4816–4825 *Nucleic Acids Research, 2006, Vol. 34, No. 17* *Published online 13 September 2006*
doi:10.1093/nar/gkl650

Amino acid residues critical for RNA-binding in the N-terminal domain of the nucleocapsid protein are essential determinants for the infectivity of coronavirus in cultured cells

Yong Wah Tan[1], Shouguo Fang[1], Hui Fan[2], Julien Lescar[2] and D.X. Liu[1,2,*]

[1]Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673 and [2]School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551

Received July 5, 2006; Accepted August 10, 2006

ABSTRACT

The N-terminal domain of the coronavirus nucleocapsid (N) protein adopts a fold resembling a right hand with a flexible, positively charged β-hairpin and a hydrophobic palm. This domain was shown to interact with the genomic RNA for coronavirus infectious bronchitis virus (IBV) and severe acute respiratory syndrome coronavirus (SARS-CoV). Based on its 3D structure, we used site-directed mutagenesis to identify residues essential for the RNA-binding activity of the IBV N protein and viral infectivity. Alanine substitution of either Arg-76 or Tyr-94 in the N-terminal domain of IBV N protein led to a significant decrease in its RNA-binding activity and a total loss of the infectivity of the viral RNA to Vero cells. In contrast, mutation of amino acid Gln-74 to an alanine, which does not affect the binding activity of the N-terminal domain, showed minimal, if any, detrimental effect on the infectivity of IBV. This study thus identifies residues critical for RNA binding on the nucleocapsid surface, and presents biochemical and genetic evidence that directly links the RNA binding capacity of the coronavirus N protein to the viral infectivity in cultured cells. This information would be useful in development of preventive and treatment approaches against coronavirus infection.

INTRODUCTION

RNA viruses, including coronaviruses, encode a group of structural phosphoproteins with basic patches on their surface. As a pivotal structural component of the virion, this group of proteins plays essential roles in packaging the RNA genome to form a ribonucleoprotein (RNP) complex resulting from assembly of the viral RNA and multiple copies of the (nucleo)capsid protein. A detailed dissection and characterization of the intrinsic RNA-binding properties of these proteins are thus essential for understanding several important processes in the life cycle of RNA viruses, including assembly of the nucleocapsid, the specific encapsidation of viral RNA and morphogenesis of virions. In this study, amino acid residues in the N-terminal domain of the nucleocapsid (N) protein of a coronavirus that are critical for its RNA-binding activity are identified and their roles in viral infectivity are analyzed, using site-directed mutagenesis based on our previous structural studies.

Coronaviruses are enveloped viruses containing the largest known single-stranded, positive-sense RNA genome of ~30 kb (1). During the coronavirus life cycle, the N protein is synthesized in large amounts and is thought to play an important role by specifically packaging the viral genome into a filamentous nucleocapsid of ~10 to 15 nm in diameter and several 100 nm in length, a macromolecular structure that is visible by using electron microscopy (2). In addition to its structural role, the N protein also participates in viral RNA transcription, replication and in modulating the metabolism of host cells (3–13). Using X-ray crystallography techniques, we recently determined a 3D structure of the N protein from coronavirus infectious bronchitis virus (IBV), a prototype coronavirus, and showed that the basic building block for nucleocapsid assembly was a dimer of the N protein (14). The 409 amino acids long IBV N protein is composed of two globular domains that are resistant to proteolysis. Amino acids 29 to 160 form the N-terminal RNA binding domain and residues 218 to 329 form its C-terminal dimerization domain (5,14). Even though the structure of the N-terminal domain was determined in the absence of a nucleic acid ligand, its overall shape and the distribution of

*To whom correspondence should be addressed. Tel: +65 65869581; Fax: +65 67791117; Email: dxliu@imcb.a-star.edu.sg

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

electrostatic charges suggest a plausible model for RNA binding. The outer surface of the N-terminal domain is enriched in aromatic and basic residues, a common feature for RNA-binding proteins (15). The shape of this domain is reminiscent of a hand having basic fingers, a hydrophobic palm and an acidic 'wrist'. The positively charged 'fingers-like' β-hairpin extension could neutralize the phosphate groups emanating from RNA, whilst the base moieties could make contact with exposed aromatic residues from the hydrophobic palm (14).

Based on this structural hypothesis, we chose to mutate a number of evolutionarily conserved residues in the N-terminal domain of the IBV N protein in order to assess their roles in viral genomic RNA binding and in viral replication. We used purified recombinant wild type and mutant proteins containing the N-terminal domain expressed in *Escherichia coli* for RNA binding assays. These studies led to the identification of a number of amino acid residues essential for the RNA-binding activity of the domain. Subsequently, we introduced mutations that either severely or mildly impair the RNA binding activities of the N-terminal domain into an infectious cDNA clone system derived from the genomic RNA of IBV to assess their effects on viral replication and infectivity. We identified several residues exposed to the solvent whose substitutions to alanine yield a decrease of RNA binding and a concomitant reduction in virus replication. Interestingly, residues Arg-76 and Tyr-94, located at the base of the positively charged flexible hairpin loop and on the hydrophobic platform, respectively, were critical for RNA binding and viral infectivity.

MATERIALS AND METHODS

Expression and purification of IBV N protein

His-tagged wild type and mutant N-terminal domains of IBV N protein were expressed in *E.coli* BL-21 by induction with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). Cells were lysed by sonication and purified by metal affinity chromatography with Protino-Ni 150 kit (Macherey Nagel).

Synthesis of RNA probes

PCR fragments covering the IBV genome from 27 100 to 27 608 nt were cloned into a plasmid in either forward or reverse orientation under the control of a T7 promoter. The Dig-labeled sense (+) and anti-sense (−) RNA probes were made *in vitro* using the DIG RNA labeling kit according to the instructions of the manufacturer (Roche).

Northwestern blot

Five micrograms of purified proteins were resolved on an SDS-15% polyacrylamide gel and transferred onto nitrocellulose membranes (Hybond C-Extra, Amersham Biosciences) using a semi-dry transfer apparatus. Membranes were washed for 10 min with the probe buffer [1× Denhardt's Reagent, 1 mM EDTA, 10 mM Tris–HCl (pH 7.5) and 50 mM NaCl], blocked for 1 h with 25 μg/ml yeast tRNA (Ambion) and subsequently incubated with 10 μg of DIG-labeled RNA probe in the same probe buffer for 1 h. Membranes were washed three times for 15 min each with the probe buffer, before proceeding to detection with CDP-Star (Roche) according to the manufacturer's instructions.

Construction of an infectious IBV clone, introduction of mutations into the clones and rescue of recombinant viruses

Construction of an infectious IBV clone was carried out essentially as described (16,17). Briefly, five fragments spanning the entire IBV genome were obtained by RT–PCR from Vero cells infected with the Vero cell-adapted IBV p65. The PCR products were purified from agarose gels and cloned into pCR-XL-TOPO (Invitrogen) or pGEM-T Easy (Promega) vectors. Subsequently, fragment A was removed from pCR-XL-TOPO by digestion with NheI and EcoRI, and subcloned into pKT0 vector. Plasmids were digested with either BsmBI (fragment A) or BsaI (fragments B, C, D and E). The digested plasmids were separated on 0.8% agarose gels containing crystal violet. Bands corresponding to each of the fragments were cut from the gels and purified with QIAquick gel extraction kit (QIAGEN Inc.). Fragments A and B, and fragments C, D and E were first ligated with T4 DNA ligase at 4°C overnight. The two reaction mixtures were then mixed and further ligated at 4°C overnight. The final ligation products were extracted with phenol/chloroform/isoamyl alcohol (25:24:1), precipitated with ethanol and detected by electrophoresis on 0.4% agarose gels.

Full-length transcripts were generated *in vitro* using the mMessage mMachine T7 kit (Ambion, Austin, TX). The N transcripts were generated by using a linearized pKT0-IBVN containing IBV N gene and the 3′-untranslated region (3′-UTR) as templates. The *in vitro* synthesized full-length and N transcripts were treated with DNase I and purified with phenol/chloroform. Vero cells were grown to 90% confluence, trypsinized, washed twice with cold phosphate-buffered saline (PBS), and resuspended in PBS. RNA transcripts were added to 400 μl of Vero cell suspension in an electroporation cuvette, and electroporated with one pulse at 450 V, 50 μF with a Bio-Rad Gene Pulser II electroporator. The transfected Vero cells were cultured overnight in 1% FBS-containing MEM in a 60 mm dish or a 6-well plate and further cultured in MEM without FBS.

Mutations were introduced into the corresponding fragments by using QuickChange site-directed mutagenesis kit (Stratagene), and confirmed by sequencing of the whole fragments.

Growth curve and plaque sizes of the recombinant viruses on Vero cells

Confluent monolayers of Vero cells on 6-well plates were infected with wild-type and mutant viruses at a multiplicity of ∼1 PFU/cell. After 1 h of incubation at 37°C, cells were washed twice with PBS and cultured in 3 ml of MEM containing 0.5% carboxymethy cellulose for 3 days. The cells were fixed and stained with 0.1% toluidine.

Vero cells were infected with wild-type and recombinant IBV, and harvested at different times post-infection. Viral stocks were prepared by freezing/thawing of the cells three times. The plaque-forming units per ml of each sample were determined by infecting Vero cells on 6-well plates in duplicate with 10-fold serial dilution of each viral stock.

Northern blot analysis

Vero cells were infected with wild type and Q74A mutant virus at a multiplicity of ~1 PFU/cell. Total RNA was extracted from cells infected with wild type or mutant viruses. Ten micrograms of RNA were added to a mixture of 1× MOPS, 37% formaldehyde and formamide and incubated at 65°C for 20 min before subjected to gel electrophoresis. The segregated RNA bands were transferred onto a Hybond N+ membrane (Amersham Biosciences) via capillary action overnight and fixed by ultraviolet (UV) crosslinking (Stratalinker). Hybridization of Dig-labeled DNA probes was carried out at 50°C in hybridization oven overnight. Membranes were washed three times for 15 min each with the probe buffer, before proceeding to detection with CDP-Star (Roche) according to the manufacturer's instructions.

Western blot analysis

Vero cells were infected with wild type and Q74A mutant virus at a multiplicity of ~1 PFU/cell. Total proteins extracted from Vero cells were lysed with 2× SDS loading buffer in the presence of 200 mM DTT plus 10 mM of iodoacetamide and subjected to SDS–PAGE. Proteins were transferred to PVDF membrane (Stratagene) and blocked overnight at 4°C in blocking buffer (5% fat free milk powder in PBST buffer). The membrane was incubated with 1:2000 diluted primary antibodies in blocking buffer for 2 h at room temperature. After washing three times with PBST, the membrane was incubated with 1:2000 diluted anti-mouse or anti-rabbit IgG antibodies conjugated with horseradish peroxidase (HRP) (DAKO) in blocking buffer for 1 h at room temperature. After washing three times with PBST, the polypeptides were detected with a chemiluminescence detection kit (ECL, Amersham Biosciences) according to the manufacturer's instructions.

Construction of plasmids

PCR products covering the IBV sequence from 25 957–26 352 nt were amplified by using the forward primer 5′-CGGGCATATGTCTTCTGGAAATGCATCTTGG-3′ and the reverse primer 5′-CGGGATCCTTACAGGGGAATGAAGTCCCAAC-3′. The PCR fragments were digested with NdeI and BamHI, and ligated into NdeI- and BamHI-digested pET-16b (Novagen). Each mutation was introduced by two-round PCR and the mutation introduced was confirmed by automated nucleotide sequencing.

RESULTS

Overview of the structure and RNA-binding properties of the N-terminal domain of coronavirus N protein

The N-terminal RNA-binding domain of N protein from both IBV and SARS-CoV contains a flexible and positively charged hairpin loop that extends much beyond the protein core and that could grasp a nucleic acid substrate by neutralizing its phosphate groups (Figure 1) (14,18). Positively charged lysine and arginine residues in the N-terminal domain of SARS-CoV N protein were proposed to bind a 32 nt stem–loop structure located at the 3′ end of the SARS-CoV RNA genome (18). Indeed, the structure of this domain possesses some features reminiscent of RNA-binding proteins sharing the RNP fold (19–22). These RNA-binding proteins, including the U1A spliceosomal protein (23) and the coat protein from MS2 bacteriophage (24), bind the RNA ligand with residues emanating from the surface of a four-stranded anti-parallel β sheet. Presumably, strands β2 and β3 as well as the flexible β-hairpin from the IBV nucleocapsid protein could fulfill a similar role by interacting with phosphate groups from an RNA ligand (Figure 1). The β3 strand is the longest in the N-terminal domain of IBV N protein and several evolutionarily conserved residues project from it (Figure 1). Nucleotide bases in single-stranded RNA have a tendency to stack either on adjacent bases or with aromatic side chains of the protein (20).

Mutational analysis of amino acid residues essential for RNA-binding in the N-terminal domain of IBV N protein

The above structural model was partially supported by the demonstration that the N-terminal domain of IBV N protein could interact with RNA fragments corresponding to the 3′ end of the viral genome (14). To more vigorously test this model, systematic mutagenesis of residues essential for the RNA-binding activity was carried out. We targeted the exposed hydrophobic residues Tyr-92 and Tyr-94 (strand β3) that could form stacking interactions with the nucleotide bases as well as several basic or polar residues emanating from the β-hairpin (Figure 2a). As summarized in Figure 2a, 10 amino acids were mutated to alanine. These residues are either exposed to the solvent (Gln-74, Arg-76, Lys-78, Lys-81, Tyr-92 and Tyr-94) or have been conserved during evolution (Tyr-70 and Arg-73) and are thus likely to be involved in RNA binding or to play an undefined functional role. In addition, mutation of Asp-111 to an alanine was included as an additional control for the RNA-binding assay (Figure 2a).

The wild type and mutant constructs, spanning 29 to 160 amino acids with a hexa-histidine tag at their N-termini, were cloned into a bacterial expression vector, expressed in *E.coli* and purified to near homogeneity. Approximately equal amounts of the purified proteins were separated on SDS–PAGE (Figure 2b, upper panels). After transfer to a Hybond C extra membrane, the RNA-binding activities of wild type and mutant proteins were analyzed by northwestern blot with two probes corresponding to the positive- and negative-sense 3′-UTR of IBV, respectively. A moderate decrease in binding to both the positive and negative sense RNA probes was consistently observed with most mutant proteins, and representative gels are shown in Figure 2b (Figure 2b, middle and lower panels). Among all mutants, mutation of Tyr-94 to an alanine residue showed the most significant reduction in binding of the domain to both positive and negative probes. The RNA binding activities of this mutant to positive and negative probes were reduced to 27 and 25%, respectively, of the wild type were observed (Figure 2b, lane 9). Relatively less reduction in the RNA-binding activity was observed for the R76A, K78A, K81A and Y92A mutant proteins which retain between 50–74% of the wild-type binding activity (Figure 2b, lanes 5–8). With the exception of the Q74A and D111A mutants,

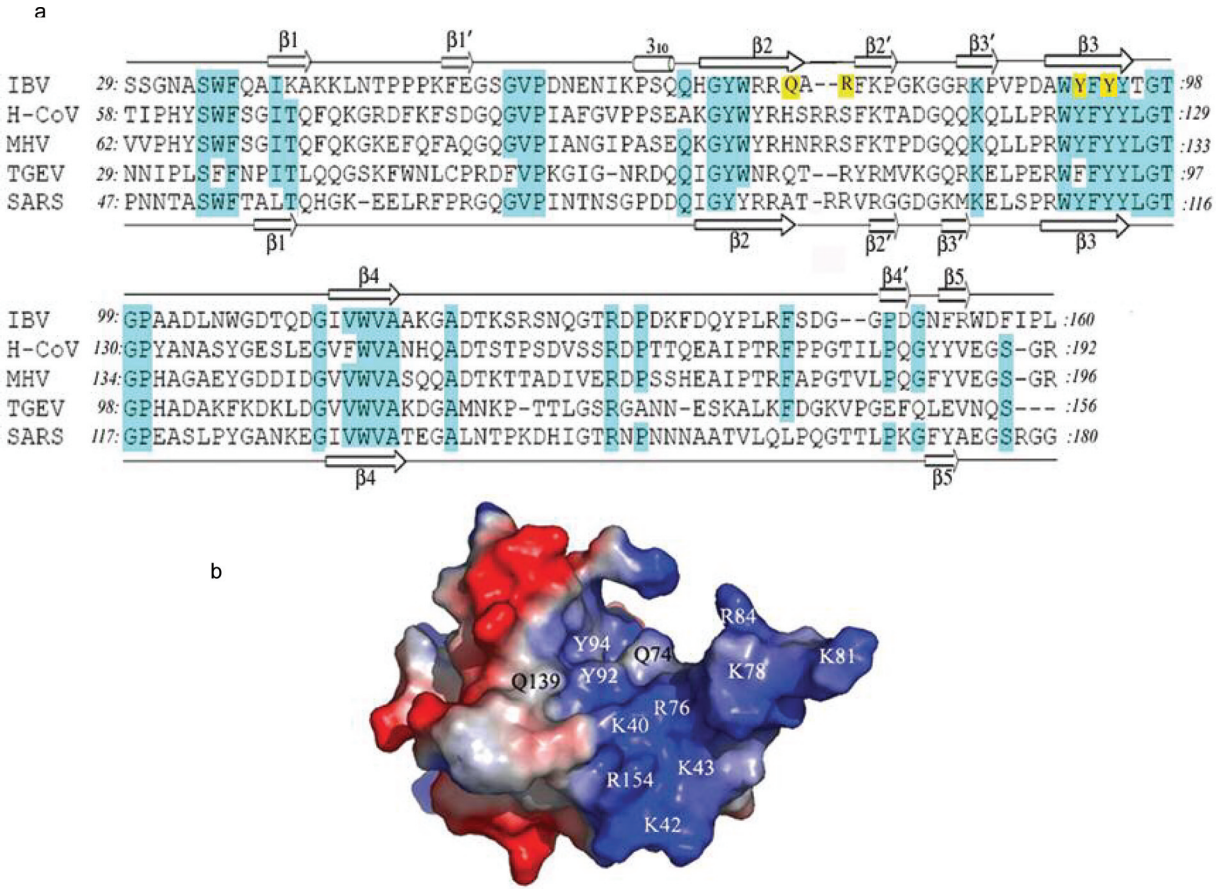

Figure 1. Structure of the N-terminal RNA binding domain of IBV N protein. (**a**) Structure-based alignment of coronavirus nucleocapsid amino acid sequences corresponding to the N-terminal RNA binding domain. Secondary structure elements are labeled above the sequence for IBV-N29–160 and below for the SARS-CoV N-terminal fragment [Huang *et al*. (18)]. Sequences of IBV (IBV, strain Beaudette, NP_040838); H-CoV (Human Coronavirus, strain HKU1, YP_173242); MHV (Murine Hepatitis Virus, strain 1, AAA46439); TGEV (Porcine Transmissible Gastroenteritis Virus, strain RM4, AAG30228) and SARS (SARS-Coronavirus, 1SSK_A) were obtained from GenBank. Conserved residues are shaded. (**b**) Surface representation of the IBV-N-terminal fragment with electrostatic potentials colored in blue (positive) and red (negative). Positively charged residues in the flexible hairpin loop and aromatic residues on the hydrophobic floor that are exposed to the solvent and which were subjected to mutagenesis in this study are labeled.

which displayed a slight increase in the RNA binding activity to both the positive- and negative-sense RNAs, (Figure 2b, lanes 4 and 10), all other mutants showed a certain degree of reduction in their binding activity to either RNA probes (Figure 2b). As a control, a circular dichroism analysis of the two corresponding R76A and Y94A single mutants was performed to check whether the mutant proteins were properly folded. The spectra obtained for these two mutants were similar to the wild-type protein, ruling out any drastic conformational change introduced by the mutations (data not shown).

As none of the single mutations introduced could totally abrogate the RNA-binding activity of the IBV N-terminal domain, two mutant constructs carrying double mutations were made. As can be seen in Figure 2b (lanes 13 and 14), mutations of both Arg-73 and Lys-78 to alanine residues (R73/K78A) reduced the RNA-binding activities to the positive- and negative-sense RNA probes between 44 and 10% of wild-type, respectively. Mutations of both Arg-76 and Tyr-94 (R76/Y94A) to alanine reduced the RNA-binding to the positive- and negative-sense RNA probes to only 7 and 19% of wild-type, respectively (Figure 2b, lanes 15 and 16). Taken together, these results are consistent with the presence of an extended RNA binding site covering a large part of the accessible surface of the N-terminal domain of the coronavirus N protein. They are also in agreement with the assumption that removal of a single positive charge at the protein surface by mutation of a lysine or arginine residue to an alanine is insufficient to totally disrupt the formation of the protein–RNA complex.

Introduction of single amino acid substitutions into the IBV genome and analysis of their effects on the recovery of infectious viruses

The R76A and Y94A mutations were first introduced into an infectious clone of IBV, and *in vitro* transcribed full-length RNA molecules derived from wild type and mutant constructs were generated by *in vitro* transcription using the T7 RNA polymerase in the presence of a cap analog. Electroporation of wild-type RNA transcripts together with an RNA fragment covering the N protein region into Vero cells showed the formation of massive CPE at 2 days post-electroporation (Figure 3, panel rIBV+N). In contrast, no

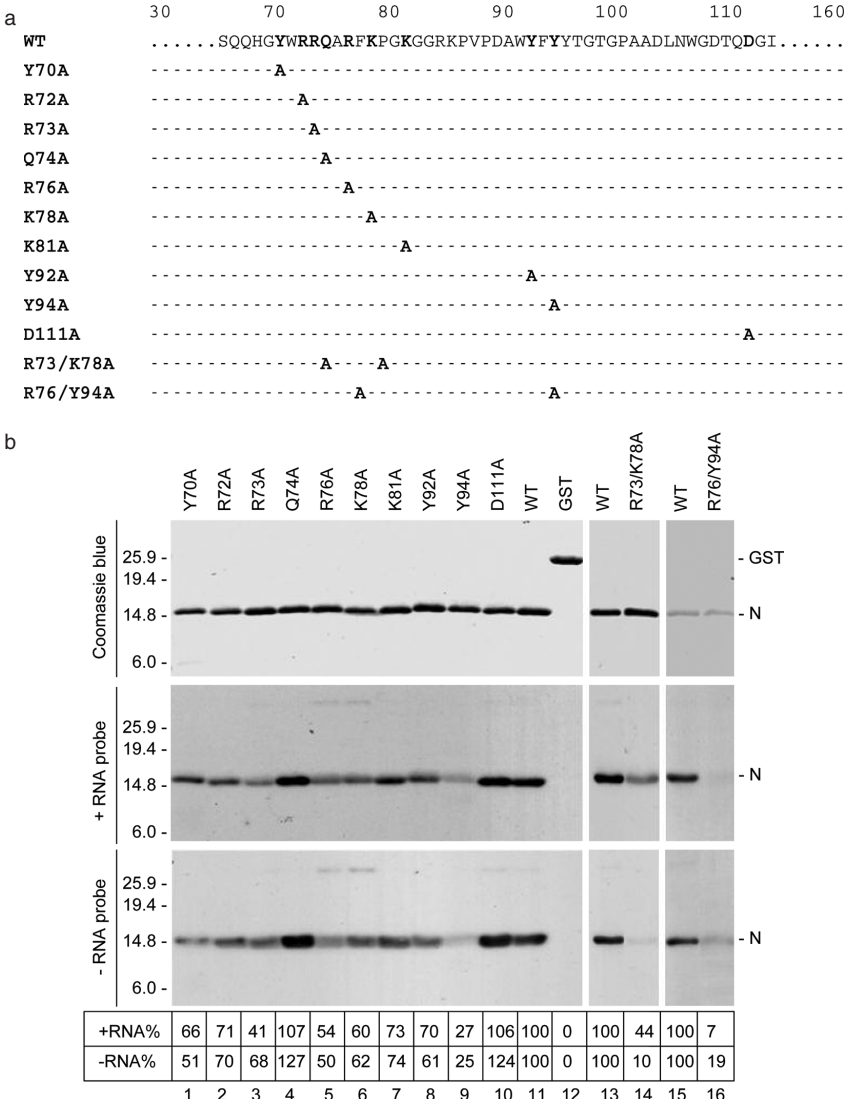

Figure 2. Northwestern analysis of the RNA binding activity of wild-type and mutant N-terminal domain of the IBV N protein. (**a**) Diagram showing the alanine mutants included in this study. (**b**) RNA binding assay for the various mutants of the IBV N protein N-terminal domain produced in this study. The binding affinities of mutated IBV N protein N-terminal domains towards positive and negative RNA ligands corresponding to 26 539–27 608 nt of the IBV genome were evaluated by northwestern blot. The mutants were separated by SDS–PAGE, transferred to Hybond C extra membrane and detected by digoxin-labelled RNA. The relative binding affinities, which were assumed to be proportional to the intensities of the detected bands, were normalized with respect to the binding activity of the wild-type domain, which was arbitrarily taken as 100.

CPE formation was observed in cells transfected with transcripts derived from either mutant up to 5 days post-electroporation, suggesting that no infectious virus could be rescued from the two mutant transcripts (Figure 3, panels R76A+N and Y94A+N). These results indicate that R76A and Y94A mutations, which severely reduce the RNA-binding activity of the N-terminal domain of the N protein, may abolish the infectivity of the *in vitro* synthesized full-length IBV transcripts.

Introduction of Q74A and Y92A mutations into the full-length infectious clone of IBV was then carried out. Typical CPEs were observed in cells transfected with Q74A mutant transcripts at 3 days post-electroporation and recombinant viruses were recovered (Figure 3, panel Q74A). In cells electroporated with Y92A transcripts, a typical CPE was observed at 2 days post-electroporation [Figure 3, panel Y92A(1)]. Incubation of the transfected cells for one more day saw the appearance of more CPEs [Figure 3, panel Y92A(2)]. However, prolonged incubation of the transfected cells up to 6 days saw no spread of the infection from the two initiated infected loci to neighboring cells. Re-infection of fresh Vero cells with either medium collected from these cells or total cell lysates prepared by freezing and thawing of the transfected cells showed no detection of the CPE formation, indicating that the minute amount of viruses rescued from the initially transfected cells could not maintain infectivity in subsequent passage.

The R76A and Y94A mutant N constructs were then used to test if they could facilitate the rescue of the *in vitro* synthesized full-length IBV transcripts in trans. As mentioned, efficient recovery of infectious virus from the *in vitro* transcribed full-length IBV RNA requires

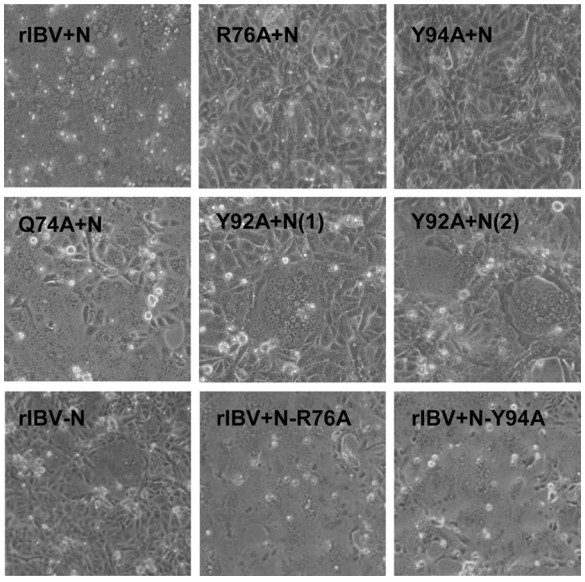

Figure 3. Introduction of single amino acid mutations into the IBV genome and recovery of infectious mutant viruses. Vero cells electroporated with the *in vitro* synthesized transcripts derived from an *in vitro* assembled full-length clone (rIBV+N), the full-length clone containing the R76A mutation (R76A+N), the full-length clone containing the Y94A mutation (Y94A+N), the full-length clone containing the Q74A mutation (Q74A+N) and the full-length clone containing the Y92A mutation [Y92A+N(1) and (2)] together with the N transcripts. The trans-effects of wild type and mutant N transcripts on rescue of infectious viruses form the *in vitro* synthesized full-length IBV transcripts were analyzed by electroporation of Vero cells with the *in vitro* synthesized full-length IBV transcripts together with wild-type (rIBV+N), R76A (rIBV+N-R76A) and Y94 (rIBV+N-Y94A) mutant N transcripts. Cells electroporated with the full-length wild-type IBV transcripts without the N-transcripts (rIBV-N) were included as a negative control. Images were taken at 3 days post-electroporation.

co-transfection of these *in vitro* transcripts together with the *in vitro* synthesized RNA covering the N gene. Electroporation of wild-type full-length IBV transcripts together with RNAs synthesized *in vitro* from wild-type and the two mutant N constructs, respectively, showed efficient recovery of the infectious virus with CPE observed on almost the whole monolayer at 3 days post-electroporation (Figure 3, panels rIBV+N-R76A and rIBV+N-Y94A). In contrast, cells transfected with the full-length IBV transcripts only (without co-electroporation with the N transcripts) showed the formation of a single CPE on the whole monolayer at 3 days post-electroporation (Figure 3, panel rIBV-N).

Analysis of negative-strand RNA replication and subgenomic RNA transcription in cells transfected with wild-type and mutant full-length transcripts

As no infectious virus was recovered from cells transfected with R76A and Y94A mutant transcripts, total RNA was extracted from cells electroporated with wild-type and mutant full-length transcripts and RT–PCR amplification of subgenomic mRNAs was carried out to check whether a low level of RNA replication and subgenomic mRNA transcription occurred in cells transfected with the mutant transcripts. The forward primer used in this reaction corresponds to the leader sequence from 26–46 nt in the genomic RNA and the downstream primers covers IBV sequences from 24 784 to 24 803 nt. If transcription of subgenomic mRNAs did occur, a 415 bp PCR product corresponding to the 5′-terminal region of the subgenomic mRNA4 and a 1010 bp fragment corresponding to the 5′-terminal region of the subgenomic mRNA3 would be expected. As shown in Figure 4a, a dominant 415 bp band and a weak 1010 bp band were observed in cells electroporated with wild-type full-length transcripts at 2 days post-electroporation (lane 2). Sequencing of the PCR fragments confirmed that they represent the correct sequences of the corresponding regions of the subgenomic mRNAs 3 and 4, respectively. The same PCR products were not detected in cells electroporated with the R76A and Y94A mutant transcripts, respectively, at 24, 48 and 72 h post-electroporation (Figure 4a, lanes 3–8). In cells electroporated with Q74A mutants at 3 days post-electroporation, RT–PCR amplification of the subgenomic mRNA 3 and 4 showed the detection of the 415 and 1010 bp bands (Figure 4a, lane 9). The same PCR products were not evident in cells electroporated with the Y92A mutant transcripts (Figure 4a, lane 10).

RT–PCR amplification of the negative strand RNA was performed to check if RNA replication occurred in these transfected cells. The primer pair was chosen so that the IBV sequence from 14 931 to 15 600 nt would be amplified by the RT–PCR. If replication of viral RNA occurred, a 670 bp PCR fragment would be expected. As shown in Figure 4b, RT–PCR fragments amplified from both positive (lanes 4, 7 and 9) and negative (lanes 3, 6 and 10) strand RNA templates were obtained from cells transfected with the mutant transcripts. The amount of the negative strand RNA was approximately half of that of the positive strand RNA (Figure 4b). As positive control, the same RT–PCR fragment for the negative strand RNA was observed in cells transfected with wild-type transcripts (Figure 4b, lanes 8 and 11). In the negative controls, total RNA extracted from cells transfected with Y94A (Figure 4b, lane 2) and R76A (Figure 4b, lane 5) was analyzed by PCR directly using the primer set for the negative strand RNA. No corresponding PCR fragment was detected from these samples, demonstrating that the detection of the negative strand RNA in cells transfected with wild-type and mutant constructs by RT–PCR is due to the replication of viral RNA. These results confirm that transcription of the negative strand RNA has taken place in cells transfected with the mutant transcripts.

The transcription of negative strand RNA in cells electroporated with wild-type, R76A, Y92A and Y94A transcripts was then quantitatively determined and compared by real time RT–PCR at 24 and 48 h post-electroporation, respectively. Compared to cells transfected with wild-type transcripts, the relative amounts of negative strand RNA in cells transfected with R76A, Y92A and Y94A were 48, 46 and 50%, respectively, at 24 h post-electroporation (Figure 4c, lanes 1–4). These figures were dropped to 41, 45 and 23%, respectively, at 48 h post-electroporation (Figure 4c, lanes 5–6). Taken together, these results demonstrate that transcription of the negative strand RNA has taken place in cells transfected with the mutant transcripts. As the level of negative strand RNA was not increased in cells transfected with the mutant transcripts over a time-course experiment, it suggests that transcription of these RNA species may

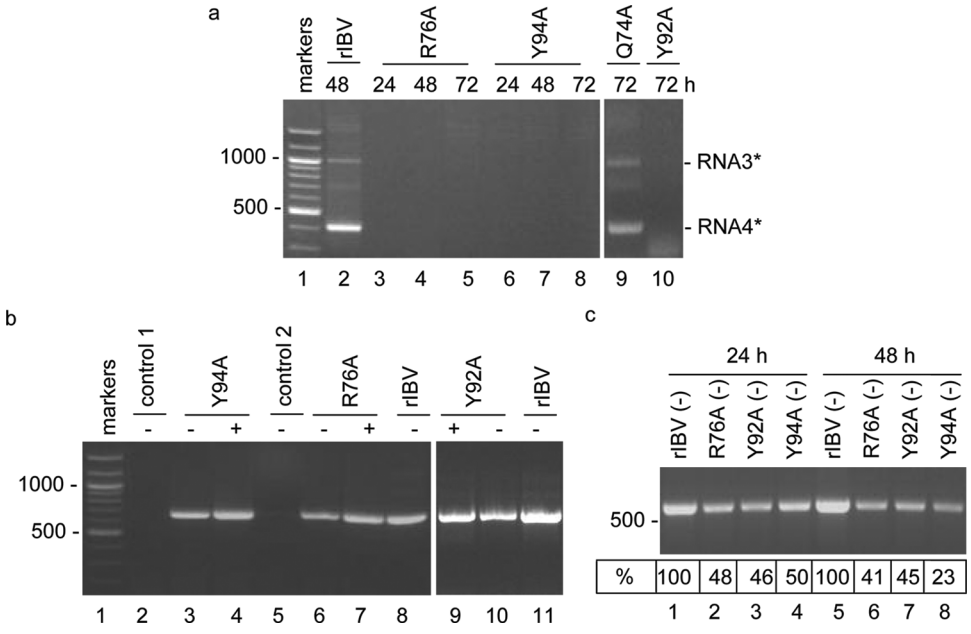

Figure 4. Analysis of negative strand RNA replication and subgenomic RNA transcription in cells transfected with wild-type and mutant full-length transcripts. (**a**) Detection of subgenomic RNA synthesis in cells electroporated with rIBV (lane 2), R76A (lanes 2–5), Y94A (lanes 6–8), Q74A (lane 9) and Y92A (lane 10) mutant transcripts. Total RNA was prepared from Vero cells electroporated with *in vitro* synthesized full-length transcripts 3 days post-electroporation. Regions corresponding to the 5′-terminal 415 and 1010 nt of the subgenomic mRNA4 and 3, respectively, were amplified by RT–PCR and analyzed on 1.2% agarose gel. Lane 1 shows DNA markers and numbers on the left indicate the length of DNA in bases. (**b**) Detection of positive and negative strand RNA synthesis in cells electroporated with rIBV (lanes 8 and 11), Y94A (lanes 3 and 4), R76A (lanes 6 and 7) and Y92A (lanes 9 and 10) mutant transcripts. Total RNA was prepared from Vero cells electroporated with *in vitro* synthesized full-length transcripts 3 days post-electroporation. Regions corresponding to 14 931–15 600 nt of the positive (+) and negative (−) sense IBV RNA were amplified by RT–PCR and analyzed on 1.2% agarose gel. The negative controls shown in lanes 2 and 5 are PCR analysis of total RNAs extracted from cells transfected with Y94A (lane 2) and R76A (lane 5) using the primer set for negative strand RNA. Lane 1 shows DNA markers. Numbers on the left indicate nucleotides in bases. (**c**) Quantitative analysis of the negative strand RNA in cells transfected with wild type (lanes 1 and 5), R76A (lanes 2 and 6), Y92A (lanes 3 and 7) and Y94A (lane 4 and 8) by real time RT-PCT at 24 (lane 1–4) and 48 (lanes 5–8) h post-electroporation. The relative ratios of the negative strand RNA in cells transfected with the mutant transcripts to those in cells transfected with wild type transcripts are shown. Numbers on the left indicate nucleotides in bases.

occur only in the initially transfected cells using the *in vitro* transcripts as templates.

Genetic stability and growth properties of the Q74A mutant virus

The growth properties of the Q74A mutant virus on Vero cells were tested by analysis of plaque sizes and growth curves of passage 3 mutant virus. Compared to cells infected with wild-type recombinant virus (rIBV), slightly smaller sized plaques were observed in cells infected with the Q74A mutant virus (Figure 5a). Analysis of the growth curves of wild-type and the mutant virus demonstrated that the mutant virus grew slightly more slowly than the wild type recombinant virus. When 10-fold more mutant virus was used, 2- to 10-fold more mutant virus was produced at 8–24 h post-infection (Figure 5a). The titers of the mutant virus dropped more rapidly than wild-type virus at 36 and 48 h post-infection (Figure 5a).

Further characterization of the mutant virus was subsequently carried out by analysis of the genomic and subgenomic RNA, and the expression of viral structural proteins S and N. Northern blot analysis of cells infected with the 5th passage of rIBV and Q74A mutant viruses showed that similar amounts of genomic and subgenomic RNAs were detected (Figure 5b). Quantitative comparison of the subgenomic RNA synthesis by densitometry analysis of the bands shown in Figure 5b demonstrated that in cells infected with Q74A mutant virus (p5), the relative amounts of mRNA 2, 3, 4, 5 and 6 are ∼98, 101, 121, 104, 106 and 101%, respectively, of those in cells infected with rIBV. However, Western blot analysis of cells infected with rIBV and Q74A mutant virus showed that more S and N proteins were detected in cells infected with wild-type virus (Figure 5c, lanes 1–14).

To confirm if the Q74A mutation is genetically stable during passage of the mutant virus in cells, viral RNA was isolated from cells infected with passage 2 and 5 virus and amplified by RT–PCR. Sequencing analysis showed that the two mutated nucleotides (CA-GC) remained stable (Figure 5d). Furthermore, no compensatory mutation was found in other region of the N protein, confirming that Q74A mutant virus is stable in cell culture system.

DISCUSSION

The intrinsic RNA-binding activity of the nucleocapsid protein of RNA viruses is essential for packaging the RNA genome into a RNP complex. In previous structural studies of the N protein from two coronaviruses, the N-terminal domain was shown to be resistant to proteolysis and to fold independently (14,15). It serves as a functional unit critical

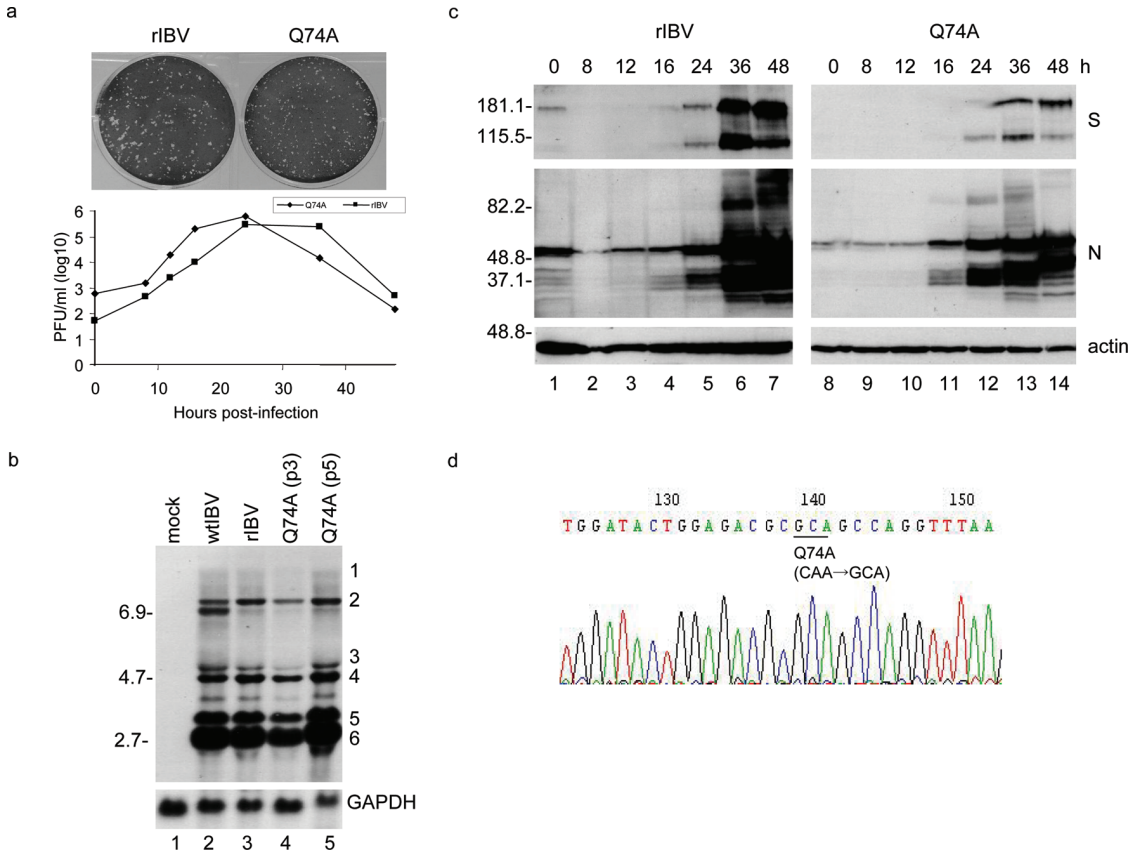

Figure 5. Analysis of the growth properties of wild type and Q74A mutant virus. (**a**) Plague sizes and one-step growth curves of rIBV and Q74A mutant viruses. Monolayers of Vero cells on a 6-well plate were infected with 100 µl of 1-, 10- and 100-fold diluted virus stock and cultured in the presence of 0.5% carboxymethy cellulose at 37°C for 3 days. The cells were fixed and stained with 0.1% toluidine. To determine the one-step growth curves of wild-type recombinant IBV and Q74A mutant viruses, Vero cells were infected with wild type and Q74A mutant viruses, and harvested at 0, 8, 12, 16, 24, 36 and 48 h post-inoculation, respectively, and viral stocks were prepared by freezing/thawing of the cells three times. The viral titer was determined by plaque assay on Vero cells. (**b**) Northern blot analysis of the genomic and subgenomic RNAs in cells infected with wild type and Q74A mutant viruses. Ten micrograms of total RNA extracted from Vero cells infected with rIBV and Q74A mutant viruses (passage 3 and 5), respectively, were separated on 1% agarose gel and transferred to a Hybond N+ membrane. Viral RNAs were probed with a Dig-labelled DNA probe corresponding to the 3-end 680 nt of the IBV genome. Total RNA extracted from mock-infected cells was included as negative control. Numbers on the left indicate nucleotides in kilobase and numbers on the right indicate the genomic and subgenomic RNA species of IBV. (**c**) Western blot analysis of viral protein expression in cells infected with wild type and Q74A mutant viruses. Vero cells infected with wild type recombinant IBV (lanes 1–7) and Q74A mutant virus (lanes 8–14) were harvested at 0, 8, 16, 24, 36 and 48 h postinfection, respectively, lysates prepared and separated on SDS-10% polyacrylamide gel. The expression of S and N proteins was analyzed by western blot with polyclonal anti-S and anti-N antibodies, respectively. The same membrane was also probed with anti-actin antibody as a loading control. Numbers on the left indicate molecular masses in kilodaltons. (**d**) Nucleotide sequencing of Q74A mutant virus. Total RNA was prepared from Vero cells infected with passage 5 of the Q74A mutant virus and the region covering the Q74A (CAA→GCA) mutation was amplified by RT–PCR and sequenced by automated nucleotide sequencing. A 30 nt region flanking the CAA→GCA mutation is shown.

for the RNA-binding activity of the protein (14). Based on these structural results, the importance of individual residues in RNA-binding was studied. Here we showed that Tyr-94 in strand β3 binds to the RNA ligand. The positively charged Arg-76 is located in the immediate vicinity of this aromatic residue at the base of the extended flexible hairpin loop, within a cluster of positively charged residues. Thus, the interaction between Arg-76 and Tyr-94 with RNA appears to play a crucial role during virus assembly. Since no single mutation could totally disrupt nucleic acid binding, other aromatic or basic residues located close to these two residues at the surface of the N-terminal domain are likely to contribute to RNA binding by creating a broad surface that makes contact with the genomic RNA. Using a simple docking experiment, the RNA binding site seems capable of accommodating up to 8–9 nt bases per N-terminal domain monomer. The Arg-76 and Tyr-94 residues identified in this paper thus provide more precise mapping of the location of the RNA-binding site on the IBV N protein.

It is interesting to note that the main effect of these mutations on viral life cycles was on subgenomic RNA transcription. However, the replication of negative strand RNA is nearly normal in cells transfected with these mutant constructs. Is it possible that these mutations will directly impair an undefined function of the N protein that is essential for subgenomic RNA transcription? Two possibilities were considered. First, the coronavirus N protein is directly involved in subgenomic RNA replication. The MHV N protein was shown to be involved in coronaviral RNA transcription (4), but the underlying mechanism was unrevealed. Alternatively, the observed phenotypic changes in subgenomic RNA replication caused by these mutations may be due to a generally reduced viral replication rate and infectivity. In fact, several other mutations that affect viral

replication cycles were found to have defects in subgenomic RNA transcription (25–27). In this study, we have repeatedly observed that detection of subgenomic RNA replication in cells infected with low multiplicity of infectivity of IBV by RT–PCR is technically challenging. Positive results can be obtained only if transcription of subgenomic RNA reaches a certainly high level. When *in vitro* transcripts containing Y92A mutation were introduced into cells, one or two CPEs were always observable in a monolayer, suggesting the presence of low level viral RNA replication. However, detection of subgenomic RNA was consistently unsuccessful in cells electroporated with this mutant transcript.

Coronavirus N protein is a multi-functional protein. The protein was shown to be able to facilitate rescue of infectious virus from several coronaviruses by co-electroporation of the *in vitro* synthesized N transcripts with the *in vitro* synthesized full-length transcripts (17,28–30). The mechanism for this trans-effect is currently unknown. In this study, the R76A and Y94A mutant N constructs could promote rescue of the infectivity of the *in vitro* synthesized IBV RNA as efficiently as did the wild type N transcripts, suggesting that the reduced RNA-binding activity in the N-terminal domain does not affect this function of the N protein. It also suggests that RNA-binding *per se* is not a major factor that contributes to this important function of the N protein.

Are the reduced replication rate and loss of the infectivity of the mutant transcripts caused by the effects of these mutations on the functionality of the N protein or on the integrity of the genomic RNA itself? The nice correlation between the reduced RNA-binding activity of the N-terminal domain from the mutant proteins and the loss of infectivity of the corresponding mutant transcripts would support that the failure to rescue the viral infectivity is due to the specific mutation that impairs the functionality of the N protein. In addition, in the cases of Q74A and R76A mutants, the two mutations are physically adjacent in the genome, but yet totally different phenotypes were observed from the two mutants. It would support that the failure to recover infectious viruses from the mutant transcripts is due to the fact that these mutations reduce/impair the RNA-binding activity of the N protein. If this were the case, however, supplement of wild type N protein in trans would rescue the infectivity of the mutant transcripts. Two experiments were conducted to test this possibility, but none of them were successful. First, co-transfection of the mutant full-length transcripts with wild-type N transcripts into cells failed to recover infectious virus from the mutant transcripts, suggesting that wild-type N RNA cannot rescue the infectivity of the mutant transcripts in trans. The second experiment was to introduce the mutant full-length transcripts into cells stably expressing a six-His-tagged IBV N protein. Once again, it failed to rescue the infectivity of the mutant transcripts. At present, it is unclear if this was due to the low-level expression of N protein in these cells or because the N protein expressed in these cells contains the six-His tag at its C-terminus.

Residues Arg-76 and Tyr-94 in the IBV N protein are well conserved across the whole family of coronaviruses. Structurally, these two residues may correspond to the Arg-94 and Tyr-112 residues in the SARS-CoV N protein (see Figure 1a). It is possible that introduction of similar alanine mutations into the SARS-CoV genome may yield a similar phenotype observed here for IBV. The fact that substitutions of Gln-74, Arg-76, Tyr-72 and Tyr-94 in the IBV genome by using an infectious clone of IBV showed nice correlation between the RNA-binding activity and viral infectivity would suggest potential strategies for intervention of viral replication. A compound which could specifically bind to this region of the N protein might interfere with RNA binding and thus inhibit virus replication. In addition, this study points to a number of residues on the surface of the N protein that contribute less to the RNA-binding activity and viral infectivity for the design of attenuated coronavirus strains as potential vaccine candidates.

ACKNOWLEDGEMENTS

This work was supported by the Agency for Science, Technology and Research, Singapore, and by a grant from the Biomedical Research Council (BMRC 03/1/22/17/220), Agency for Science, Technology and Research, Singapore. Funding to pay the Open Access publication charges for this article was provided by the Agency for Science, Technology and Research, Singapore.

Conflict of interest statement. None declared.

REFERENCES

1. Lai,M.M. and Cavanagh,D. (1997) The molecular biology of coronaviruses. *Adv. Virus Res.*, **48**, 1–100.
2. De Hann,C.A. and Rottier,P.J. (2005) Molecular interactions in the assembly of coronaviruses. *Adv. Virus Res.*, **64**, 165–230.
3. Almazan,F., Galan,C. and Enjuanes,L. (2004) The nucleoprotein is required for efficient coronavirus genome replication. *J. Virol.*, **78**, 12683–12688.
4. Baric,R.S., Nelson,G.W., Fleming,J.O., Deans,R.J., Keck,J.G., Casteel,N. and Stohlman,S.A. (1988) Interactions between coronavirus nucleocapsid protein and viral RNAs: implications for viral transcription. *J. Virol.*, **62**, 4280–4287.
5. Chang,C.K., Sue,S.C., Yu,T.H., Hsieh,C.M., Tsai,C.K., Chiang,Y.C., Lee,S.J., Hsiao,H.H., Wu,W.J., Chang,C.F. *et al.* (2005) The dimer interface of the SARS coronavirus nucleocapsid protein adapts a porcine respiratory and reproductive syndrome virus-like structure. *FEBS Lett.*, **579**, 5663–5668.
6. Chang,R.Y. and Brian,D.A. (1996) *cis*-Requirement for N-specific protein sequence in bovine coronavirus defective interfering RNA replication. *J. Virol.*, **70**, 2201–2207.
7. Eleouet,J.F., Slee,E.A., Saurini,F., Castagne,N., Poncet,D., Garrido,C., Solary,E. and Martin,S.J. (2000) The viral nucleocapsid protein of transmissible gastroenteritis coronavirus (TGEV) is cleaved by caspase-6 and -7 during TGEV-induced apoptosis. *J. Virol.*, **74**, 3975–3983.
8. He,R., Leeson,A., Andonov,A., Li,Y., Bastien,N., Cao,J., Osiowy,C., Dobie,F., Cutts,T., Ballantine,M. and Li,X. (2003) Activation of AP-1 signal transduction pathway by SARS coronavirus nucleocapsid protein. *Biochem. Biophys. Res. Commun.*, **311**, 870–876.
9. He,R., Dobie,F., Ballantine,M., Leeson,A., Li,Y., Bastien,N., Cutts,T., Andonov,A., Cao,J., Booth,T.F. *et al.* (2004) Analysis of multimerization of the SARS coronavirus nucleocapsid protein. *Biochem. Biophys. Res. Commun.*, **316**, 476–483.
10. Li,F.Q., Xiao,H., Tam,J.P. and Liu,D.X. (2005) Sumoylation of the nucleocapsid protein of severe acute respiratory syndrome associated coronavirus. *FEBS Lett.*, **579**, 2387–2396.
11. Schelle,B., Karl,N., Ludewig,B., Siddell,S.G. and Thiel,V. (2005) Selective replication of coronavirus genomes that express nucleocapsid protein. *J. Virol.*, **79**, 6620–6630.
12. Surjit,M., Liu,B., Jameel,S., Chow,V.T. and Lal,S.K. (2004) The SARS coronavirus nucleocapsid protein induces actin reorganization and apoptosis in COS-1 cells in the absence of growth factors. *Biochem. J.*, **383**, 13–18.

13. Surjit,M., Liu,B., Chow,V.T K. and Lal,S.K. (2006) The nucleocapsid protein of severe acute respiratory syndrome-coronavirus inhibits the activity of cyclin-dependent kinase complex and blocks S phase progression in mammalian cells. *J. Biol. Chem.*, **281**, 10669–10681.
14. Fan,H., Ooi,A., Tan,Y.W., Wang,S., Fang,S., Liu,D.X. and Lescar,J. (2005) Crystal structure of the N-terminal domain from the nucleocapsid protein of coronavirus infectious bronchitis virus and multimerization properties. *Structure*, **13**, 1859–1868.
15. Hall,K. (2002) RNA-protein interactions. *Curr. Opinion Struct. Biol.*, **12**, 283–288.
16. Fang,S.G., Chen,B., Tay,F.P.L. and Liu,D.X. (2006) An arginine-to-proline mutation in a domain with undefined function within the RNA helicase protein (NSP13) is lethal to the coronavirus infectious bronchitis virus in cultured cells. *Virology*, in press.
17. Youn,S., Leibowitz,J.L. and Collisson,E.W. (2005) *In vitro* assembled, recombinant infectious bronchitis viruses demonstrate that the 5a open reading frame is not essential for replication. *Virology*, **332**, 206–215.
18. Huang,Q., Yu,L., Petros,A.M., Gunasekera,A., Liu,Z., Xu,N., Hajduk,P., Mack,J., Fesik,S.W. and Olejniczak,E.T. (2004) Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. *Biochemistry*, **43**, 6059–6063.
19. Antson,A.A., Burns,J.E., Moroz,O.V., Scott,D.J., Sanders,C.M., Bronstein,I.B., Dodson,G.G., Wilson,K.S. and Maitland,N.J. (2000) Structure of the intact transactivation domain of the human papillomavirus E2 protein. *Nature*, **403**, 805–809.
20. Antson,A.A. (2000) Single stranded RNA binding proteins. *Curr. Opin. Struct. Biol.*, **10**, 87–94.
21. Landsman,D. (1992) RNP-1, an RNA-binding motif is conserved in the DNA-binding cold shock domain. *Nucleic Acids Res.*, **20**, 2861–2864.
22. Nagai,K. (1996) RNA–protein complexes. *Curr. Opin. Struct. Biol.*, **6**, 53–61.
23. Oubridge,C., Ito,N., Evans,P.R., Teo,C.H. and Nagai,K. (1994) Crystal structure at 1.92 A resolution of the RNA binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature*, **372**, 432–438.
24. Valegard,K., Murray,J.B., Stockley,P.G., Stonejhouse,N.J. and Liljas,L. (1994) Crystal structure of an RNA bacteriophage coat protein operator complex. *Nature*, **372**, 623–626.
25. Seybert,A., Posthuma,C.C., van Dinten,L.C., Snijder,E.J., Gorbalenya,A.E. and Ziebuhr,J. (2005) A complex zinc finger controls the enzymatic activities of nidovirus helicases. *J. Virol.*, **79**, 696–704.
26. van Dinten,L.C., den Boon,J.A., Wassenaar,A.L., Spaan,W.J. and Snijder,E.J. (1997) An infectious arterivirus cDNA clone: identification of a replicase point mutation that abolishes discontinuous mRNA transcription. *Proc. Natl Acad. Sci. USA*, **94**, 991–996.
27. van Dinten,L., van Tol,C.H., Gorbalenya,A.E. and Snijder,E.J. (2000) The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. *J. Virol.*, **74**, 5213–5223.
28. Casais,R., Thiel,V., Siddell,S.G., Cavanagh,D. and Britton,P. (2001) Reverse genetics system for the avian coronavirus infectious bronchitis virus. *J. Virol.*, **75**, 12359–12369.
29. Yount,B., Curtis,K.M. and Baric,R.S. (2000) Strategy for systematic assembly of large RNA and DNA genomes: transmissible gastroenteritis virus model. *J. Virol.*, **74**, 10600–10611.
30. Yount,B., Denison,M.R., Weiss,S.R. and Baric,R.S. (2002) Systematic assembly of a full-length infectious cDNA of mouse hepatitis virus strain A59. *J. Virol.*, **76**, 11065–11078.