

Phosphorylation of the arginine/serine dipeptide-rich motif of the severe acute respiratory syndrome coronavirus nucleocapsid protein modulates its multimerization, translation inhibitory activity and cellular localization

Tsui-Yi Peng^{1,2}, Kuan-Rong Lee² and Woan-Yuh Tarn¹

1 Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan

2 Institute of Molecular Medicine, National Tsing Hua University, Hsin-Chu, Taiwan

Keywords

coronavirus; nucleocapsid protein; phosphorylation; RS domain; stress granules

Correspondence

W.-Y. Tarn, Institute of Biomedical Sciences, Academia Sinica, 128 Academy Road, Section 2, Nankang, Taipei 11529, Taiwan

Fax: +886 2 2782 9142

Tel: +886 2 2652 3052

E-mail: wtarn@ibms.sinica.edu.tw

(Received 15 April 2008, revised 17 June 2008, accepted 19 June 2008)

doi:10.1111/j.1742-4658.2008.06564.x

Coronavirus nucleocapsid protein is abundant in infected cells and participates in viral RNA replication and transcription. The central domain of the nucleocapsid protein contains several arginine/serine (RS) dipeptides, the biological significance of which has not been well investigated. In the present study, we demonstrate that the severe acute respiratory syndrome coronavirus nucleocapsid protein is phosphorylated primarily within the RS-rich region in cells and by SR protein kinase 1 *in vitro*. The nucleocapsid protein could suppress translation and its RS motif is essential for such an activity. Moreover, phosphorylation of the RS motif could modulate the translation inhibitory activity of the nucleocapsid protein. We further found that RS motif phosphorylation did not significantly affect RNA binding of the nucleocapsid protein but impaired its multimerization ability. We observed that the nucleocapsid protein could translocate to cytoplasmic stress granules in response to cellular stress. Deletion or mutations of the RS motif enhanced stress granule localization of the nucleocapsid protein, whereas overexpression of SR protein kinase 1 inhibited nucleocapsid protein localization to stress granules. The nucleocapsid protein lacking the RS motif formed high-order RNP complexes, which may also account for its enhanced stress granule localization. Taken together, phosphorylation of the severe acute respiratory syndrome-CoV nucleocapsid protein modulates its activity in translation control and also interferes with its oligomerization and aggregation in stress granules.

An outbreak of severe acute respiratory syndrome (SARS) occurred primarily in Asia in 2003. The causative agent of SARS is a coronavirus-related virus [1,2]. The genome sequence of the SARS virus is only moderately similar to that of other species of coronaviruses [3] and, thus, the SARS virus represents a distinct member of the coronaviruses.

Coronaviruses are enveloped viruses with positive-stranded, capped and polyadenylated RNA genomes

of approximately 30 kb [3,4]. The 5' two-thirds of the genome encode the replicase-transcription complex [3,4]. During viral replication, a nested set of subgenomic mRNAs encoding structural proteins including the nucleocapsid (N) is synthesized via a discontinuous transcription mechanism [5,6]. The N protein is the most abundant viral protein produced throughout viral infection and may exert several distinct functions [7]. The N protein is primarily involved in encapsidation

Abbreviations

GST, glutathione S-transferase; HBV, hepatitis B virus; MHV, mouse hepatitis virus; N, nucleocapsid; NΔRS, RS-deleted mutant; PABP1, poly(A)-binding protein 1; RS, arginine/serine; SARS, severe acute respiratory syndrome; SG, stress granule; SRPK, SR protein kinase.

and packaging of viral genomic RNA [8–10]. Moreover, it binds to the 5' and/or 3' end of the genomic RNA [11–13] and may participate in viral genome replication and subgenomic mRNA transcription [14,15]. However, other evidence suggests that the N protein is dispensable for these processes [16]. In addition, the mouse hepatitis virus (MHV) N protein stimulates translation of a reporter mRNA containing an intact MHV 5'-untranslated region, probably by binding to a tandem repeat of UCYAA in the leader sequence [17]. By contrast, recent evidence indicates that the SARS-CoV N protein interferes with translation through its interaction with cellular translation elongation factor 1 α [18]. In addition, the SARS-CoV N protein can inhibit the activity of cellular cyclin-dependent kinases and thereby perturb S phase progression of virus-infected cells [19,20]. Therefore, coronavirus N proteins may affect various cellular functions.

Within various coronaviruses, the N protein varies from 377 to 455 amino acid residues in length. Although the sequence conservation between N proteins is relatively low [3,21], they are likely to adopt a common secondary structure essentially consisting of two functional domains. The N-terminal domain interacts with RNA through a structural module rich in positively charged residues [22–24]. The RNA binding capacity of the N protein is critical for viral infectivity [24]. The C-terminal domain folds into a β -sheet platform engaged in homodimerization [24,25] and may also confer the RNA binding activity [10]. Moreover, RNA binding may promote multimerization of the N protein, implicating a nucleocapsid formation mechanism [10,23,26].

Between the two functional domains is a structurally flexible segment containing several arginine/serine (RS)-dipeptides. This RS-rich motif is characteristic of cellular precursor mRNA (pre-mRNA) splicing factors, termed SR proteins [27]. The RS domain is dynamically phosphorylated by several SR protein specific kinases, such as those of the SR protein kinase (SRPK) and Clk families [28]. Phosphorylation of the RS domain modulates the activity, protein–protein interactions and subcellular localization of SR proteins [29]. Coronavirus N proteins are phosphorylated in host cells and in virions [25,30,31] and it has been reported that phosphorylation affects the RNA binding specificity and nucleocytoplasmic shuttling of the N proteins [25,32]. Indeed phosphorylation can occur within the RS motif of coronavirus N proteins [19,33] and this motif may play a role in C-terminal domain dimerization [26]. Nevertheless, whether phosphorylation of the RS motif can modulate the functions of N proteins remains to be examined in detail.

Coronavirus N proteins localize to both the cytoplasm and the nucleolus in virus-infected cells [34–36] and can shuttle between the nucleus and the cytoplasm [37]. Nucleolar localization of N protein requires regions in the protein that are rich in arginine residues and is likely cell cycle-dependent [20,35,36]. The avian infectious bronchitis virus N protein indeed interacts with and colocalizes with the nucleolar proteins nucleolin and fibrillarin [38,39]. However, the ability of nucleolar localization varies between N proteins of different coronaviruses [36]. The SARS-CoV N protein is poorly localized to the nucleolus [36]. In the present study, we found that the SARS-CoV N protein appeared in cytoplasmic stress granules (SGs). When eukaryotic cells encounter environmental stress, mRNA metabolism is reprogrammed to adapt to stress-induced damage. Translationally stalled mRNAs together with a number of translation initiation factors and RNA-binding proteins are deposited into SGs [40]. Formation of SGs can also be induced by overexpression of the prion-like RNA binding protein TIA-1 [41]. Upon stress induction, TIA-1 forms aggregates in SGs and may play a role in translation inhibition [41].

In the present study, we examined phosphorylation of the SARS-CoV N protein. Our data provide evidence that phosphorylation of the N protein primarily occurs within its RS-rich motif and may affect its oligomerization, translation inhibitory activity and subcellular localization.

Results

Phosphorylation of the RS-rich motif of the SARS-CoV N protein

Coronavirus N proteins are phosphoproteins [30,32]. The N protein of all coronaviruses, including the SARS-CoV, contains an RS-rich motif (Fig. 1) that likely provides potential phosphorylation sites for multiple cellular kinases [32]. We predicted that N proteins, due to their similarity with cellular SR proteins in the RS-rich motif, may serve as a substrate of SR protein specific kinases. To study phosphorylation of SARS-CoV N protein RS domain, we overexpressed FLAG-tagged N protein and the RS-deleted mutant (N Δ RS) in HEK293 cells. Transfected cells were incubated with [³²P]orthophosphate for labeling. Anti-FLAG immunoprecipitation of the full-length N protein from the cell lysate revealed a ³²P-labeled band at approximately 52 kDa (Fig. 2A, lane 1), similar to previous reports [42], indicating that the SARS-CoV N protein was phosphorylated *in vivo*. However,

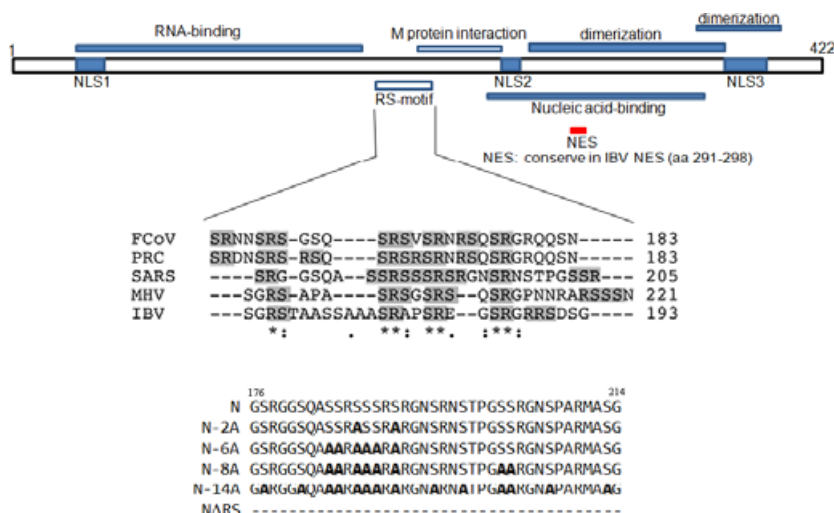


Fig. 1. Schematic representation of the domain structure of the SARS-CoV N protein and RS-rich motif sequence alignment of the coronavirus N proteins. Functional motifs and domains are depicted as previously described. Alignment of the RS-rich motifs was performed using CLUSTALW of the European Molecular Biology Laboratory's European Bioinformatics Institute (Hinxton, UK). The arginine and serine residues of the RS motif are highlighted in gray. The accession number of the indicated coronavirus N proteins is: feline coronavirus (FCoV; BAC01161), porcine respiratory coronavirus (PRC; CAA80841), mouse MHV (P03416), human SARS coronavirus (AAP37024) and avian infectious bronchitis virus (AAA46214). Bottom: serine-to-alanine mutants of the SARS-CoV N protein that were used in the study.

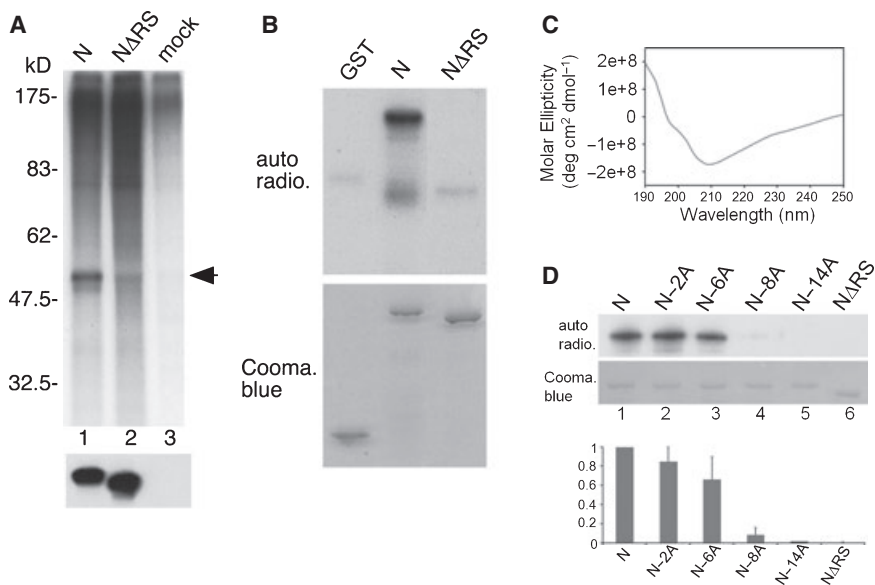


Fig. 2. Phosphorylation of the SARS-CoV N protein within the RS-rich motif. (A) HEK293 cells that transiently expressed FLAG-tagged full-length N protein (lane 1) or NΔRS (lane 2) were fed [³²P]orthophosphate. Immunoprecipitation of FLAG-tagged proteins was performed using anti-FLAG; full-length N protein is indicated by the arrow. Lane 3 shows mock-transfection. The lower panel shows anti-SARS-CoV N immunoblotting. (B) Recombinant GST and GST-N (wild-type and ΔRS) proteins (lower: Coomassie blue staining) were phosphorylated by purified SRPK1 in reactions containing [³²P]ATP (upper: autoradiography). (C) The CD spectrum of purified GST-NΔRS was monitored in the range 190–250 nm. The γ -coordinate shows $\Delta\epsilon$. (D) Wild-type and mutant GST-N proteins and GST-NΔRS were *in vitro* phosphorylated by SRPK1 as in (B) (upper: autoradiography; lower: Coomassie blue staining). Values below the gel represent relative phosphorylation levels; the data were obtained from two to three independent experiments.

inefficient labeling of NΔRS (Fig. 2A, lane 2) suggested that the RS motif is the major phosphorylation site for the SARS-CoV N protein in cells. We next examined

whether the SR protein kinase SRPK1 can phosphorylate the N protein within its RS-rich motif. Full-length N and NΔRS were each fused to glutathione

S-transferase (GST) and subjected to *in vitro* phosphorylation using purified SRPK1. Figure 2B shows that only full-length N protein was phosphorylated by SRPK1. However, neither another SR protein kinase, Clk1, nor protein kinase A was able to phosphorylate the N protein *in vitro* (see supplementary Fig. S1). To avoid the possibility that inefficient phosphorylation of NΔRS resulted from its improper folding, recombinant GST-NΔRS protein was analyzed by CD spectroscopy. The CD spectrum of purified NΔRS suggested that this truncated protein still adopted an ordered conformation (Fig. 2C) and was also similar to that of the full-length N protein (data not shown). Taken together, these data suggest that the RS-rich region of the SARS-CoV N protein is possibly phosphorylated by SRPK1.

Phosphorylation of multiple serines within the RS motif

The RS motif of the SARS-CoV N protein is divergent from that of canonical SR proteins and contains fewer RS dipeptides. To determine which serines might be the major phosphorylation sites, we made a series of serine to alanine substitution mutants and investigated their phosphorylation using SRPK1. As shown in Fig. 2D, increasing the number of alanine substitution gradually decreased the phosphorylation level of the N protein. This result indicated that multiple serines are phosphorylated, and was consistent with the observations made for other SR proteins [43]. However, because N-8A was much poorly phosphorylated compared to N-6A (Fig. 2D, lanes 3 and 4), S²⁰³S²⁰⁴ might serve as the primary site of SRPK1-mediated phosphorylation.

RS motif phosphorylation modulates the activity of the N protein in translation suppression

Previous reports indicate that MHV infection induces host translational shut-off [44]. Coronavirus N proteins are primarily distributed throughout the cytoplasm, with a higher concentration within nucleoli [21,35,39] and, thus, have the potential to interfere with ribosome biogenesis or translation in host cells. To test whether the SARS-CoV N protein plays a role in translation control and whether phosphorylation modulates its activity, we performed an *in vitro* translation assay. Using a firefly luciferase reporter, we titrated recombinant GST-N or GST-NΔRS protein into the reticulocyte lysate. Both the protein level and activity of the luciferase were measured, which may directly reflect the translation activity because luciferase mRNA levels were similar between treatments

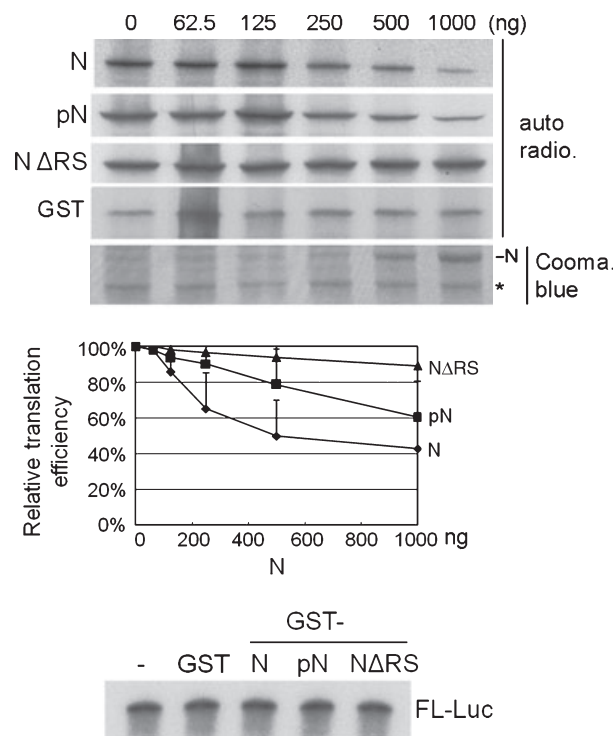


Fig. 3. The translation inhibition activity of the SARS-CoV N protein is modulated by phosphorylation. Translation of an *in vitro* transcribed firefly luciferase mRNA was performed in reticulocyte lysate in the presence of different amounts of nonphosphorylated (N) or phosphorylated (pN) GST-N, GST-NΔRS or GST protein. Representative autoradiograms show the resulting firefly luciferase protein; Coomassie blue staining shows titrated N protein (N) and a reticulocyte lysate protein (*) as loading control. The graph shows relative translation efficiency obtained by comparison with the reaction without N protein; data are the mean \pm SD values are from three independent experiments. *In vitro* translation reactions contained 1 μ g of indicated recombinant protein as well as ³²P-labeled luciferase reporter mRNA as tracer. After incubation, the level and the integrity of radioisotope labeled RNA were examined on a denaturing 4% polyacrylamide gel.

(Fig. 3, bottom). The GST-N protein suppressed luciferase translation in a dose-dependent manner but this translation suppressive effect was attenuated upon phosphorylation by SRPK1 (Fig. 3). GST-NΔRS or GST control had no significant effects on translation of the luciferase mRNA. Therefore, the SARS-CoV N protein might possess translation suppression activity that requires its RS motif and, thus, could be modulated by phosphorylation.

Effect of RS motif phosphorylation on oligomerization of the N protein

To better understand the effect of RS motif phosphorylation on the biological function of the N protein, we

next examined the RNA binding activity of the SARS-CoV N protein. His-tagged N protein was used to avoid dimerization caused by GST. Recombinant His-N protein was subjected to *in vitro* phosphoryla-

tion by SRPK1. Figure 4A shows that His-N was ^{32}P -labeled after phosphorylation and protein phosphatase treatment removed ^{32}P phosphates (Fig. 4A, lanes 2 and 3). Moreover, phosphorylation resulted in

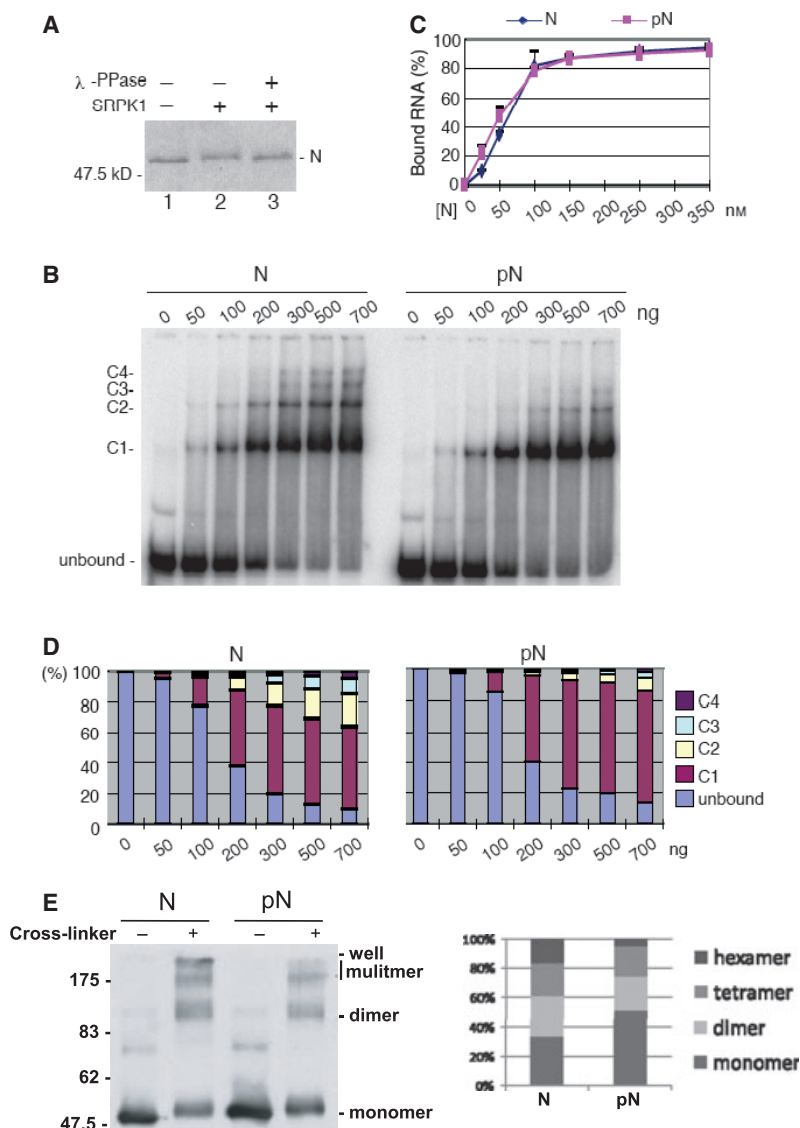


Fig. 4. Effects of RS motif phosphorylation on the RNA binding activity and multimerization of the SARS-CoV N protein. (A) Recombinant His-tagged N protein was phosphorylated by SRPK1 (lanes 2 and 3) or mock-phosphorylated (lane 1) in the presence (upper panel) or absence (lower panel) of $[\gamma\text{-}^{32}\text{P}]\text{ATP}$. Phosphorylated N protein was subsequently treated with λ -protein phosphatase (lane 3) or mock-treated (lane 2). (B) An increasing amount of mock- (N) or SRPK1- (pN) phosphorylated N protein was incubated with an approximately 110 nucleotide ^{32}P -labeled RNA probe, and binding was analyzed by electrophoresis on a nondenaturing polyacrylamide gel. C1, C2, C3 and C4 denote RNP complexes that may contain two, three, four and six copies of the N protein, respectively. (C) The RNA binding efficiency of N protein is represented as a percentage of bound RNA (i.e. the percentage of bound RNA = 100% - percentage of free probe). The apparent K_d was calculated as $\frac{1}{2} V_{\text{max}}$. Each SD was obtained from four independent experiments. (D) The relative abundance (percentage) of unbound RNA and distinct RNA/N protein complexes was calculated as $100 \times (\text{arbitrary unit of each band in individual lane divided by the unit of the unbound RNA detected in the absence of the N protein})$. The results are representative of four independent experiments. (E) Chemical crosslinking of nonphosphorylated (N) and phosphorylated (pN) N proteins (lanes 2 and 4). Lanes 1 and 3 are the mock reactions without crosslinker. The right-hand panel shows the relative abundance (percentage) of monomer and crosslinked forms. Percentage was calculated as $100 \times (\text{arbitrary unit of each form divided by the sum units of all forms})$.

slight mobility shift of the N protein (lane 2), possibly indicating its stoichiometric phosphorylation. Next, electrophoretic mobility shift assay showed that non-phosphorylated His-N bound an approximately 110 nucleotide RNA probe with an apparent dissociation constant of 52.9 nM, comparable to that reported previously [45,46]. Moreover, His-N appeared to form oligomers in a concentration-dependent manner (Fig. 4B,C). The phosphorylated N protein also bound this RNA probe and its dissociation constant was determined to be 66.7 nM, which was similar to that of nonphosphorylated N protein (Fig. 4C). However, formation of high-order N protein RNP complexes appeared to be impaired when the N protein was phosphorylated (Fig. 4D). Chemical crosslinking of the N protein confirmed that phosphorylated N was less capable of forming oligomers than the nonphosphorylated one (Fig. 4E). Therefore, it is likely that phosphorylation of the RS motif interferes with oligomerization of the N protein.

Translocation of the N protein to stress granules is modulated by RS motif phosphorylation

Because the RS domain can modulate subcellular localization of cellular SR proteins [45], we next examined whether the RS motif of SARS-CoV N protein has this activity. When the FLAG-tagged N Δ RS fusion protein was transiently expressed in HeLa cells, approximately 5% of transfected cells showed a punctate staining pattern (Fig. 5A). This granule-like localization pattern was also observed with the full-length N protein, albeit rarely (approximately 1% of the transfected cells). Although this granule staining pattern was observed only in a few percent of N or N Δ RS-protein expressing cells under normal cell conditions, it was greatly enhanced upon arsenite treatment (> 95% transfected cells; Fig. 5B). Indeed both N and N Δ RS colocalized with endogenous poly(A)-binding protein 1 (PABP1) and transiently expressed TIA-1 (Fig. 5B), both of which are SG components [40].

To distinguish whether the RS motif deletion or a lack of phosphorylation enhanced N protein localization in SGs, we examined the cellular localization of two RS motif mutants, N-6A and N-14A. Both mutants showed a higher tendency towards SG localization than the wild-type N (Fig. 5A), suggesting that SG localization of the N protein primarily resulted from its hypophosphorylation. Moreover, the N-terminal (N^{NT}) but not the C-terminal (N^{CT}) part of the N protein appeared to be responsible for granule localization (Fig. 5A). We apparently reasoned that the N-terminal domain contains the RS motif and confers RNA

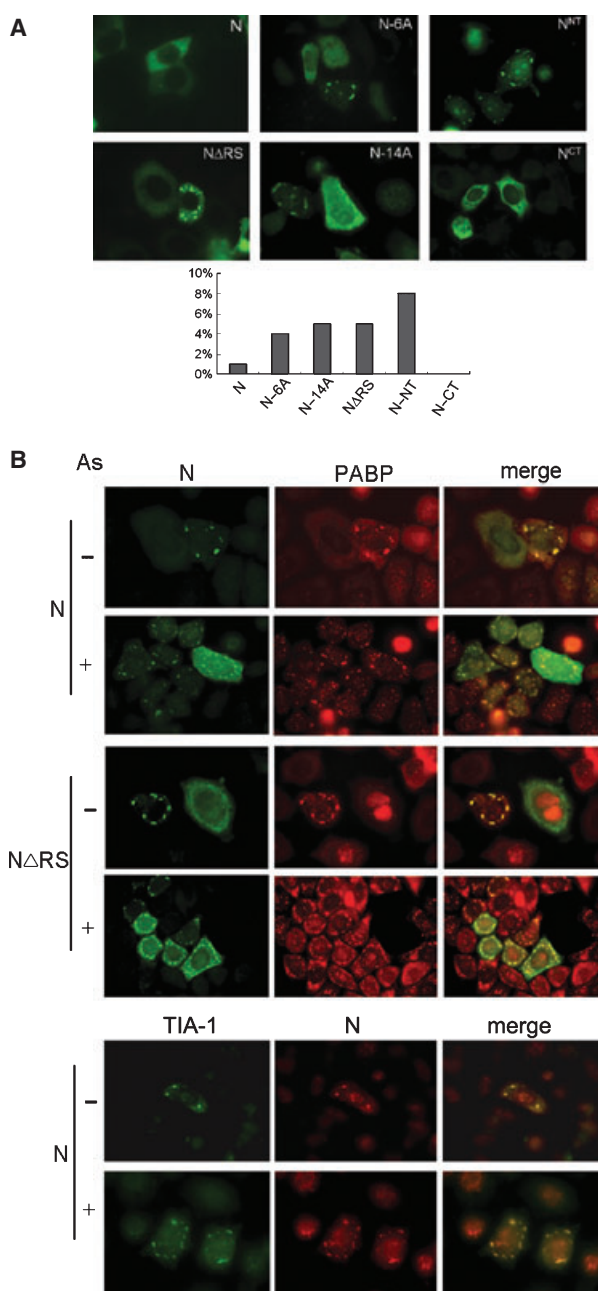


Fig. 5. Translocation of the SARS-CoV N protein to cytoplasmic granules can be induced by cell stress and modulated by phosphorylation. (A) Expression vector encoding FLAG-tagged full-length (N), RS motif-deleted (N Δ RS), two serine-to-alanine mutants (N-6A and N-14A), N-terminal-half (N^{NT}) or C-terminal-half (N^{CT}) N protein was transiently transfected into HeLa cells. Upper panel: representative fluorescence images. Lower panel: percentage of granule-positive cells; approximately 100 transfected cells were counted for each protein. (B) HeLa cells transiently expressing HA-tagged N or N Δ RS or coexpressing HA-N and GFP-TIA-1 were mock treated (-) or treated (+) with 0.5 mM arsenite for 1 h. Double immunofluorescence was performed using anti-HA and anti-PABP. A merged image is shown in the right-hand panel.

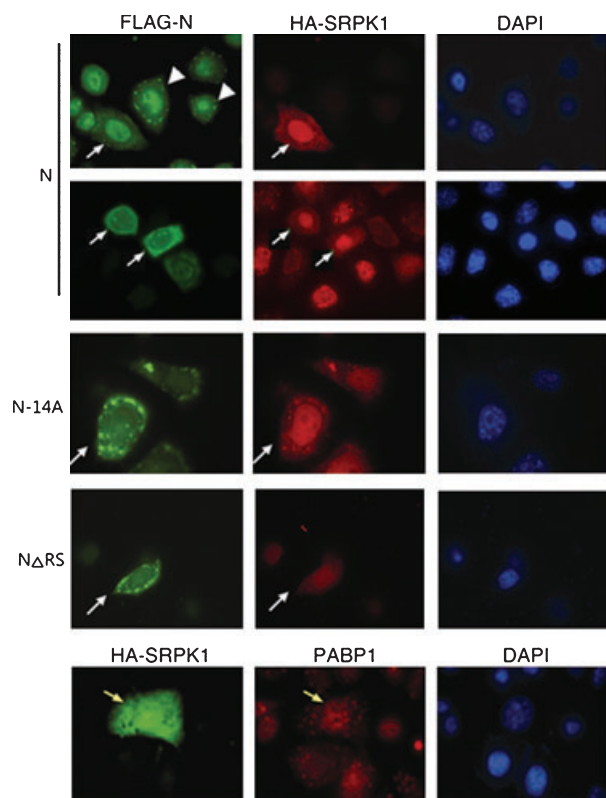


Fig. 6. Overexpression of SRPK1 prevents N protein translocation to stress granules. HeLa cells were transiently cotransfected with vectors encoding FLAG-tagged full-length N, N-14A or N Δ RS and HA-tagged SRPK1, and treated with arsenite as in Fig. 4B. Immunofluorescence using anti-HA and anti-FLAG was performed; two representative images are shown for the N protein. Arrowheads indicate cells that expressed FLAG-N protein alone, and white arrows indicate cells expressing both FLAG-N and HA-SRPK1. Cell nuclei were stained with 4',6'-diamidino-2-phenylindole (DAPI). The lower panel shows double immunofluorescence of HA-SRPK1-overexpressing HeLa cells using anti-HA and anti-PABP. Yellow arrows indicate cells that overexpressed HA-SRPK1.

binding ability [10,22] and is therefore capable of forming granules. Next, we evaluated SRPK1-mediated RS motif phosphorylation in modulating SG localization or retention of the N protein. FLAG-tagged N protein and HA-tagged SRPK1 were transiently coexpressed in HeLa cells. In the presence of overexpressed SRPK1, the N protein was unable to localize to SGs, even after arsenite treatment of the cells (Fig. 6, white arrows). However, SRPK1 overexpression could not disperse RS deletion or alanine substitution mutants to the cytoplasm (N Δ RS and N-14A; Fig. 6). Under this condition, PABP1, similar to N Δ RS and N-14A, showed a granular pattern (Fig. 6, lower panel), indicating that SG assembly is not disturbed by overexpression of SRPK1. Therefore, phosphorylation of the

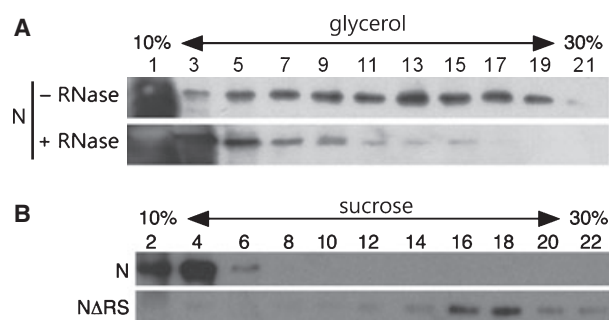


Fig. 7. The SARS-CoV N protein forms RNPs in cell. (A) Mock- or RNase-treated HEK293 cell lysate containing HA-tagged N protein was fractionated on a 10–30% glycerol density gradient. (B) Lysate containing full-length N or N Δ RS protein was fractionated on a 10–30% sucrose density gradient. N protein was detected by immunoblotting with anti-SARS-CoV N serum.

RS motif might diminish N protein oligomerization (Fig. 4) and thereby prevent its aggregation in SGs.

Together, the SARS-CoV N protein could target to SGs, reflecting its role in translation suppression. Moreover, phosphorylation of the RS motif modulates the ability of the N protein to form SGs.

RS motif deletion induces the N protein to form large RNP complexes

The SARS-CoV N protein might regulate translation and could target to SGs; therefore, we evaluated whether it forms RNPs in host cells. Using glycerol gradient sedimentation, we observed that the N protein formed RNPs in cells because it was moved to lighter density fractions after RNase treatment (Fig. 7A). Compared to full-length N, N Δ RS even migrated in heavier fractions of the sucrose density gradient (Fig. 7B). The high-order N Δ RS complexes were also sensitive to RNase (data not shown). Therefore, removal of the RS motif from the N protein induced larger RNP formation, which may account for N Δ RS aggregation in SGs. The above data show that RS motif deletion induced high-order N protein-containing RNP formation. We inferred that this might result from hypophosphorylation of the N Δ RS protein.

Discussion

The RS domain is a characteristic feature of cellular pre-mRNA splicing factors [27,29]. Several viral proteins also contain various numbers of repeated RS dipeptides. The transactivator E2 protein of cutaneous papillomaviruses has a relatively long RS domain, which functions to recruit cellular splicing factors for

cotranscriptional splicing regulation [47]. The core protein of hepatitis B virus (HBV) has an arginine-rich domain at the C-terminus that bears a few RS dipeptides. The HBV core protein can be phosphorylated by SRPK1 and SRPK2 [48]. Similar to the HBV core, coronavirus N proteins contain a short RS-rich motif (Fig. 1) and the SARS-CoV N protein might be phosphorylated by SRPK1 (Fig. 2). We have examined whether the SARS-CoV N protein plays a role in pre-mRNA splicing due to the presence of the RS motif but, so far, we do not have any evidence to support this hypothesis (data not shown). In the present study, we provide evidence that the SARS-CoV N protein could suppress translation at least *in vitro* (Fig. 3). The potential role of the N protein in translation control might correlate with its localization in the cytoplasmic SGs (Fig. 5) and is also in line with recent reports that coronavirus infection could cause translational shut-off in host cells and that the SARS-CoV N protein may execute this activity via its interact with elongation factor 1 α [18,44]. Although coronaviral N protein largely forms helical nucleocapsids with the viral RNA genome during infection [10], how it participates in translation control in host cells and whether it has any substrate specificity or functions under certain cellular conditions remain to be studied in the future.

Previous reports have suggested that the SARS-CoV N protein can act as a substrate of various kinases, such as cyclin-dependent kinases, glycogen synthase kinase, creatine kinase II and mitogen-activated protein kinase [32]. Our data show that multiple serine residues within the RS motif could be *in vitro* phosphorylated by SRPK1 (Fig. 2). The SARS-CoV N protein is primarily distributed in the cytoplasm, coincident with the cellular localization of SRPK1. Coexpression of SRPK1 could modulate cellular localization of the N protein, suggesting that the N protein is a substrate of SRPK1 in cells (Fig. 6). Phosphorylation of the transmissible gastroenteritis virus N protein also occurs on a moderately conserved serine within the RS motif, although which kinases could phosphorylate this serine is as yet unknown [33]. In the present study, we provide evidence that SRPK1-mediated RS motif phosphorylation influences the biochemical and biological activities of the SARS-CoV N protein. First, the potential translation suppression activity of the SARS-CoV N protein might be modulated by phosphorylation (Fig. 3). Moreover, phosphorylation may also impact on its oligomerization, cellular localization and perhaps RNP complex formation (Figs 4–7). The questions of whether SRPK1 phosphorylates the SARS-CoV N protein in cells particularly during viral

infection and where this phosphorylation occurs remain to be answered.

A mammalian two-hybrid assay previously showed that the RS motif is directly involved in N protein self-interaction [42]. However, other evidence indicated that the RS motif interferes with SARS-CoV N protein multimerization but this activity requires its C-terminal domain [26]. Our data show that RS motif phosphorylation partially impaired N protein multimerization (Fig. 4). Perhaps such phosphorylation modulates the balance between N protein self-association and dissociation, which thereby impacts on its cellular functions. Multimerization of the N protein is necessary for nucleocapsid formation and assembly of the viral particles [42]. Thus, whether phosphorylation of the RS motif in virions could modulate N protein function in encapsulation of genomic RNA remains to be investigated. Moreover, we observed that deletion of the RS motif greatly enhanced association of the SARS-CoV N protein with cellular RNPs (Fig. 7). Perhaps RS motif phosphorylation prevents nonspecific binding of the N protein to cellular RNP complexes and thus aids viral genome packaging into capsids; this possibility also remains to be tested.

During infection, coronaviral N protein participates in virus replication that probably occurs at the sites associated with ER-derived membrane tubules and vesicles [49]. Subsequently, viral nucleocapsids are transported to the budding sites in the Golgi region for viral particle formation. Although overexpressed, most coronavirus N proteins are located in the cytoplasm as well as in the nucleolus [34,35]. Nevertheless, the SARS-CoV N protein does not localize substantially to the nucleolus [36,50], as also observed in the present study (Fig. 5). It has been proposed that the signals for nuclear and nucleolar targeting of the SARS-CoV N protein are poorly accessible to the nuclear import machinery due to phosphorylation regulation or conformational constraints [36,50]. Nevertheless, the present study has revealed for the first time that overexpressed SARS-CoV N protein might localize to SGs in HeLa cells (Fig. 5). Such an SG localization pattern was enhanced by deletion or phosphorylation site mutations of the RS motif and was obvious in stress-treated cells (Fig. 5). SGs contain mRNPs whose translation is temporarily blocked [40]. Therefore, the N protein may sequester cellular mRNPs in SGs and inhibit their translation, possibly during viral infection. Nevertheless, the evidence demonstrating that RS motif phosphorylation reduced oligomerization of the N protein and prevented its aggregation in SGs is likely to be in accordance with the attenuation of its translation suppression activity.

Experimental procedures

Plasmid construction

The cDNA encoding the SARS-CoV N protein was kindly provided by K. Peck (Academia Sinica, Taipei, Taiwan). We generated the NΔRS cDNA by ligating the PCR fragments encoding amino acid residues 1–175 and 215–422, respectively. The full-length N and NΔRS cDNAs were each cloned into pcDNA3 (Invitrogen, Carlsbad, CA, USA) in-frame with the FLAG-epitope tag, and also into pCEP4 (Invitrogen) to generate the HA-tagged proteins. All the N protein mutants were generated from the FLAG-N construct using the QuikChange site-directed mutagenesis system (Stratagene, La Jolla, CA, USA); the sequences of these mutants were verified. The cDNAs encoding the N-terminal (residues 1–214) and C-terminal (residues 215–422) domain of the N protein and N-6A and N-14A were individually cloned into pCDNA3 (Invitrogen) in-frame fusion with the pre-engineered FLAG-tag. The pET11D-His-N vector was obtained from T. H. Huang (Institute of Biomedical Sciences, Academia Sinica) and used for production of the His-tagged full-length N protein in *Escherichia coli*. The NΔRS cDNA was appropriately cloned into pET15b (Novagen, Madison, WI, USA) for production of recombinant NΔRS. The wild-type and mutant N and NΔRS cDNAs were subcloned into pGEX-5X (GE Healthcare, Piscataway, NJ, USA) using *EcoRI* and *SalI* sites to generate the GST-fusion proteins. Subsequently, the cDNAs encoding mutant N proteins were cloned into pGEX-5X. The pET15b-FLAG used for *in vitro* transcription of an RNA probe was constructed by insertion of the FLAG-epitope coding sequence into *NheI* and *BamHI*. The *in vitro* translation reporter pFL-SV40 was constructed by replacing the renilla luciferase of pRL-SV40 (Promega, Madison, WI, USA) with the firefly luciferase.

Cell culture, transfection and indirect immunofluorescence

HeLa and HEK293 cells were grown in DMEM (Gibco, Grand Island, NY, USA) supplemented with 10% fetal calf serum and penicillin/streptomycin. Transfection was performed using Lipofectamine 2000 (Invitrogen) as recommended by the manufacturer. For stress treatment, HeLa cells were cultured in the presence of 0.5 mM arsenite for 1 h. The procedure for indirect immunofluorescence was essentially as described previously [47]. Polyclonal antibody against the HA epitope was from BAbCO (Richmond, CA, USA). Monoclonal anti-FLAG and anti-PABP were from Sigma (St Louis, MO, USA). Fluorescein isothiocyanate and rhodamine conjugated secondary antibodies were from Cappel Laboratories Cochranville, PA, USA. Immunostained cells were visualized with an Axiovert 200 microscope (Carl Zeiss Inc., Oberkochen, Germany).

Recombinant proteins

The His-tagged SARS-CoV N and NΔRS proteins were overproduced in *E. coli* BL21 (DE3). The bacterial lysate was prepared in a buffer containing 50 mM sodium phosphate (pH 8.0), 300 mM NaCl and 6 M urea, and was subsequently passed through His•Bind Resin (Novagen) for purification of His-tagged proteins. Bound proteins were eluted using the above buffer containing 250 mM imidazole. The eluate was dialyzed against a buffer containing 50 mM sodium phosphate (pH 7.4), 100 mM NaCl, 1 mM EDTA and 0.01% Na₃N. GST and GST-fusion to N, NΔRS and all mutant proteins were overproduced in *E. coli* strain BL21 and purified over glutathione-Sepharose beads (GE Healthcare) as recommended by the manufacturer. Purified GST fusion proteins were dialyzed against buffer D (20 mM Hepes, pH 7.9, 50 mM KCl, 0.5 mM dithiothreitol, 0.2 mM EDTA and 20% glycerol).

Phosphorylation

For *in vivo* phosphorylation, 3×10^6 transfected HeLa cells expressing FLAG-N or NΔRS in a 60 mm diameter plate were incubated in sodium phosphate-deficient DMEM (Invitrogen) supplemented with 0.75 mCi [³²P]orthophosphate (Amersham, Little Chalfont, UK) for 2.5 h. FLAG-tagged protein was immunoprecipitated from the cell lysates using anti-FLAG M2 agarose (Sigma) in a buffer containing 10 mM sodium phosphate (pH 7.2), 150 mM NaCl, 2 mM EDTA, 1% NP-40 and a mixture of protease inhibitors (Roche, Indianapolis, IN, USA), which was used as recommended by the manufacturer. *In vitro* phosphorylation of the N protein using recombinant GST-SRPK1 was essentially as described previously [42]; the reactions contained 5 μM ATP with or without additional 40 nM [λ-³²P]ATP. Dephosphorylation was performed using 200 U λ-protein phosphatase (New England Biolabs, Beverly, MA, USA).

CD spectrometry

Purified recombinant GST-NΔRS (3 μM) in 20 mM potassium acetate, 5 mM sodium acetate, 2 mM magnesium acetate and 1 mM EGTA was subjected to far-UV CD analysis using a Jasco J-720 spectropolarimeter (Jasco Inc., Easton, MD, USA). The measurement was performed in the range 190–250 nm in a 1 mm path length cuvette at room temperature. The data were recorded at 1 nm intervals.

Electrophoretic mobility shift assay

The approximately 110 nucleotide RNA probe was *in vitro* transcribed by T7 RNA polymerase using *BamHI*-digested pET15b-FLAG as template. The RNA was uniformly labeled with [α-³²P]UTP with a specificity activity at approximately 1.4×10^4 c.p.m. ng⁻¹. Recombinant His-N

protein was incubated with 5×10^4 c.p.m. of ^{32}P -labeled RNA in a 20 μL reaction containing 10 mM Hepes (pH 7.9), 50 μM EDTA, 10% glycerol, 1 mM dithiothreitol, 5 mM MgCl_2 , 0.1 mg of BSA, 2.5 μg of tRNA and 10 U of RNasin (Promega) at 25 °C for 15 min. Samples were fractionated on a 6% polyacrylamide nondenaturing gel in 0.5 \times TBE buffer (45 mM Tris-HCl, 45 mM boric acid, 1 mM EDTA, pH 8.0). Quantification was performed using Typhoon9410 Variable Mode Imager (Amersham).

Chemical crosslinking

The crosslinker disuccinimidyl suberate (Sigma) was prepared in *N,N*-dimethylformamide (Sigma) and used for chemical crosslinking of recombinant His-tagged N protein. The reaction mixtures contained 0.35 mM phosphorylated or nonphosphorylated N protein and 5 mM crosslinker in the NMR buffer (5 mM Hepes, 100 mM NaCl, 2 mM KCl, 1 mM MgCl_2 , 2 mM CaCl_2 and 0.5 mM EDTA, pH 7.8). The reaction was performed at 4 °C for 1 h and stopped by 100 mM glycine. Proteins were fractionated by SDS/PAGE and detected by immunoblotting using anti-SARS-CoV N serum (Imgenex, San Diego, CA, USA). Quantification was performed using IMAGE J software (National Institutes of Health, Bethesda MD, USA).

In vitro translation

The TNT coupled reticulocyte lysate system (Promega) was used for *in vitro* translation of a firefly luciferase reporter mRNA that contained 68 and 42 nucleotides in the 5' and 3' UTR, respectively, and was *in vitro* synthesized by T7 RNA polymerase from the template pFL-SV40. Each 10 μL of translation reaction contained 100 ng of the luciferase mRNA and different amounts of recombinant GST-N or NARS protein. The resulting luciferase activity was assessed by the luciferase assay system (Promega). To visualize luciferase protein, [^{35}S]methionin was added into the reaction according to the manufacturer's recommendation.

Sucrose and glycerol gradient sedimentation

HEK293 cells were transiently transfected with the vector expressing HA-tagged N or NARS protein. The cell lysate was then prepared in 10 mM Tris-HCl (pH 7.4), 150 mM NaCl and 3 mM MgCl_2 for sucrose gradient or in 20 mM Hepes (pH 7.9), 100 mM KCl and 1 mM MgCl_2 for glycerol gradient; both buffers additionally contained 100 $\mu\text{g}\cdot\text{mL}^{-1}$ cycloheximide, 35 $\mu\text{g}\cdot\text{mL}^{-1}$ digitonin and 20 $\text{U}\cdot\text{mL}^{-1}$ RNasin (Promega). Density gradient sedimentation was performed in a Beckman SW41 rotor (Beckman-Coulter, Fullerton, CA, USA) at 4 °C; for sucrose and glycerol gradient sedimentation, the centrifugation condition was 30 000 *g* for 5 h and 74 000 *g* for 16 h, respectively.

Proteins were precipitated by 20% trichloroacetic acid from each fraction and analyzed by immunoblotting using anti-SARS-CoV N serum.

Acknowledgements

We thank Tai-Huang Huang and Konan Peck for the SARS-CoV N protein cDNAs, plasmids and recombinant proteins, and Chwan-Deng Hsiao and Yi-Wei Chang for CD analysis. We thank Ru-Inn Lin and Wei-Lun Chang for their initial experimental assistance and Dr Tim C. Taylor for editing the manuscript. This work was supported by the National Science Council of Taiwan (NSC 95-3112-B001-007).

References

- 1 Drosten C, Gunther S, Preiser W, van der Werf S, Brodt HR, Becker S, Rabenau H, Panning M, Kolesnikova L, Fouchier RA *et al.* (2003) Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* **348**, 1967–1976.
- 2 Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, Tong S, Urbani C, Comer JA, Lim W *et al.* (2003) A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med* **348**, 1953–1966.
- 3 Marra MA, Jones SJ, Astell CR, Holt RA, Brooks-Wilson A, Butterfield YS, Khattri J, Asano JK, Barber SA, Chan SY *et al.* (2003) The genome sequence of the SARS-associated coronavirus. *Science* **300**, 1399–1404.
- 4 Lai MM & Cavanagh D (1997) The molecular biology of coronaviruses. *Adv Virus Res* **48**, 1–100.
- 5 Zuniga S, Sola I, Alonso S & Enjuanes L (2004) Sequence motifs involved in the regulation of discontinuous coronavirus subgenomic RNA synthesis. *J Virol* **78**, 980–994.
- 6 Tan YJ, Lim SG & Hong W (2006) Understanding the accessory viral proteins unique to the severe acute respiratory syndrome (SARS) coronavirus. *Antiviral Res* **72**, 78–88.
- 7 Hiscox JA, Cavanagh D & Britton P (1995) Quantification of individual subgenomic mRNA species during replication of the coronavirus transmissible gastroenteritis virus. *Virus Res* **36**, 119–130.
- 8 Risco C, Anton IM, Enjuanes L & Carrascosa JL (1996) The transmissible gastroenteritis coronavirus contains a spherical core shell consisting of M and N proteins. *J Virol* **70**, 4773–4777.
- 9 Hsieh PK, Chang SC, Huang CC, Lee TT, Hsiao CW, Kou YH, Chen IY, Chang CK, Huang TH & Chang MF (2005) Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into

- virus-like particles is nucleocapsid dependent. *J Virol* **79**, 13848–13855.
- 10 Chen CY, Chang CK, Chang YW, Sue SC, Bai HI, Riang L, Hsiao CD & Huang TH (2007) Structure of the SARS coronavirus nucleocapsid protein RNA-binding dimerization domain suggests a mechanism for helical packaging of viral RNA. *J Mol Biol* **368**, 1075–1086.
 - 11 Zhou ML, Williams AK, Chung SI, Wang L & Collisson EW (1996) Infectious bronchitis virus nucleocapsid protein binds RNA sequences in the 3' terminus of the genome. *Virology* **217**, 191–199.
 - 12 Nelson GW, Stohlman SA & Tahara SM (2000) High affinity interaction between nucleocapsid protein and leader/intergenic sequence of mouse hepatitis virus RNA. *J Gen Virol* **81**, 181–188.
 - 13 Zhou M & Collisson EW (2000) The amino and carboxyl domains of the infectious bronchitis virus nucleocapsid protein interact with 3' genomic RNA. *Virus Res* **67**, 31–39.
 - 14 Stohlman SA, Baric RS, Nelson GN, Soe LH, Welter LM & Deans RJ (1988) Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J Virol* **62**, 4288–4295.
 - 15 Chang RY & Brian DA (1996) Cis requirement for N-specific protein sequence in bovine coronavirus defective interfering RNA replication. *J Virol* **70**, 2201–2207.
 - 16 Molenkamp R, van Tol H, Rozier BCD, van der Meer Y, Spaan WJM & Snijder EJ (2000) The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. *J Gen Virol* **81**, 2491–2496.
 - 17 Tahara SM, Dietlin TA, Nelson GW, Stohlman SA & Manno DJ (1998) Mouse hepatitis virus nucleocapsid protein as a translational effector of viral mRNAs. *Adv Exp Med Biol* **440**, 313–318.
 - 18 Zhou B, Wang Q, Liu X, Li P, Ma Q & Cao C (2008) Nucleocapsid protein of SARS-CoV inhibits cell cytokinesis and proliferation by interacting with translation elongation factor 1 α . *J Virol* **82**, 6962–6971.
 - 19 Surjit M, Liu B, Chow VT & Lal SK (2006) The nucleocapsid protein of severe acute respiratory syndrome-coronavirus inhibits the activity of cyclin-cyclin-dependent kinase complex and blocks S phase progression in mammalian cells. *J Biol Chem* **281**, 10669–10681.
 - 20 Cawood R, Harrison SM, Dove BK, Reed ML & Hiscox JA (2007) Cell cycle dependent nucleolar localization of the coronavirus nucleocapsid protein. *Cell Cycle* **6**, 863–867.
 - 21 Chen H, Wurm T, Britton P, Brooks G & Hiscox JA (2002) Interaction of the coronavirus nucleoprotein with nucleolar antigens and the host cell. *J Virol* **76**, 5233–5250.
 - 22 Huang Q, Yu L, Petros AM, Gunasekera A, Liu Z, Xu N, Hajduk P, Mack J, Fesik SW & Olejniczak ET (2004) Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. *Biochemistry* **43**, 6059–6063.
 - 23 Fan H, Ooi A, Tan YW, Wang S, Fang S, Liu DX & Lescar J (2005) The nucleocapsid protein of coronavirus infectious bronchitis virus: crystal structure of its N-terminal domain and multimerization properties. *Structure* **13**, 1859–1868.
 - 24 Tan YW, Fang S, Fan H, Lescar J & Liu DX (2006) 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. *Nucleic Acids Res* **34**, 4816–4825.
 - 25 Chen H, Gill A, Dove BK, Emmett SR, Kemp CF, Ritchie MA, Dee M & Hiscox JA (2005) Mass spectroscopic characterization of the coronavirus infectious bronchitis virus nucleoprotein and elucidation of the role of phosphorylation in RNA binding by using surface plasmon resonance. *J Virol* **79**, 1164–1179.
 - 26 Luo H, Ye F, Chen K, Shen X & Jiang H (2005) RS-rich motif plays a pivotal role in recombinant SARS coronavirus nucleocapsid protein multimerization. *Biochemistry* **44**, 15351–15358.
 - 27 Fu XD (1995) The superfamily of arginine/serine-rich splicing factors. *RNA* **1**, 663–680.
 - 28 Stojdl DF & Bell JC (1999) SR protein kinases: the splice of life. *Biochem Cell Biol* **77**, 293–298.
 - 29 Graveley BR (2000) Sorting out the complexity of SR protein functions. *RNA* **6**, 1197–1211.
 - 30 Wootton SK, Rowland RR & Yoo D (2002) Phosphorylation of the porcine reproductive and respiratory syndrome virus nucleocapsid protein. *J Virol* **76**, 10569–10576.
 - 31 White TC, Yi Z & Hogue BG (2007) Identification of mouse hepatitis coronavirus A59 nucleocapsid protein phosphorylation sites. *Virus Res* **126**, 139–148.
 - 32 Surjit M, Kumar R, Mishra RN, Reddy MK, Chow VT & Lal SK (2005) The severe acute respiratory syndrome coronavirus nucleocapsid protein is phosphorylated and localizes in the cytoplasm by 14-3-3-mediated translocation. *J Virol* **79**, 11476–11486.
 - 33 Calvo E, Escors D, Lopez JA, Gonzalez JM, Alvarez A, Arza E & Enjuanes L (2005) Phosphorylation and subcellular localization of transmissible gastroenteritis virus nucleocapsid protein in infected cells. *J Gen Virol* **86**, 2255–2267.
 - 34 Hiscox JA, Wurm T, Wilson L, Britton P, Cavanagh D & Brooks G (2001) The coronavirus infectious bronchitis virus nucleoprotein localizes to the nucleolus. *J Virol* **75**, 506–512.
 - 35 Wurm T, Chen H, Hodgson T, Britton P, Brooks G & Hiscox JA (2001) Localization to the nucleolus is a common feature of coronavirus nucleoproteins, and the protein may disrupt host cell division. *J Virol* **75**, 9345–9356.

- 36 You J, Dove BK, Enjuanes L, DeDiego ML, Alvarez E, Howell G, Heinen P, Zambon M & Hiscox JA (2005) Subcellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein. *J Gen Virol* **86**, 3303–3310.
- 37 Timani KA, Liao Q, Ye L, Zeng Y, Liu J, Zheng Y, Ye L, Yang X, Lingbao K, Gao J *et al.* (2005) Nuclear/nucleolar localization properties of C-terminal nucleocapsid protein of SARS coronavirus. *Virus Res* **114**, 23–34.
- 38 Dove BK, You JH, Reed ML, Emmett SR, Brooks G & Hiscox JA (2006) Changes in nucleolar morphology and proteins during infection with the coronavirus infectious bronchitis virus. *Cell Microbiol* **8**, 1147–1157.
- 39 Reed ML, Dove BK, Jackson RM, Collins R, Brooks G & Hiscox JA (2006) Delineation and modeling of a nucleolar retention signal in the coronavirus nucleocapsid protein. *Traffic* **7**, 833–848.
- 40 Kedersha N & Anderson P (2002) Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochem Soc Trans* **30**, 963–969.
- 41 Gilks N, Kedersha N, Ayodele M, Shen L, Stoecklin G, Dember LM & Anderson P (2004) Stress granule assembly is mediated by prion-like aggregation of TIA-1. *Mol Biol Cell* **15**, 5383–5398.
- 42 He R, Dobie F, Ballantine M, Leeson A, Li Y, Bastien N, Cutts T, Andonov A, Cao J, Booth TF *et al.* (2004) Analysis of multimerization of the SARS coronavirus nucleocapsid protein. *Biochem Biophys Res Commun* **316**, 476–483.
- 43 Ma CT, Velazquez-Dones A, Hagopian JC, Ghosh G, Fu XD & Admas JA (2008) Ordered multi-site phosphorylation of the splicing factor ASF/SF2 by SRPK1. *J Mol Biol* **376**, 55–68.
- 44 Raaben M, Groot Koerkamp MJ, Rottier PJ & de Haan CA (2007) Mouse hepatitis coronavirus replication induces host translational shutoff and mRNA decay, with concomitant formation of stress granules and processing bodies. *Cell Microbiol* **9**, 2218–2229.
- 45 Scott AT & Julio AA (2005) Dephosphorylation shows SR protein the way out. *Molecular Cell* **20**, 499–501.
- 46 Anderson P & Kedersha N (2002) Visibly stressed: the role of eIF2, TIA-1, and stress granules in protein translation. *Cell Stress Chaperones* **7**, 213–221.
- 47 Lai MC, Teh BH & Tarn WY (1999) A human papillomavirus E2 transcriptional activator: the interactions with cellular splicing factors and potential function in pre-mRNA processing. *J Biol Chem* **274**, 11832–11841.
- 48 Daub H, Blencke S, Habenberger P, Kurtenbach A, Dennenmoser J, Wissing J, Ullrich A & Cotton M (2002) Identification of SRPK1 and SRPK2 as the major cellular protein kinases phosphorylating hepatitis B virus core protein. *J Virol* **76**, 8124–8137.
- 49 Stertz S, Reichelt M, Spiegel M, Kuri T, Martinez-Sobrido L, Garcia-Sastre A, Weber F & Kochs G (2007) The intracellular sites of early replication and budding of SARS-coronavirus. *Virology* **361**, 304–315.
- 50 Rowland RR, Chauhan V, Fang Y, Pekosz A, Kerrigan M & Burton MD (2005) Intracellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein: absence of nucleolar accumulation during infection and after expression as a recombinant protein in vero cells. *J Virol* **79**, 11507–11512.

Supplementary material

The following supplementary material is available online:

Fig. S1. The SARS N protein is phosphorylated by SRPK1 but not by Clk1 and PKA.

This material is available as part of the online article from <http://www.blackwell-synergy.com>

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.