

STRUCTURE, EXPRESSION, AND INTRACELLULAR LOCALIZATION OF THE SARS-CoV ACCESSORY PROTEINS 7a AND 7b

Andrew Pekosz, Scott R. Schaecher, Michael S. Diamond,
Daved H. Fremont, Amy C. Sims, and Ralph S. Baric*

1. INTRODUCTION

The virus responsible for severe acute respiratory distress syndrome (SARS) has been identified as a coronavirus (SARS-CoV), and the genomic structure of several human and animal isolates has been determined.^{1,2} Consistent with other members of the *Nidoviridae*, the genome consists of several genes that are highly conserved and represent proteins with essential, basic functions in the viral life cycle.³ In addition to these essential genes, SARS-CoV encodes several accessory genes that are believed to be nonessential but important for viral replication *in vivo*.

Among these accessory genes, gene 7 encodes two major open reading frames, ORF7a and ORF7b. ORF7a is predicted to be a 122-amino-acid type I integral membrane protein with a cleavable N-terminal signal peptide and a C-terminal membrane spanning domain. ORF7b is predicted to encode a 44-amino-acid, extremely hydrophobic protein. The stop codon for ORF7a overlaps the start codon of ORF7b.

Previously, we and others have confirmed the expression of ORF7a in SARS-CoV infected cells;⁴⁻⁶ however, there is a discrepancy in the subcellular localization ascribed to the protein. ORF7a has been demonstrated in one study to colocalize with endoplasmic reticulum–Golgi intermediate compartment (ERGIC)-specific markers and possess an endoplasmic reticulum–retrieval motif in the cytoplasmic tail.⁷

However, experiments in our laboratories have localized ORF7a with Golgi-specific markers.⁶ In addition, overexpression of ORF7a has been reported to induce apoptosis in a variety of transformed cell lines.⁷ As of yet, no particular function has been assigned to the protein but the X-ray crystallographic structure of the extracellular domain has been determined.⁶ The expression of the ORF7b protein has yet to be demonstrated in SARS-CoV infected cells.

We have continued our studies on the expression and function of the proteins encoded by gene 7 and identified an endoplasmic reticulum (ER) export signal residing in

*Andrew Pekosz, Scott R. Schaecher, Michael S. Diamond, Daved H. Fremont, Washington University School of Medicine, St. Louis, Missouri 63110. Amy C. Sims, Ralph S. Baric, University of North Carolina, Chapel Hill, North Carolina 27599-7435.

the ORF7a cytoplasmic tail. In addition, we have verified the expression of ORF7b in SARS-CoV-infected cells. Finally, we have begun to characterize the replication and cytopathicity of a recombinant SARS-CoV encoding GFP in place of gene 7.

2. MATERIALS AND METHODS

2.1. Cell Culture, Transfection, and Mutagenesis

Vero cells were cultured at 37°C in a humidified incubator with 5% CO₂ as previously described.⁸ The cells were plated onto glass coverslips and incubated overnight at 37°C. Transfections were performed using LT-1 (Mirrus) transfection reagent at a ratio of 1 µg plasmid DNA to 4 µl of transfection reagent as previously described.^{6,9} Cells were analyzed at 18–20 hours post transfection.

The ORF7a-GFP fusion protein has been described previously.⁶ The localization of the ORF7a-GFP fusion protein is indistinguishable from that of the ORF7a protein (data not shown). Site-directed mutagenesis was performed using standard PCR techniques.

2.2. Confocal Microscopy

Indirect immunofluorescence for the Golgi protein Golgin-97 was carried out using a mouse monoclonal antibody (Molecular Probes, diluted 1:50). Coverslips were incubated with phosphate buffered saline (PBS) containing 3% normal goat serum (Sigma), 0.5% bovine serum albumin (BSA; Sigma), and 0.1% saponin (Sigma) for 30 minutes in primary antibody for 1 hour and in secondary antibody (goat anti-mouse IgG conjugated to AlexaFluor 594 (Molecular Probes; diluted 1:500) containing To-Pro-3 (Molecular Probes, diluted 1:500). A Zeiss LSM510 confocal microscope was used to acquire all images.

2.3. Viruses and Infections

Recombinant SARS-CoV containing a complete genome (rSARS WT) or with the gene 7 replaced with the GFP open reading frame (rSARS GFPΔORF7) were generated by reverse genetics.¹⁰ Virus was diluted in DMEM containing 5% fetal bovine serum (FBS) and Vero cells were infected at a multiplicity of infection (MOI) of approximately 0.01 for 1 hour at 37°C. After extensive washing, the cells were incubated with DMEM containing 5% FBS at 37°C. At the indicated times post infection, the virus-infected cell supernatant was removed and stored at -70°C. Infectious virus was quantified by plaque assay on Vero cells.⁶

2.4. Western Blotting

Vero cells were infected with an MOI of approximately 5.0 of rSARS WT or rSARS GFPΔORF7. At 24 hours postinfection, the cells were lysed with 1% sodium dodecyl sulfate (SDS) in water. Polypeptides were separated on a 17.5% acrylamide/4M urea SDS-PAGE and transferred to PVDF membranes. The membranes were blocked in PBS

containing 0.3% Tween-20 and 5% dry milk powder, followed by antibodies recognizing ORF7b (rabbit polyclonal; diluted 1:1,000), the SARS-CoV nucleocapsid (N) protein (mouse monoclonal 87-A1; diluted 1:2,000) or beta-actin (mouse monoclonal; diluted 1:500; Abcam). After incubation with secondary antibodies conjugated to horseradish peroxidase that recognized either rabbit (diluted 1:7,500; Jackson Laboratories) or mouse (diluted 1:7,500; Jackson Laboratories) IgG, the membranes were developed using Amersham ECL-Plus substrate and quantitated by phosphorimager analysis.

3. RESULTS AND DISCUSSION

3.1. Intracellular Localization of ORF7a

We have demonstrated previously that ORF7a does not reach the plasma membrane after expression from cDNA or in SARS-CoV-infected cells.⁶ However, there are contradictory data published with respect to the precise intracellular localization of the protein. The localization of the ORF7a protein was determined by transfecting Vero cells with a cDNA expression plasmid encoding an ORF7a-GFP fusion protein and co-localizing the protein with the Golgi-localized protein Golgin-97 (Figure 1). An identical localization pattern was seen in SARS-CoV virus-infected cells and cells transfected with the ORF7a cDNA alone. The data indicate the ORF7a protein targets primarily to the Golgi apparatus and not to the endoplasmic reticulum or intermediate compartment.

3.2. Identification of an ER Export Motif

The cytoplasmic tail of the ORF7a protein contains a motif R/K-X-R/K (corresponding to amino acids 118–120) that has been demonstrated to mediate COPII-dependent transport out of the ER.¹¹ To ascertain whether this sequence plays a role in ORF7a export out of the ER, the cDNA encoding the ORF7a-GFP fusion protein was altered from codons encoding Lys to ones encoding Ala at amino acids 118 and 120. The intracellular localization of the mutated proteins with respect to Golgin-97 was then determined. Figure 1 shows that mutation of these amino acids leads to a loss of Golgin-97 co-localization. The protein appears to co-localize with markers for the ER (data not shown), indicating the introduced mutations may be hindering the transport of the protein out of the ER. The OR7aF-GFP K118A K120A still reacts with a conformation-dependent monoclonal antibody,⁶ indicating the mutations have not resulted in misfolding of the protein (data not shown). Taken together, the data in Figure 1 indicates the ORF7a protein is targeted to the Golgi apparatus and utilizes the COPII transport machinery to exit the ER.

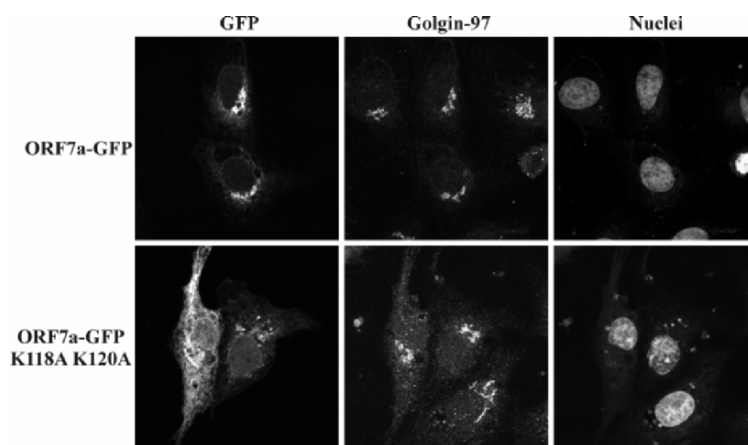


Figure 1. Intracellular localization of ORF7a-GFP. Vero cells expressing ORF7a-GFP were immunostained for the Golgi marker Golgin-97. The ORF7a-GFP protein co-localizes with Golgin-97. Substituting alanine for lysine at positions 118 and 120 of ORF7a-GFP eliminates a COPII transport sequence and results in a loss of Golgi localization.

3.3. Vero Cell Growth Kinetics of SARS-CoV Lacking Gene 7

The location of gene 7 in the SARS-CoV genome and its lack of conservation among other coronaviruses implies that the gene may be an accessory gene that is not essential for virus replication but important for infection of the natural host for the virus. To test this hypothesis, a recombinant SARS-CoV was generated in which the gene 7 coding region was replaced with that of GFP (rSARS GFP Δ ORF7ab). The replication of this virus was compared with its recombinant parental virus by infecting Vero cells at an MOI-0.01, harvesting infected cell supernatants at various times postinfection and determining the infectious virus titer by plaque assay on Vero cells. The parental recombinant and the gene 7 replacement virus displayed identical replication kinetics and reached comparable peak titers (Figure 2), indicating the gene 7 products were not essential for SARS-CoV replication.

3.4. Expression of ORF 7b in SARS-CoV-Infected Cells

Sequence analysis indicates the SARS-CoV gene 7 potentially encodes for two proteins: ORF7a and ORF7b. The start codon of ORF7b is located far downstream of the 5' end of the only mRNA associated with the gene 7 transcription start site, and therefore is not in a very good context to initiate translation.^{3,12} Nevertheless, we cloned the

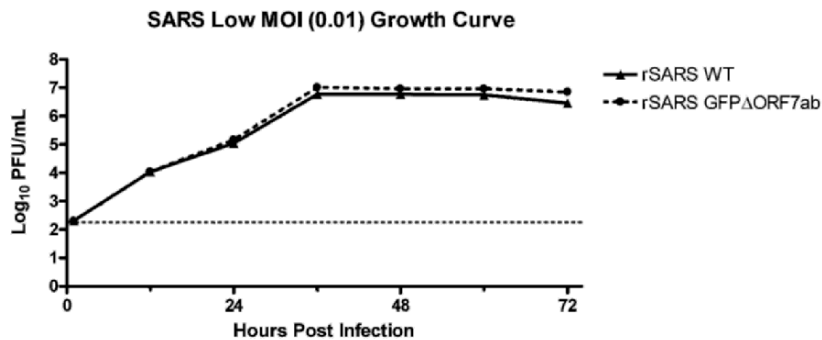


Figure 2. Replication kinetics of wild-type recombinant SARS-CoV compared with a recombinant virus containing GFP in place of gene 7.

predicted ORF7b open reading frame into a bacterial expression vector as a GST-fusion protein and immunized rabbits with the resulting protein. The antiserum was used to probe lysates from mock, SARS-CoV, or SARS-CoV GFP Δ ORF7ab cells. A band corresponding to the predicted molecular weight of the ORF7b protein was detected in SARS-CoV-infected cells but not in mock- or SARS-CoV GFP Δ ORF7ab-infected cells (Figure 3), indicating the 7b protein was in fact synthesized in virus-infected cells. The blot was probed with antibodies to the N protein in order to verify the expression of a known viral protein, as well as antibodies to beta-actin, which serves as the protein loading control. The data indicate that the SARS-CoV gene 7 does indeed encode two distinct proteins, ORF7a and ORF7b.

4. CONCLUSIONS

Our data indicate the two proteins predicted to be encoded by gene 7 of SARS-CoV are in fact authentic viral proteins. ORF7a is a type I transmembrane protein that localizes to the Golgi and requires COPII-mediated export out of the ER. As we have not yet ascribed a particular function to this protein, it is not yet clear why the protein possess these two properties. It is intriguing to speculate that ORF7a is perhaps interacting and retaining host cell proteins in the Golgi, in order to prevent their transport to the plasma membrane or other organelles. ORF7b is a highly hydrophobic protein, and preliminary data suggests it associates with cellular membranes (data not shown). It will also be important to determine whether ORF7a or ORF7b are structural components of the virion or simply nonstructural proteins.

Our data on the replication of a gene 7 replacement SARS-CoV indicate the gene 7 products are not essential, nor do they alter virus replication on Vero cells. It is important to emphasize that these results have only been demonstrated in Vero cells, and it is possible that infection and replication of the SARS-CoV gene 7 replacement virus may be altered on other cell lines. It is more likely that these proteins are required for efficient replication of SARS-CoV in the yet to be identified natural host of the virus. The use of animal models such as nonhuman primates, ferrets, mice, or hamsters may also be useful in elucidating a role for the gene 7 protein products in viral pathogenesis.

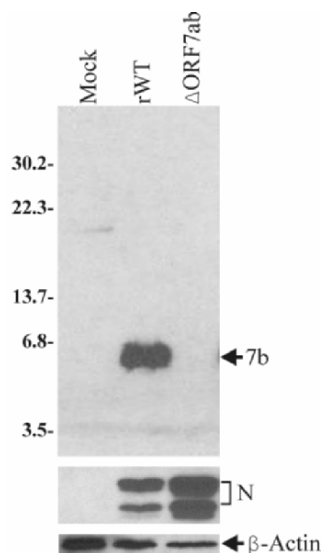


Figure 3. Western blotting for ORF7b expression in SARS-CoV-infected cells.

5. REFERENCES

1. Drosten, C., Gunther, S., et al., 2003, Identification of a novel coronavirus in patients with severe acute respiratory syndrome, *N. Engl. J. Med.* **348**:1967-1976.
2. Rota, P. A., Oberste, M. S., et al., 2003, Characterization of a novel coronavirus associated with severe acute respiratory syndrome, *Science* **300**:1394-1399.
3. Snijder, E. J., Bredenbeek, P. J., et al., 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.
4. Fielding, B. C., Tan, Y.-J., et al., 2004, Characterization of a unique group-specific protein (U122) of the severe acute respiratory syndrome coronavirus, *J. Virol.* **78**:7311-7318.
5. Chen, Y. Y., Shuang, B., et al., 2005, The protein X4 of severe acute respiratory syndrome-associated coronavirus is expressed on both virus-infected cells and lung tissue of severe acute respiratory syndrome patients and inhibits growth of Balb/c 3T3 cell line, *Chin. Med. J. (Engl.)* **118**:267-274.
6. Nelson, C. A., Pekosz, A., et al., 2005, Structure and intracellular targeting of the SARS-coronavirus Orf7a accessory protein, *Structure (Camb.)* **13**:75-85.
7. Tan, Y.-J., Fielding, B. C., et al., 2004, Overexpression of 7a, a protein specifically encoded by the severe acute respiratory syndrome coronavirus, induces apoptosis via a caspase-dependent pathway, *J. Virol.* **78**:14043-14047.
8. McCown, M., Diamond, M. S., et al., 2003, The utility of siRNA transcripts produced by RNA polymerase I in down regulating viral gene expression and replication of negative- and positive-strand RNA viruses, *Virology* **313**:514-524.
9. Pekosz, A., and Lamb, R. A., 1999, Cell surface expression of biologically active influenza C virus HEF glycoprotein expressed from cDNA, *J. Virol.* **73**:8808-8812.
10. Yount, B., Curtis, K. M., et al., 2003, Reverse genetics with a full-length infectious cDNA of severe acute respiratory syndrome coronavirus, *Proc. Natl. Acad. Sci. USA* **100**:12995-13000.
11. Giraud, C. G., and Maccioni, H. J., 2003, Endoplasmic reticulum export of glycosyltransferases depends on interaction of a cytoplasmic dibasic motif with Sar1, *Mol. Biol. Cell* **14**:3753-3766.
12. Hussain, S., Pan, J. A., et al., 2005, Identification of novel subgenomic RNAs and noncanonical transcription initiation signals of severe acute respiratory syndrome coronavirus, *J. Virol.* **79**:5288-5295.