The 3a Accessory Protein of SARS Coronavirus Specifically Interacts with the 5′UTR of Its Genomic RNA, Using a Unique 75 Amino Acid Interaction Domain

Kulbhushan Sharma, Milan Surjit, Namita Satija, Boping Liu, Vincent T. K. Chow, and Sunil K. Lal

Virology Group, International Centre for Genetic Engineering & Biotechnology, Aruna Asaf Ali Road, New Delhi 110067, India, and Microbiology Department, National University of Singapore, Singapore

Received October 3, 2006; Revised Manuscript Received April 3, 2007

ABSTRACT: More than four years have passed since the outbreak of the severe acute respiratory syndrome (SARS) epidemic, and still very little is known about the molecular biology and pathogenesis of this deadly virus. Among the accessory proteins of the SARS coronavirus (SARS-CoV), the 3a protein has been shown to interact with the spike, envelope, and membrane glycoprotein and has recently been established to be a structural component of capsid. Recent studies suggest that the 3a protein may function as an ion channel and may promote virus release. In order to further characterize the functional properties of this protein, we initiated studies to check its RNA binding activity. Using the yeast three-hybrid system, electrophoretic mobility shift assay (EMSA), and ultraviolet (UV) cross-linking techniques, we have shown that the 3a protein is capable of binding specifically to the 5′ untranslated region (5′UTR) of the SARS virus genomic RNA. Further, we have mapped the interaction domain of the 3a protein responsible for this RNA–protein interaction using a series of deletion mutants and defined it to the central 75 amino acid region. This RNA binding motif of 3a does not share homology with any other known RNA binding protein and may have an important role in viral capsid assembly and pathogenesis.

The virus responsible for severe acute respiratory syndrome (SARS) is a novel human coronavirus (SARS-CoV) that has become a serious public health concern since its identification after the epidemic in 2003 which resulted in more than 8400 cases and around 800 deaths in different regions of the world. SARS-CoV is an enveloped, positive-stranded RNA virus comprising approximately 30,000 nucleotides. Like other members of the genus coronavirus, e.g., murine hepatitis virus (MHV) and bovine coronavirus (BCoV), the SARS-CoV genome contains six major open reading frames (ORFs) that encode the two replicase polyprotein and four structural proteins such as the spike (S) glycoprotein, the envelope (E) protein, the membrane (M) glycoprotein, and the nucleocapsid (N) protein (1, 2, 4, 5). Earlier studies indicate SARS-CoV to be different from other coronaviruses and constitute a fourth group in the Coronavirus genus (2, 3, 6, 7). However, many studies have identified the group 2 coronaviruses as the closest relatives of SARS-CoV (8–11). In contrast to other human coronaviruses that cause benign infection such as common cold, SARS-CoV is highly virulent and induces acute atypical pneumonia associated with a high mortality rate (~3–6%) that is even higher (~43–55%) in the elderly population (6). Therefore, it is critical from both medical and research points of view to find out the factors that enable SARS-CoV to behave so distinctly from other human coronaviruses.

Apart from its divergence from other coronaviruses in amino acid sequence of other known structural proteins, the SARS-CoV also possesses nine unique ORFs that encode accessory proteins. Given that these proteins are novel in nature and uniquely present only in the SARS-CoV genome, it is tempting to speculate that they might be playing a significant regulatory role during viral pathogenesis. However, to date only the 7a, 6, and 3a proteins have been studied to some extent (3, 12–20).

The 3a protein (also termed as X1 or U274) is translated from the ORF3 of the SARS-CoV genome (Figure 1A). It is the largest of all predicted SARS-CoV accessory proteins consisting of 274 amino acids (3) and lacks significant homology with any other known protein, although some regions show low-level homology to cytochrome b561, outer membrane porin of bacteria, and the calcium pump of malaria parasite (19). This protein has been experimentally demonstrated to be expressed in SARS-CoV infected patients as well as in VeroE6 cells infected with SARS-CoV using immunohistochemistry and immunofluorescence assays (19). Sera from SARS patients could detect recombinant 3a protein expressed in Escherichia coli, further proving that this protein is indeed expressed during the SARS-CoV life cycle (19). 3a is a type III membrane protein that associates with the viral membrane and is distributed over the cytoplasm in a fine punctate pattern (19). An elegant study done by Tan et al. has shown that this protein gets transported to the
cell surface followed by endocytosis mediated by the tyrosine 
based sorting motif present in its cytoplasmic domain (17).
This protein has also been shown to localize at both the 
perinuclear region and the plasma membrane (17). Interaction 
studies have shown that 3a interacts with the spike, envelope, 
nucleocapsid, and membrane proteins of SARS-CoV (17). 
Based on its ability to get endocytosed from the cell surface 
and interact with the S protein, it has been proposed that 
this protein might be regulating the trafficking of the S 
protein during viral infection (18). Moreover, 3a has also 
been detected in purified virions further supporting its status 
as a structural protein (13). The 3a protein gets efficiently 
released in membranous structures in SARS-CoV infected 
cells and when heterologously expressed in some other cell 
lines (21). Overexpressed 3a protein was found to be secreted 
into the culture medium in association with M and E, and 
was able to form virus-like particles (3). Also, 3a has been 
shown to induce apoptosis in Vero E6 cells (13). In a study 
by Lu et al., the 3a protein forms homodimers and homotet-
ramers (20). Its expression has been confirmed in SARS 
infected patients, and it has been shown to function as an 
ion channel that modulates virus release (20). An elegant 
model describing the localization of the 3a protein has been 
proposed which shows that the 3a protein is a transmembrane 
protein (20). Even the orientation of the protein has been 
studied and the N-terminus region of the protein has been 
shown as extracellular whereas the C-terminus region is 
intracellular (20). Thus, in addition to the basic structural 
proteins of SARS coronavirus, the 3a accessory protein also 
appears to serve multiple and diverse functions during SARS-
CoV infection.

In this report, we provide multiple experimental evidence 
in support of an additional function for the 3a protein: the 
ability to specifically bind to its viral genomic RNA. Using 

---

**Figure 1:** The SARS genome and the yeast three-hybrid constructs used to study RNA–protein interactions. (A) Genes and genome organization of SARS-CoV. Both 5′ UTR and 3′ UTR are shown. Triangles show viral protease cleavage points and their resulting nonstructural proteins (nsp). Star represents the 3a (also called X1) ORF. (B) Fusion RNA constructs designed to express fusion transcripts in yeast cells. MS2 vector RNA coding region was cloned with two different SARS genomic regions, i.e., 5′ UTR and 3′ UTR. (C) Schematic diagram showing the Gal4-AD fused in-frame to the 3a gene thus expressing fusion proteins in yeast cells. (D) Schematic diagram of the yeast three-hybrid system showing the different fusion-RNA and hybrid-protein constructs being examined. P: promoter. Ter: terminator.
the yeast three-hybrid assay, electrophoretic mobility shift assay (EMSA), and UV–cross-linking assay, the 3a protein was found to interact with the 5′ untranslated region (UTR) of the SARS-CoV genome. However, 3′ UTR as well as a nonspecific RNA sequence (iron responsive element) failed to associate with the 3a protein, thus indicating that the 3a protein specifically associates only with the 5′ region of the genome. Further, deletion mapping revealed that the central 75 amino acid (aa 125–200 from the N-terminus) region of the 3a protein was responsible for this RNA protein interaction. Interestingly, this region does not share homology with any other known RNA binding protein. Thus, we propose that the 3a protein contains a unique RNA binding motif. The possible significance of this RNA–protein interaction during the viral life cycle is discussed.

**EXPERIMENTAL PROCEDURES**

**Plasmids and Reagents.** Yeast strains and plasmid constructs used in this study are summarized in Table 1. The 3a region of SARS-CoV genome (GenBank accession number NC_004718; Figure 1A) was PCR amplified and cloned into the pGEMTeasy vector. In order to express N-terminal HA tagged 3a protein, pGEMTeasy-3a was digested with EcoRI, Klenow treated, and cloned into the pIII MS2-1 vector (Figure 1C). All the deletions of 3a were PCR amplified using specific primers having sites for Smal and EcoRI at the 5′ and 3′ ends respectively and were cloned into the pCR-XL-TOPO vector. These constructs were then digested with Smal—EcoRI, and the fragment was ligated into pGAD vector and digested with the same set of enzymes.

### Table 1: Yeast Strains, Plasmids, Recombinant Plasmid Constructs and Primers Used in This Study

<table>
<thead>
<tr>
<th>strain/plasmid construct/primers</th>
<th>genotype:description</th>
</tr>
</thead>
<tbody>
<tr>
<td>L40-coat</td>
<td>MATa, ura3-52 leu2-3, 112, his3Δ200, trp1Δ1, ade2, LYS2::(LexA-op)-His3, ura3-::(LexA-op)-LacZ, LexA-MS2 coat (TRP1)</td>
</tr>
<tr>
<td>pGAD</td>
<td>GAL4 AD vector [GAL4 (768–881)]; LEU2, 2 μm, Ampr</td>
</tr>
<tr>
<td>pGAD-3a (201–275)</td>
<td>derived from pIII Ex426RP</td>
</tr>
<tr>
<td>pGAD-3a (1–70)</td>
<td>Smal—EcoRI site</td>
</tr>
<tr>
<td>pGAD-3a (1–70)</td>
<td>Smal—EcoRI site</td>
</tr>
<tr>
<td>pGAD-3a (1–124)</td>
<td>Smal—EcoRI site</td>
</tr>
<tr>
<td>pGAD-3a (125–275)</td>
<td>Smal—EcoRI site</td>
</tr>
<tr>
<td>pGAD-3a (125–200)</td>
<td>Smal—EcoRI site</td>
</tr>
<tr>
<td>pGAD-3a (601–825)</td>
<td>Smal—EcoRI site</td>
</tr>
<tr>
<td>pGAD-3a (601–825)</td>
<td>Smal—EcoRI site</td>
</tr>
<tr>
<td>pCR-XL-TOPO-5′UTR</td>
<td>Smal—EcoRI site</td>
</tr>
<tr>
<td>pCR-XL-TOPO-3′UTR</td>
<td>Smal—EcoRI site</td>
</tr>
<tr>
<td>pCR-XL-TOPO-5′UTR</td>
<td>Smal—EcoRI site</td>
</tr>
<tr>
<td>pCR-XL-TOPO-3′UTR</td>
<td>Smal—EcoRI site</td>
</tr>
</tbody>
</table>

**Primer Sequences:**

F1: 5′ CGAATTCTAGGATGTTTATGAGATTTTTTAC 3′
F2: 5′ GGAAATTCATGAGATTTTATGAGATTTTTTAC 3′
F3: 5′ GGAAATTCTAGGATTGAGTTTATGAGATTTTTTAC 3′
F4: 5′ GGAAATTCATGAGATTTTATGAGATTTTTTAC 3′
F5: 5′ GGAAATTCATGAGATTTTATGAGATTTTTTAC 3′
F6: 5′ CCGGATATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTA AGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGTTATTTAAGT
SARS-CoV genomic sequence spanning 1-264nt from the 5’ end (5’UTR) was PCR amplified and cloned into the pCR-XL-TOPO vector. The pCR-XL-TOPO-5’UTR was then digested with Smal and Sphl and was cloned into the Smal–Sphl site of pHIMS2-2 vector (Figure 1B). Similarly, the 29389–29751 nucleotide region (3’UTR) of SARS-CoV genomic sequence was PCR amplified and cloned into pCR-XL-TOPO vector to produce pCR-XL-TOPO-3’UTR (Figure 1B). The pCR-XL-TOPO-3’UTR was digested with BamHI, Klenow treated, and then digested by EcoRV and cloned into Smal site in pIII MS2.1. All clones were checked by restriction mapping and sequencing.

Yeast Three-Hybrid Assay. The GAL-4 based yeast three-hybrid system (Figure 1D), kindly provided by Dr. M. Wickens (University of Wisconsin, Madison, WI) was used in the experiment (22). To prepare competent cells, L-40 strain was inoculated into 5 mL of YPD media and was grown on a 30 °C shaker at 200 rpm for 24 h. Next, 300 μL of this culture was inoculated into 100 mL of YPD media and was grown on a 30 °C shaker at 200 rpm until A600 was 0.8. The culture was then centrifuged at 3000 rpm for 2 min at room temperature, and the pellet was washed with 10 mL of LiAc (0.1M). The pellet was resuspended in 1 mL of LiAc (0.1 M).

To transform into host yeast cells, appropriate fusion constructs (10 μg) were mixed with 4 μg of carrier DNA in a microcentrifuge tube. To this, 100 μL of L-40 competent cells were added, and the mixture was kept for 5 min at room temperature. Next, 280 μL of 50% PEG-3350 was added to the reaction mixture, and it was kept for 45 min at 30 °C. Next, 43 μL of dimethyl sulfoxide (DMSO) was added and mixed by inversion. Then, a heat shock was given at 42 °C for 5 min. A pulse spin was given, and the pellet was washed with 500 μL of sterile water. It was finally dissolved into 100 μL of water and plated on YPD. The plates were kept at 30 °C for 3 days. Next, from each YPD plate, 10 individual colonies were streaked on LeuUraHis− plates only or LeuUraHis− plates containing 50 mM 3-amino-1,2,3-triazole (3-AT) and the plates were kept for 3 days at 30 °C. The filter β-gal assay was performed by streaking doubly transformed yeast colonies onto filter paper and allowing them to grow for 2 days on selection medium. Yeast cells were permeabilized by freezing yeast-impregnated filters in liquid nitrogen and thawing at room temperature. The filter was placed over a second filter that was presoaked in 0.1 M carbonate (1 M) was added. Reaction tubes were centrifuged at 13 000 rpm for 10 min, and the supernatant was transferred to cuvettes to calculate β-gal expression levels, which were determined by measuring A420 (26, 29, 30).

Protein Expression and Immunoprecipitation. The full-length 3a protein (pGAD-3a encoding 275 amino acids of 3a with N-terminal HA tag) and its deletion mutants (Table 1) were expressed in separate reactions using the in vitro coupled transcription—translation rabbit reticulocyte lysate (TNT) system (Promega, Madison, WI). The reaction was assembled according to the manufacturer’s protocol. In parallel, one control reaction with an empty pGAD vector was also assembled and labeled as mock lysate. All reactions were then incubated for 90 min at 30 °C. Protein either was labeled with S35 or was unlabeled. The expression of unlabeled protein was detected by immunoprecipitation followed by chemiluminescence (20X LumiGLO Reagent, cell signaling). For immunoprecipitation, a 10 μL aliquot of the reaction products was incubated with 1 μg of anti-HA antibody in 500 μL of PBS (pH 7.0) and kept overnight on a rocking shaker at 4 °C. The next day, 100 μL of 10% protein-A sepharose beads (Amersham, Buckinghamshire, U.K.) was added to the reaction and the reaction mix was kept on a rocking shaker for 90 min at 4 °C. The beads were then washed thrice with PBS (pH 7.0), resuspended in 10 μL of SDS PAGE loading buffer (50 mM Tris-HCl pH 6.8, 5% β-mercaptoethanol, 2% SDS, 0.1% bromo-phenol blue, and 10% glycerol), and boiled. The supernatant was subjected to SDS PAGE, and the protein expression was analyzed either by autoradiography for labeled protein or by chemiluminescence for the unlabeled protein (30). Chemiluminescence was performed using 20X LumiGLO Reagent (cell signaling) as per the manufacturer’s protocol.

Gel Shift Assay (EMSA). pCR-XL-TOPO-5’UTR and pCR-XL-TOPO-3’UTR were linearized by HinIII and used to transcribe 32P labeled RNA using the T7 Maxi-script kit (Ambion, Austin, TX) and were purified using the RNeasy kit (Qiagen, Valencia, CA). Integrity of RNA transcripts was checked on a 6% urea–acrylamide gel. Unlabeled full-length 3a protein and its deletions were produced using the in vitro coupled transcription—translation rabbit reticulocyte system (Promega, Madison, WI) and verified by immunoprecipitation using anti-HA antibody.

To perform EMSA, unlabeled 3a protein and its deletion mutants or mock translated lysate were mixed with 75 000 cpm of 32P labeled RNA and 25 μg of yeast t-RNA, and the reaction was incubated on ice for 20 min in 20 μL of binding buffer having 10 mM Tris-HCl (pH 7.5), 50 mM KCl, 1 mM DTT, and 10% glycerol. Reaction products were analyzed on native 6% polyacrylamide gel followed by autoradiography.

A supershift assay was performed by adding 1 μg of anti-HA antibody in addition to the reaction mix used for EMSA. All other steps involved were exactly similar as described for EMSA. As a control for supershift assay, a nonspecific antibody was used.
For competition binding assay, unlabeled RNA was transcribed using Ambion’s T7 Maxi script kit and checked on formaldehyde-agarose gel (data not shown). A 50-fold excess of unlabeled RNA was incubated along with the regular reaction mix, and gel shift assay was performed as described above (30).

**UV—Cross-Linking Assay.** Here, we have used immunoprecipitated 3a protein in order to reduce the background in the gel. To perform the UV—cross-linking assay, 15 μL of the unlabeled TNT reaction product was incubated with 1 μg of anti-HA antibody in 500 μL of PBS (pH 7.0) and kept overnight at 4 °C with gentle rocking. Next, 100 μL of 10% protein-A sepharose beads (Amersham, Buckinghamshire, U.K.) was added to the reaction and the reaction mix was kept for 90 min at 4 °C on a rocking shaker. The beads were then washed five times with PBS (pH 7.0) and dissolved in 15 μL of binding buffer as described for the EMSA experiment above. To this reaction mix was added 75 000 cpm of 32P labeled RNA and 25 μg of yeast t-RNA. The reactions were incubated at room temperature for 20 min and kept on ice, where they were irradiated for 20 min with UV light (254 nm) having energy value of 1200 μJ/cm2. Next, RNase A was added at a concentration of 1.25 μg/mL to digest excess RNA and the reaction mixture was incubated at 37 °C for 60 min. The reaction product was separated on 15% SDS—PAGE and was analyzed by typhoon variable mode imager (Amersham, Buckinghamshire, U.K.).

**RESULTS**

3a Protein Interacts with the 5′ End of SARS Genomic RNA. Since the 3a protein has been suggested to be a structural component of the SARS virus particle, we wished to check whether 3a also associated with the viral genomic RNA. To begin with, we tested the ability of the 3a protein to associate with the 5′- and 3′UTRs of the SARS-CoV genomic RNA. Our primary experimental setup involved testing the interactions of 3a with 5′UTR and 3′UTR of the SARS-CoV genome, using the yeast three-hybrid method (22, 30, 31). In the yeast three-hybrid assay, plasmid constructs expressing a fusion RNA and a fusion protein of interest are transformed into a genetically engineered host strain that bears HIS3 and lacZ reporter genes downstream of a GAL4 responsive promoter. In the co-transformed yeast cell, the fusion RNA molecule bridges the two-hybrid proteins, one containing a DNA-binding domain (BD) and the other containing a transcriptional GAL4 activation domain (AD), to which the test protein is fused, in-frame, thus resulting in the transcriptional activation of HIS3 and lacZ reporter genes downstream of the BD (Figure 1D).

To study the interaction of the 3a protein with genomic RNA by the yeast three-hybrid system, we designed constructs fusing MS2-RNA with the 5′UTR (1—264 nt) and the 3′UTR (29389—29751 nt) of SARS-CoV (Figure 1A, 1B, and 1D, Table 1), so as to express fusion RNA transcripts in the L-40 coat yeast cells (Table 1). The full-length 3a gene of the SARS-CoV was cloned in-frame with the Gal4 AD to express the AD-3a protein (Figure 1C) (26, 27, 29). Expression of the full-length, 35S labeled 3a protein (tagged with an HA epitope at the N-terminus) in pGAD vector was performed using the in vitro coupled transcription—translation rabbit reticulocyte lysate (TNT) system, followed by immunoprecipitation using anti-HA antibodies (Figure 2A). The 3a protein was found to correspond to the correct molecular size, i.e., approximately 35 kDa. The expression of unlabeled 3a protein was analyzed and confirmed by immunoprecipitation followed by chemiluminescence (data not shown).

The yeast L-40-coat host strain was systematically cotransformed with the AD-3a construct along with MS2-5′ SARS or MS2-3′ SARS or other control plasmids. Cotransformants were selected on Leuura− dropout medium and tested again for their ability to grow in LeuuraHis− dropout media containing increasing concentrations of 3-AT. As a second marker to detect positive RNA—protein interactions, the cotransformants that showed positive on the LeuuraHis+ 3-AT plates were assayed for β-galactosidase activity using O-nitrophenyl-d-galactoside as described in Experimental Procedures (24, 25, 28). Results of the yeast three-hybrid assay are summarized in Figure 2B. As expected, the host strain (L40-coat) as well as single transformants did not show any HIS3 or β-gal reporter activity. Similarly, cotransformants bearing empty vectors (MS2- + AD-) showed negative growth on LeuuraHis− dropout media and β-galactosidase assays. As a positive control for the assay, cotransformants expressing the iron response element (pIII MS2-IRE) and iron regulatory protein (AD-IRP) were tested for their ability to grow on LeuuraHis− dropout medium and assayed for β-galactosidase activity (22, 32, 33). These cells showed positive histidine prototrophy and produced a distinct blue color on the β-galactosidase filter and bright yellow color for the liquid β-galactosidase assays (Figure 2B). Further, the strength of interaction was measured by growing these cells in different concentrations of 3-AT, which is known to be a competitive inhibitor of the HIS3 protein, thus enhancing the stringency of selection (34). Interestingly, of the two genomic regions (5′UTR and 3′UTR) which were tested for their ability to associate with the 3a protein, only the 5′UTR showed histidine prototrophy as well as strong β-galactosidase activity. The interaction between 5′UTR and 3a protein appeared to be relatively stronger than the positive control as judged by liquid β-galactosidase assay (Figure 2B). In contrast, the 3′UTR did not show any significant reporter activity, further suggesting that the 5′UTR specifically associates with the 3a protein.

In order to confirm the results obtained from the yeast three-hybrid assay, we verified the ability of the 3a protein to bind to the 5′UTR of the SARS genome by electrophoretic mobility shift assay (EMSA) and further confirmed the same by conducting a UV—cross-linking assay as well. Radiolabeled transcripts were prepared as described in Experimental Procedures. The integrity of the purified 32p labeled transcripts was then analyzed on a 6% urea—polyacrylamide gel (data not shown). Unlabeled 3a protein was produced in vitro and EMSA was performed as described in Experimental Procedures (30, 35). Reaction products were analyzed by electrophoresis on native 6% PAGE gel and visualized by autoradiography. As a control experiment, mock lysate was incubated along with the labeled probe in a similar manner to rule out any possible binding of endogenous protein from the in vitro coupled transcription—translation rabbit reticulocyte lysate system. Figure 2C shows the results of this control experiment where the 5′UTR transcript binds with the 3a protein showing a clear mobility shift (lane 3). In
FIGURE 2: Full-length 3a protein interacts specifically with the 5′UTR. (A) The S35 labeled 3a protein tagged with an HA epitope was expressed in vitro using a rabbit reticulocyte coupled transcription—translation system (lane 2) and verified further by immunoprecipitation (lane 4) using anti-HA antibodies. The reaction mixtures were analyzed on a 15% SDS-polyacrylamide gel. Mock lysate was used as a negative control (lanes 1 and 3). The arrow shows 3a protein expression at approximately 35 kDa. M represents mock lysate. (B) Results from the three-hybrid analysis showing 5′UTR genomic RNA interacting with the 3a protein. YPD: yeast extract peptone dextrose media (nonselective). Leu−, Ura−, and LU− represent SD-Leu− (synthetic dextrose complete media lacking leucine), SD-Ura− (synthetic dextrose complete media lacking uracil), and SD-Leu− Ura− (synthetic dextrose complete media lacking leucine and uracil); LUHis− +3-AT (synthetic dextrose complete media lacking histidine, leucine, and uracil with 3-aminotriazole) with 0 and 50 mM 3-aminotriazole (3-AT) added. βF represents results from the β-galactosidase filter assay. L40-coat is the untransformed yeast host strain. MS2-IRE/AD-IRP is used as a positive control in the assay (22, 33). (C) EMSA to verify the 3a-5′UTR interaction. The 5′UTR and 3′UTR were 32P radiolabeled. Arrow shows the unlabeled 3a protein bound to 32P labeled 5′UTR (lane 3) and not to 32P labeled 3′UTR (lane 4). Lanes 1 and 2 are negative controls with mock translated lysate added to the reactions. * refers to 32P labeled transcripts. (D) Supershift assay to confirm the results of mobility shift assay. The 5′UTR and 3′UTR were 32P radiolabeled. Lane 3 shows the supershift observed by adding anti-HA antibody as compared to lane 1 where no antibody was added. As a control reaction, a nonspecific antibody was added instead of anti-HA antibody, and as expected, there was no supershift observed as shown in lane 2. Lane 3 shows a control reaction where mock lysate was added instead of 3a protein. The 3a protein being used in this experiment is unlabeled. * refers to 32P labeled transcript. (E) Competitor binding assay to prove specificity of the RNA—protein interaction. The 3a protein being used in this experiment is unlabeled. Lane 4 contains a nonspecific competitor, 3′UTR RNA transcript. Lane 3 contains a 50-fold excess of unlabeled 5′UTR RNA transcript. Lanes 1 and 2 are negative and positive controls, respectively. * refers to 32P labeled transcript. Non-radioactive transcripts 3′UTR and 5′UTR were used in 50-fold higher concentrations. Arrow shows 3a protein bound to 32P radiolabeled 5′UTR.
contrast, the 5′UTR transcript incubated with the mock translated lysate (lane 1) or the 3′UTR transcript incubated with the mock lysate or lysate expressing the 3a protein (lane 2 and 4 respectively) all failed to show any shift in migration.

To finally prove that the mobility shift is a result of binding by the 3a protein, we performed a supershift assay where we added anti-HA antibody to the above reaction mix. The results of the supershift assay are shown in Figure 2D. A clear difference in mobility was observed for the complex having anti-HA antibody in addition to the 3a protein and 5′UTR (lane 3). Full-length 3a protein interacted with 5′UTR as a positive control (lane 1). A nonspecific antibody when used along with 3a and the 5′UTR showed negative supershift (lane 2), and a negative shift was also observed when the 3a protein was missing from the same reaction as in lane 3 (lane 4). This data, along with the mobility shift assay results, supports our hypothesis that the 5′UTR specifically associates with the 3a protein.

To further confirm the results of the gel-shift assay, a competitor-binding assay was performed. A 50-fold excess of the unlabeled 5′UTR transcript (specific competitor) was incubated along with the regular reaction mixture described above. Also, as a nonspecific competitor, unlabeled 3′UTR was used in 50-fold higher concentration in a parallel reaction. Results obtained from this experiment showed that the unlabeled 3′UTR did not compete for binding with the labeled 5′UTR–3a RNA–protein complex (Figure 2E, lane 4), whereas unlabeled 5′UTR was able to compete with the labeled probe for binding to the 3a protein, thus resulting in a significant decrease of the radioactive signal (Figure 2E, lane 1 and 2) respectively. The 3a protein being used in this experiment is unlabeled. * refers to 32P labeled transcript. The 3a protein has an ability to interact with the 5′UTR–3a RNA–protein complex (Figure 2E, lane 4) whereas no interaction was seen for unlabeled 3a with 32P labeled 5′UTR (lane 4) whereas no interaction was seen for unlabeled 3a with 32P labeled 3′UTR (lane 3). Lanes 1 and 2 are negative controls with mock translated lysate added to the reaction. * refers to 32P labeled transcripts. Arrow shows the binding of 3a protein to labeled RNA. (B) UV–cross-linking competitor binding assay to prove specificity of the RNA–protein interaction. Lane 3 contains a nonspecific competitor, 3′UTR RNA transcript. Lane 4 contains a 50-fold excess of unlabeled 5′UTR of SARS-CoV RNA transcript. Lanes 1 and 2 are negative and positive controls, respectively. The 3a protein being used in this experiment is unlabeled. * refers to 32P labeled transcript. The 3a protein has been shown by the arrow. Non-radioactive transcripts 3′UTR and 5′UTR were used in 50-fold higher concentrations.

**Amino Acids 125–200 of the 3a Protein Constitute the Binding Domain Responsible for Interaction with the 5′UTR of the SARS-CoV RNA.** We designed six deletion mutants of the 3a protein in an attempt to identify the region responsible for its RNA binding property. All 3a deletions were cloned in-frame with the Gal4 AD into the vector pGAD, as described in Table 1. Expression of various HA tagged deletion mutants of the 3a protein using a TNT kit were verified by immunoprecipitation using anti-HA antibodies followed by chemiluminescence (Figure 4A). Next, these deletions were used in a yeast three-hybrid assay to test for interaction with the 5′UTR region. Of the different mutant constructs, only AD-3a (125–275) and AD-3a (125–200) showed positive on the assay. AD-3a (1–70), AD-3a (1–124), AD-3a (70–124), and AD-3a (201–275) lost their ability to interact with the 5′UTR region (Figure 4B). As a nonspecific RNA control, 3′UTR was tested for interaction...
with all the 3a deletions, which gave negative results (data not shown). The three-hybrid analysis, thus, gave us a fairly good idea of the region of 3a imparting RNA binding activity to this protein.

In order to compare the relative strengths of interaction of the different 3a deletions with the 5'UTR, a liquid β-galactosidase assay was performed using yeast cells cotransformed with 5'UTR and different 3a deletion constructs. Standard deviation (not shown) was variable but did not exceed 25% of the experimental value. Results of this experiment revealed that, of all the 3a deletions, only 3a (125–275) and 3a (125–200) showed significant β-galactosidase activity in the assay (Figure 4C). When compared, although less than

---

**FIGURE 4:** The RNA binding domain for 3a protein is present in the amino acid 125–200 region. (A) All deletions of the 3a protein were HA-tagged and expressed in vitro using a coupled transcription–translation system and further checked by immunoprecipitation using anti-HA antibodies followed by chemiluminescence. Result analysis was done on an 18% SDS–polyacrylamide gel. (B) Mapping the interaction domain for the 3a protein by yeast three-hybrid assay. Checkered boxes represent the AD region which was fused in-frame with the 3a protein and its deletions, shown in boxed regions filled with horizontal lines. Open boxes represent regions that were deleted from 3a. The numbers above the boxes fused with AD regions represent the first and last amino acid of the regions included in the 3a deletion constructs. The numbers above the boxes fused with MS2 regions represent the first and last nucleotide used to make the constructs (Table 1). Boxes filled with vertical lines represent the MS2 regions fused with the 1–264nt 5'UTR region from the SARS-CoV genome which is shown as boxes with squares. YPD: yeast extract peptone dextrose media (nonselective). Lu represents SD-Leu-Ura-His synthetic media. βF represents results from the β-galactosidase filter assay. (C) Comparison of interaction strengths of the 3a protein deletions using the yeast three-hybrid liquid β-galactosidase assay. The column graph represents relative β-galactosidase units from the liquid β-gal assay for interaction of both full-length 3a protein and its deletions, tested with 5'UTR of SARS-CoV. Standard deviation (not shown) was variable but did not exceed 25% of the experimental value. Values are given in arbitrary units. Relative β-galactosidase units for appropriate positive and negative controls are shown. S represents UTR of SARS-CoV here.
the full-length 3a protein, the relative α-galactosidase activity of 3a (125–200) interaction was found to be almost equal to that of 3a (125–275), indicating that these two deletion mutants were able to associate with 5′UTR, with equal efficiency. This data also suggests that deletion of the C-terminal 75 amino acids from the 3a (125–275) deletion does not adversely affect its RNA binding activity. Therefore, we propose that the 3a protein of SARS-CoV is capable of binding to the 5′UTR region of genomic RNA through its central 125–200 amino acid interaction domain.

Results obtained from the yeast three-hybrid system and liquid α-galactosidase assay were verified by EMSA. For this, all unlabeled 3a deletions were expressed in a coupled transcription–translation system. These unlabeled deletions were subsequently used for EMSA with 32P radiolabeled 5′UTR transcript (Figure 5A). 5′UTR with full-length 3a protein was used as a positive control, and the 5′UTR with mock translated lysate served as a negative control for this experiment. Results thus obtained from this assay clearly showed that the full-length 3a (lane 8) as well its deletions

Figure 5: Mapping the interaction domain for the 3a protein. (A) EMSA showing that of all the deletion mutants of 3a, only 3a (125–275) and 3a (125–200) showed positive interaction with the 5′UTR region of SARS-CoV RNA (lanes 2 and 6, respectively). Arrow shows unlabeled protein bound to 32P labeled RNA. Lanes 7 and 8 are negative and positive controls, respectively. * refers to 32P labeled transcript. (B) UV-cross-linking assay to show the interaction domain for the 3a protein. A positive interaction with 5′UTR of SARS-CoV RNA was seen for 3a (125–200) and 3a (125–275) (lanes 8 and 6, respectively) when results were analyzed on a 15% SDS–polyacrylamide gel. All other deletion mutants proved negative by the assay. Lanes 1 and 2 represents negative and positive controls respectively. * refers to 32P labeled transcript.
3a (125–275) (lane 2) and 3a (125–200) (lane 6) were capable of interacting with the 5′UTR region of the SARS genome. These results fall in line with the yeast three-hybrid interaction domain mapping results.

To further validate the above observation, a UV–cross-linking assay was performed. Again, 32P labeled 5′UTR transcript as well as all deletions of 3a used for this experiment were expressed and checked, as for EMSA. The cross-linking assay was performed using immunoprecipitated, unlabeled 3a protein as discussed before, along with appropriate controls (Figure 5B). Samples were resolved by SDS–PAGE, and bands were detected by phosphorimaging. As observed in Figure 5B, binding was detected only in the cases of full-length 3a (lane 2), 3a (125–275) (lane 6), and 3a (125–200) (lane 8). These results were in perfect agreement with our yeast three-hybrid assay and EMSA results.

DISCUSSION

This study uncovers the RNA binding activity of the 3a protein of SARS-CoV. All three assays, i.e., yeast three-hybrid, in vitro gel shift assays, and UV–cross-linking assays, showed a positive interaction between the 3a protein and the 5′UTR region of the SARS-CoV genome. The fact that the observed phenomenon was not an artifact of the experimental design was ruled out in multiple ways. First, the 3′UTR as well as the iron responsive element (IRE) containing RNA was unable to show any interaction with the 3a protein. Similarly, iron regulatory protein (IRP) was unable to interact with the 3′UTR or the 5′UTR, thus proving that the interaction between the 5′UTR and the 3a protein is a specific property exclusive to both partners. The RNA binding motif was mapped to the central 75 amino acids of the 3a protein. This region was found to consist of sheets and turns, when analyzed by Macvector program (Figure 6). The secondary structure plot shown in Figure 6 has two parts: running average plots for the helix, coil, and turn structures; and structure prediction bars across the top of the plot. Both Robson–Garnier and Chou–Fasman methods are used to construct this plot (38–43). A consensus plot has been generated using information from both Robson–Garnier and Chou–Fasman plots. The running average values are the products of the normalized values from each individual plot, and structure prediction bars are only drawn where both methods agree. This region shares no homology with any known protein and hence may represent a novel RNA binding motif. Future studies using a mutagenesis approach will be useful in characterizing this motif.

Genomic RNA binding by viral proteins is a common phenomenon in RNA viruses. This serves two essential functions during the viral life cycle, viz., viral replication and packaging of the genome. Generally, the encapsidation signal in viruses is strategically placed such that only the full-length genome gets packaged into the viral capsid. In coronaviruses and many other RNA viruses, the nucleocapsid protein serves the function of binding the genomic RNA and encapsidating it. The genome incorporation signal in mouse hepatitis virus lies at about 20 kilobases from the 5′UTR (23). In some of the viruses like HIV, the responsibility of genome packaging is shared by more than one domain of a structural protein (44). In the case of mouse hepatitis virus, in addition to the nucleocapsid protein, another structural protein (M protein) has also been shown to cooperate in efficient full-length genome packaging although that cooperativity is mediated through the interaction of M with the nucleocapsid–RNA complex (45). These interactions may act as checkpoint mechanisms to enhance efficient genome packaging and prohibit incorporation of truncated or subgenomic sequences. A contemporary work by Lu et al. has shown that 3a is a transmembrane protein (20). The probable explanation for accessibility of a transmembrane protein to RNA may be that either the 3a protein spans the membrane only once or, in addition to plasma membrane, it may also be localized at some other regions where it is accessible to RNA. Alternatively, the possibility of a transient RNA–protein interaction, where the protein is present in cytoplasm for a limited period of time which is enough for the interaction, cannot be ignored. As already discussed, the 3a protein has been shown to interact with spike, envelope, nucleocapsid, and membrane proteins (17). All these interactions, along with this new interaction, may form complexes that may prove as important players in viral assembly and/or viral release. An in vivo study needs to be performed on live virus to conclude the exact role of these interactions.

Since 3a has been shown to be a structural protein and is able to bind the 5′UTR of the genome, it appears to be an
attractive candidate that might cooperate with the nucleocapsid protein of the SARS-CoV during genome packaging. Additionally, as 5′UTR contains sequences required for replication, an alternative role of this interaction may be during replication. Moreover, a recent report has shown that, although deletion of the group-specific ORFs of SARS-CoV did not influence replication efficiency of virus to a significant level, the greatest reduction in virus growth occurred following ORF3a deletion (46). Thus, this interaction may be a significant component of the virus-replication machinery. If this is true, this interaction may, in the future, prove to be an important target for viral capsid assembly arrest or replication arrest.

ACKNOWLEDGMENT

Technical assistance by Purnima Kumar is gratefully acknowledged. Collaborative support from Microbiology Department, National University of Singapore, is gratefully acknowledged. We thank Dr. M. Wickens for the GAL-4 based yeast three-hybrid system. K.S. is a Senior Research Fellow of the University Grants Commission (UGC), India.

REFERENCES


BI062057P