JOURNAL OF VIROLOGY, Apr. 2009, p. 3007–3018 Vol. 83, No. 7
0022-538X/09/$08.00+0 doi:10.1128/JVI.01505-08

Severe Acute Respiratory Syndrome Coronavirus nsp9 Dimerization Is Essential for Efficient Viral Growth▽

Zachary J. Miknis,[1] Eric F. Donaldson,[2] Timothy C. Umland,[1,3] Ryan A. Rimmer,[3] Ralph S. Baric,[2] and L. Wayne Schultz[1,3]*

Department of Structural Biology, State University of New York at Buffalo, Buffalo, New York 14203[1]; Department of Microbiology and Immunology, University of North Carolina at Chapel-Hill, Chapel-Hill, North Carolina 27599[2]; and Hauptman-Woodward Medical Research Institute, Buffalo, New York 14203[3]

Received 17 July 2008/Accepted 9 January 2009

The severe acute respiratory syndrome coronavirus (SARS-CoV) devotes a significant portion of its genome to producing nonstructural proteins required for viral replication. SARS-CoV nonstructural protein 9 (nsp9) was identified as an essential protein with RNA/DNA-binding activity, and yet its biological function within the replication complex remains unknown. Nsp9 forms a dimer through the interaction of parallel α-helices containing the protein-protein interaction motif GXXXG. In order to study the role of the nsp9 dimer in viral reproduction, residues G100 and G104 at the helix interface were targeted for mutation. Multi-angle light scattering measurements indicated that G100E, G104E, and G104V mutants are monomeric in solution, thereby disrupting the dimer. However, electrophoretic mobility assays revealed that the mutants bound RNA with similar affinity. Further experiments using fluorescence anisotropy showed a 10-fold reduction in RNA binding in the G100E and G104E mutants, whereas the G104V mutant had only a 4-fold reduction. The structure of G104E nsp9 was determined to 2.6-Å resolution, revealing significant changes at the dimer interface. The nsp9 mutations were introduced into SARS-CoV using a reverse genetics approach, and the G100E and G104E mutations were found to be lethal to the virus. The G104V mutant produced highly debilitated virus and eventually reverted back to the wild-type protein sequence through a codon transversion. Together, these data indicate that dimerization of SARS-CoV nsp9 at the GXXXG motif is not critical for RNA binding but is necessary for viral replication.

The discovery of a novel coronavirus (CoV) as the causative agent of severe acute respiratory syndrome (SARS; SARS-CoV) has highlighted the need for a better understanding of CoV replication (19). After emerging in late 2002 from the Guangdong Province in China, SARS-CoV was rapidly isolated, and its genome sequenced to reveal a new CoV that was phylogenetically distinct, suggesting a new classification as a type IIb CoV (15, 32, 43, 50). Genomic comparison to the closely related murine hepatitis virus (MHV), human CoV OC43 (HCoV-OC43), and bovine CoV revealed a highly conserved genomic structure with many regions nearly identical (52). After strict quarantine controls were initiated, SARS-CoV was contained, with approximately 8,000 individuals being clinically infected resulting in close to 800 deaths (www.who.int/csr/sars/en). Recently, the natural reservoir for SARS-CoV was reported as the Chinese horseshoe bat, indicating that (i) the disease is still circulating in animals and (ii) a future re-emergence from a zoonotic source is possible, underscoring the importance of continued study of this virus (36, 38).

CoVs devote a significant portion of their positive-sense single-stranded RNA (ssRNA) genome to proteins related to viral RNA replication. SARS-CoV has a genome of ~29,700 nucleotides, of which more than 21,000 code for the 16 nonstructural proteins (nsp's) (43, 50). The replicase genes open reading frame 1a (ORF1a) and ORF1b are translated into large polyproteins pp1a (4,300 amino acids [aa]), and through a −1 ribosomal frameshift mechanism, a fusion protein known as pp1ab (7,000 aa) (58). Posttranslational processing of the polyproteins by two distinct viral proteinases, the papain-like proteinase and a 3C-like proteinase ($3CL^{pro}$; also known as M^{pro}), yields 16 mature nsp's, many of which interact to form the replication complex responsible for the synthesis of negative-strand template, positive-strand genomic, and all subgenomic RNAs (sgRNA) with additional roles related to cellular processes (23, 30, 37, 52, 54, 68). The precise stoichiometry of the replicase complex is unknown, but yeast two hybrid screens, glutathione *S*-transferase pull-down assays and X-ray crystallography, have revealed a number of protein-protein interactions between the various nsps (27, 60, 67).

In general, proteins form multimers for a variety of reasons including stability, allostery, and to ensure accurate translation of genetic information (22). Translation of large polypeptides can result in errors in protein sequence and therefore large complexes are usually assembled from multiple small proteins. Viruses both violate and adhere to this rule by first producing large polyproteins and then processing them into individual proteins that assemble to form active replication complexes. For the CoVs, the protein components and stoichiometry of the positive- and negative-strand replication complexes remain unknown. Crystal structures of individual SARS-CoV nsp's have revealed several different multimeric states for the same protein. For example, the structure of nsp10 is reported as a dimer and as a dodecamer by separate groups (29, 55), although genetic data suggest that the dodecamer structure is

* Corresponding author. Mailing address: Department of Structural Biology, SUNY at Buffalo, Hauptman-Woodward Institute, 700 Ellicott St., Buffalo, NY 14203. Phone: (716) 898-8640. Fax: (716) 898-8660. E-mail: schultz@hwi.buffalo.edu.

▽ Published ahead of print on 19 January 2009.

not essential for in vitro replication (14). Biochemical studies of nsp7 and nsp8 indicate that, independently, they are dimeric in solution but the crystal structure of the nsp7 and nsp8 complex is a hexadecamer (67). For nsp9, there are four crystallographic structures which report a variety of dimeric interfaces (16, 49, 57). In one case, two interfaces are present in a single crystal form, suggesting that a tetrameric complex that incorporates both interfaces may be possible. Although the information provided by these structures is significant, the biological relevance to the replication complex has not been established.

SARS-CoV nsp9 has been shown to have RNA and DNA-binding ability through a variety of methods (16, 49, 57). nsp9 from MHV-A59 has been shown through immunofluorescence studies to localize with, among others, the helicase (nsp13), nucleocapsid (N protein) and $3CL^{pro}$ (nsp5) (4, 5). nsp9 also localizes to late endosomes at sites of replication with nsp7, nsp8, and nsp10 and is likely a member of the replication complex (4). An nsp9 knockout in MHV-A59 is not viable, while fusion of a nsp9-10 polyprotein is viable but attenuated in growth, suggesting that the mature form of nsp9 plays a critical role in viral replication (13).

Several crystallographic structures of nsp9 have shown that it is composed of seven beta strands and a single alpha helix (16, 49, 57). The fold of nsp9 is reported to be similar in fold to domains I and II of the $3CL^{pro}$ encoded within the SARS-CoV genome (16); however, no significant sequence similarity exists between the two. The presumed biological dimer utilizes the interaction between the lone helices of each monomer to form the parallel helix-helix dimer (Fig. 1A). A second dimeric form has a beta-sheet interface stabilized by main chain atom interactions within the sheet regions of each monomer (Fig. 1B). A recent crystal structure of HCoV 229E nsp9 reveals an antiparallel helix-helix dimer formed by a disulfide bond at Cys69 (Fig. 1C). Upon generating a multiple sequence alignment (Fig. 1D) of various CoV nsp9 proteins, it was found that while the helix-helix interface found in Fig. 1A contained multiple conserved residues along the dimer interface and buried ~1,000 $Å^2$ of surface-exposed area, the sheet-sheet dimer in Fig. 1B contained no conserved residues in the dimer region and buried only ~500 $Å^2$. A survey of dimer interfaces found in protein structures suggests for a protein of ~15 kDa a buried surface area of ~1,000 $Å^2$ would be expected (2, 28).

The common protein-protein interaction motif GXXXG (20, 31) is conserved at the dimer interface, allowing the helices to closely pack at the positions of G100 and G104 (Fig. 1A). Therefore, we designed mutagenesis experiments to investigate the stability and function of that dimer. Mutants G100E, G104E, and G104V were created to disrupt the dimeric interface and were characterized by size exclusion chromatography (SEC), multi-angle light scattering (MALS), and circular dichroism (CD) spectroscopy. Effects on ssRNA binding were assessed by RNA electrophoretic mobility shift assay (EMSA) and fluorescence anisotropy (FA). The crystal structure of G104E was solved and refined to 2.6-Å resolution, and changes in the helix-helix interface were observed.

The development of reverse genetics approaches to studying CoVs has also allowed us to reintroduce the G100E, G104E, and G104V mutations into SARS-CoV (Urbani strain) to study their effects on the virus in vivo (65). The G100E and G104E mutations were lethal to the virus, while the G104V mutation produced a highly debilitated growth phenotype with eventual transversion of the codon at position 104 from GTG (Val) to GGG (Gly), indicating that dimerization of nsp9 at the GXXXG interface is required for efficient viral growth.

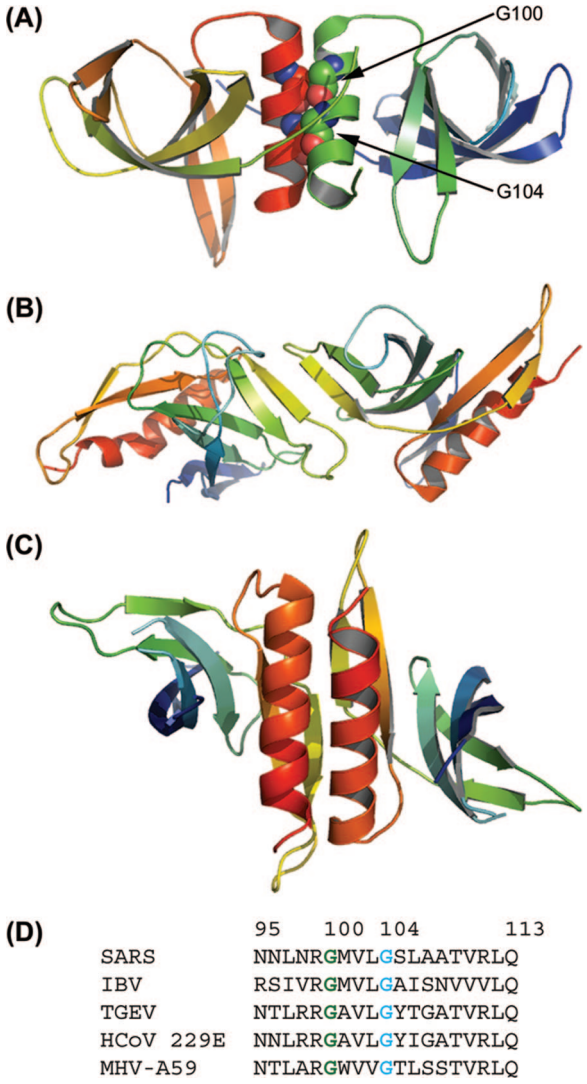


FIG. 1. Dimer arrangements in nsp9. (A) nsp9 crystal structure (1QZ8) showing presumed biological dimer and helix-helix interface. The individual monomers are colored in blue/green and red/yellow, respectively. The positions of G100 and G104 are depicted as space-filling models for both monomers. G100E and G104E are labeled in one monomer. (B) Alternate nsp9 dimer (1UW7) stabilized through sheet regions. Each monomer is colored from the N terminus (blue) to the C terminus (red). (C) Antiparallel helix-helix dimer of HCoV-229E nsp9 stabilized by a disulfide linkage at C69. Each monomer is colored from the N terminus (blue) to the C terminus (red). (D) Multiple sequence alignment of CoV nsp9 homologs showing absolute conservation of glycines equivalent to G100 (green) and G104 (cyan) in SARS-CoV. Images were prepared by using PyMol.

MATERIALS AND METHODS

Mutant generation. SARS-CoV nsp9-pET23d(+) plasmid containing the coding sequence of nsp9 in addition to a C-terminal His_6 tag was a gift from Mark Denison (Vanderbilt University). Nsp9 G100E, G104E, and G104V

mutants were generated by using a QuikChange site-directed mutagenesis kit (Qiagen) according to the manufacturer's protocol using the following primers: 9G100EF, 5′-AACAACCTAAATAGAG**AA**ATGGTGCTGGGCAGTT TAGC-3′; 9G100ER, 5′-GCTAAACTGCCCAGCACCAT**TT**CTCTATTTA GGTTG-3′; 9G104EF, 5′-GAGGTATGGTGCTGG**AA**AGTTTAGCTGCTA C-3′; 9G104ER, 5′-GTAGCAGCTAAACT**TT**CCAGCACCATACCTC-3′; 9G104VF, 5′-GAGGTATGGTGCTGG**TG**AGTTTAGCTGCTAC-3′; and 9G104VR, 5′-GTAGCAGCTAAACT**CA**CCAGCACCATACCTC-3′ (bold facing indicates sites of mutation). Mutagenesis products were transformed into DH5α cells (Invitrogen, Carlsbad, CA) and plated on LB agar plates (supplemented with 100 μg of ampicillin/ml) from which colonies were selected and grown in 5-ml cultures of LB supplemented with 100 μg of ampicillin/ml. Plasmid was purified by using a Qiagen spin miniprep kit according to the manufacturer's protocol. The presence of the mutations within the plasmids was confirmed through DNA sequencing on an ABI Prism 3130XL genetic analyzer (Roswell Park Cancer Institute Biopolymer Facility).

Expression and purification. Plasmids were each transformed into BL21(DE3) (Novagen) cells and screened for expression conditions. Soluble protein was obtained through the addition of IPTG (isopropyl-β-D-thiogalactopyranoside) to 1 mM in 1 liter of culture grown in LB (supplemented with 100 μg of ampicillin/ml and grown to an optical density at 600 nm of ~0.6 at 37°C), followed by shaking at 37°C for 3 h. Cells were collected by centrifugation and stored at −80°C.

All purification steps were carried out at 4°C. Cell pellets from 1-liter cultures were thawed, resuspended in 15 ml of buffer A (50 mM Tris [pH 8.0], 300 mM NaCl, 10 mM imidazole) and 1 ml of protease inhibitor cocktail (Roche), and then lysed by using a single pass through a Microfluidizer (Microfluidics Co.) at ~18,000 lb/in^2. The resulting lysate was centrifuged at 35,000 rpm at 4°C for 30 min in a 45 Ti rotor (Beckman). Supernatant was filtered (0.22-μm pore size) and applied to a 5-ml HisTrap (GE Healthcare) column pre-equilibrated in buffer A. Protein was washed with buffer A followed by 12% buffer B (50 mM Tris [pH 8.0], 300 mM NaCl, 300 mM imidazole) and eluted from the column in a linear gradient from 12 to 100% buffer B. Fractions were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and nsp9 containing fractions were pooled and dialyzed against 1 liter of buffer C (10 mM Tris [pH 8], 150 mM NaCl, 1 mM EDTA, 5 mM dithiothreitol). Concentrated protein was subsequently applied to a Superdex 75 HL 16/60 column pre-equilibrated in buffer C. Fractions were analyzed by SDS-PAGE with appropriate fractions pooled and dialyzed against either 1 liter of buffer D (50 mM MES [morpholineethanesulfonic acid; pH 5.6], 15 mM NaCl, 0.5 mM TCEP [tri(2-carboxyethyl) phosphine HCl]) or 1 liter of buffer C. Final samples were stored at 4°C and assayed for protein concentration at 280 nm using an extinction coefficient of 13,075 M^{-1} cm^{-1}calculated by ProtParam (63) using the coding sequence of the mature nsp9 product with a C-terminal tag (LEHHHHHH).

SEC and SEC-MALS. Aliquots of elution fractions from immobilized metal affinity chromatography (IMAC) purification were applied to a Superdex 75 HR 10/30 column (GE Healthcare) preequilibrated in buffer E (10 mM $NaPO_4$ [pH 8], 100 mM NaCl). The Superdex column was attached on an AKTA Purifier coupled with a Wyatt Systems MiniDAWN (three detectors) and Optilab DSP (Wyatt Systems) immediately downstream of the column. Protein samples were applied and eluted from the column at 0.5 ml/min with simultaneous collection of light scattering and refractive index data, using a laser excitation wavelength of 690 nm and light scattering detectors at 45°, 90°, and 135°. Peak analysis was performed by using the ASTRA software provided with the system.

CD spectroscopy. Protein samples were passed over a Superdex 75 HL 16/60 column in buffer F (25 mM $NaPO_4$ [pH 8], 150 mM NaCl) to remove any nonspecific aggregates. Appropriate fractions were pooled and dialyzed against 1 liter of buffer G (10 mM $NaPO_4$ [pH 8.0]). CD spectra of nsp9 samples were gathered from 200 to 255 nm in a 1-mm-pathlength quartz cuvette at 20°C (1-nm steps, 50-nm/s scan speed, 4-s response time, three scans) using a Jasco J-715 Spectrapolarimeter (Department of Pharmaceutical Sciences, University at Buffalo). Secondary structure assignment was achieved through use of the K2d web server (www.embl-heidelberg.de/~andrade/k2d) and the DICHROWEB server (www.cryst.bbk.ac.uk/cdweb/html/home.html) (1, 40, 62).

For thermal denaturation experiments using CD spectroscopy, a similar protocol was followed. Samples were monitored while heating from 20 to 95°C using a Peltier device. Spectra were gathered at 5° intervals, with peak monitoring at 205 nm. The data were analyzed using Origin 7.0.

RNA EMSA. A labeled ssRNA oligonucleotide (Integrated DNA Technologies) was used for gel shift assays (Biotin-5′-CGACUCAUGGACCUUGGCA G-3′). Oligonucleotide was resuspended in RNase-free water at 1 μM and stored at −20°C. 5 μg of nsp9 samples (in buffer D) were mixed with 1 μM RNA, 1 μM ssDNA oligonucleotide (3CL NdeI Forward2 [5′-GGTGGTCATATGAGTGG TTTTAGGAAAATGGCATT-3′]), and RNase inhibitor (SUPERaseIN; Ambion), followed by incubation for 60 min at room temperature. After incubation, the samples were cross-linked at 254 nm for 15 min. High-density Tris-borate-EDTA (TBE) loading buffer (Invitrogen) was added to the nucleic acid-protein mixtures, which were loaded onto Novex 4 to 20% TBE polyacrylamide gels (Invitrogen) and resolved at 80 V for 2 h in 0.5× TBE buffer. The gels were then transferred to positively charged nylon membrane (Hybond) in 0.5× TBE at 25 V for 30′ in a semidry transfer apparatus (Bio-Rad). Membranes were cross-linked for 5 min in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at 254 nm. The membranes were then developed using a BrightStar biodetection kit (Ambion) according to the manufacturer's protocol. Extended wash steps were incorporated to help further decrease the nonspecific background of the membrane.

Attempts were also made to compete the RNA probe off of nsp9 using an ssDNA template which was added at equivalent and 10-fold-higher levels compared to the RNA probe. Developed membrane was exposed to Hyperfilm ECL (GE Healthcare) and analyzed for electrophoretic mobility.

FA. FA measurements were performed on a Fluoromax-4 spectrophotometer (Jobin Yvon Horiba) equipped with a temperature-controlled cell and polarizing filters. All experiments were carried out at 22°C. 5′-Fluorescein-labeled RNA (5′ FAM-CGACUCAUGGACCUUGGCAG-3′; IDT) was used in all experiments. FAM-RNA was dissolved in buffer (20 mM Tris [pH 7.2]) to a final concentration of 58 nM in a 1-ml quartz cuvette. Small aliquots (1 to 2 μl) of nsp9 or nsp9 mutants (in 10 mM MES [pH 5.6]) were added to the cuvette covering a protein concentration range of 1 to 2,000 nM, followed by incubation with stirring for 5 min between measurements. FA was measured by exciting at 490 nm and monitoring the emission at 515 nm. Slits for the excitation and emission were set to 2.5 nm. The integration time was 2.0 s for the anisotropy measurements. FA data collection was controlled by using the FA module of the Fluoressence software (version 2.1.5.0). Each anisotropy value is the average of 10 individual anisotropy measurements. The relative standard deviation was <2% for all measurements. The data were analyzed using nonlinear regression fitting of the data to the following equation (42) for a single site binding model using Prism software:

$$A = A_f + (A_b - A_f) \times \left[\frac{1 + K_a[L_T] + K_a[R_T] - \sqrt{\{(1 + K_a[L_T] + K_a[R_T])^2 - 4[L_T][R_T]K_a^2\}}}{2K_a[R_T]} \right]$$

ITC. Dilution isothermal titration calorimetry (ITC) is a technique amenable to determining protein dimerization constants (7, 9). Purified protein samples were subjected to buffer exchange five times in an Amicon centrifugal concentrator to exchange protein in appropriate buffer (20 mM MES [pH 5.6], 15 mM NaCl, and 0.1 mM TCEP for wild type, G100E, and G104E; 20 mM Na TAPS {N-[Tris(hydroxymethyl)methyl]-3-aminopropanesulfonic acid} [pH 8.5], 100 mM $MgCl_2$, and 0.1 mM TCEP for G104V). Protein samples and corresponding buffer were filtered to 0.2 μm (Acrodisc syringe filter, low protein binding HT Tuffryn membrane; PALL Life Sciences) and thoroughly degassed.

Buffer corresponding to the sample to be studied was loaded into the reference and sample cells of the VP-ITC (Microcal). Typically, 28 injections of 10 μl of protein were made into the sample cell filled with buffer. Concentrations of nsp9 in the injection syringe ranged from 189 μM to 6.3 mM, and concentration ranges in the sample cell were monitored from ~0.001 mM to ~0.3 mM (near the published dimerization K_d). Additional measurements on samples were made up to a final concentration of ~1 mM within the sample cell. Samples which showed a saturation curve were fit to a simple dissociation model using the included analysis software (Origin 7) specifically designed for dilution ITC data analysis.

Mutant design for reverse genetics. To determine the effects of mutations G100E, G104E, and G104V on SARS-CoV replication, the three mutations were individually engineered into the molecular clone of SARS-CoV, replacing the wild-type codons as follows: G100E, GGT to GAG; G104V, GGC to GTG; and G104E, GGC to GAA. This was accomplished by using the "No See'm" (66) approach whereby primers were designed utilizing type IIS restriction endonuclease sites which allowed sticky ends to be generated with non-native nucleotides present. This was used to introduce the mutations into the codons described above, using the wild-type SARS-D fragment from the infectious clone of SARS-CoV strain Urbani (icSARS) as the backbone. The primers used were as follows: G100E mutant, G100EF (5′-AGAGAGATGGTGCTGGGCAGT-3′) and G100ER (5′-ACTGCCCAGCACCATCTCTCT-3′); G104E mutant, G104EF (5′-AGAGGTATGGTGCTGGAAAGT-3′) and G104ER (5′-ACTTTCCAGC ACCATACCTCT-3′); and mutant G104V, G104VF (5′-AGAGGTATGGTGC TGGTGAGT-3′) and G104VR (5′-ACTCACCAGCACCATACCTCT-3′). The

primers used to amplify the nsp9 cDNA were M13R (5′-AGGAAACAGCTATGAC-3′) and nsp9R (5′-AAGTTCCTTGAAACTGAGACG-3′). Each mutation was engineered in this way into the SARS-D fragment and the full-length mutated D fragments were digested, purified, and ligated back into the original vector. Transformation was conducted by heat shock at 42°C for 2 min, and transformed *E. coli* was plated on LB plates with appropriate selection. Plates were grown at room temperature for 48 h, after which five colonies were picked and amplified in 5 ml of LB broth incubated at room temperature, under appropriate selection, and with agitation. After 24 h, the plasmids were purified and digested with restriction endonuclease BglI to determine whether the DNA resembled wild type by restriction screening. Plasmids that digested correctly were amplified in 20 ml of LB broth with the appropriate antibiotic and grown for 24 h at room temperature with agitation. The plasmids were again purified, and three of each mutant were sequenced to verify that the mutation was incorporated into the wild-type SARS-D fragment.

Assembling the mutant viruses. The infectious clone of the Urbani strain of SARS-CoV was used as the backbone for this project, and mutated SARS-D fragments containing each of the mutations of interest were assembled into the full-length infectious cDNA as described previously (64, 66). Briefly, icSARS fragments A through F were amplified in *E. coli* (TOP-10; Invitrogen), purified, and screened by restriction digestion. Large stocks of each fragment that screened correctly were established and digested with the appropriate enzyme as follows. icSARS fragments B, C, D and E were digested with BglI. icSARS fragments A and F were digested with EcoRI and NotI, respectively. icSARS fragments A and F were then dephosphorylated and digested with BglI.

The digested bands were subjected to electrophoresis, excised, and gel purified (Qiagen), and ligation reactions were set up using equivalent molar ratios of each fragment and ligase (Roche). Wild-type fragments were ligated to generate a full-length cDNA of wild-type SARS-CoV, and the SARS-D fragment bearing each mutation was used to generate full-length cDNAs for each mutant. The ligation reactions were purified by chloroform extraction and isopropanol precipitation, and nucleocapsid cDNA and full-length viral genomic cDNA were used as templates for in vitro transcription reactions (Ambion). Nucleocapsid and full-length viral genomic transcripts were then electroporated into Vero cells. Transfections were monitored for 3 days and passaged if necessary. All recombinant icSARS strains were propagated on Vero E6 cells in Eagle minimal essential medium (Invitrogen) supplemented with 10% fetal calf serum (HyClone, Logan, UT), kanamycin (0.25 μg/ml), and gentamicin (0.05 μg/ml) at 37°C in a humidified CO_2 incubator. All work was performed in a biological safety cabinet in a biosafety level 3 laboratory containing redundant exhaust fans. Personnel were equipped with powered air-purifying respirators with high-efficiency particulate air and organic vapor filters (3M, St. Paul, MN), wore Tyvek suits (DuPont, Research Triangle Park, NC), and were double gloved.

Verification of viral replication. The cells were analyzed daily for cytopathic effect (CPE), and viral subgenomic transcription was verified by reverse transcription-PCR (RT-PCR) of leader containing transcripts of the S gene. Briefly, primers pairs were designed to amplify the leader RNA sequence and ~300-nucleotide domain into the S gene. Total RNA was harvested from transfected cells by using TriZol reagent, and purified RNA was reverse transcribed using random hexamer primers and Superscript III reverse transcriptase, according to the manufacturer's protocol. S gene specific leader containing cDNA was then amplified by PCR.

Analysis of mutant G104V RNA replication. The SARS-CoV infectious clone bearing the G104V mutation was assembled in triplicate as described above and transfected via three independent cuvettes containing 8×10^6 Vero cells. The transfected cells from each cuvette were then divided into three T25 flasks, and medium was added; therefore, three independent transfections were each divided into three flasks for a total of nine flasks. One flask from each transfection was harvested at 12, 24, and 36 h posttransfection. As a control, icSARS was also transfected into cells and harvested in the same way. In each case, transfected cells were isolated in TRIzol reagent (Invitrogen) using the manufacturer's protocol for isolation of total RNA by isopropanol precipitation and eluted to 50 μl in nuclease-free water. All samples were then incubated for 1 h at 37°C with DNase (Applied Biosystems, Foster City, CA) according to manufacturer's protocol to eliminate any residual input cDNA.

Total RNA from each time point and transfection was reverse transcribed to cDNA using SuperScript III (Invitrogen) with modifications to the protocol as follows. Random hexamers (300 ng) and total RNA (5 μl) were incubated for 10 min at 70°C. The remaining reagents were then added according to the manufacturer's recommendations, and the reaction was incubated at 55°C for 1 h, followed by 20 min at 70°C to deactivate the reverse transcriptase. Quantitative real-time PCR was conducted using SmartCycler II (Cepheid, Sunnyvale, CA) with SYBR green (diluted to 0.25×; Cepheid) to detect subgenomic cDNA with primers (7.5 pmol) optimized to detect from the leader sequence to the 5′ end of the N gene (forward N1S, AAAGCCAACCAACCTCGATC; reverse N1A, GCGTCCTCCATTCTGGTTAT) or genomic cDNA with primers (7.5 pmol) designed to detect genomic ORF1a cDNA (forward G104Vf, AGAGGTATGGTGCTGGTGAGT; reverse nsp9r, AAGTTCCTTGAAACTGAGACG). The cDNA from the RT reaction mixtures of each virus was used, at a volume of 5 μl for each reaction, with a total reaction mixture volume of 25 μl. Omnimix beads (Cepheid) containing all reagents except SYBR green, primer, and template were used to standardize the reaction conditions, and template concentrations were normalized by concentrations of the housekeeping gene GAPDH (glyceraldehyde-3-phosphate dehydrogenase; forward GAPDHF, CATGGGGAAGGTGAAGGTCG; reverse GAPDHR, TTGATGGTACATGACAAGGTGC). In addition, all products were verified by melting curve analysis.

Crystallization, cryopreservation, and diffraction. Purified nsp9 G104E was crystallized by the hanging-drop vapor diffusion method utilizing a Hangman system (41). Protein was concentrated to 5.8 mg/ml and mixed in a 1:1 ratio with crystallization cocktail (1.7 to 1.8 M ammonium sulfate, 0.1 M citrate phosphate buffer [pH 4 to 4.3]). Hexagonal rods generally appeared within 1 week and grew to a maximum of 0.3 by 0.1 by 0.1 mm.

Crystals were transferred for 30 min to 1 h into cryobuffer (2 M ammonium sulfate, 0.1 M citrate phosphate buffer [pH 4.0 to 4.2], 25% [vol/vol] glycerol). Cryoprotected crystals were mounted onto CryoLoops (Hampton Research), quickly plunged into liquid nitrogen, and stored in a cryodewar.

Diffraction experiments were remotely conducted on beamline 11-1 of the Stanford Synchrotron Radiation Laboratory (SSRL) using the Blue-Ice interface and Web-Ice analysis software (21, 44). The data were collected at a wavelength of 0.979 Å using a 20-s exposure and 1° oscillation per frame on a MAR-325 charge-coupled device detector. The data were indexed, integrated, merged, and scaled to 2.6-Å resolution, as implemented within the HKL2000 package (47).

Phasing and refinement. Processed data was input in the CaspR (10) server for molecular replacement trials using a variety of search models, including the previously solved nsp9 structures (Protein Data Bank codes 1QZ8 and 1UW7) (16, 57). The best solution was found using monomeric 1QZ8 (16) as a search model and resulted in a model with four copies of nsp9 within the asymmetric unit.

The nsp9 G104E model was subjected to a single round of rigid body refinement as implemented within Refmac5 as part of the CCP4 package (3). The model was then refined using alternating cycles of manual fitting in Coot and TURBO-FRODO (8, 17) and simulated annealing in CNS 1.2 (6). Composite simulated annealing omit maps were generated at regular intervals to guide the modeling efforts. After several rounds of rebuilding group B factors were refined in conjunction with 50 cycles of energy minimization (CNS). In the final rounds of rebuilding and refinement, waters, ions, and cryoprotectant molecules were added to the model through the use of an F_o-F_c map and a $2F_o$-F_c omit map contoured to 3 σ and 1 σ, respectively. Model quality was monitored by using CNS and Molprobity (11). The final structure containing four monomers of nsp9, 134 waters, eight PO_4, and 16 glycerols has been deposited into the Protein Data Bank (PDB) with the accession code 3EE7. The data and refinement statistics are listed in Table 1.

Structure analysis. The individual monomers of nsp9 G104E were aligned (least-squares superposition) with one another and with monomers from 1UW7 (57) and 1QZ8 (16). Dimer-only structures were also created from the G104E structure and superimposed on the analogous dimers from 1QZ8 (16) or 1UW7 (57). In all cases, a C_α root mean square deviation (RMSD) was calculated by using either PyMol (12) or Coot. In the case of features such as the symmetry related helix-helix dimer formed by G104E, manual inspection (Coot) was made of the helices compared to 1QZ8 (16). Further examination of the helix-helix dimerization interface was accomplished through ProMotif, as found within the PDBSum webserver (26, 33–35).

RESULTS

Expression and purification. Variants of nsp9 were generated with the purpose of disrupting the helix-helix dimer found in the crystal structure of nsp9 as solved by Egloff et al. (16) to assess its biological relevance. Mutations at G100 and G104 were created and verified through DNA sequencing. Wild-type nsp9, along with the G100E, G104E, and G104V mutants, were found to induce to high levels of soluble expression upon induction with IPTG.

TABLE 1. X-ray data and refinement statistics

General category and specific parameter	Value[a]
Crystal data	
Space group	P31
a (Å)	91.836
b (Å)	91.836
c (Å)	84.217
α, β, γ (°)	90, 90, 120
Mol/AU[b]	4
Vm[c]	3.81
% Solvent	67.76
Data collection	
Resolution range (Å)	79.566–2.6 (2.69–2.60)
Wavelength (Å)	0.9792
No. of observations	235,950
No, of unique reflections	24,437
I/σI	34.01 (2.75)
Redundancy	3.7 (2.8)
Data completeness	98.8
R_{merge}[d]	5.3 (39.4)
Refinement	
No. of reflections used	24,148
Resolution range (Å)	79.566–2.6
Free set size (%)	9.80
No. of atoms (protein)	3,462
No. of atoms (nonprotein)	252
R_{work}[e]	0.2138
R_{free}[f]	0.2694
Mean B value (Å^2)	70.66
RMSD	
Bond length (Å)	0.006402
Bond angle (°)	1.69421
Ramachandran plot	
Favored (%)	91.8
Allowed (%)	7.3
Disallowed (%)	0.9

[a] Numbers in parentheses are for the highest resolution shell.
[b] Molecules per asymmetric unit.
[c] Matthews coefficient.
[d] $R_{merge} = \frac{\Sigma_{hkl}(\Sigma_i(|I_{hkl,i} - \langle I_{hkl}\rangle|))}{\Sigma_{hkl,i}\langle I_{hkl}\rangle}$, where $I_{hkl,i}$ is the intensity of an individual measurement of the reflection with Miller indices h, k, and l, and $\langle I_{hkl}\rangle$ is the mean intensity of that reflection.
[e] $R_{work} = \frac{\Sigma_{hkl}(||Fobs_{hkl}| - |Fcalc_{hkl}||)}{|Fobs_{hkl}|}$, where $|Fobs_{hkl}|$ and $|Fcalc_{hkl}|$ are the observed and calculated structure amplitudes.
[f] R_{free} is R_{work} calculated with the free set data only, which were omitted from refinement.

Wild-type, G100E, G104E, and G104V forms of nsp9 were purified using IMAC (HisTrap HP; GE Healthcare), which yielded protein of 85% purity, aliquots of which were used for size estimation by SEC and SEC-MALS. Further purification to obtain stocks for crystallization, EMSA, and FA analyses was obtained by applying protein from the IMAC column to a preparative gel filtration column (Superdex 75 HL 16/60; GE Healthcare). The final protein yield was of ~98% purity as judged by SDS-PAGE stained with silver (data not shown). Due to the presence of a noncleavable C-terminal His_6 tag and linker, the pure nsp9 monomer had an expected molecular mass of 13.4 kDa. The structure of the G104E mutant in addition to the wild-type nsp9 structures found in the PDB suggested that the location of the C-terminal His_6 tag does not affect the stability of the helix-helix dimer.

TABLE 2. nsp9 sizes estimated by SEC-MALS and SEC

Protein	Size (kDa)		
	Monomer	Superdex 75	SEC-MALS
nsp9 wild type	13.4	26.4	25.5
nsp9 G100E	13.5	18.5	15
nsp9 G104E	13.5	17.8	16.3
nsp9 G104V	13.5	17.26	13.64

Size estimation using SEC and SEC-MALS. For analysis of the mutants' oligomeric state, aliquots of peak fractions from IMAC were immediately applied to an analytical gel filtration column. The elution from the Superdex column was directed to a Wyatt Systems Mini-DAWN coupled with an Optilab DSP to allow for simultaneous collection of SEC data in conjunction with MALS data.

While SEC is commonly used to estimate the molecular weight of a sample and thus the oligomerization state, the molecular weight estimates are highly dependent on the conformation of a protein and any interactions between the protein and the chromatographic matrix. In effect, a protein with a nonglobular structure or one that interacts with the SEC matrix by another property other than size (e.g., electrostatic, hydrophobic, etc.) will elute at a volume corresponding to a protein with a much larger size. Size estimates made by MALS are independent of nonideal SEC behavior and have been shown to give a more accurate size determination for a variety of proteins than is possible using SEC alone (18). The addition of the SEC to the system has the added benefit of effectively separating all species within a sample prior to their analysis by MALS, leading to data which is more amenable to size assignment and has an expected error of only ±5% (18).

Nsp9 wild type eluted as a single peak and light scattering estimated a molecular size of 25.5 kDa, corresponding to a particle size very close to the predicted value of 26.8 kDa for a dimer. In contrast, the SEC-MALS data on the nsp9 mutants gave molecular size estimates in good agreement with the protein being monomeric (Table 2). Molecular size estimates were also made using elution data from SEC alone by calculating a calibration curve using well-characterized proteins, also shown in Table 2. Sutton et al. reported K_d values of 0.16 to 0.46 mM for the formation of the dimer (57). However, our gel filtration experiments on wild-type nsp9 using an analytical column indicate that dimerization occurs at protein concentrations between 40 and 105 μM. Similar experiments showed G100E, G104E, and G104V mutants were monomeric at or above concentrations where the wild-type protein was dimeric.

The transition of nsp9 wild type from dimer to monomer could not be detected using ITC. The concentration range required to capture the dissociation from dimer to monomer was too low to generate any substantial change in enthalpy above the background. The G100E and G104E samples behaved in ITC experiments as monomers and showed no signs of a monomer-dimer equilibrium. A weak dimerization constant of 0.871 mM was determined for G104V.

Secondary structure analysis. Analysis of nsp9 samples through the use of CD spectroscopy showed that the overall

TABLE 3. nsp9 structural features as measured by CD spectroscopy

Sample	%			Mean T_m ± SD (°C)
	Sheet content	Helix	Random	
nsp9 wild type	39	9	51	62.5 ± 1.6
nsp9 G100E	46	8	46	60.5 ± 3.8
nsp9 G104E	39	9	51	75 ± 0.7
nsp9 G104V	44	9	47	58.6 ± 2.3
Expected value	41	6	52	NA[a]

[a] NA, not applicable.

secondary structure composition was maintained in all samples. As shown in Table 3, the secondary structure distribution in the wild-type, G100E, G104E, and G104V proteins was estimated from the CD data. Given that structure assignment for all samples was nearly identical to percentages predicted from the crystal structure, it was determined that the disruption of dimerization seen in the SEC and SEC-MALS data was not due to protein misfolding or the C-terminal His tag. In particular, the amount of α-helix remained similar to the wild type, indicating that the mutations had not affected the helical structure of the dimerization interface. In addition, the crystal structure of the G104E mutant (discussed below) showed that secondary structure was not disrupted by the mutation. Measurement of the melting temperature of the samples using CD spectroscopy revealed that wild-type, G100E, and G104V proteins have similar melting temperatures (58.6 to 62.5°C), while G104E protein had a melting temperature that was significantly higher (75°C). While it is unclear why G104E is more stable, it does indicate that the individual mutations do not destabilize nsp9.

Gel shift assays. In order to determine whether the mutations and subsequent disruption of dimerization altered the ability of nsp9 to bind to RNA, we performed RNA EMSAs using standard techniques. To maximize the visualization of RNA-protein complexes, samples were UV cross-linked prior to separation of complexes from free RNA through electrophoresis (24, 25). RNA EMSA experiments revealed that wild-type nsp9 exhibited clear binding with the RNA templates (Fig. 2). Bovine serum albumin controls did not bind RNA (data not shown). Importantly, the G100E, G104E, and G104V mutants were also able to bind ssRNA (Fig. 2). To date, the RNA-binding surface for nsp9 has not been determined; however, the ability of the mutants to bind suggests that the dimer interface is not critical for RNA binding. Changes in mobility were seen between the wild type and the mutants, but this could be a result of the change in protein charge or size of the complexes due to the monomeric or dimeric state of the protein. Unlabeled ssDNA probe at a 10-fold excess was unable to compete with the ssRNA for binding and resulted in better resolution of the nsp9/RNA bands (Fig. 3).

FA. While EMSA allowed for the visualization of the interactions between the RNA and nsp9, it did not allow for quantitative measurements of the binding affinities and comparisons to known data. Consequently, RNA-binding measurements were made using a fluorescently labeled 20-mer of ssRNA on a fluorometer equipped for measurement of FA data (Fig. 4).

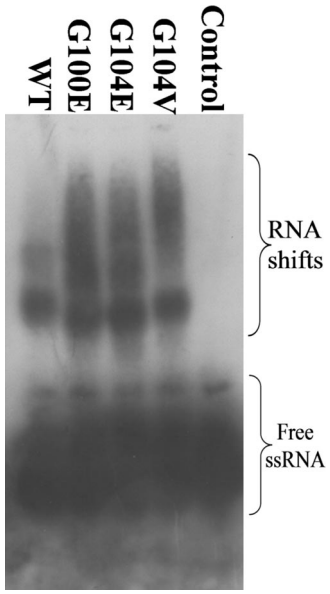


FIG. 2. Wild-type nsp9 along with mutants bind ssRNA. Lane 1, nsp9 wild type/RNA (WT); lane 2, G100E/RNA; lane 3, G104E/RNA; lane 4, G104V/RNA; lane 5, ssRNA. The positions of free probe and protein/RNA shifts are indicated. Protein-RNA complexes were resolved on a 4 to 20% TBE gel and observed after detection of the biotinylated ssRNA.

Wild type bound to ssRNA with a K_d of 55 nM compared to K_ds of 250, 620, and 320 nM for G104V, G104E, and G100E, respectively.

Reverse genetics. The three mutants were designed for and incorporated into the wild-type SARS-D fragment of icSARS. The three mutant SARS-D fragments were assembled, amplified, sequence verified, and placed into the infectious clone of SARS-CoV. The full-length cDNAs were then transcribed and

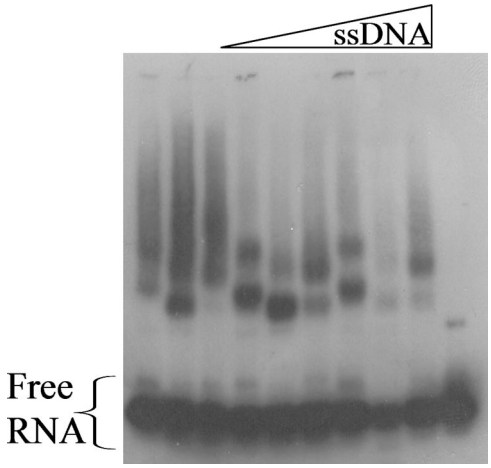
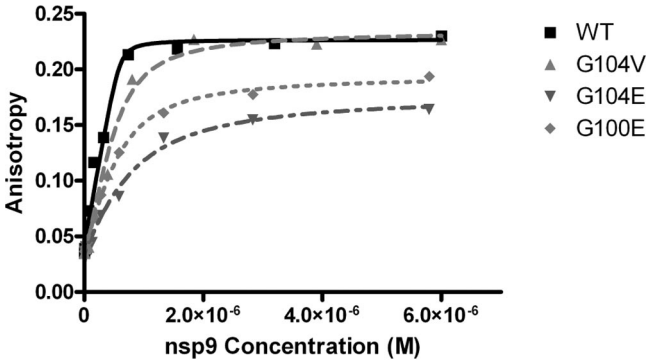


FIG. 3. ssDNA does not prevent a ssRNA from binding to nsp9. Wild-type nsp9 was loaded into lanes 1, 4, and 7. nsp9 G100E is found in lanes 2, 5, and 8. G104E nsp9 is loaded into lanes 3, 6, and 9. Lanes 1 to 3 contain only RNA and protein, lanes 4 to 6 contain equivalent amounts of ssRNA and ssDNA, and lanes 7 to 9 contain an 10-fold excess of ssDNA over ssRNA. Lane 10 is an RNA-only loading control.



Protein	K_d
WT	55 nM
G100E	320 nM
G104E	620 nM
G104V	250 nM

FIG. 4. FA measurement of RNA binding by wild-type (WT) (■), G100E (◆), G104E (▼), or G104V (▲) nsp9. Each measurement is an average of 10 individual measurements. The data were fitted to the equation given in Materials and Methods to generate the curves depicted in the graph and derive binding equilibrium K_d values.

transfected into Vero cells and monitored for CPE. In all cases, detection of CPE was not evident. To determine whether low-level RNA synthesis occurred, RT-PCR was conducted to assay for leader containing transcripts using primers designed to detect subgenomic S, but no bands were detected, suggesting either that no sgRNA synthesis was occurring or that transcript levels were below the levels of detectability.

Mutants G100E and G104E were introduced into the infectious clone three times, and in all three cases no recombinant mutant virus was detected or rescued. However, recombinant wild-type icSARS prepared in parallel did produce CPE and sgRNA synthesis. For mutant G104V, the experiment was repeated four times. The first two times there was no evidence of viral replication by cytopathology or by RT-PCR. The third time, however, CPE was evident, and subgenomic S was detected by RT-PCR. In addition, the recombinant virus was successfully passaged in Vero cells and also produced wild type-sized plaques. Total viral RNA was harvested and the region flanking the mutation site was amplified by RT-PCR, electroporated on a 0.8% agarose gel, gel purified, and sequenced. Interestingly, sequence analysis revealed that the G104V mutation had reverted to wild type-SARS-CoV at that position. The mutated codon introduced was GTG which encodes a Val, and this reverted to GGG which encodes a Gly. Interestingly, the wild-type virus had a Gly at that position, but the codon was GGC. Since two changes were originally introduced and only one change was sufficient to revert it to a wild-type sequence at that site, the other mutation was still present, indicating that it was indeed a revertant and not an accidental contamination with wild type-SARS D fragment. This suggests that the G104V mutant likely replicates very inefficiently, which allowed the revertant to evolve. The G104V mutant was generated and tested a fourth time, with results identical to those of the first two trials, demonstrating the rarity of this event.

Crystallization and structure determination. Crystals of nsp9 G104E grew in a hexagonal, rod-like morphology and diffracted to a maximum of 2.6-Å resolution on beamline 11-1 of SSRL. No significant disruptions in the secondary structure were seen, with nsp9 G104E exhibiting seven beta strands and one alpha helix as previously observed in 1UW7 (57) and 1QZ8 (16). As determined by molecular replacement (Fig. 5A), four monomers of nsp9 G104E are present within the asymmetric unit. In four regions residues could not be modeled due to the lack of $2F_o$-F_c omit density (chain A residues 36 and 37 and residues 60 and 61, chain C residues 37 and 38, and chain D residue 37). This arrangement forms a helix-like bundle of the individual monomers through two separate motifs, the previously observed sheet-sheet dimer interface and a new loop-sheet interface, with each monomer oriented roughly perpendicular to each other along the helix axis. The helix-helix dimer interface is generated through crystallographic symmetry (Fig. 5B); however, the interface is structurally distorted and will be discussed below. The helical arrangement of nsp9 G104E monomers forms a continuous strand parallel to the b axis of each unit cell throughout the crystal by crystallographic symmetry, which continually exposes a charged surface along the edge of the helix (Fig. 5C). At the center of the asymmetric unit is the previously observed sheet-sheet dimer interface which has a C_α RMSD of 1.92 Å compared to the sheet-sheet dimer reported in 1UW7 (57). The G104E sheet-sheet dimer introduces an additional six hydrogen bonds and 11 nonbonded interactions compared to 1UW7 (57). A loop-sheet interface is present at either end of the central sheet-sheet dimer and results in the generation of the helical motif in the asymmetric unit. This interface is formed through the association of the loop between beta strands 6 and 7, with strand 6 of an opposing monomer. Residues Asp78-C and Asp78-B make hydrogen bonds with residues Asp47-D/Lsy86-D and Asp47-A/Lsy86-A, respectively. The two loop-sheet interactions present in the tetramer each bury roughly 800 $Å^2$ of surface area. Moderate sequence conservation is found in these residues.

In the G104E mutant structure, the helix-helix dimer is present as in both 1QZ8 (16) and 1UW7 (57) and is thought to be the biologically relevant dimer, as seen in Fig. 5B. The helix-helix interface is disturbed but not disrupted completely, with the presence of Glu instead of Gly at the interface causing the C_α positions at position 104 to move to 6.56 Å apart, compared to 4.76 Å in 1QZ8 (16) and 5.4 Å in 1UW7 (57), also shown in a superposition in Fig. 5B. The distances between other atoms in the helices are essentially maintained, compared to 1QZ8 (16) and 1UW7 (57), with an overall RMSD over the helices of 0.66 Å. In addition, the helical crossing angles are markedly different for G104E (−50.9°) compared to 1QZ8 (16) and 1UW7 (57) (−41.1° and −38.7°, respectively). The helix-helix interface is held together by hydrogen bonds that form between either the OE1 or OE2 atoms of the Glu side chain and the main-chain nitrogen of Leu 9 of the opposing strand (Fig. 5D). A number of phosphate ions were found within the structure coordinated by surface charged residues thought to be involved in nucleotide binding. Specific overlaps upon superposition of 1QZ8 (16) with our structure occur at

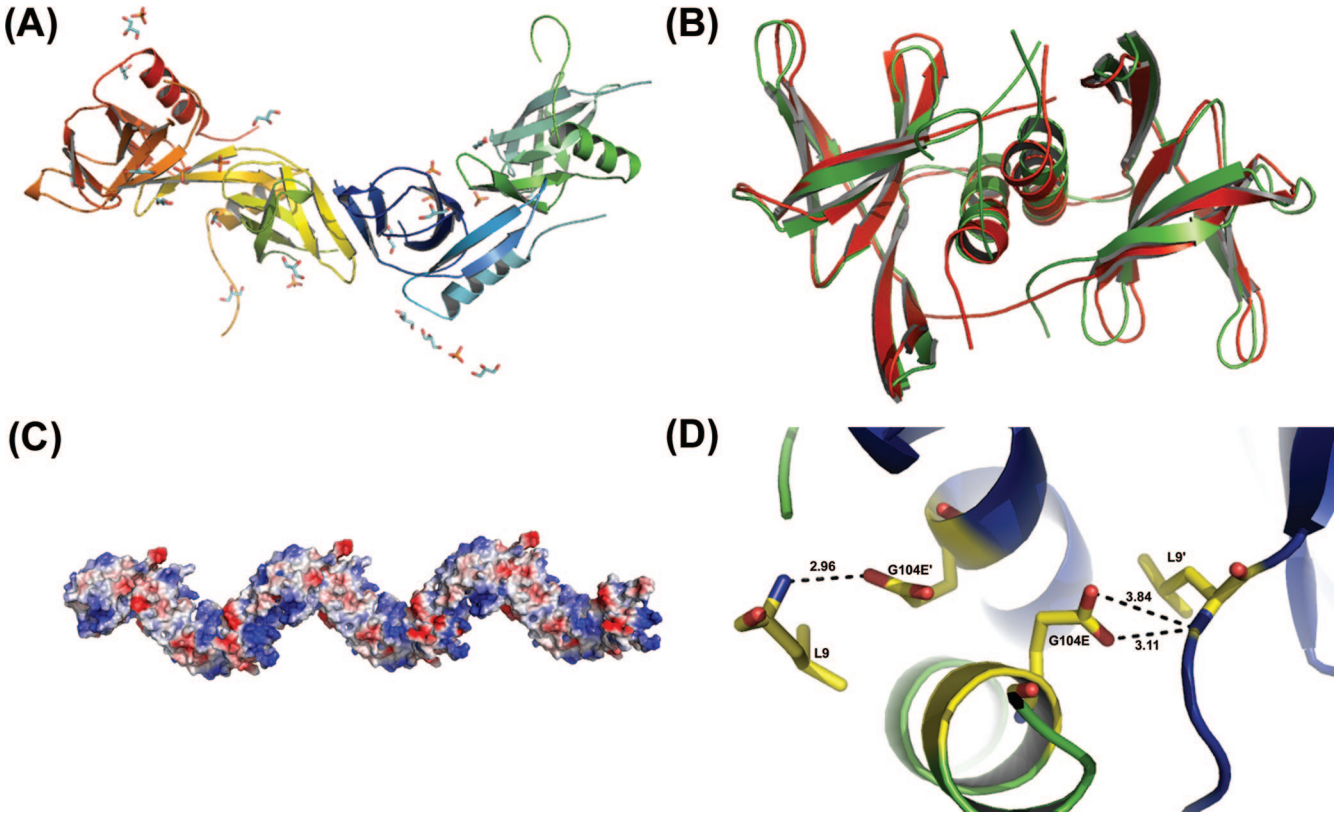


FIG. 5. Tetrameric arrangement of G104E crystals. (A) G104E structure exhibiting sheet-sheet dimer (center of structure, yellow and blue monomers) and loop-sheet interface (flanking sheet-sheet dimer on either side, red/yellow or blue/green monomers) present within the asymmetric unit containing four monomers of nsp9. (B) Helix-helix dimer in the G104E structure, similar to the helix-helix dimer in 1QZ8. Alignment of the two structures is depicted with 1QZ8 in red and G104E in green. Significant deviations in both crossing angles and buried surface area arise as a result of the G104E mutation. (C) Electrostatic surface of helix-like arrangement of G104E monomers along the b axis of a unit cell. Colors: blue, negative surface; white, neutral surface; red, positive surface. (D) Interactions stabilizing the helix-helix dimer. Primary contacts are formed between G104E and L9 of a symmetry related monomer. Distances between atoms are in angstroms. The L9′ and G104E′ designations indicate residues are a from symmetry-related monomer. Images were prepared by using PyMol.

sulfate 693 of 1QZ8 (16) with glycerols 3 and 8 and phosphates 1 and 2 of G104E. If the phosphates are taken as representative binding areas for backbone phosphates of target nucleic acid, then the nearby residues are potential sites to target for mutagenesis examining RNA or DNA binding. The overall symmetry motifs generated by crystallographic symmetry in 1QZ8 (16) and 1UW7 (57) are not replicated in our structure due to the presence of the loop-sheet interaction and the 3_1 screw-axis.

DISCUSSION

SARS-CoV, like other CoVs, produces a number of small highly conserved RNA-binding proteins essential to viral replication. We have used a combination of structure guided mutagenesis and reverse genetics to evaluate the role of nsp9 in replication. In the present study, G100 and G104 of a GXXXG dimerization motif were mutated to determine whether the nsp9 homodimer could be disrupted at the conserved helix-helix interface and whether the homodimer was relevant to virus replication. nsp9 mutants were analyzed in vitro by SEC-MALS, CD spectroscopy, EMSA, and FA. The crystal structure of G104E nsp9 was solved, and disruption of the helix-helix interface was observed. In addition, each nsp9 mutant was introduced into the virus using reverse genetics, and the in vivo effects were monitored. The biochemical data strongly suggest that mutation of either glycine results in the production of monomeric nsp9 in vitro and nonviable virus in vivo.

nsp9 mutations disrupt dimerization. An investigation of a nonredundant PDB set revealed that the GXXXG motif being examined in nsp9 was well represented among soluble proteins (31) but has not been extensively studied for its contributions to homo- and heterodimerization of soluble proteins containing the motif. The motif was also commonly found in transmembrane helices, where it was postulated to allow for close approach of adjacent helices for subsequent dimerization. In several examples, mutation of either Gly in the motif lead to decreased or complete loss of dimerization of transmembrane helices (45, 51, 56). The SEC-MALS and SEC analyses in conjunction with EMSA data showed that G100E, G104E, and G104V mutants were monomeric in solution, while wild-type nsp9 was dimeric, indicating that the GXXXG motif was important for maintaining the nsp9 dimer. This also indicates that the helix-helix dimer and not the sheet-sheet dimer is the major dimer in solution.

RNA binding by nsp9. The results from the EMSA of wild-type nsp9 and the G100E, G104E and G104V mutants show that each protein binds RNA, indicating the dimeric state of the protein is not critical for RNA binding. nsp9 apparently utilizes RNA as the preferred ligand. An unlabeled DNA oligonucleotide could not compete with labeled RNA during an EMSA, which confirms reports by Egloff et al., who suggested that nsp9 preferred RNA over DNA (16).

In order to distinguish quantitatively between wild-type and mutant RNA binding affinities, FA measurements were used. FA measurements using an ssRNA 20-mer give K_d values of 55 nM for wild type and a range from 250 to 650 nM for the mutants, which is comparable to the values obtained by Egloff et al. in which a larger fragment of RNA (560 bases) was reported to have a binding constant of 400 nM, using an assay monitoring the change in tryptophan fluorescence (16). The equilibrium binding constants presented here were determined at a lower ionic strength (15 mM NaCl) and may be lower due to coulombic effects. The G100 and G104 mutants have K_d values higher than that of the wild type ranging from 250 to 650 nM. Disruption of the nsp9 dimer does not abrogate binding of RNA but does result in a 5- to 12-fold decrease in affinity.

The relatively low affinity binding constants may reflect that, as part of the replication complex, nsp9 may not have a specific binding sequence but may act in conjunction with nsp7 and nsp8 as a processivity factor. Among the SARS-CoV nsp's, nsp9 has already been identified to have binding in low micromolar affinity for ssRNA and ssDNA (16, 57), nsp10 was reported to have high micromolar affinity for dsRNA (29), while nsp8 and the nsp7-8 hexadecamer were also shown to bind double-stranded RNA and double-stranded DNA with moderate efficiency (67). The interaction of these RNA-binding proteins with relatively weak binding affinities suggests that the presence of a complete replication complex may be necessary for high-affinity binding. nsp9 has been shown in several cases to interact with these proteins through several different methods (57). A specific target RNA sequence for nsp9 within the SARS-CoV genome has not been identified, although interactions with a stem-loop in the 3′ region of the genome were recently reported (69).

Local conditions could also affect binding of nucleic acid to these proteins. CoVs use large membrane-associated replicase complexes found within viral generated double-membrane vesicles to achieve viral replication (53); thus, it can be assumed that the local concentration of RNA around nsp9 would be markedly higher than nanomolar levels. Alternatively, multiple weak binding sites in nsp9 may contribute to the formation of the viral replication complex without altering the individual nsp9 proteins. In addition, the presence of the replication complex may serve to introduce protein-RNA interaction sites that are not available for the individual nsp's.

Dimerization is critical for viral function. Several experiments suggest that the nsp9 dimer is required for viral growth. In vitro, the SEC-MALS data show that the dimerization of nsp9 was successfully disrupted with the G100E, G104E, and G104V mutations. In vivo, the G100E and G104E mutations in nsp9 were lethal to efficient virus replication in cell culture, based on detection for subgenomic mRNAs. In contrast, the G104V mutation displayed a single example of delayed growth kinetics with reversion to an alternate codon for wild-type virus production.

Analysis of the genomic RNA and subgenomic mRNA using real-time PCR suggested that, while subgenomic synthesis was not detected at any time, genomic RNA increased slightly at early time points before decreasing at later times during infection. This increase was not significantly different than background (data not shown), but this observation leaves open the possibility that genomic plus-strand RNA synthesis occurred at a reduced rate in the G104V mutant. van Marle et al. constructed a mutant incorporating a single point mutation in a nsp that only made genome length molecules but not mRNA in the related equine arteritis virus (59). Low-level genomic replication similar to that observed in equine arteritis virus may have provided the opportunity for the reversion to occur in the G104V mutation. An additional possibility would be that an error in the T7 transcription step could have introduced the revertant nucleotide in the G104V mutant; however, the error rate for the T7 RNA polymerase is one error per 20,000 nucleotides and, when this error rate is factored in with a genome size of 29,751 nucleotides, the probability of an error occurring in that specific nucleotide would be ~1/500,000,000.

Functionally, it would seem that during infection nsp9 must form a dimer to properly bind and orient RNA for subsequent use by the replicase machinery despite the ability of the mutants to bind RNA in vitro. It is unlikely that the nsp9 dimer mutations have an effect on processing of the ORF1a and ORF1ab polyproteins. The MHV nsp9-10 fusion studied by Deming et al. (13) was shown to be functionally competent for viral replication, indicating that nsp9-10 did not need to be processed for replication to occur. There is some evidence using light scattering that an nsp9-10 construct is dimeric in solution (Z. J. Miknis and L. W. Schultz, unpublished results). Further modeling indicates that the C termini of nsp9 can line up with the N termini of nsp10 and still maintain the dimer structures of both proteins. Therefore, the nsp9-10 fusion may be functionally identical to its wild-type processed counterparts.

Interestingly, the G104V mutation was still able to produce revertant viruses, despite SEC and SEC-MALS data indicating that it is monomeric. ITC data suggest that G104V can form a dimer, but ~8-fold more weakly than the wild type, as estimated by SEC. Dimerization has been shown in the case of the RNA-binding protein NS1 of influenza virus to be necessary for RNA binding (61), as well as for the She2P protein from yeast (46). The hepatitis A virus 3C protease has affinity for viral RNA that is increased from millimolar to micromolar levels upon dimerization of the enzyme (48). G104V may be able to form a weakly associated helix-helix dimer that is still able to function at a reduced level within the virus, allowing sufficient virus replication to evolve a revertant strain.

While the nsp9 dimerization mutants are able to bind to RNA effectively, the mutations could also disrupt interactions that allow them to form hetero-oligomers with other SARS proteins. In rabies virus, the nucleoprotein shows no sequence specificity for RNA; however, when it is incorporated into the ribonucleoprotein complex that is ultimately responsible for RNA transcription, the specificity is gained for genomic RNA (39). A similar mechanism may exist in SARS, where nsp9 can

switch from nonspecific binding to specific binding upon formation of the replicase complex.

Structural analysis. The crystal structure of G104E shows the clear formation of three interaction motifs defined by a helix-helix dimer, sheet-sheet dimer, and a loop-sheet interface. This is in contrast to SEC and SEC-MALS, where no G104E dimers or tetramers were seen in solution, even in experiments performed at protein concentrations used in the crystallizations. The concentration of nsp9 G104E in the crystal may be high enough to favor dimer, or there may be enough energy gained from formation of the crystal lattice and other crystal contacts to force dimerization. Hydrogen bonds form between the OE1 or OE2 atoms of the G104E side chain across the dimer interface with the main-chain nitrogen of Leu 9. These hydrogen bonds serve to allow the formation of the dimer in the crystal, the equivalent of which are not accessible in the G104V and G100E mutants. The structure does show significant disruption of the dimer interface, including a wider separation between the two halves of the dimer and a steeper helix crossing angle. Together, these differences may be enough to disrupt the formation of a dimer in solution.

The G104E mutant showed a higher T_m than did the wild type, the G100E mutant, or the G104V mutant as measured by CD spectroscopy. It is unclear why the G104E mutant displays a T_m ~10°C higher than that for the other mutants and the wild type. The G104E was also the only sample that produced crystals which may be a reflection of its greater thermal stability compared to the wild type and the G100E or G104V mutant.

A recent report has described an antiparallel helix-helix dimer of nsp9 from HCoV-229E that is held together by a disulfide bridge (49). For the experiments described here, reducing agents were maintained in wild-type and mutant nsp9 samples in all experiments. However, wild-type nsp9 samples showed no evidence of disulfide bond formation and could be dissociated into monomers under oxidizing conditions in solution, as monitored by high-resolution SEC. This may represent a significant difference between the function of nsp9 in SARS-CoV and HCoV-229E.

As suggested by Sutton et al. (57), it is likely that the sheet-sheet dimer formed in the center of the tetramer is not biologically relevant. The new loop-sheet interface and the helix-helix dimer bury roughly equivalent surface areas (~800 Å^2) and make approximately the same number of intermolecular contacts. The new loop-sheet interface does have contacts from residues D78, D47, and K86 that are moderately conserved among various CoVs. The sheet-sheet dimer has a significantly smaller buried surface area but has more H-bonds and nonbonded contacts for stabilization.

Interestingly, the formation of the sheet-sheet dimer flanked by the loop-sheet interface within the tetramer results in the formation through crystallographic symmetry of a continuous helical arrangement of nsp9 along the b axis of the crystal. This arrangement of tetramers exposes a near-continuous patch of positively charged surface wrapping around the overall helical structure, suggesting the potential for binding of long stretches of RNA. Although this tetramer was not observed in solution, it does suggest that nsp9 is capable of forming polymeric structures that might act as scaffolds for binding genomic RNA.

Summary. Crystallography is an essential tool for the study of protein structure and function. However, the biological implications of structures are often not clear. In the case of the SARS-CoV nsp9 RNA-binding protein, several different dimer interfaces have been proposed based on interactions seen in the crystal. We have shown that the conserved helix-helix dimer interface containing a GXXXG protein-protein interaction motif is biologically relevant to SARS-CoV replication. Disruption of this interface by site-directed mutagenesis of the glycine residues in the GXXXG motif resulted in monomeric forms of nsp9 in solution. Subsequent introduction of mutations into SARS-CoV by reverse genetics showed that the formation of the nsp9 dimer was necessary for viral viability. A single mutant G104V reverted to wild-type nsp9, indicating significant evolutionary and functional pressure on the dimerization interface. The RNA-binding affinity of monomeric nsp9 mutants was reduced but not eliminated, indicating that the dimer retains a slight advantage over the monomer in RNA binding. The inability of the nsp9 monomers to function in vivo may not reflect an RNA-binding ability but rather the correct positioning of RNA in the replication complex requiring a properly dimerized nsp9. This view is supported by the crystal structure of G104E nsp9 in which the helix-helix dimer interface is significantly disrupted from wild-type nsp9. Future experiments combining crystallography and in vitro biophysical analysis of protein-protein and protein-RNA interactions will be necessary to unravel the complicated structure of the CoV replication complex. Collectively, our data add to a growing body of literature that implicates nsp9 as a key ingredient that intimately engages other proteins in the replicase complex to mediate efficient virus transcription and replication.

ACKNOWLEDGMENTS

We thank Wendy Franke, Carleen Pope, and Peggy Cegielski for technical assistance and Mark Denison (Vanderbilt University) for the gift of the original nsp9 plasmid.

Portions of this research were carried out at the SSRL, a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program, and the National Institute of General Medical Sciences. The work described here was funded in part by NIAID grant R21AI061057 (to L.W.S. and T.C.U.), the John R. Oishei Foundation (to L.W.S. and T.C.U.), the William G. McGowan Charitable Fund (to L.W.S. and T.C.U.), and the State University of New York at Buffalo (Z.J.M.). This study was also supported by NIH grants AI23946 and AI059136 to R.S.B.

REFERENCES

1. **Andrade, M. A., P. Chacon, J. J. Merelo, and F. Moran.** 1993. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. Protein Eng. **6:**383–390.
2. **Bahadur, R. P., P. Chakrabarti, F. Rodier, and J. Janin.** 2003. Dissecting subunit interfaces in homodimeric proteins. Proteins **53:**708–719.
3. **Bailey, S.** 1994. The Ccp4 Suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. **50:**760–763.
4. **Bost, A. G., R. H. Carnahan, X. T. Lu, and M. R. Denison.** 2000. Four proteins processed from the replicase gene polyprotein of mouse hepatitis virus colocalize in the cell periphery and adjacent to sites of virion assembly. J. Virol. **74:**3379–3387.
5. **Brockway, S. M., C. T. Clay, X. T. Lu, and M. R. Denison.** 2003. Characterization of the expression, intracellular localization, and replication complex association of the putative mouse hepatitis virus RNA-dependent RNA polymerase. J. Virol. **77:**10515–10527.
6. **Brunger, A. T., P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W.**

Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren. 1998. Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. **54:**905–921.

7. **Burrows, S. D., M. L. Doyle, K. P. Murphy, S. G. Franklin, J. R. White, I. Brooks, D. E. McNulty, M. O. Scott, J. R. Knutson, D. Porter, P. R. Young, and P. Hensley.** 1994. Determination of the monomer-dimer equilibrium of interleukin-8 reveals it is a monomer at physiological concentrations. Biochemistry **33:**12741–12745.
8. **Cambillau, C., et al.** 1997. TURBO-FRODO. http://www.afmb.univ-mrs.fr/-TURBO-.
9. **Chen, S., L. L. Chen, J. Z. Tan, J. Chen, L. Du, T. Sun, J. H. Shen, K. X. Chen, H. L. Jiang, and X. Shen.** 2005. Severe acute respiratory syndrome coronavirus 3C-like proteinase n terminus is indispensable for proteolytic activity but not for enzyme dimerization: biochemical and thermodynamic investigation in conjunction with molecular dynamics simulations. J. Biol. Chem. **280:**164–173.
10. **Claude, J. B., K. Suhre, C. Notredame, J. M. Claverie, and C. Abergel.** 2004. CaspR: a web server for automated molecular replacement using homology modeling. Nucleic Acids Res. **32:**W606–W609.
11. **Davis, I. W., A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall III, J. Snoeyink, J. S. Richardson, and D. C. Richardson.** 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. **35:**W375–W383.
12. **DeLano, W. L.** 2002. The PyMOL molecular graphics system. DeLano Scientific, San Carlos, CA.
13. **Deming, D. J., R. L. Graham, M. R. Denison, and R. S. Baric.** 2006. MHV-A59 ORF1a replicase protein nsp7-nsp10 processing in replication. Adv. Exp. Med. Biol. **581:**101–104.
14. **Donaldson, E. F., A. C. Sims, R. L. Graham, M. R. Denison, and R. S. Baric.** 2007. Murine hepatitis virus replicase protein nsp10 is a critical regulator of viral RNA synthesis. J. Virol. **81:**6356–6368.
15. **Drosten, C., S. Gunther, W. Preiser, S. van der Werf, H. R. Brodt, S. Becker, H. Rabenau, M. Panning, L. Kolesnikova, R. A. M. Fouchier, A. Berger, A. M. Burguiere, J. Cinatl, M. Eickmann, N. Escriou, K. Grywna, S. Kramme, J. C. Manuguerra, S. Muller, V. Rickerts, M. Sturmer, S. Vieth, H. D. Klenk, A. D. M. E. Osterhaus, H. Schmitz, and H. W. Doerr.** 2003. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. N. Engl. J. Med. **348:**1967–1976.
16. **Egloff, M. P., F. Ferron, V. Campanacci, S. Longhi, C. Rancurel, H. Dutartre, E. J. Snijder, A. E. Gorbalenya, C. Cambillau, and B. Canard.** 2004. The severe acute respiratory syndrome-coronavirus replicative protein nsp9 is a single-stranded RNA-binding subunit unique in the RNA virus world. Proc. Natl. Acad. Sci. USA **101:**3792–3796.
17. **Emsley, P., and K. Cowtan.** 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. **60:**2126–2132.
18. **Folta-Stogniew, E., and K. R. Williams.** 1999. Determination of molecular masses of proteins in solution: implementation of an HPLC size exclusion chromatography and laser light scattering service in a core laboratory. J. Biomol. Tech. **10:**51–63.
19. **Fouchier, R. A. M., T. Kuiken, M. Schutten, G. van Amerongen, J. van Doornum, B. G. van den Hoogen, M. Peiris, W. Lim, K. Stohr, and A. D. M. E. Osterhaus.** 2003. Aetiology: Koch's postulates fulfilled for SARS virus. Nature **423:**240–241.
20. **Gimpelev, M., L. R. Forrest, D. Murray, and B. Honig.** 2004. Helical packing patterns in membrane and soluble proteins. Biophys. J. **87:**4075–4086.
21. **Gonzalez, A., P. Moorhead, S. McPhillips, and N. Sauter.** 2005. New tools to integrate data analysis and data collection at SSRL: Web-Ice. Acta Crystallogr. A Found. Crystallogr. **61:**C486.
22. **Goodsell, D. S., and A. J. Olson.** 2000. Structural symmetry and protein function. Annu. Rev. Biophys. Biomol. Struct. **29:**105–153.
23. **Graham, R. L., J. S. Sparks, L. D. Eckerle, A. C. Sims, and M. R. Denison.** 2008. SARS coronavirus replicase proteins in pathogenesis. Virus Res. **133:**88–100.
24. **Greenberg, J. R.** 1979. Ultraviolet light-induced cross-linking of messenger RNA to proteins. Nucleic Acids Res. **6:**715–732.
25. **Gu, H. D., and D. R. Schoenberg.** 2003. U2AF modulates poly(A) length control by the poly(A)-limiting element. Nucleic Acids Res. **31:**6264–6271.
26. **Hutchinson, E. G., and J. M. Thornton.** 1996. PROMOTIF: a program to identify and analyze structural motifs in proteins. Protein Sci. **5:**212–220.
27. **Imbert, I., E. J. Snijder, M. Dimitrova, J. C. Guillemot, P. Lecine, and B. Canard.** 2008. The SARS-Coronavirus PLnc domain of nsp3 as a replication/transcription scaffolding protein. Virus Res. **133:**136–148.
28. **Jones, S., and J. M. Thornton.** 1995. Protein-protein interactions: a review of protein dimer structures. Prog. Biophys. Mol. Biol. **63:**31–65.
29. **Joseph, J. S., K. S. Saikatendu, V. Subramanian, B. W. Neuman, A. Brooun, M. Griffith, K. Moy, M. K. Yadav, J. Velasquez, M. J. Buchmeier, R. C. Stevens, and P. Kuhn.** 2006. Crystal structure of nonstructural protein 10 from the severe acute respiratory syndrome coronavirus reveals a novel fold with two zinc-binding motifs. J. Virol. **80:**7894–7901.
30. **Kamitani, W., K. Narayanan, C. Huang, K. Lokugamage, T. Ikegami, N. Ito, H. Kubo, and S. Makino.** 2006. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation. Proc. Natl. Acad. Sci. USA **103:**12885–12890.
31. **Kleiger, G., R. Grothe, P. Mallick, and D. Eisenberg.** 2002. GXXXG and AXXXA: common alpha-helical interaction motifs in proteins, particularly in extremophiles. Biochemistry **41:**5990–5997.
32. **Ksiazek, T. G., D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret, S. Emery, S. X. 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, L. J. Anderson, and S. W. Grp.** 2003. A novel coronavirus associated with severe acute respiratory syndrome. N. Engl. J. Med. **348:**1953–1966.
33. **Laskowski, R. A.** 2001. PDBsum: summaries and analyses of PDB structures. Nucleic Acids Res. **29:**221–222.
34. **Laskowski, R. A., V. V. Chistyakov, and J. M. Thornton.** 2005. PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. Nucleic Acids Res. **33:**D266–D268.
35. **Laskowski, R. A., E. G. Hutchinson, A. D. Michie, A. C. Wallace, M. L. Jones, and J. M. Thornton.** 1997. PDBsum: a Web-based database of summaries and analyses of all PDB structures. Trends Biochem. Sci. **22:**488–490.
36. **Lau, S. K., P. C. Woo, K. S. Li, Y. Huang, H. W. Tsoi, B. H. Wong, S. S. Wong, S. Y. Leung, K. H. Chan, and K. Y. Yuen.** 2005. Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. Proc. Natl. Acad. Sci. USA **102:**14040–14045.
37. **Law, A. H., D. C. Lee, B. K. Cheung, H. C. Yim, and A. S. Lau.** 2007. Role for nonstructural protein 1 of severe acute respiratory syndrome coronavirus in chemokine dysregulation. J. Virol. **81:**416–422.
38. **Li, W., Z. Shi, M. Yu, W. Ren, C. Smith, J. H. Epstein, H. Wang, G. Crameri, Z. Hu, H. Zhang, J. Zhang, J. McEachern, H. Field, P. Daszak, B. T. Eaton, S. Zhang, and L. F. Wang.** 2005. Bats are natural reservoirs of SARS-like coronaviruses. Science **310:**676–679.
39. **Liu, P., J. Yang, X. Wu, and Z. F. Fu.** 2004. Interactions amongst rabies virus nucleoprotein, phosphoprotein and genomic RNA in virus-infected and transfected cells. J. Gen. Virol. **85:**3725–3734.
40. **Lobley, A., L. Whitmore, and B. A. Wallace.** 2002. DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. Bioinformatics **18:**211–212.
41. **Luft, J. R., and G. T. Detitta.** 1992. Hangman: a macromolecular hanging-drop vapor-diffusion technique. J. Appl. Crystallogr. **25:**324–325.
42. **Lundblad, J. R., M. Laurance, and R. H. Goodman.** 1996. Fluorescence polarization analysis of protein-DNA and protein-protein interactions. Mol. Endocrinol. **10:**607–612.
43. **Marra, M. A., S. J. M. Jones, C. R. Astell, R. A. Holt, A. Brooks-Wilson, Y. S. N. Butterfield, J. Khattra, J. K. Asano, S. A. Barber, S. Y. Chan, A. Cloutier, S. M. Coughlin, D. Freeman, N. Girn, O. L. Griffin, 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. E. Stott, G. S. Yang, F. Plummer, A. Andonov, 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. Krajden, M. Petric, D. M. Skowronski, C. Upton, and R. L. Roper.** 2003. The genome sequence of the SARS-associated coronavirus. Science **300:**1399–1404.
44. **McPhillips, T. M., S. E. McPhillips, H. J. Chiu, A. E. Cohen, A. M. Deacon, P. J. Ellis, E. Garman, A. Gonzalez, N. K. Sauter, R. P. Phizackerley, S. M. Soltis, and P. Kuhn.** 2002. Blu-Ice and the distributed control system: software for data acquisition and instrument control at macromolecular crystallography beamlines. J. Synchrotron Radiat. **9:**401–406.
45. **Munter, L. M., P. Voigt, A. Harmeier, D. Kaden, K. E. Gottschalk, C. Weise, R. Pipkorn, M. Schaefer, D. Langosch, and G. Multhaup.** 2007. GxxxG motifs within the amyloid precursor protein transmembrane sequence are critical for the etiology of Aβ42. EMBO J. **26:**1702–1712.
46. **Niessing, D., S. Huttelmaier, D. Zenklusen, R. H. Singer, and S. K. Burley.** 2004. She2p is a novel RNA binding protein with a basic helical hairpin motif. Cell **119:**491–502.
47. **Otwinowski, Z., and W. Minor.** 1997. Processing of X-ray diffraction data collected in oscillation mode. Macromol. Crystallogr. A **276:**307–326.
48. **Peters, H., Y. Y. Kusov, S. Meyer, A. J. Benie, E. Bauml, M. Wolff, C. Rademacher, T. Peters, and V. Gauss-Muller.** 2005. Hepatitis A virus proteinase 3C binding to viral RNA: correlation with substrate binding and enzyme dimerization. Biochem. J. **385:**363–370.
49. **Ponnusamy, R., R. Moll, T. Weimar, J. R. Mesters, and R. Hilgenfeld.** 2008. Variable oligomerization modes in coronavirus nonstructural protein 9. J. Mol. Biol. **383:**1081–1096.
50. **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. X. Tong, A. Tamin, L. Lowe, M. Frace, J. L. DeRisi, Q. Chen, D. Wang, D. D. Erdman, T. C. T. 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. M. E. 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.

51. **Schneider, D., and D. M. Engelman.** 2004. Motifs of two small residues can assist but are not sufficient to mediate transmembrane helix interactions. J. Mol. Biol. **343:**799–804.
52. **Snijder, E. J., P. J. Bredenbeek, J. C. Dobbe, V. Thiel, J. Ziebuhr, L. L. M. Poon, Y. Guan, M. Rozanov, W. J. M. Spaan, and A. E. Gorbalenya.** 2003. Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J. Mol. Biol. **331:**991–1004.
53. **Snijder, E. J., Y. van der Meer, J. Zevenhoven-Dobbe, J. J. Onderwater, J. van der Meulen, H. K. Koerten, and A. M. Mommaas.** 2006. Ultrastructure and origin of membrane vesicles associated with the severe acute respiratory syndrome coronavirus replication complex. J. Virol. **80:**5927–5940.
54. **Sperry, S. M., L. Kazi, R. L. Graham, R. S. Baric, S. R. Weiss, and M. R. Denison.** 2005. Single-amino-acid substitutions in open reading frame (ORF) 1b-nsp14 and ORF 2a proteins of the coronavirus mouse hepatitis virus are attenuating in mice. J. Virol. **79:**3391–3400.
55. **Su, D., Z. Lou, F. Sun, Y. Zhai, H. Yang, R. Zhang, A. Joachimiak, X. C. Zhang, M. Bartlam, and Z. Rao.** 2006. Dodecamer structure of severe acute respiratory syndrome coronavirus nonstructural protein nsp10. J. Virol. **80:** 7902–7908.
56. **Sulistijo, E. S., and K. R. MacKenzie.** 2006. Sequence dependence of BNIP3 transmembrane domain dimerization implicates side chain hydrogen bonding and a tandem GxxxG motif in specific helix-helix interactions. J. Mol. Biol. **364:**974–990.
57. **Sutton, G., E. Fry, L. Carter, S. Sainsbury, T. Walter, J. Nettleship, N. Berrow, R. Owens, R. Gilbert, A. Davidson, S. Siddell, L. L. M. Poon, J. Diprose, D. Alderton, M. Walsh, J. M. Grimes, and D. I. Stuart.** 2004. The nsp9 replicase protein of SARS-coronavirus, structure and functional insights. Structure **12:**341–353.
58. **Thiel, V., K. A. Ivanov, A. Putics, T. Hertzig, B. Schelle, S. Bayer, B. Weissbrich, E. J. Snijder, H. Rabenau, H. W. Doerr, A. E. Gorbalenya, and J. Ziebuhr.** 2003. Mechanisms and enzymes involved in SARS coronavirus genome expression. J. Gen. Virol. **84:**2305–2315.
59. **van Marle, G., L. C. van Dinten, W. J. Spaan, W. Luytjes, and E. J. Snijder.** 1999. Characterization of an equine arteritis virus replicase mutant defective in subgenomic mRNA synthesis. J. Virol. **73:**5274–5281.
60. **von Brunn, A., C. Teepe, J. C. Simpson, R. Pepperkok, C. C. Friedel, R. Zimmer, R. Roberts, R. Baric, and J. Haas.** 2007. Analysis of intraviral protein-protein interactions of the SARS coronavirus ORFeome. PLoS ONE **2:**e459.
61. **Wang, W., K. Riedel, P. Lynch, C. Y. Chien, G. T. Montelione, and R. M. Krug.** 1999. RNA binding by the novel helical domain of the influenza virus NS1 protein requires its dimer structure and a small number of specific basic amino acids. RNA **5:**195–205.
62. **Whitmore, L., and B. A. Wallace.** 2004. DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res. **32:**W668–W673.
63. **Wilkins, M. R., E. Gasteiger, A. Bairoch, J. C. Sanchez, K. L. Williams, R. D. Appel, and D. F. Hochstrasser.** 1999. Protein identification and analysis tools in the ExPASy server. Methods Mol. Biol. **112:**531–552.
64. **Yount, B., K. M. Curtis, and R. S. Baric.** 2000. Strategy for systematic assembly of large RNA and DNA genomes: transmissible gastroenteritis virus model. J. Virol. **74:**10600–10611.
65. **Yount, B., K. M. Curtis, E. A. Fritz, L. E. Hensley, P. B. Jahrling, E. Prentice, M. R. Denison, T. W. Geisbert, and R. S. Baric.** 2003. Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus. Proc. Natl. Acad. Sci. USA **100:**12995–13000.
66. **Yount, B., M. R. Denison, S. R. Weiss, and R. S. Baric.** 2002. Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. J. Virol. **76:**11065–11078.
67. **Zhai, Y., F. Sun, X. Li, H. Pang, X. Xu, M. Bartlam, and Z. Rao.** 2005. Insights into SARS-CoV transcription and replication from the structure of the nsp7-nsp8 hexadecamer. Nat. Struct. Mol. Biol. **12:**980–986.
68. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya.** 2000. Virus-encoded proteinases and proteolytic processing in the *Nidovirales*. J. General Virology **81:**853–879.
69. **Zust, R., T. B. Miller, S. J. Goebel, V. Thiel, and P. S. Masters.** 2008. Genetic interactions between an essential 3′ *cis*-acting RNA pseudoknot, replicase gene products, and the extreme 3′ end of the mouse coronavirus genome. J. Virol. **82:**1214–1228.