

Synthesis and Characterization of a Native, Oligomeric Form of Recombinant Severe Acute Respiratory Syndrome Coronavirus Spike Glycoprotein

Hyun Chul Song,¹ Mi-Young Seo,¹ Konrad Stadler,² Byoung J. Yoo,³ Qui-Lim Choo,¹ Stephen R. Coates,¹ Yasushi Uematsu,² Takashi Harada,² Catherine E. Greer,¹ John M. Polo,¹ Piero Pileri,² Markus Eickmann,⁴ Rino Rappuoli,^{1,2} Sergio Abrignani,² Michael Houghton,¹ and Jang H. Han^{1*}

*Vaccines Research, Chiron Corporation, Emeryville, California*¹; *IRIS, Chiron Vaccines, Siena, Italy*²; *Division of Natural Sciences, Daegu University, Kyungbuk, Korea*³; and *Institute for Virology, University of Marburg, Marburg, Germany*⁴

Received 30 March 2004/Accepted 18 May 2004

We have expressed and characterized the severe acute respiratory syndrome coronavirus (SARS-CoV) spike protein in cDNA-transfected mammalian cells. The full-length spike protein (S) was newly synthesized as an endoglycosidase H (endo H)-sensitive glycoprotein (gp170) that is further modified into an endo H-resistant glycoprotein (gp180) in the Golgi apparatus. No substantial proteolytic cleavage of S was observed, suggesting that S is not processed into head (S1) and stalk (S2) domains as observed for certain other coronaviruses. While the expressed full-length S glycoprotein was exclusively cell associated, a truncation of S by excluding the C-terminal transmembrane and cytoplasmic tail domains resulted in the expression of an endoplasmic reticulum-localized glycoprotein (gp160) as well as a Golgi-specific form (gp170) which was ultimately secreted into the cell culture medium. Chemical cross-linking, thermal denaturation, and size fractionation analyses suggested that the full-length S glycoprotein of SARS-CoV forms a higher order structure of ~500 kDa, which is consistent with it being an S homotrimer. The latter was also observed in purified virions. The intracellular form of the C-terminally truncated S protein (but not the secreted form) also forms trimers, but with much less efficiency than full-length S. Deglycosylation of the full-length homotrimer with peptide N-glycosidase-F under native conditions abolished recognition of the protein by virus-neutralizing antisera raised against purified virions, suggesting the importance of the carbohydrate in the correct folding of the S protein. These data should aid in the design of recombinant vaccine antigens to prevent the spread of this emerging pathogen.

A new infectious disease, known as severe acute respiratory syndrome (SARS), first appeared in China in 2003 and has so far resulted in a cumulative total of 8,098 probable SARS cases and more than 774 deaths in 32 countries (33). Phylogenetic sequence analysis of the SARS coronavirus (SARS-CoV) genome demonstrated that it is not closely related to any of the three previously known coronavirus groups, nor is it a reassortant of known coronaviruses (17, 21, 27). However, the S1 domain of the SARS-CoV spike protein has significant similarity with those of group 2 coronaviruses in terms of the number and spatial arrangement of cysteine residues, suggesting that SARS-CoV is more closely related to the group 2 coronaviruses (10). Recently, a virus with very close sequence identity with SARS-CoV was identified from civet cats, raccoon dogs, and the Chinese ferret badger, indicating that the SARS viral origin was zoonotic (11, 13, 22). After the onset of the SARS epidemic, aggressive quarantine measures successfully suppressed the emergence of SARS. In 2004, however, several cases of community-acquired SARS have been reported in China, indicating that the transmission of SARS still continues and that a vaccine is urgently required.

Coronaviruses are enveloped viruses with large, positive-

stranded RNA genomes (27 to 31 kb). Their genomes have five major open reading frames that encode the replicase polyproteins, the spike (S), envelope (E), and matrix (M) glycoproteins, the nucleocapsid protein (N), and other small proteins with unknown functions (6, 18). Among these proteins, the S protein is responsible for the binding of virus to cellular receptors and is a likely target for eliciting viral neutralizing antibodies (2). The S polypeptide is N-glycosylated cotranslationally in the endoplasmic reticulum (ER) and further processed in the Golgi apparatus (2). While the S glycoproteins of human coronavirus 229E, transmissible gastroenteritis virus (TGEV), porcine respiratory coronavirus, feline infectious peritonitis virus, and canine coronavirus remain as single glycoproteins, those of mouse hepatitis virus, bovine coronaviruses, and human coronavirus OC43 are proteolytically cleaved into two subunits, S1 and S2, during the cellular transport process. Earlier studies of TGEV indicated that the S protein is an oligomer composed of three copies of the monomeric S glycoprotein (7). Such a quaternary structure has been reported for other enveloped RNA viruses and has been demonstrated to be important for eliciting neutralizing antibodies against hemagglutinin A (HA) of influenza virus (30), the gp120-gp41 heterodimer of human immunodeficiency virus (34), and the G protein of vesicular stomatitis virus (9).

The amino acid sequence of the SARS-CoV S glycoprotein, as deduced from the published genome sequence, reveals that

* Corresponding author. Mailing address: Vaccines Research, Chiron Corporation, 4560 Horton St., Emeryville, CA 94608. Phone: (510) 923-2937. Fax: (510) 923-2586. E-mail: jang_han@chiron.com.

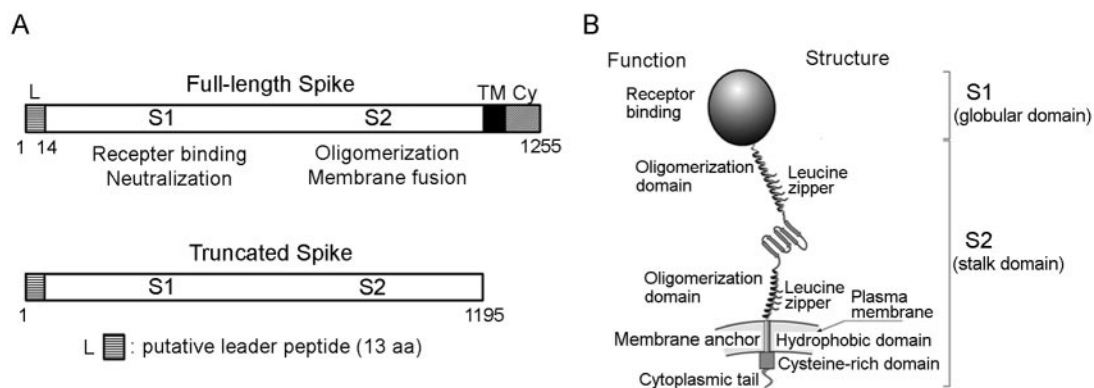


FIG. 1. (A) Structural components of SARS-CoV spike glycoprotein and expression construct. L denotes the leader peptide (residues 1 to 13). The cDNA fragments encompassing the full-length spike coding sequence and the spike sequence with deletions of the TM and Cys-rich regions were cloned into an expression vector, pCMVIII, to create nSh and nSh Δ TC, respectively. All proteins were tagged with six histidine residues at the end of the C terminus. (B) Schematic diagram of SARS-CoV spike glycoprotein. The hypothetical features of the SARS-CoV S protein were drawn based on information from other coronaviruses.

it has a 1,255-amino-acid precursor polypeptide with 23 potential N-linked glycosylation sites (10, 21, 27, 35). As judged by comparison with other coronaviruses, the sequence of the SARS-CoV S protein may comprise the following hypothetical features (Fig. 1): a 13-amino-acid cleavable secretory signal (21), a putative S1 globular domain with a potential receptor binding site (19, 29, 32, 35), a putative S2 stalk domain with a fusion peptide and heptad repeat motifs that could form an alpha-helical structure which may interact with other subunits to form a coiled-coil oligomeric structure (1, 15, 20), and a hydrophobic transmembrane (TM) domain near the C terminus that could be responsible for anchoring the S protein to the virion lipid envelope (1). This potential membrane anchor region is immediately followed by a cysteine-rich (Cy) domain, a feature common to all other coronaviruses, which may stabilize protein-lipid interactions (4).

For this investigation, we expressed various forms of the SARS-CoV spike glycoprotein in several mammalian cell lines in order to determine the physicochemical and functional properties of this glycoprotein and its domains. Our data indicate that the SARS-CoV spike glycoprotein is expressed as a plasma membrane-associated, uncleaved, homotrimeric form for which the quaternary structure is similar to that of TGEV. A C-terminally truncated form can be secreted into the cell medium primarily in a nonoligomeric form. Since the S protein is a likely candidate for inducing protective immunity, the elucidation of this native form of the recombinant S glycoprotein is likely to be crucial for the development of effective vaccines.

MATERIALS AND METHODS

Construction of expression plasmids. cDNA fragments containing the S protein open reading frame of 1,255 amino acids were amplified by reverse transcription-PCR from SARS viral RNA (Frankfurt isolate, accession number AY310120) grown in Vero cells (10). The amplified PCR fragments were cloned into the pBlueScript vector (Stratagene, La Jolla, Calif.) and then sequenced, and a consensus spike sequence was assembled to create a full-length SARS spike clone, called pBS-nSh. The insert of this plasmid was recloned via XhoI and NotI into a mammalian expression vector, pCMVIII (28), to create the construct nSh (Fig. 1). A PCR fragment containing a spike protein of 1,195 amino acids, which had a deletion of the TM domain and the cysteine-rich cytoplasmic tail (Cy), was amplified and cloned into the pCMVIII vector to

generate the construct nSh Δ TC (Fig. 1). Both constructs were tagged with six histidine residues at the C terminus in order to aid in their characterization. The XhoI/NotI fragment without a histidine tag was also subcloned into the alphavirus replicon vector backbone pVCRchim2.1 (23) for use in the production of an alphavirus replicon particle chimera that expresses the S protein. The production and characterization of the replication-defective alphavirus vector particles were performed essentially as described previously (23, 24). The resultant alphavirus vector particles were named VEE/SIN-SSP.

Cells and transfection. COS7 cells and BHK21 cells were maintained in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum at 37°C and 5% CO₂ in air. COS7 cells were transfected with expression plasmids (nSh and nSh Δ TC) by use of a transfection kit (TransIT-COS) according to the manufacturer's protocol (Mirus, Madison, Wis.). The cells were washed once with ice-cold phosphate-buffered saline (PBS) and lysed with 1× lysis buffer (20 mM MOPS, 10 mM NaCl, 1.5 mM MgCl₂, and 1% Triton X-100) containing complete mini protease inhibitor (Roche, Indianapolis, Ind.). After a 30-min incubation on ice, the debris was cleared by centrifugation. The cleared lysate was either purified or used directly for Western blotting.

Purification of secreted spike proteins from medium. The medium from transfected cells was collected and subjected to centrifugation at 12,000 × g for 10 min to remove cellular debris. The cleared medium was applied to a concanavalin A-agarose column (Vector Laboratories, Burlingame, Calif.). The column was washed extensively with 20 mM sodium phosphate buffer, and then the bound proteins were eluted with 1 M methyl α -D-mannopyranoside (Sigma, St. Louis, Mo.) and 1 M NaCl in 20 mM sodium phosphate buffer (26). Column fractions containing SARS-CoV spike proteins were applied to the MagneHis protein purification system according to the protocol suggested by the manufacturer (Promega, Madison, Wis.).

Western blot analysis. Proteins were separated by sodium dodecyl sulfate-4 to 20% polyacrylamide gel electrophoresis (SDS-4 to 20% PAGE) and then transferred electrophoretically to a nitrocellulose membrane (Invitrogen, Carlsbad, Calif.). The membrane was blocked with blocking buffer (5% skim milk and 0.1% Tween 20 in PBS), incubated with the indicated antibody at room temperature for 1 h, washed, probed with a horseradish peroxidase-conjugated secondary antibody (Biosource, Camarillo, Calif.) followed by chemiluminescence (ECL system; Amersham, Piscataway, N.J.), and exposed on X-ray films. The antibodies used were a mouse anti-histidine monoclonal antibody (anti-His-tag MAb; Novagen, Darmstadt, Germany), a rabbit polyclonal antipeptide antibody against the SARS-CoV spike protein (SARS-Sm antibody; Abgent, San Diego, Calif.), and rabbit anti-SARS sera (2BE) obtained by the immunization of rabbits with inactivated and purified SARS-CoV virions. This last antibody has a cell culture neutralizing titer of 1/500 (K. Stadler, unpublished data). Unless stated otherwise, antibodies were used at 1/1,000 for the anti-histidine and SARS-Sm antibodies and at 1/10,000 for anti-SARS rabbit sera.

Indirect immunofluorescence assay. At 48 h posttransfection, cells were directly fixed with 2% paraformaldehyde without detergent for cell surface staining or were treated with a detergent mix (Cytofix/Cytoperm solution; BD Biosciences, San Jose, Calif.) for intracellular staining. Fixed cells were then stained

with rabbit anti-SARS sera (2BE) and a fluorescein isothiocyanate-conjugated antibody (Molecular Probes, Eugene, Oreg.).

Radiolabeling and immunoprecipitation. BHK21 cells were infected with VEE/SIN-SSP replicon particles at a multiplicity of infection (MOI) of 5. After incubation for 6 h, the cells were incubated with methionine- and cysteine-free Dulbecco's modified Eagle's medium (GIBCO BRL, Carlsbad, Calif.) for 1 h at 37°C and then pulse labeled with 300 μ Ci of L-[³⁵S]methionine-cysteine (Amersham) for 1 h at 37°C. At the end of the pulse period, the cells were washed once with serum-free medium and chased at 37°C for the indicated time with complete growth medium containing 5% fetal bovine serum. The labeled cells were then washed once with PBS (pH 7.4) and lysed with 1 \times lysis buffer as described above. The cleared cell lysates were incubated overnight with rabbit anti-SARS sera (2BE) at 4°C and then were incubated with protein A-Sepharose (Amersham) for 2 h at 4°C. The beads were collected by centrifugation (3,000 \times g for 2 min at 4°C), washed three times with TPBS buffer (0.1% Tween 20 in PBS), and washed once with 120 mM Tris, pH 7.0. The samples were resuspended in SDS loading buffer with 50 mM dithiothreitol (DTT), denatured by boiling for 5 min, and either digested with endoglycosidase H (endo H) or used directly for SDS-PAGE and autoradiography.

Endo H digestion of spike proteins. The cell lysate, purified secreted spike protein, or eluted immune complex from immunoprecipitation was diluted with sample buffer (50 mM sodium phosphate, 0.1% SDS, 50 mM DTT, pH 6.0) and boiled for 5 min. After denaturation, the samples were further diluted with 0.75% Triton X-100 and then treated with endo H according to the manufacturer's protocol (Calbiochem) for 3 h at 37°C. Enzyme-treated samples were added to gel loading buffer containing 1% SDS and DTT and then were analyzed by SDS-8% PAGE.

PNGase F treatment of spike proteins. The cell lysates were diluted in 0.5% SDS and 1% β -mercaptoethanol and denatured at 100°C for 10 min. After a twofold dilution with 1% NP-40 in 50 mM sodium phosphate (pH 7.5), the samples were treated with peptide *N*-glycosidase F (PNGase F; NEB, Beverly, Mass.) at 37°C for 1 h. Enzyme-treated samples were analyzed by SDS-4 to 12% PAGE under reducing conditions. For partial digestion with PNGase F, the cell lysates were diluted with 50 mM sodium phosphate (pH 6.0) containing 0.75% Triton X-100 and were treated with PNGase F (Calbiochem) at 37°C for 3 h. Enzyme-treated samples were analyzed by SDS-4 to 12% PAGE under nonreducing conditions.

Inactivated SARS-CoV preparation. Vero cells were infected with SARS-CoV (FRA isolate, accession number AY310120) at an MOI of 0.001. The supernatant was harvested at 48 h postinfection and centrifuged for 10 min at 10,000 \times g at 4°C to remove cellular debris, and the virus was inactivated with β -propiolactone. The virus was then isolated by passage over an affinity matrix and was further purified by sucrose gradient centrifugation (15 to 40% [wt/wt] in PBS, pH 7.2) in an SW-28 rotor at 20,000 rpm for 2 h at 4°C.

Cross-linking experiments. Aliquots of inactivated virions from sucrose gradient fractions were treated with 10% SDS at a 1% final concentration and diluted twofold with 0.2 M triethanolamine-HCl (pH 8) (Sigma); dimethyl suberimidate (DMS; Pierce Biotechnology, Inc., Rockford, Ill.) was then added from a freshly prepared solution (10 mg/ml in 0.2 M triethanolamine-HCl) to a final concentration of 3.3 mg/ml. After 2 h at room temperature, the samples were concentrated with Centricon-30 columns (Millipore, Billerica, Mass.) and analyzed by silver staining after electrophoresis through a 4% polyacrylamide gel.

RESULTS

Characteristics of recombinant SARS spike proteins expressed in mammalian cells. We first cloned a cDNA fragment encompassing the full-length SARS-CoV S protein of 1,255 amino acids into the Bluescript transcription vector. In vitro transcription, followed by translation in a rabbit reticulocyte lysate with or without membranes, resulted in the production of a single polypeptide with an estimated molecular mass of \sim 140 kDa (data not shown). A cDNA fragment of the full-length sequence as well as a construct with a deletion of both the predicted TM and the downstream Cys-rich domains was then inserted into the expression vector pCMVIII, creating constructs termed nSh and nSh Δ TC, respectively (Fig. 1A). Both constructs were tagged with six histidine residues at the C

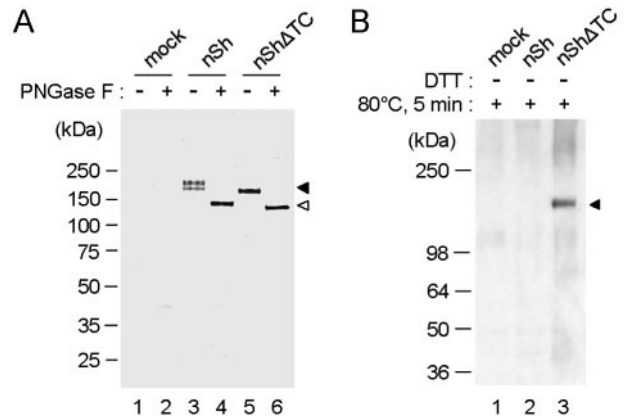


FIG. 2. Western blot analysis of SARS-CoV spike proteins expressed in COS7 cells. (A) COS7 cells were transfected with the indicated plasmid constructs, and the expressed proteins in cell lysates at 48 h posttransfection were analyzed by SDS-4 to 20% PAGE under reducing and denaturing conditions and then visualized with an anti-histidine MAb. (B) COS7 cells were transfected with the indicated plasmids, and proteins were collected from the cell culture medium at 48 h posttransfection and purified first with a concanavalin A column and then with His-tag magnetic beads as described in Materials and Methods. Purified proteins were analyzed by SDS-PAGE (4 to 20% polyacrylamide) and visualized with an anti-SARS rabbit serum (2BE).

terminus in order to aid the characterization of the recombinant full-length spike (S) or deleted spike (Sd) proteins.

We evaluated the profile of spike protein expression in COS7 cells at 48 h posttransfection by Western blot analyses with an anti-histidine MAb or with a polyclonal rabbit antiserum obtained after immunizations with inactivated and purified SARS-CoV virus preparations (Fig. 2). The S protein was detected in cell lysates as a doublet with an estimated molecular mass of \sim 170 to \sim 180 kDa when the lysate was boiled and analyzed under reducing SDS-PAGE conditions (Fig. 2A, lane 3). This doublet appeared to result from differential glycosylation of one polypeptide product since pretreatment of the cell lysate with PNGase F reduced the doublet to a single species of \sim 140 kDa (Fig. 2A, lane 4). This is the expected size predicted from the genomic sequence for a full-length, intact polypeptide product. This experiment indicates that the full-length SARS-CoV S protein is expressed in mammalian cells as a single, uncleaved polypeptide, but in two differentially glycosylated forms, termed gp170 and gp180, respectively. Unlike the two S glycoforms encoded by the full-length sequence, neither of which was secreted, the Sd protein product was detected both in cell lysates (Fig. 2A, lane 5) and in the cell culture medium (Fig. 2B, lane 3) as a single species of \sim 160 kDa.

Processing of full-length SARS-CoV spike protein. In order to further characterize the intracellular processing of the S protein, we infected BHK21 cells with replication-defective alphavirus particles expressing the full-length S protein. At 6 h postinfection at an MOI of 5, the infected cells were pulse labeled for 1 h with L-[³⁵S]methionine-cysteine and then chased for 2 or 4 h. The ³⁵S-labeled S protein was immunoprecipitated with the rabbit antisera raised against the inactivated and purified virus and then was digested with endo H. Both digested and undigested proteins were boiled in SDS and

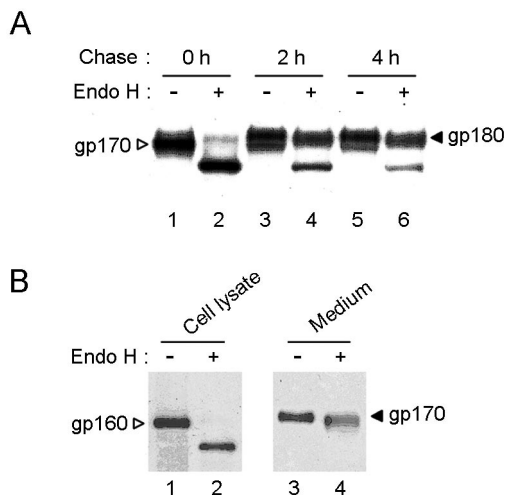


FIG. 3. (A) Pulse-chase analysis of processing of full-length SARS-CoV spike (S) proteins. BHK21 cells were infected with recombinant alphavirus particles carrying an expression cassette for the spike protein. After 6 h of incubation, the infected cells were pulse labeled for 1 h with L-[³⁵S]methionine-cysteine and then chased for the indicated times. The ³⁵S-labeled S protein was immunoprecipitated with an anti-SARS rabbit serum (2BE) and digested with endo H. Both digested and undigested S proteins were analyzed by SDS-8% PAGE. The positions of the two S glycoproteins are marked with arrowheads. (B) Endo H sensitivity of C-terminally truncated spike protein (Sd) found in the cell lysate (lanes 1 and 2) and the culture medium (lanes 3 and 4). The positions of internal and secreted Sd proteins are marked with arrowheads.

analyzed by reducing SDS-PAGE (Fig. 3A). After a 1-h pulse, the S protein was apparent as a single gp170 component that was endo H sensitive (Fig. 3A, lanes 1 and 2). After a 2-h chase, a new species (gp180) was present along with gp170 in approximately equal proportions (Fig. 3A, lane 3). After a 4-h

chase, the gp180 species was the dominant S protein component (Fig. 3A, lane 5), which was then endo H resistant (Fig. 3A, lanes 5 and 6). These data are consistent with gp170 being an ER-resident glycoprotein containing high levels of mannose chains and with gp180 corresponding to a Golgi-processed glycoprotein containing endo H-resistant complex oligosaccharides.

We next tested the endo H sensitivity of the C-terminally deleted Sd protein purified from the cell culture medium. As shown in Fig. 3B, the Sd protein observed within cell lysates was found to be endo H sensitive (Fig. 3B, lanes 1 and 2), while the secreted Sd in the cell culture medium was endo H resistant (Fig. 3B, lanes 3 and 4). This result is consistent with this glycoprotein being synthesized in an immature form in the ER prior to transfer to the Golgi, where the complex carbohydrate is added and from which the protein is then secreted.

Oligomeric nature of full-length SARS-CoV spike proteins.

As previously described, the S protein expressed in COS7 cells was detected as a gp170-gp180 doublet by Western blot analyses of cell lysates that were fully denatured by boiling in the presence of DTT. However, the majority of the S protein was detected as a high-molecular-mass glycoprotein in the 440- to 669-kDa range when the same cell lysate was not heat denatured prior to Western blot analysis after SDS-PAGE (Fig. 4A, lane 2). As shown in Fig. 4, this ~500-kDa species was resistant to a 50 mM DTT treatment (Fig. 4A, lane 3) and was not dissociated into the monomeric form unless the lysate was first heat denatured at 100°C (lane 4). In contrast, the oligomeric form of a test protein (thyroglobulin), for which the quaternary structure is formed by disulfide linkage, was converted into a monomeric subunit form by the 50 mM DTT treatment (data not shown). These data suggest that the ~500-kDa oligomeric form of the S protein is not disulfide linked but is heat labile. To confirm the heat sensitivity of the ~500-kDa species of the

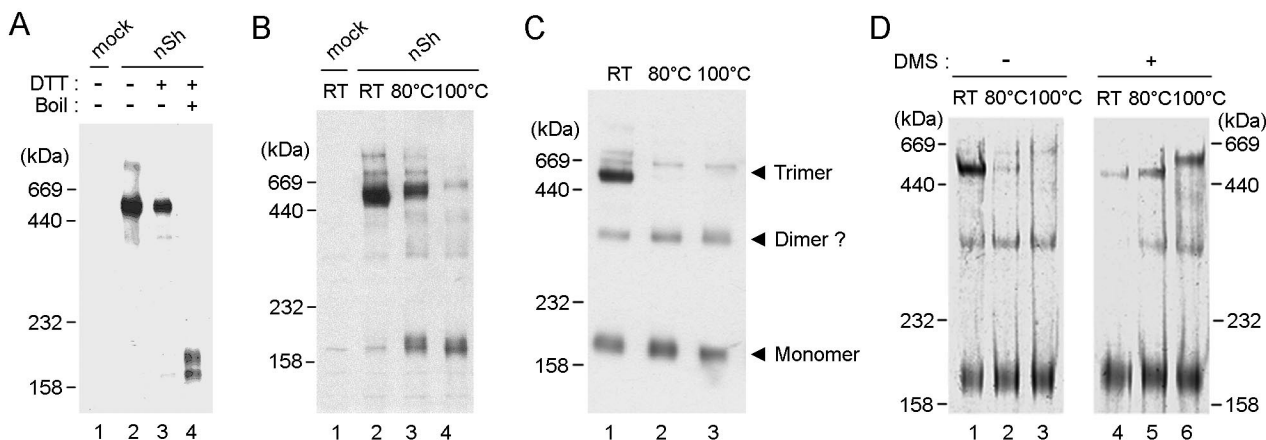


FIG. 4. Oligomeric status of SARS-CoV spike protein. (A) Recombinant S protein oligomer in COS7 cells. Cells were transfected with the full-length spike construct (nSh). The cell lysates were treated with DTT and/or heat as indicated above each lane. The different forms of the S protein in treated and untreated samples were visualized by SDS-PAGE (4% polyacrylamide) and Western blot analysis with an anti-histidine MAbs. (B) Effect of heat denaturation on oligomeric status of recombinant S protein in the absence of DTT. The COS7 cell lysates were heated before electrophoresis as indicated, and the S proteins were visualized as described for panel A. (C) Effect of heat denaturation on oligomeric status of spike protein in SARS virion particles. SARS-CoV virions were grown in Vero cells and purified as described in Materials and Methods. The proteins were solubilized from the virion particles with SDS, heat denatured as indicated, and visualized as described for panel A, except that a rabbit antiserum (2BE) against the purified virus was used as a probe. (D) Analysis of oligomeric status of SARS virion spike protein by cross-linking experiment. Solubilized SARS virion proteins were treated with DMS. Both untreated and DMS-treated virion proteins were heat denatured in the absence of DTT and visualized by 4% PAGE followed by silver staining.

S protein, we repeated the heat denaturation experiment without DTT. As shown in Fig. 4B, heat denaturation of the ~500-kDa protein at 100°C alone was sufficient to convert it to gp170 and gp180 monomeric forms (Fig. 4B, lane 4). With an 80°C heat denaturation step, both the ~500-kDa and monomeric forms were detectable in similar proportions (Fig. 4B, lane 3).

In order to investigate further whether this ~500-kDa species represents an S protein oligomer in its native conformation, we performed comparative analyses with a virion-derived S glycoprotein derived from Vero cell cultures. The purified virions were solubilized in 1% SDS prior to Western blot analyses after SDS-PAGE. The presence of the ~500-kDa spike protein oligomer was confirmed in the virion particles (Fig. 4C, lane 1). In addition, heat denaturation of solubilized virions produced the same oligomer-to-monomer conversion as that seen with the full-length recombinant S protein (Fig. 4C, lanes 2 and 3). We further analyzed the oligomeric nature of virion S with a cross-linking experiment. For this analysis, we heat denatured both untreated and cross-linked virion proteins with DMS and analyzed the heat effect on the maintenance of the oligomer structure by SDS-PAGE and silver staining. In the absence of cross-linking, heat denaturation resulted in the replacement of the ~500-kDa spike protein species with the monomer species (Fig. 4D, lanes 1 to 3). In contrast, in the cross-linked proteins, the levels of the ~500-kDa and monomer species did not change significantly after heating (Fig. 4D, lanes 4 to 6). These data support the fact that the ~500-kDa protein is an oligomer of S monomer proteins that are bound noncovalently. After cross-linking and boiling, the ~500-kDa species migrated as a somewhat slower diffuse form than the untreated form (Fig. 4D, lane 6). This mobility shift was probably due to a structural change resulting from boiling. In addition, we detected a minor protein species of ~300 kDa, which may represent a nondissociated S dimer.

We attempted to estimate more precisely the size of the recombinant ~500-kDa S species expressed in COS7 cells. To this end, we fractionated the COS7 cell lysate containing the S protein oligomer by size-exclusion column chromatography. The major portion of the ~500-kDa oligomer was coeluted with a 572-kDa marker protein (data not shown). Taken together, these experiments suggest that the ~500-kDa S species seen in COS7 cell lysates is probably a homotrimer of the S protein monomer. We cannot rule out the formal possibility that S has been oligomerized with some other cellular proteins. Purification and further analysis of the recombinant oligomeric S protein are needed to confirm this hypothesis.

Oligomeric status of spike protein with a deletion of the TM domain and Cys-rich region (Sd). We next examined the oligomeric status of the Sd protein expressed in COS7 cells. As shown in Fig. 5, the recombinant Sd proteins present in cell lysates were also detected in high-molecular-mass forms of ~500-kDa when the lysate was not heated prior to SDS-PAGE and Western blot analysis (Fig. 5A, lane 1). However, a heat sensitivity test with this ~500-kDa protein showed that the intracellular Sd oligomer was more heat labile than the full-length S oligomer, as demonstrated by the >90% conversion of all of the ~500-kDa species into monomeric Sd forms at 80°C (Fig. 5A, lane 2). Also, as shown in Fig. 5B, the majority of the secreted Sd protein was found in the monomeric form, with the ~500-kDa species being barely detectable (and only detectable

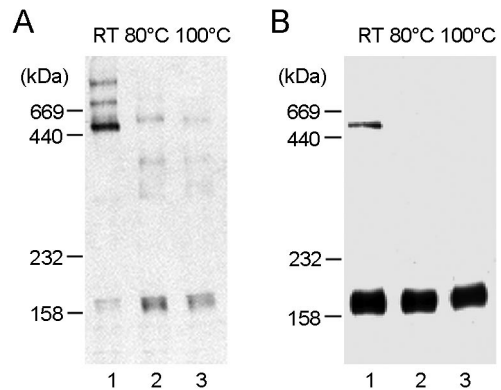


FIG. 5. Analysis of oligomeric status of truncated spike protein by heat denaturation. The truncated spike protein within COS7 cell lysates (A) or that secreted into the culture medium (B) was heat denatured in the absence of DTT as indicated and was visualized by Western blot analysis as described in the legend for Fig. 4.

when the protein was loaded in excess for Western blot analysis) (Fig. 5B, lane 1). At a temperature above 80°C, all secreted Sd proteins were detected as monomers (Fig. 5B, lanes 2 and 3). While these data indicate the capacity of the Sd protein to form oligomers, it appears that the C-terminal TM and/or cytoplasmic tail region may be required for stabilization of the trimer.

Influence of sugar moiety in S oligomers on recognition by neutralizing antisera raised against purified SARS virus. We next tested whether deglycosylation of the ~500-kDa (gp540) oligomer of the S protein affects antibody binding. We therefore treated the recombinant COS7 cell lysate with PNGase F under nondenaturing conditions without SDS and DTT, as described in Materials and Methods, and performed Western blot analyses. As shown in Fig. 6, deglycosylation did not affect the binding of an anti-histidine MAb to the treated S oligomer (lanes 2 and 3). However, it compromised the reactivity with the rabbit antiserum raised against the purified virus (lane 6). This antiserum bound to virion-derived S in Western blot anal-

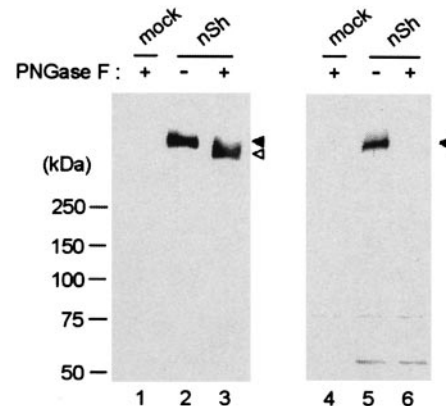


FIG. 6. Reactivity of deglycosylated full-length spike oligomer with conformational and nonconformational antibodies. The full-length recombinant spike oligomer was partially deglycosylated with PNGase F under nondenaturing conditions and visualized by Western blot analysis with an anti-histidine MAb (lanes 1 to 3) or a rabbit antiserum against purified SARS-CoV (lanes 4 to 6).

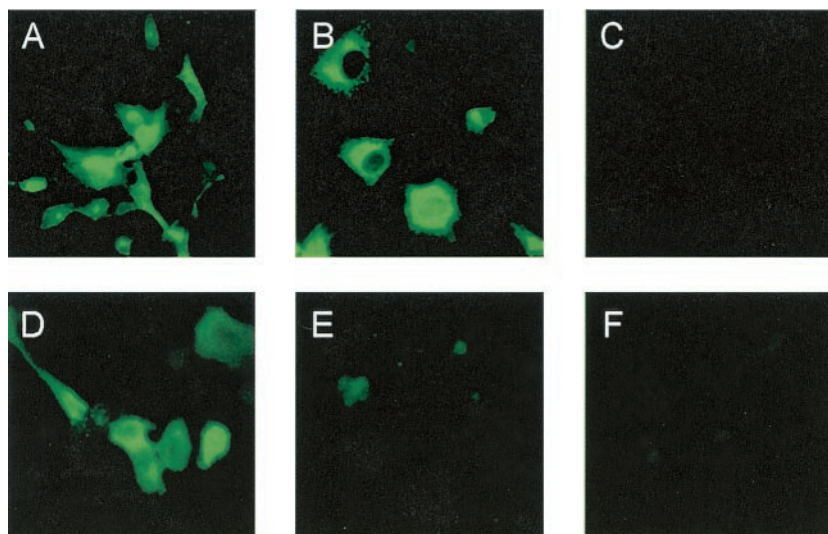


FIG. 7. Intracellular and cell surface expression of recombinant full-length (A and D) or truncated (B and E) spike protein in COS7 cells. At 48 h posttransfection, the cells were treated with a detergent mix (Cytofix/Cytoperm kit) for intracellular immunofluorescence (A, B, and C) or were fixed with 2% paraformaldehyde without detergent for cell surface immunofluorescence observation (D, E, and F) at a magnification of $\times 400$. The cells were then stained with a rabbit anti-SARS serum (2BE) and a fluorescein isothiocyanate-conjugated antibody. Mock-transfected cells (C and F) were included as a control.

yses only when DTT was omitted from the sample for SDS-PAGE, indicating that it recognizes primarily a discontinuous conformational epitope(s) (data not shown). This antiserum has also been shown to have a high titer of viral neutralizing antibodies (data not shown). Its lack of binding to deglycosylated recombinant S suggests that the carbohydrate may actively contribute to the higher order native structure of the S polypeptide oligomer.

Cellular location of recombinant SARS-CoV S proteins. The cellular location of the S and Sd proteins in COS7 cells was analyzed by indirect immunofluorescence microscopy. Upon the probing of fixed cells with rabbit antiserum against the whole inactivated virus, the nSh-transfected cells showed a cytoplasmic staining pattern (Fig. 7A), as did nSh Δ TC-transfected cells (Fig. 7B), although in the latter case, a more punctate pattern was observed, possibly indicative of stronger Golgi staining. While the complete S protein was also observed on the surfaces of transfected and unfixed cells (Fig. 7D), the Sd protein was undetectable on the cell surface (Fig. 7E). These results indicate the role played by the TM and Cys-rich domains in anchoring the S protein to the plasma membrane. Although the TM region alone is likely responsible for membrane anchorage, the potential role of the Cys-rich region remains to be determined.

DISCUSSION

This work demonstrates that the SARS recombinant full-length S protein is an N-linked glycoprotein with an estimated molecular weight of 170,000 to 180,000 Da. Deglycosylation with PNGase F resulted in a polypeptide of the expected size for the uncleaved encoded polypeptide (140 kDa). Both transient and stable expression of the full-length SARS-CoV S gene in a variety of mammalian cells, including the COS7, 293, BHK21, and Huh7 cell lines, consistently produced an S pro-

tein doublet (gp170-gp180) that was detected in Western blot analyses. Pulse-chase analyses of transfected cells demonstrated that the SARS-CoV S protein was initially synthesized as an endo H-sensitive gp170 species followed by the gradual appearance of an endo H-resistant gp180 form, presumably as a result of the addition of a complex carbohydrate within the Golgi apparatus.

The recombinant S protein was not secreted into the cell culture medium unless the C-terminal 60 amino acids containing the TM region and the Cys-rich tail were deleted. It is known that coronaviral S protein secretion is significantly increased during a viral infection (25), so the secretion of the full-length spike protein in our system could have been disrupted due to the absence of viral infection and morphogenesis.

In the present study, we analyzed the quaternary structure of the full-length recombinant S protein by cross-linking treatment, heat denaturation, and size fractionation analyses. Our data are consistent with the recombinant S protein existing as a homotrimer of ~ 500 kDa. Similar analyses of virion-derived S yielded the same results. Such a trimeric structure has been reported for other enveloped RNA viruses, including HA of influenza virus (30), the E1-E2 heterodimer of alphaviruses (14), and the G protein of vesicular stomatitis virus (9). Incubations under reducing conditions indicated that the SARS-CoV S trimeric structure is noncovalently associated and is very stable. S oligomers present in the cell lysate were shown to be resistant to reduction by 50 mM DTT, a detergent treatment with 1% SDS, and heat denaturation up to 60°C. Incubation at a temperature higher than 80°C resulted in the dissociation of the trimeric complex, as evidenced by the decrease in trimer bands with a concomitant increase in monomer bands. The temperature-induced appearance of the high-mannosylated gp170 (ER monomer form) as well as the complex-glycosylated gp180 (Golgi monomer form) suggests that trim-

erization may occur before the transport of the monomer spike protein to the medial Golgi apparatus. This is consistent with other reports for TGEV, influenza virus HA, and vesicular stomatitis virus G proteins (5, 7, 12, 16, 36). For these proteins, trimerization was reported to take place before the addition of complex oligosaccharides in the Golgi apparatus.

The C-terminally truncated form of S was found in the cell lysate in both oligomeric and monomeric forms at similar frequencies. The truncated protein secreted into the medium was found to be fully glycosylated and essentially all monomeric. We conclude that the C-terminal 60 amino acids of the S glycoprotein contain a membrane anchor region(s) that affects the efficiency of trimerization. For S protein trimerization, it is possible that the C-terminal region is required to initiate the event, with the triple-stranded coiled-coil structures elsewhere in the S molecule providing a further stabilizing force, as seen for the HA oligomer of influenza virus (8, 31, 36).

It should be pointed out that our data are based on the expression of recombinant S glycoproteins and that we therefore may be missing other possible modifications to S that are made in virus-infected cells. Further characterization of S and other viral proteins from virus-infected cells is therefore important.

The epidemic potential and severe pathogenicity of SARS-CoV highlight the need for an effective vaccine. Attenuated live viral vaccines have been effective against avian coronavirus infections (3). Due to its severe pathogenicity and the risks associated with attenuation, current approaches to SARS-CoV vaccine development may focus on the use of inactivated viral vaccines, recombinant polypeptide subunit vaccines, or antigen delivery via safe viral vectors. Although an inactivated viral vaccine could be proven effective against SARS-CoV, a recombinant antigen vaccine would be potentially easier and safer to prepare and use. It is generally believed that polymeric antigens are more immunogenic and effective in the induction of protective immunity. The oligomeric and apparently native form of recombinant S described and characterized here represents a promising vaccine candidate. Animal immunogenicity and viral challenge studies are in progress.

REFERENCES

- Bosch, B. J., R. Van der Zee, C. A. de Haan, and P. J. Rottier. 2003. The coronavirus spike protein is a class I virus fusion protein: structural and functional characterization of the fusion core complex. *J. Virol.* **77**:8801–8811.
- Cavanagh, D. 1995. The coronavirus surface glycoprotein, p. 73–113. *In* S. G. Siddell (ed.), *The Coronaviridae*. Plenum Press, New York, N.Y.
- Cavanagh, D. 2003. Severe acute respiratory syndrome vaccine development: experiences of vaccination against avian infectious bronchitis coronavirus. *Avian Pathol.* **32**:567–582.
- Chang, K. W., Y. Sheng, and J. L. Gombold. 2000. Coronavirus-induced membrane fusion requires the cysteine-rich domain in the spike protein. *Virology* **269**:212–224.
- Copeland, C. S., R. W. Doms, E. M. Bolzau, R. G. Webster, and A. Helenius. 1986. Assembly of influenza hemagglutinin trimers and its role in intracellular transport. *J. Cell Biol.* **103**:1179–1191.
- De Haan, C. A. M., H. Volders, C. A. Koetznar, P. S. Masters, and P. J. M. Rottier. 2002. Coronaviruses maintain viability despite dramatic rearrangements of the strictly conserved genome organization. *J. Virol.* **76**:12491–12502.
- Delmas, B., and H. Laude. 1990. Assembly of coronavirus spike protein into trimers and its role in epitope expression. *J. Virol.* **64**:5367–5375.
- Doms, R. W., and A. Helenius. 1986. Quaternary structure of influenza virus hemagglutinin after acid treatment. *J. Virol.* **60**:833–839.
- Doms, R. W., D. S. Keller, A. Helenius, and W. E. Balch. 1987. Role for adenosine triphosphate in regulating the assembly and transport of vesicular stomatitis virus G protein trimers. *J. Cell Biol.* **105**:1957–1969.
- Eickmann, M., S. Becker, H. D. Klenk, H. W. Doerr, K. Stadler, S. Censini, S. Guidotti, V. Masignani, M. Scarselli, M. Mora, C. Donati, J. H. Han, H. C. Song, S. Abrignani, A. Covacci, and R. Rappuoli. 2003. Phylogeny of the SARS coronavirus. *Science* **302**:1504–1505.
- Enserink, M. 2003. Infectious diseases. Clues to the animal origins of SARS. *Science* **300**:1351.
- Gething, M. J., K. McCammon, and J. Sambrook. 1986. Expression of wild-type and mutant forms of influenza hemagglutinin: the role of folding in intracellular transport. *Cell* **46**:939–950.
- Guan, Y., B. J. Zheng, Y. Q. He, X. L. Liu, Z. X. Zhuang, C. L. Cheung, S. W. Luo, P. H. Li, L. J. Zhang, Y. J. Guan, K. M. Butt, K. L. Wong, K. W. Chan, W. Lim, K. F. Shortridge, K. Y. Yuen, J. S. Peiris, and L. L. Poon. 2003. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* **302**:276–278.
- Harrison, S. C. 1986. Alphavirus structure, p. 21–34. *In* S. Schlesinger and M. J. Schlesinger (ed.), *The Togaviridae and Flaviviridae*. Plenum Publishing Corp., New York, N.Y.
- Kliger, Y., and E. Y. Levanon. 2003. Cloaked similarity between HIV-1 and SARS-CoV suggests an anti-SARS strategy. *BMC Microbiol.* **3**:20.
- Kreis, T. E., and H. F. Lodish. 1986. Oligomerization is essential for transport of vesicular stomatitis viral glycoprotein to the cell surface. *Cell* **46**:929–937.
- Ksiazek, T. G., D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J. A. Comer, W. Lim, P. E. Rollin, S. F. Dowell, A. E. Ling, C. D. Humphrey, W. J. Shieh, J. Guarner, C. D. Paddock, P. Rota, B. Fields, J. DeRisi, J. Y. Yang, N. Cox, J. M. Hughes, J. W. LeDuc, W. J. Bellini, and L. J. Anderson (SARS Working Group). 2003. A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* **348**:1953–1966.
- Lai, M. M. C., and K. V. Holmes. 2001. Coronaviridae: the viruses and their replication, p. 1163–1186. *In* D. M. Knipe, P. M. Howley, D. E. Griffin, M. A. Martin, R. A. Lamb, and B. Roizman (ed.), *Fields virology*. Lippincott Williams and Wilkins, Philadelphia, Pa.
- Li, W., M. J. Moore, N. Vasilieva, J. Sui, S. K. Wong, M. A. Berne, M. Somasundaran, J. L. Sullivan, K. Luzuriaga, T. C. Greenough, H. Choe, and M. Farzan. 2003. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**:450–454.
- Luo, Z., A. M. Matthews, and S. R. Weiss. 1999. Amino acid substitutions within the leucine zipper domain of the murine coronavirus spike protein cause defects in oligomerization and the ability to induce cell-to-cell fusion. *J. Virol.* **73**:8152–8159.
- Marra, M. A., S. J. Jones, C. R. Astell, R. A. Holt, A. Brooks-Wilson, Y. S. Butterfield, J. Khattri, J. K. Asano, S. A. Barber, S. Y. Chan, A. Cloutier, S. M. Coughlin, D. Freeman, N. Girm, O. L. Griffith, S. R. Leach, M. Mayo, H. McDonald, S. B. Montgomery, P. K. Pandoh, A. S. Petrescu, A. G. Robertson, J. E. Schein, A. Siddiqui, D. E. Smailus, J. M. Stott, G. S. Yang, F. Plummer, A. Anonov, H. Artsob, N. Bastien, K. Bernard, T. F. Booth, D. Bowness, M. Czub, M. Drebot, L. Fernando, R. Flick, M. Garbutt, M. Gray, A. Grolla, S. Jones, H. Feldmann, A. Meyers, A. Kabani, Y. Li, S. Normand, U. Stroher, G. A. Tipples, S. Tyler, R. Vogrig, D. Ward, B. Watson, R. C. Brunham, M. Kraiden, M. Petric, D. M. Skowronski, C. Upton, and R. L. Roper. 2003. The genome sequence of the SARS-associated coronavirus. *Science* **300**:1399–1404.
- Martina, B. E., B. L. Haagmans, T. Kuiken, R. A. Fouchier, G. F. Rimmelzwaan, G. Van Amerongen, J. S. Peiris, W. Lim, and A. D. Osterhaus. 2003. Virology: SARS virus infection of cats and ferrets. *Nature* **425**:915.
- Perri, S., C. E. Greer, K. Thudium, B. Doe, H. Legg, H. Liu, R. E. Romero, Z. Tang, Q. Bin, T. W. Dubensky, Jr., M. Vajdy, G. R. Otten, and J. M. Polo. 2003. An alphavirus replicon particle chimera derived from Venezuelan equine encephalitis and Sindbis viruses is a potent gene-based vaccine delivery vector. *J. Virol.* **77**:10394–10403.
- Polo, J. M., B. A. Belli, D. A. Driver, I. Frolov, S. Sherrill, M. J. Hariharan, K. Townsend, S. Perri, S. J. Mento, D. J. Jolly, S. M. Chang, S. Schlesinger, and T. W. Dubensky, Jr. 1999. Stable alphavirus packaging cell lines for Sindbis virus and Semliki Forest virus-derived vectors. *Proc. Natl. Acad. Sci. USA* **96**:4598–4603.
- Pulford, D. J., and P. Britton. 1991. Intracellular processing of the porcine coronavirus transmissible gastroenteritis virus spike protein expressed by recombinant vaccinia virus. *Virology* **182**:765–773.
- Ralston, R., K. Thudium, K. Berger, C. Kuo, B. Gervase, J. Hall, M. Selby, G. Kuo, M. Houghton, and Q. L. Choo. 1993. Characterization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia viruses. *J. Virol.* **67**:6753–6761.
- Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli, J. P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M. H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J. L. DeRisi, Q. Chen, D. Wang, D. D. Erdman, T. C. Peret, C. Burns, T. G. Ksiazek, P. E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A. D. Osterhaus, C. Drosten, M. A. Pallansch, L. J. Anderson, and W. J. Bellini. 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* **300**:1394–1399.
- Srivastava, I. K., L. Stamatatos, E. Kan, M. Vajdy, Y. Lian, S. Hilt, L.

- Martin, C. Vita, P. Zhu, K. H. Roux, L. Vojtech, D. C. Montefiori, J. Donnelly, J. B. Ulmer, and S. W. Barnett. 2003. Purification, characterization, and immunogenicity of a soluble trimeric envelope protein containing a partial deletion of the V2 loop derived from SF162, an R5-tropic human immunodeficiency virus type 1 isolate. *J. Virol.* **77**:11244–11259.
29. Wang, P., J. Chen, A. Zheng, Y. Nie, X. Shi, W. Wang, G. Wang, M. Luo, H. Liu, L. Tan, X. Song, Z. Wang, X. Yin, X. Qu, X. Wang, T. Qing, M. Ding, and H. Deng. 2004. Expression cloning of functional receptor used by SARS coronavirus. *Biochem. Biophys. Res. Commun.* **315**:439–444.
30. Wiley, D. C., J. J. Skehel, and M. Waterfield. 1977. Evidence from studies with a cross-linking reagent that the haemagglutinin of influenza virus is a trimer. *Virology* **79**:446–448.
31. Wilson, I. A., J. J. Skehel, and D. C. Wiley. 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289**:366–373.
32. Wong, S. K., W. Li, M. J. Moore, H. Choe, and M. Farzan. 2004. A 193-amino acid fragment of the SARS coronavirus S protein efficiently binds angiotensin-converting enzyme 2. *J. Biol. Chem.* **279**:3197–3201.
33. World Health Organization. 23 September 2003, posting date. Cumulative number of reported probable cases of severe acute respiratory syndrome (SARS). [Online.] http://www.who.int/csr/sars/country/table2003_09_23/en.
34. Wyatt, R., P. D. Kwong, E. Desjardins, R. W. Sweet, J. Robinson, W. A. Hendrickson, and J. G. Sodroski. 1998. The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* **393**:705–711.
35. Xiao, X., S. Chakraborti, A. S. Dimitrov, K. Gramatikoff, and D. S. Dimitrov. 2003. The SARS-CoV S glycoprotein: expression and functional characterization. *Biochem. Biophys. Res. Commun.* **312**:1159–1164.
36. Yewdell, J. W., A. Yellen, and T. Bachi. 1988. Monoclonal antibodies localize events in the folding, assembly, and intracellular transport of the influenza virus hemagglutinin glycoprotein. *Cell* **52**:843–852.