Reduced Incorporation of SARS-CoV Spike Protein into Viral Particles Due to Amino Acid Substitutions within the Receptor Binding Domain

Shu-Ming Li1,2, Gui-Mei Li2, Shota Nakamura1, Kazuyoshi Ikuta2,3 and Takaaki Nakaya1*

1International Research Center for Infectious Diseases and 2Department of Virology, Research Institute for Microbial Diseases, Osaka University, Osaka 565-0871, Japan, and 3Thailand-Japan Research Collaboration Center on Emerging and Re-emerging Infections, Nonthaburi, Thailand

(Received December 6, 2007. Accepted January 25, 2008)

SUMMARY: Cell clone #21 is a long-term producer of the infectious SARS-coronavirus, although the incorporation rate of spike (S) protein into virions is significantly lower. Sequencing analysis of the viral structural proteins revealed four and one amino acid substitutions in the S and membrane (M) proteins, respectively. We demonstrated, using a viral-like particle formation system, that the S mutations were involved in the lower incorporation of the S protein into virions, although the M mutation that disrupts the glycosylation was not present in this phenotype. Further mutational experiments identified two substitutions, Y442C and L472F, within the receptor binding domain that could be critical for the reduced S incorporation, as well as reduced binding affinity between the S protein and ACE2 receptor. Thus, these two amino acid substitutions might lead to a conformational change in the S protein, resulting in reduced incorporation into viral particles.

INTRODUCTION

The severe acute respiratory syndrome (SARS)-coronavirus (CoV) was identified as the etiological agent for an acute infectious respiratory disorder (1). SARS-CoV is an enveloped, positive-strand RNA virus with a ~30 kb genome that encodes replicase (1a and 1b), spike (S), envelope (E), membrane (M), nucleocapsid (N) and several accessory proteins (2,3). The virus particle consists of four structural components which are S, E, M and N proteins. The generation of viral-like particles (VLPs) bearing four structural proteins has been reported (4,5). It was demonstrated that the S, M and N proteins were necessary for pseudovirus assembly (4). In addition, it was established using a VLP system that the N protein played an essential role on the packaging of SARS-CoV RNA (5). Thus, the VLP system is very useful for understanding virion assembly and morphology at the molecular level.

S protein is a type I integral membrane glycoprotein that makes up the crown-like appearance of the viral particles (6). It was reported that angiotensin-converting enzyme 2 (ACE2) is the functional receptor for SARS-CoV (7). The interaction of the receptor-binding domain (RBD) of S protein with ACE2 was demonstrated by crystal structure analysis (8). Further studies showed that the RBD contains major determinants for viral entry and neutralization (9,10).

We previously reported that a total of four (#13, #18, #21 and #34) of 87 cell clones isolated from persistently infected cells were shown to be viral RNA-positive (11). However, several passages of subsequent culturing cleared the viral RNA from cell clones #13, #18 and #34, and only #21 was a long-term producer of infectious viral particles at a high rate for more than one year (11). Interestingly, electron-microscopic as well as Western blotting analyses demonstrated that the incorporated number of S proteins on the viral particles from #21 cells was lower than that from acutely infected cells (11). Sequencing analysis of the viral genome of #21-derived virus revealed several amino acid substitutions in S and M genes (GenBank accession no. AB257344). In this study, we focused on the relation of the amino acid substitutions in the structural proteins with the reduced S incorporation into virions in #21.

MATERIALS AND METHODS

Virus and cells: The Vero E6 cell line was used for the propagation of SARS-CoV (Frankfurt-1 strain) (12). A #21 cell clone obtained from Vero E6 cells infected with the Frankfurt-1 strain was previously prepared (11). Vero E6, #21 and HEp-2 cells were maintained in MEM (GIBCO BRL, Carlsbad, Calif., USA) supplemented with 10% fetal bovine serum (ICN Flow), 100 U/ml penicillin and 100 μg/ml streptomycin (GIBCO BRL) and passed every 3 days.

Reverse transcription (RT)-polymerase chain reaction (PCR) for SARS-CoV genome RNA: Total RNA was extracted from #21 cells with TRIzol (Invitrogen, Carlsbad, Calif., USA) and subjected to RT-PCR. The RNA was reverse-transcribed with SuperScript III reverse transcriptase (Invitrogen) using random primers (Invitrogen). The PCR amplification was carried out as described previously (11).

Direct sequencing of PCR products: To perform genome sequencing, ~3 kb DNA fragments were obtained by RT-PCR. The PCR products were sequenced with a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, Calif., USA) according to the manufacturer’s protocol.

Western blotting: The cell fractions were subjected to Western blotting and probed with an anti-V5 monoclonal antibody (Invitrogen) for 1 h at 37°C, then with horseradish peroxidase-conjugated anti-mouse secondary antibody for 1 h at 37°C (Jackson ImmunoResearch Laboratories, West Grove, Pa., USA).

Generation of VLPs by co-expression with S, E, M and...
N-expressing plasmids: The whole ORF region without the termination sequences of S, E, M and N were amplified using RT-PCR (S: forward primer 5'-GAATTCTCATGTTTTA TTTCTTATATTTC-3' and reverse primer 5'-TCTAGATG TGATAATGTTGACACC-3'; E: forward primer 5'-GA ATTCATGCTACATTGTTCGGA-3' and reverse primer 5'-TCTAGAGACCGAAGATCAGGAAGCTC-3'; M: forward primer 5'-GAATTCTAGGCAAGACACCGG-3' and reverse primer 5'-CTCGAGCAGCTGAGAACGAAGATG-3'; N: forward primer 5'-GAATTCATGCTAAGATGCGTTGAGCACC-3' and reverse primer 5'-TCTAGAGACCGAAGATCAGGAAGCTC-3'). The PCR products of S, E and N were cloned into the EcoRI-XbaI sites of the pcDNA3.1/V5-His expression plasmid (Invitrogen). The PCR product of M was cloned into the EcoRI-XhoI sites of the same expression plasmid. These plasmids express C-terminal V5-His fusion proteins of S, E, M or N. The wild-type M (MWT) and S (S21) genes were prepared from Vero E6 cells acutely infected with wild-type virus. Mutant M (M#21: T6A) and S (S21: Y442C, L472F, V594F and P794S) genes were prepared from #21. Four types of chimeric S protein (Smut1: Y442C and L472F, Smut2: V594F and P794S, Smut3: Y442C and L472F) were generated by the GeneTailor\textsuperscript{TM} Site-Directed Mutagenesis System (Invitrogen) according to the manufacturer\textquotesingle s protocol. The E and N genes with the termination sequence were also PCR-amplified and cloned into pEF-BOS expression vector (13), and the M gene with the termination sequence was cloned into pcDNA3.1 vector (Invitrogen). A monolayer of Vero E6 cells in a 100-mm dish was infected with recombinant vaccinia virus expressing T7 RNA polymerase (14) at a multiplicity of infection (moi) of 1 for 1 h and were transfected with S, E, M and N expression plasmids (6 µg in each) using Lipofectamine 2000 (Invitrogen). The transfected cells were cultured for 4 days and were harvested for Western blotting analysis using an anti-V5 monoclonal antibody. SARS-CoV VLPs were partially purified from the culture supernatant of the transfected cells by ultracentrifugation through 20% sucrose in PBS with a Beckmann SW28 rotor at 25,000 rpm for 2 h at 4°C. VLPs collected in the pellet were resuspended in 100 µl of PBS for Western blotting with the anti-V5 monoclonal antibody. To separate various VLPs with different compositions, the virus suspension was further loaded on a discontinuous sucrose gradient consisting of 20, 30, 50 and 60% sucrose as described by Hsiiesh et al. (5) and then centrifuged in a Beckmann SW28 rotor at 25,000 rpm for 2 h at 4°C. VLPs and the deleted nucleotides were found to be inserted into ORF 7a and 7b (nt 27,886-28,103) with recombinant vaccinia virus expressing T7 RNA polymerase at an moi of 1 for 1 h, and were transfected with WT-M and WT-S expression plasmids using Lipofectamine 2000. The proteins were detected by Western blotting with an anti-V5 antibody. For the expression of S proteins, Vero E6 cells were infected with recombinant vaccinia virus expressing T7 RNA polymerase at an moi of 1 for 1 h, and were transfected with S-WT, S21 and Smut1-4 expression plasmids using Lipofectamine 2000. The reactivity levels of an anti-S monoclonal antibody 3A2 (11) with these proteins were compared by Western blotting.

**Immunofluorescence assay (IFA):** Fixed cells with 2% paraformaldehyde were permeabilized by treatment with 0.1% Triton X-100. Cells were reacted with an anti-V5 anti-body for 30 min at 37°C, then incubated with fluorescein isothiocyanate-conjugated anti-mouse IgG (Jackson Immuno-Research Laboratories) for 45 min at 37°C and observed by confocal microscopy (FLUOVIEW FV1000; Olympus, Tokyo, Japan).

### Structural analysis of S_{WT} and S_{21} proteins using a computer program
The RBD of the S_{WT} protein with the substitutions Y442C and L472F was modeled by replacing the side-chains of S_{WT} protein (PDB ID: 2AJF) using PyMOL v.0.99rc6 (http://www.pymol.org).

### RESULTS

**Sequencing analysis of SARS-CoV S, E, M and N genes in the #21 cell clone:** The SARS-CoV genome in #21 was independently amplified at the S, E, M and N genes by RT-PCR, and individual PCR products were subjected to direct sequencing. In total, six amino acid substitutions were identified: Y442C, L472F, V594F, H641Y and P794S in S and T6A in M, as summarized in Fig. 1A. Since H641Y was also detected in the original wild-type virus (Frankfurt-1 strain) from acutely infected cells (data not shown), the other four substitutions in S could be present in the phenotype of the #21-derived virus. One amino acid substitution, T6A at the N-terminus of the M protein, resulted in disruption of the putative N-glycosylation site (NGTG, to tNGA). In contrast, there were no mutations in the E and N proteins (data not shown). All of the mutation sites that were found in the #21-derived virus are summarized in Table 1.

**Formation of VLPs with S, E, M and N proteins:** We previously reported that a smaller amount of S glycoprotein was incorporated into #21-derived virions in comparison with the wild-type virus, as shown by electron-microscopic and Western blotting analyses (11). Therefore, we examined whether the above-described mutations were related to the lower incorporation of S protein into virions using a VLP system co-expressing S, E, M and N proteins. Reconstitution analysis using VLPs confirmed that a smaller amount of S protein was incorporated into the #21 VLPs than into the wild-type VLPs, although the intracellular expression level of each structural protein was comparable (Figs. 1B and 1C). In addition, since the amino acid substitution T6A in the M protein abolished the N-glycosylation in the M protein as determined by Western blotting (Fig. 2A), we next investigated whether this mutation was also involved in the reduced incorporation of the S protein into virions. However, there was no difference in the amount of S protein incorporated into VLPs between the M21 and MWT proteins (Fig. 2B). In contrast to this finding with the M protein, when S21 was co-expressed with the MWT, E and N proteins, the level of S incorporation was reduced to a level similar to that of the #21 VLPs (Fig. 2B). We therefore concluded that a mutation(s) in the S protein determines the reduced incorporation of S protein into virions. Further mutational analysis of the S gene

**Table 1. Summary of amino acid substitutions in #21-derived virus**

<table>
<thead>
<tr>
<th>ORF</th>
<th>Amino acids</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF1a</td>
<td>4 amino acids</td>
</tr>
<tr>
<td>L2430F</td>
<td>S321IL, D359A, M3964V</td>
</tr>
<tr>
<td>S protein</td>
<td>5 amino acids</td>
</tr>
<tr>
<td>Y442C, L472F, V594F, H641Y, P794S</td>
<td></td>
</tr>
<tr>
<td>M protein</td>
<td>1 amino acid</td>
</tr>
<tr>
<td>T6A</td>
<td></td>
</tr>
<tr>
<td>MWT-ORF8</td>
<td>218-t deletion and insertion</td>
</tr>
</tbody>
</table>

1) Deletion of 218-nt (nt 27,886-28,103) was found in ORF 8a and 8b and the deleted nucleotides were found to be inserted into ORF 7a (nt 27,571-27,572).

---

**RESULTS**

**Sequencing analysis of SARS-CoV S, E, M and N genes in the #21 cell clone:** The SARS-CoV genome in #21 was independently amplified at the S, E, M and N genes by RT-PCR, and individual PCR products were subjected to direct sequencing. In total, six amino acid substitutions were identified: Y442C, L472F, V594F, H641Y and P794S in S and T6A in M, as summarized in Fig. 1A. Since H641Y was also detected in the original wild-type virus (Frankfurt-1 strain) from acutely infected cells (data not shown), the other four substitutions in S could be present in the phenotype of the #21-derived virus. One amino acid substitution, T6A at the N-terminus of the M protein, resulted in disruption of the putative N-glycosylation site (N GTG, to tNGA). In contrast, there were no mutations in the E and N proteins (data not shown). All of the mutation sites that were found in the #21-derived virus are summarized in Table 1.

**Formation of VLPs with S, E, M and N proteins:** We previously reported that a smaller amount of S glycoprotein was incorporated into #21-derived virions in comparison with the wild-type virus, as shown by electron-microscopic and Western blotting analyses (11). Therefore, we examined whether the above-described mutations were related to the lower incorporation of S protein into virions using a VLP system co-expressing S, E, M and N proteins. Reconstitution analysis using VLPs confirmed that a smaller amount of S protein was incorporated into the #21 VLPs than into the wild-type VLPs, although the intracellular expression level of each structural protein was comparable (Figs. 1B and 1C). In addition, since the amino acid substitution T6A in the M protein abolished the N-glycosylation in the M protein as determined by Western blotting (Fig. 2A), we next investigated whether this mutation was also involved in the reduced incorporation of the S protein into virions. However, there was no difference in the amount of S protein incorporated into VLPs between the M21 and MWT proteins (Fig. 2B). In contrast to this finding with the M protein, when S21 was co-expressed with the MWT, E and N proteins, the level of S incorporation was reduced to a level similar to that of the #21 VLPs (Fig. 2B). We therefore concluded that a mutation(s) in the S protein determines the reduced incorporation of S protein into virions. Further mutational analysis of the S gene

**Table 1. Summary of amino acid substitutions in #21-derived virus**

<table>
<thead>
<tr>
<th>ORF</th>
<th>Amino acids</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF1a</td>
<td>4 amino acids</td>
</tr>
<tr>
<td>L2430F</td>
<td>S321IL, D359A, M3964V</td>
</tr>
<tr>
<td>S protein</td>
<td>5 amino acids</td>
</tr>
<tr>
<td>Y442C, L472F, V594F, H641Y, P794S</td>
<td></td>
</tr>
<tr>
<td>M protein</td>
<td>1 amino acid</td>
</tr>
<tr>
<td>T6A</td>
<td></td>
</tr>
<tr>
<td>MWT-ORF8</td>
<td>218-t deletion and insertion</td>
</tr>
</tbody>
</table>

1) Deletion of 218-nt (nt 27,886-28,103) was found in ORF 8a and 8b and the deleted nucleotides were found to be inserted into ORF 7a (nt 27,571-27,572).
showed that the first two mutations (Smut1: Y442C and L472F) were critical for determining the incorporation efficiency (Fig. 2C). However, a similar lower level of incorporation of S protein into virions was also observed in the single amino acid mutation (Smut3: Y442C and Smut4: L472F) (Fig. 2C). Thus, the incorporation of S protein into virions was shown to be critically affected by two amino acids at residues 442 and 472. We also investigated the distribution of intracellular S protein in SWT- or S#21-transfected cells together with M, N and E expression vectors. As shown in Fig. 3, S WT protein...
was located mainly in the plasma membrane at 48 h post-transfection, whereas S<sub>21</sub> protein was detected not only in the plasma membrane but also in the cytoplasmic region. This tendency was also observed at 72 h post-transfection (Fig. 3).

**Binding affinity of mutant S proteins with receptor ACE2, as well as with an anti-S neutralizing antibody:** Structural analysis of the S protein using a computer program indicated that Y442C and L472F substitutions might reduce the binding affinity between the S protein and receptor ACE2 (Fig. 4A). In particular, Y442C substitution might change the orientation of the side-chain, leading to weak S-ACE2 (His<sup>34</sup>) binding (Fig. 4A). In addition, extra Cys<sup>442</sup> may generate a novel S-S bond in the S protein, resulting in drastic conformational change. Furthermore, it has been observed using another computer model that a naturally occurring L472P mutation may contribute to the attenuation by reducing the S-ACE2 contact surface (8). Further structural studies are necessary to confirm this prediction.

It was reported that the epitope on the RBD that bound to the anti-S human monoclonal antibody (80R) overlaps very closely with the ACE2 binding site, and further, that the S-ACE2 and S-80R interfaces share many common S amino acid residues, including Tyr<sup>442</sup> and Leu<sup>472</sup> (15). In addition, the overall structure of the other neutralizing antibody (m396)-RBD interface was also shown to be not significantly different from that of the ACE2-RBD interface (16). Neutralizing determinants of the m396 antibody are located contiguously in one major segment of the β6-β7 loop, whereas receptor ACE2 has determinants over most of the same extended loop appearing at the top of the RBD (16). Thus, we tried to investigate the binding affinity of mutant S proteins with an anti-S monoclonal antibody. The result showed that the anti-S monoclonal antibody (3A2), possessing neutralizing activity (Fig. 4B), reduced the binding efficiency to the S<sub>21</sub> protein by Western blotting (Fig. 4C). By using four types of chimeric proteins (Sm<sub>mut1</sub> to Sm<sub>mut4</sub>), we demonstrated that this antibody recognized an epitope containing Tyr<sup>442</sup> and that the substitution of Tyr by Cys reduced the binding efficiency to 3A2 (Fig. 4C). These results suggest that a conformational change in the RBD caused by the Y442C substitution led to a reduction of the binding affinity to this anti-S neutralizing antibody, which is consistent with the data obtained using the above computer analysis (Fig. 4A).

**DISCUSSION**

According to the genome sequence data of the structural genes S, E, M and N in #21-derived SARS-CoV, we identified a total of four and one amino acid substitutions in the S and M proteins, respectively, and no mutation in the N and E proteins. Among them, two amino acid substitutions, Y442C and L472F, of the S protein were shown to be related to the
reduced incorporation of this protein into virions that could affect the #21 phenotype, i.e., the reduced affinities to the receptor ACE2 and anti-S-neutralizing antibody, and the different intracellular distribution. The SARS-CoV M protein is exclusively N-glycosylated in asparagine at amino acid residue 4, and a nonglycosylated M is selectively incorporated into virions (17). It was also shown that glycosylation of the CoV M protein is not required for virus assembly, nor for the interaction between M and S proteins to occur (18). Our results obtained using the VLP system were consistent with the above observations, although the function of the glycosylation of the SARS-CoV M protein is still not known, as has been described (18).

Mutational analysis using the VLP system demonstrated that the S protein itself determined the efficiency of its incorporation into the pseudoparticles, as shown by the significant effect on the incorporation efficiency of the S protein into virions by amino acid substitutions Y442C and L472F. It is not clear why a small amount of N protein was detected in VLPs in our system. Extra amino acids, including a V5-tag at the C terminus of the N protein, may be one reason for the above. Otherwise, it is possible that the V5-tag of the M protein affects the N protein assembly.

Interestingly, Tyr442 and Leu472 were shown to be located in the RBD (8). Therefore, the Tyr442 to Cys442 substitution in the S21 protein might affect the RBD structure in #21-derived viral particles. In fact, the substitution impaired the binding affinity of S21 protein and the anti–S neutralizing monoclonal antibody (3A2). In addition, the Leu472 to Pro472 substitution often appeared in SARS-CoV isolates from patients during the 2003-2004 outbreak, and this substitution was shown by conformational analysis to contribute to the viral attenuation by reducing the S-ACE2 contact surface (8). Consequently, the two amino acid substitutions Y442C and L472F identified in #21 in this study could alter the antigenic structure, as well as the S-ACE2 binding affinity. However, the molecular mechanism by which these two substitutions lead to reduced incorporation of S protein into virions remains unclear. As for influenza virus HA protein, a single amino acid substitution at HA1 resulted in the reduction of and slowed rates of transport to the medial Golgi apparatus and cell surface due to partial impairment of folding (19). In this study, the intracellular location of S of strain #21 was slightly different from that of S (Fig. 3). Thus, the two amino acid substitutions found in #21-derived SARS-CoV may have led to improper conformation of the S protein in the endoplasmic reticulum/Golgi network, and this could have contributed to the decrease in S proteins on the cell surface and viral membrane.

Based on this evidence, the modification of SARS-CoV S structure may be related to the smaller amount of S protein in the released virions. It has been reported that M protein plays a pivotal role in virion assembly through its ability to interact with S protein (4). Although an average of 65 trimeric spikes per SARS-CoV particle was estimated by cryo-electron microscopy (20), the detailed machinery of SARS-CoV S protein incorporation into particles remains to be clarified.

ACKNOWLEDGMENTS

We are grateful to D. John Ziebuhr, University of Würzburg, Germany for giving us the Frankfurt strain of SARS-CoV through Dr. Fumihiro Taguchi, National Institute of Infectious Diseases, Tokyo, Japan. We thank Dr. Makiko Yamashita and Mr. Masanobu Yamate for helpful discussions.

This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), the Program of the Founding Research Center for Emerging and Reemerging Infectious Diseases launched by a project commissioned by the MEXT (K. I. and T. N.), and the project for the International Research Center for Infectious Diseases, Research Institute for Microbial Diseases, Osaka University from the MEXT of Japan (T. N.).

REFERENCES