Iwatani, Y. (2016) Structural insights into HIV-1 Vif-APOBEC3F interaction. *J. Virol.* **90**, 1034–1047

61. Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12

62. Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915

63. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079

64. Liao, Y., Smyth, G. K., and Shi, W. (2014) featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930

Kazuhiro Matsuoka received his Ph.D at Ehime University in Japan and is now a research fellow in the Iwatani lab at National Hospital Organization Nagoya Medical Center. He studies on biochemical and virological functions of an HIV-accessory protein Vif and the host antiviral APOBEC3 enzymes. During the COVID-19 pandemic, he applied his knowledge to discover the functional roles of the SARS-CoV-2 accessory protein ORF8 as well.