

## SR-Rich Motif Plays a Pivotal Role in Recombinant SARS Coronavirus Nucleocapsid Protein Multimerization<sup>†</sup>

Haibin Luo,<sup>‡</sup> Fei Ye,<sup>‡</sup> Kaixian Chen,<sup>‡</sup> Xu Shen,<sup>\*,‡,§</sup> and Hualiang Jiang<sup>\*,‡,§</sup>

Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Shanghai Institutes for Biological Sciences, Graduate School of the Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 201203, China, and School of Pharmacy, East China University of Science and Technology, Shanghai 200237, China

Received June 13, 2005; Revised Manuscript Received September 28, 2005

**ABSTRACT:** The nucleocapsid (N) protein of SARS coronavirus (SARS-CoV) is reported to function in encapsidating the viral genomic RNA into helical nucleocapsid, and its self-association is believed to be vital in coating the viral genomic RNA. Characterization of SARS-CoV N multimerization may thereby help us better understand the coronavirus assembly. In the current work, using the yeast two-hybrid technique, an unexpected interaction between residues 1–210 and 211–290 (central region) of the SARS-CoV N protein was detected, and SPR results further revealed that the SR-rich motif (amino acids 183–197) of SARS-CoV N protein is responsible for such an interaction. Chemical cross-linking and gel-filtration analyses indicated that the residues 283–422 of the SARS-CoV N protein have multimeric ability, although the full-length N protein is prone to exist predominantly as dimers. In addition, the multimeric ability of the C-terminal domain of SARS-CoV N protein could be weakened by the SR-rich motif interaction with the central region (amino acids 211–290). All of these data suggested that the SR-rich motif of the SARS-CoV N protein might play an import role in the transformation of the SARS-CoV N protein between the dimer and multimer during its binding to its central region for self-association or dissociation. This current paper will hopefully provide some new ideas in studying SARS-CoV N multimerization.

Between the end of the year 2002 and June of the year 2003, the severe acute respiratory syndrome (SARS) broke out in China and more than 30 other countries. It has been known that SARS coronavirus (SARS-CoV)<sup>1</sup> is responsible for the SARS infection (1–4), and further research showed that SARS-CoV is a brand new type of *Coronaviridae* (5) and moderately related to the other known coronaviruses (6). SARS-CoV has 14 open-reading frames (ORFs) that encode the replicase polyproteins (ORF1a and 1b), structural proteins (S, E, M, and N), and accessory proteins (3a, 3b, 6, 7a, 7b, 8a, 8b, and 9b) (6, 7). The N protein of coronavirus is the most abundant structural protein and probably functions in recognizing RNA to lead to the formation of ribonucleoprotein (RNP) (8) or wraps the large single-stranded viral genomic RNA to form a helical nucleocapsid (9). Recently, we discovered that the SARS-CoV N protein could interact

with human cyclophilin A and heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1); these facts have provided some useful hints in understanding the possible SARS-CoV infection pathway and mRNA synthesis (10, 11).

Self-association of the N protein has been found to be an important step within virus particles assembly (9, 12–15). SARS-CoV N is a 46-kDa structural protein, which has been reported to be able to form multimers but exist predominantly as dimers (16–18). Using the two-hybrid method, He et al. proposed that the fragment contributing to SARS-CoV N protein self-association is the SR-rich motif (amino acids 183–197) (13) (Scheme 1), although no further detailed description was given. In fact, the study of the SR-rich motif in the SARS-CoV N protein has been of special interest. Some papers have even suggested that this SR-rich motif might be also responsible for SARS-CoV N interactions with the SARS-CoV membrane protein (19) and hnRNP A1 (11). However, also using the two-hybrid method, Surjit et al. demonstrated that the C-terminal 209 residues are attributed to the SARS-CoV N protein dimerization (18). Moreover, residues 279–370 in C terminus were also recently discovered to self-associate as only dimers (17). All of these above-mentioned facts seem to have made the SARS-CoV N multimerization mechanism a little bewildering, and it thereby needs more characterization.

In the current work, through some related biochemical and biophysical investigations, the SR-rich motif (amino acids 183–197, Scheme 1) was proven to play a pivotal role in SARS-CoV N multimerization. It is found that the SR-rich

<sup>†</sup> State Key Program of Basic Research of China (Grants 2002CB512802, 2002CB512807, and 2004CB518905), the National Natural Science Foundation of China (Grants 20372069 and 20472095), Shanghai Basic Research Project from the Shanghai Science and Technology Commission (Grants 02DJ14070, 03DZ19228, and 03DZ19212), the 863 Hi-Tech Program (Grants 2002AA233011 and 2005AA235030), and Sino-European Project on SARS Diagnostics and Antivirals (proposal/contract number 003831).

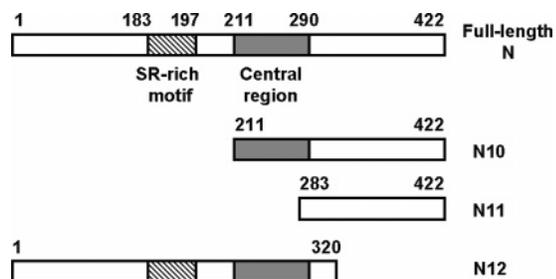
\* To whom correspondence should be addressed. Telephone and Fax: 86-21-50806918. E-mail: xshen@mail.shcnc.ac.cn (X.S.); hljiang@mail.shcnc.ac.cn (H.J.).

<sup>‡</sup> Chinese Academy of Sciences.

<sup>§</sup> East China University of Science and Technology.

<sup>1</sup> Abbreviations: SARS-CoV, severe acute respiratory syndrome coronavirus; SARS-CoV N, the nucleocapsid protein of SARS-CoV; SPR, surface plasmon resonance.

Scheme 1



motif could bind to the central region (amino acids 211–290, Scheme 1) of the SARS-CoV N protein, and this binding might be responsible for the SARS-CoV N protein dimerization. Although the C-terminal domain (amino acids 283–422, Scheme 1) was determined to exhibit strongly multimeric ability, the full-length SARS-CoV N protein is prone to forming dimers. Such a dimeric rather than multimeric formation preference for the SARS-CoV N protein was elucidated by the fact that the central domain of the SARS-CoV N protein might bind to the SR-rich motif during self-association, which makes the N protein fail to form multimers through the C-terminal domain. All of these data implied that the SARS-CoV N protein multimerization is closely related to the SR-rich motif interaction with the central region.

## EXPERIMENTAL PROCEDURES

**Chemicals.** NTA affinity resin and the molecular-weight marker were purchased from Amersham Bioscience. IPTG was purchased from Promega. The SR peptide (AS-SRSSRSRGNRSRN, SR-rich motif of SARS-CoV N) and the control peptide (AGGAGAGGEA) were synthesized by Sangon (China). All other chemicals were from Sigma in their analytical grade.

Cloning, expression, and purification of the His-tagged full-length SARS-CoV N, N10, N11, and N12 proteins. The recombinant SARS-CoV N protein was prepared on the basis of the published procedure reported by Luo et al. (16), and the cloning, expression, and purification of the SARS-CoV N truncated mutant N10 (amino acids 211–422), N11 (amino acids 283–422), and N12 (amino acids 1–320) proteins (Scheme 1) were performed by using the modified protocol similar to that of SARS-CoV N protein. The primers used for amplification are listed as follows: N10 (sense), 5'-AA-TTGGATCCATGGCTAGCGGAGGTGGTGAAC-3'; N10 (anti-sense), 5'-ACGGGTCGACTTATGCCTGAGTTGAAT-CAGCAG-3'; N11 (sense), 5'-AATTGGATCCACCC-AAGGAAATTCGGGGACCAA-3'; N11 (anti-sense), 5'-ACGGGTCGACTTATGCCTGAGTTGAATCAGCAG-3'; N12 (sense), 5'-ATTAGGATCCTCTGATAATGGACCC-CAATCA-3'; N12 (anti-sense), 5'-ACGGGTCGACTTATC-GAGCAGGAGAATTTCC-3'.

**Recombinant Vector Construction for the Yeast Two-Hybrid Assay.** The yeast vectors of pGADT7 and pGBKT7 were purchased from Clontech (Palo Alto, CA). The cDNA of SARS-CoV N truncated mutants were amplified by PCR using the corresponding primers (Table 1). The amplified products were digested with *EcoRI* and *BamHI* (NEB, Cottdale, AL) and then cloned separately into the pGADT7 or pGBKT7 vectors.

**Yeast Transformation and Culture.** Competent cells of yeast strain AH109 were obtained from Clontech (Palo Alto, CA), and transformation was performed according to the protocol of the manufacturer. Briefly, 500 ng of plasmid DNA was added into 50  $\mu$ L of the competent cells and then mixed with 36  $\mu$ L of 1 M lithium acetate and 240  $\mu$ L of 50% poly(ethylene glycol) (MW3350) at 30  $^{\circ}$ C for 30 min followed by heat-shock at 42  $^{\circ}$ C for 30 min and subsequently spread on a drop-out-agar plate in the absence of tryptophan and leucine (SD-TL). The plates were incubated at 30  $^{\circ}$ C for 48 h for yeast growth. PCR was used to confirm transformation with the target gene. The positive clone was inoculated to the SD-TL solid medium for storage.

**$\alpha$ -Galactosidase Activity Assay.** The quantitative  $\alpha$ -galactosidase activity assays were carried out by using *p*-nitrophenyl  $\alpha$ -D-galactopyranoside (PNP- $\alpha$ -Gal) as the substrate according to the Clontech manual. A positive clone was inoculated from the above-mentioned plate to 3 mL of SD-TL liquid medium, and then, the yeast cells were cultured overnight (16–18 h) at 30  $^{\circ}$ C with shaking (250 rpm). After the OD<sub>600</sub> of the medium reached  $\sim$ 0.8–1.2, the cells were centrifuged at 4000 rpm for 5 min and 16  $\mu$ L of supernatants was transferred into a 96-well plate. Then, 48  $\mu$ L of assay buffer was added into each sample. The assay buffer was newly prepared by combining 2 volumes of NaOAc buffer (0.5 M NaOAc at pH 4.5) with 1 volume of PNP- $\alpha$ -Gal solution (100 mM PNP- $\alpha$ -Gal) before use. After 60 min of incubation at 30  $^{\circ}$ C, the reaction was terminated by the addition of 136  $\mu$ L of stop buffer (1 M Na<sub>2</sub>CO<sub>3</sub>) and the OD<sub>410</sub> of the reaction mixture was measured by a microplate spectrophotometer (Bio-Rad). The  $\alpha$ -galactosidase activity was calculated according to the following formula:

$$\alpha\text{-galactosidase activity [milliunits/(mL} \times \text{cell)]} = \frac{\text{OD}_{410} V_f 1000}{\epsilon b t V_i \text{OD}_{600}}$$

where  $t$  is elapsed time (in minutes) of incubation,  $V_f$  is the final volume of the assay (200  $\mu$ L),  $V_i$  is the volume of the culture medium supernatant added (16  $\mu$ L), and OD<sub>600</sub> is the cell density at the start of the assay,  $\epsilon b$  is equal to 10.5 (mL/mol) here.

**Chemical Cross-Linking Assay.** To investigate the multimeric features of N10 and N11 (in PBS at pH 7.4), the chemical cross-linking experiments were carried out. The total volume of each reaction is 30  $\mu$ L. The protein solutions were mixed with glutaraldehyde and reacted at 16  $^{\circ}$ C for 20 min. The reaction was stopped by adding 20  $\mu$ L of SDS-PAGE loading buffer and heated at 99  $^{\circ}$ C for 5 min.

**Gel-Filtration Assay.** The gel-filtration assay was performed for further investigating protein multimeric features. During the assay, an AKTA-purify system was carried out by using a HiLoad 16/60 superdex 200 prep grad column (separation range of 10–600 kDa, Amersham Biosciences). The column was equilibrated at a flow rate of 1 mL/min with PBS buffer at room temperature. The protein sample (500  $\mu$ L) was injected at a given concentration and detected by following the absorbance at 280 nm. The column was calibrated using a mixture of bovine serum albumin (67 kDa), ovalbumin (45 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa) as a reference. These standard proteins were purchased from Amersham Biosciences.

Table 1: Primers Used in the Yeast Two-Hybrid Assay

$\Delta N_N$ , N (sense)	GGTTGAATTCCCAGATGACCAAATTGGCTACTACC
$\Delta N_N$ (anti-sense)	GGAAGGATCCCTTATCGAGCAGGAGAATTTCCCC
$\Delta N_C$ , $\Delta N_5$ (sense)	AATTGAATTCATGGCTAGCGGAGGTGGTAAAA
$\Delta N_C$ , N, $\Delta N_4$ (anti-sense)	AATTGGATCCCTT ATGCCTGAGTTGAATCAGCAG
$\Delta N_4$ (sense)	CGGTGAATTCGACCTAATCAGACAAGGAACTGA
$\Delta N_5$ (anti-sense)	GGCCGGATCCCTTAGTTGTCTTTGAATTGTGGAT

**Surface Plasmon Resonance (SPR) Technology-Based Analysis.** SPR technology-based Biacore 3000 was used to study the interaction between the SR peptide (ASSRSSRS-RGNSRN) and the immobilized N10 (Scheme 1). CM5 research-grade sensor chips was purchased from Biacore AB. *N*-Hydroxysuccinimide (NHS), *N*-ethyl-*N'*-(3-dimethylamino-propyl)carbodiimide (EDC), and ethanolamine coupling reagents were used to immobilize the ligand to the sensor surface using a standard amine-coupling procedure. The running and sample buffers were HBS-EP (10 mM HEPES at pH 7.4, 150 mM NaCl, 3 mM EDTA, and 0.005% P<sub>20</sub>), and all of the buffers were filtered by a 0.22  $\mu$ m filter and degassed before use. The SR peptide was injected (120 s, 30  $\mu$ L/s) followed by a dissociation phase (120 s). Each experiment was repeated at least twice, and the data were obtained at 25 °C. An untreated flow cell was used as a negative control. The binding affinity of the SR peptide to the immobilized N10 was thus calculated by a steady-state fitting model with the BIAevaluation 3.1 software.

## RESULTS

**Yeast Two-Hybrid Technology-Based Analysis of the Interactions among the Different SARS-CoV N Truncated Mutants.** During mapping the interaction regions involved in the SARS-CoV N protein self-association by using the yeast two-hybrid system, we carried out a different strategy from what has ever been reported (13, 18), in which the research merely focused on the truncated mutant interactions with the full-length SARS-CoV N protein. In our current work, the cDNAs of the two truncated mutants,  $\Delta N_N$  (amino acids 1–210) and  $\Delta N_C$  (amino acids 211–422) (Figure 1A), were separately cloned into pGADT7 and pGBKT7 plasmids to analyze the three different combinational interactions:  $\Delta N_N/\Delta N_N$ ,  $\Delta N_C/\Delta N_C$ , and  $\Delta N_N/\Delta N_C$ . In addition, the full-length SARS-CoV N gene was also cloned into pGADT7 and pGBKT7 plasmids for evaluation of the assay system reliability in this work, and the empty pGADT7 and pGBKT7 plasmids were transformed into a yeast cell as a negative control.

The protein–protein interactions were studied by monitoring the  $\alpha$ -galactosidase activity of different yeast cells. As shown in Figure 1B, the  $\alpha$ -galactosidase activity of cells with pGADT7-N and pGBKT7-N is over 2.5 milliunits, which follows the fact that the SARS-CoV N protein could self-associate as indicated by the published results (16–18), further confirming that our yeast two-hybrid system is reliable enough. As also indicated in Figure 1B, the  $\Delta N_N/\Delta N_N$  interaction could not be detected and  $\Delta N_C/\Delta N_C$  showed a stronger interaction than even N/N. Furthermore, it is remarkable that  $\Delta N_N$  could strongly bind to  $\Delta N_C$ .

To get more information about the  $\Delta N_C/\Delta N_C$  interaction, two smaller truncated mutants,  $\Delta N_4$  (amino acids 291–422) and  $\Delta N_5$  (amino acids 211–350) (Figure 1A) were cloned into the pGBKT7 plasmid as pGBKT7- $\Delta N_4$  and pGBKT7-

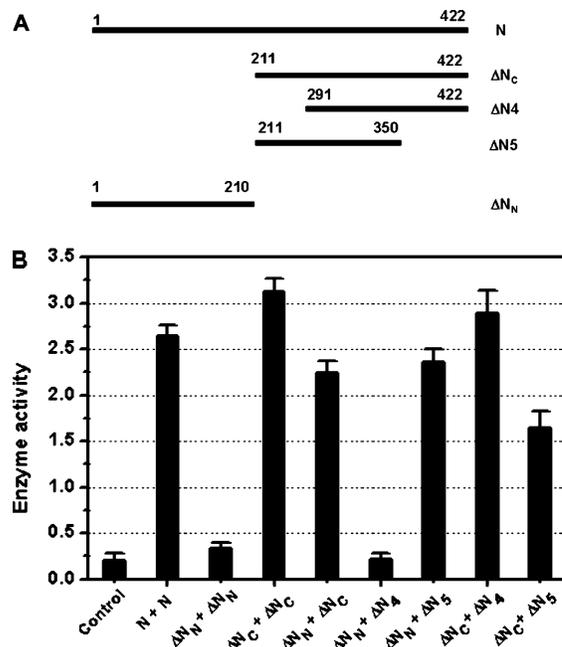


FIGURE 1: Mapping the self-association domain of SARS-CoV N protein. Schematic description of the truncated mutants (A) and the yeast two-hybrid assay results for their interactions (B). The empty vectors pGBKT7 and pGADT7 co-transformed yeast cells were used as a negative control. Every experiment was repeated for at least 3 times, and the data were obtained by the average.

$\Delta N_5$ . These two plasmids were then separately transformed with pGADT7- $\Delta N_C$  into a yeast cell to investigate the related protein–protein interactions. As shown in Figure 1B, both  $\Delta N_4$  and  $\Delta N_5$  could interact with  $\Delta N_C$ .  $\Delta N_C/\Delta N_4$  showed a similar affinity comparing with  $\Delta N_C/\Delta N_C$ , while the  $\Delta N_C/\Delta N_5$  interaction is weaker than  $\Delta N_C/\Delta N_C$  according to their  $\alpha$ -galactosidase activities, thereby suggested that residues 291–422 of the SARS-CoV N protein might have the similar self-association capability relative to  $\Delta N_C$  and residues 351–422 were involved in the self-association of  $\Delta N_C$ .

To investigate the involved region of  $\Delta N_C$  within the  $\Delta N_N/\Delta N_C$  interaction that contains the full-length SARS-CoV N protein self-association information, pGBKT7- $\Delta N_4$  and pGBKT7- $\Delta N_5$  were separately transformed with pGADT7- $\Delta N_N$  into a yeast cell. All of the results are listed in Figure 1B; it is revealed that  $\Delta N_5$  might interact with  $\Delta N_N$  and  $\Delta N_4$  failed. Thereby, this indicated that the central region (amino acids 211–290, Scheme 1) of the SARS-CoV N protein was involved in the  $\Delta N_N/\Delta N_C$  interaction, further providing the possibility that the full-length SARS-CoV N protein self-association might relate to this central region.

**SR-Rich Motif of  $\Delta N_N$  Contributes to the  $\Delta N_N/\Delta N_C$  Interaction.** Recently, He et al. reported that the SR-rich motif (amino acids 183–197) is responsible for the SARS-CoV N protein self-association (13), although there was a lack of detailed description for this case. Therefore, according to this fact and our above result that  $\Delta N_N$  (amino acids

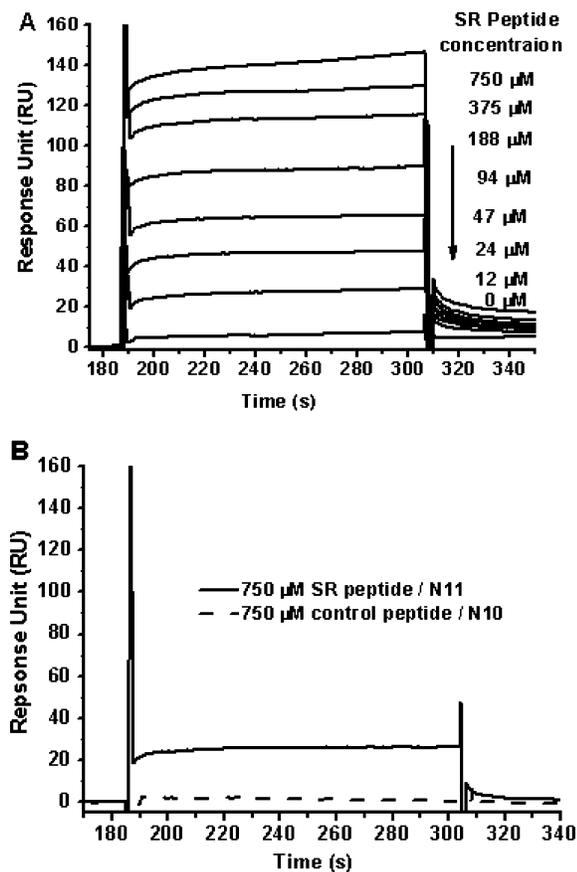


FIGURE 2: Validation and quantitative evaluation of the interaction between N10 and the SR peptide (ASSRSSSRGNSRN). (A) Binding curves of different concentrations (labeled in the right) of the SR peptide to the immobilized N10. (B) Running curve (---) of 750  $\mu$ M control peptide interacting with the immobilized N10; the running curve (—) of 750  $\mu$ M SR peptide binding to immobilized N11 (The immobilized level was 3000 RU).

1–210, Figure 1) fails to interact with itself, it is supposed that the SR-rich motif in  $\Delta N_N$  might contribute to the  $\Delta N_N/\Delta N_C$  interaction. To prove this, we synthesized the SR peptide (ASSRSSSRGNSRN, residues 183–197 of the N protein) that involves the SR-rich motif. In the Biacore assays, the N10 protein (amino acids 211–422, Scheme 1) containing the central region (amino acids 211–290) was covalently immobilized on the CM5 sensor chip by the level of 3500 RU (response unit), and a series of SR peptides with different concentrations were then injected to the surface of the sensor chip. As shown in Figure 2A, with the increase of the peptide injected to the surface, the binding response in RU increases up to saturation, thus indicating that this peptide binds to the N10 protein. The binding affinity of the SR peptide against the N10 protein was thus evaluated as  $K_D$  at  $60.5 \pm 5.3 \mu$ M by the steady-state fitting model. To further confirm the specificity of the SR peptide, a control peptide (AGGAGAGGEA) was used as a negative control. As shown in Figure 2B (---), 750  $\mu$ M of the control peptide hardly binds to the immobilized N10, thus suggesting that the interaction between the SR peptide and N10 was specific. Moreover, the fact that 750  $\mu$ M SR peptide showed a much weaker binding ability to the immobilized N11 (—, Figure 2B) than the immobilized N10 (Figure 2A) suggested the possible involvement of the central domain (amino acids 211–290, Scheme 1) in the  $\Delta N_N/\Delta N_C$  interaction.

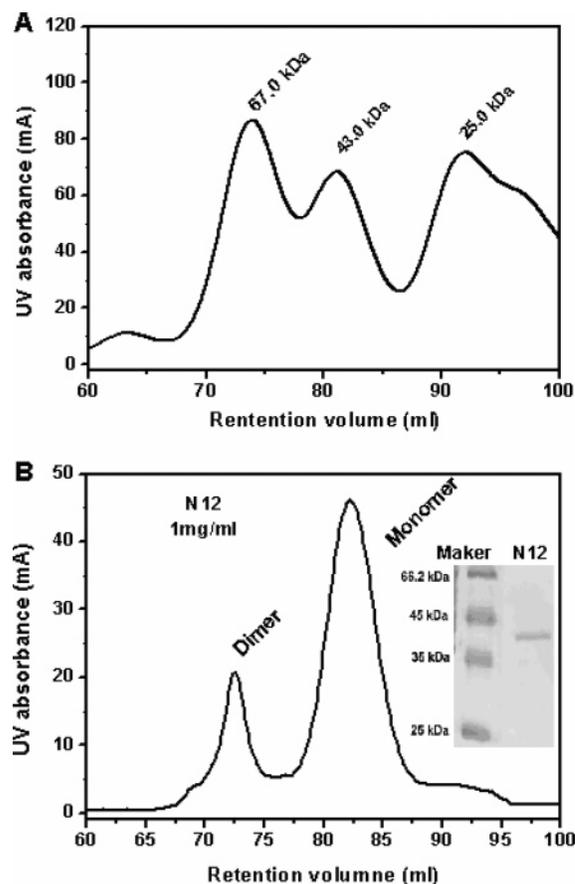


FIGURE 3: Gel-filtration assays of N12. (A) Retention volume graph of the three standard proteins: bovine serum albumin (67 kDa), ovalbumin (45 kDa) and chymotrypsinogen A (25 kDa). (B) Retention volume graph of N12 (amino acids 1–320) at 1 mg/mL. The inset shows the purity of N12.

*Dimerization of N12 (Amino Acids 1–320) Validates the  $\Delta N_N/\Delta N_C$  Interaction.* To further prove the  $\Delta N_N/\Delta N_C$  interaction, N12 (amino acids 1–320, Scheme 1), a fragment covering the SR-rich domain and the central domain, was cloned, expressed, and purified as a His-tagged protein in the *Escherichia coli* system. The purified N12 was soluble and pure as a single band as confirmed on SDS-PAGE (inset in Figure 3B). The purified N12 was then applied to the gel-filtration assay, in which the standard marker was used as a reference (Figure 3A). As shown in Figure 3B, the N12 dimerization was validated according to the two peaks in the elution curve. This result further indicates the existence of the  $\Delta N_N/\Delta N_C$  interaction.

*Residues 211–422 of SARS-CoV N Have the Capability of Multimerization.* From yeast two-hybrid assays, we have just discovered the interaction fact for  $\Delta N_C/\Delta N_C$ , which showed that residues 211–422 of SARS-CoV N might possess self-association capability (Figure 1), but it is still unknown whether this fragment could form a dimer or multimer. To further characterize the multimeric features of this fragment, the His-tagged N10 (amino acids 211–422, Scheme 1) was generated in the *E. coli* system. The purified N10 (lane 2 of Figure 4A) was mixed with different concentrations of glutaraldehyde and then detected by SDS-PAGE. As shown in Figure 4A, the amount of dimers or multimers of N10 increased with the increase of the glutaraldehyde concentration, indicating that residues 211–422 have the capability of forming dimers and multimers.

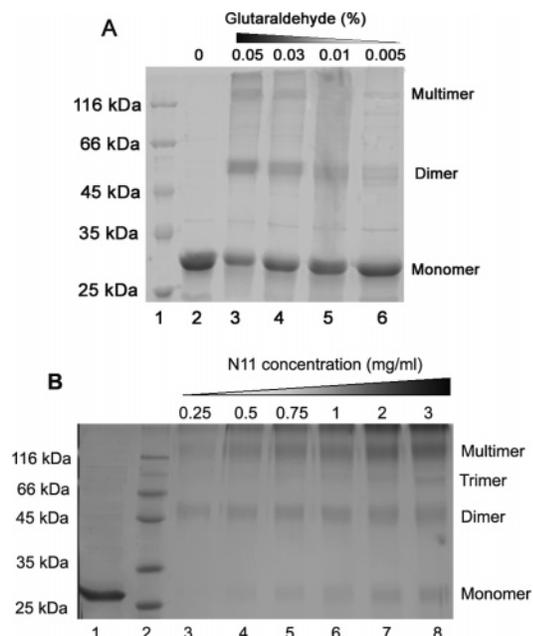


FIGURE 4: Chemical cross-linking assay based on N10 and N11 multimerization analysis. (A) Marker is in lane 1. N10 (1.5 mg/mL) reacted with different concentrations of glutaraldehyde with the results in lanes 2–6. (B) Lane 1, the purified N11; lane 2, marker; lanes 3–8, N11 at different concentrations. The final concentration of glutaraldehyde in each sample was 0.08%.

*C-Terminal 140 Amino Acids (Amino Acids 283–422, N11) of SARS-CoV N Could Form Multimers.* As presented above, the  $\Delta N4/\Delta N_C$  interaction was detected by the yeast two-hybrid assay, which suggested that residues 291–422 of SARS-CoV N also possess self-association capability (Figure 1). However, it is unclear whether this shorter fragment could also form multimers such as N10 (amino acids 210–422). To study the multimeric ability of this shorter fragment, the His-tagged N11 (amino acids 283–422, Scheme 1) was generated in the *E. coli* system and purified through a NTA column. N11 was soluble, and its purity was checked by SDS-PAGE (lane 1 of Figure 4B). To synchronously investigate whether N11 multimerization is concentration-dependent, different concentrations of purified N11 were cultured in glutaraldehyde (0.08%). As indicated in Figure 4B, it is found that the amount of dimers and multimers of N11 increased with the increase of the N11 concentration. This result thus demonstrated that N11 (amino acids 283–422) has the capability of forming dimers and multimers, and its multimerization is concentration-dependent.

*The Full-Length SARS-CoV N Protein Is Prone to Dimers Compared with N11.* To compare the multimerization capability of N11 with that of the full-length SARS-CoV N, we applied these two proteins to a gel-filtration column using the same molar concentration. On the basis of the standard proteins as markers (Figure 5A), the full-length SARS-CoV N protein was found to exist mostly as dimers (Figure 5B, ---), according to the recently reported results (16, 17). However, N11 typically forms multimers as shown in Figure 5B (—). Theoretically, the SARS-CoV N protein should also exist predominantly as multimers because the N11 protein is easy to form multimers. However, the fact is that the full-length SARS-CoV N protein prefers to form dimers instead

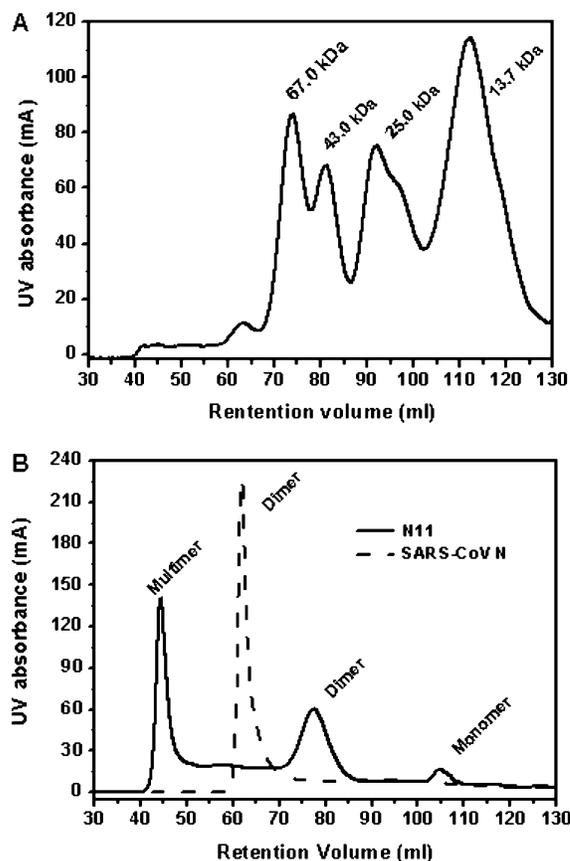
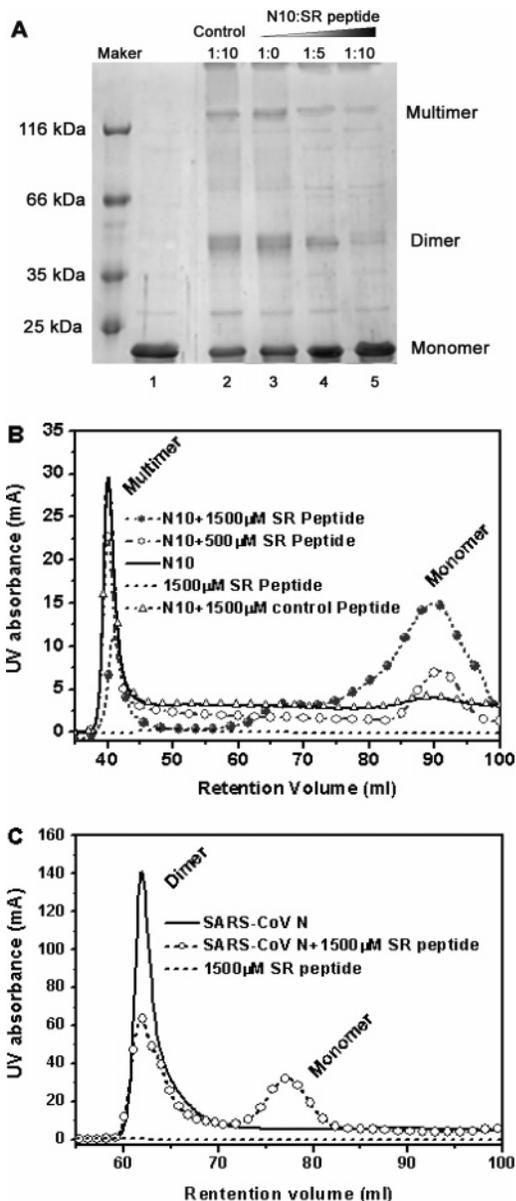


FIGURE 5: Protein gel-filtration assays for N11 and SARS-CoV N. (A) Retention volume graph of the four standard proteins: bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa). (B) The retention volume graphs of the SARS-CoV N protein at 7 mg/mL (---) and N11 at 2.5 mg/mL (—).

of multimers (Figure 4B, ---). What results in such an unexpected result? In comparison with the full-length SARS-CoV N protein, the N11 protein (amino acids 283–422) is shorter and lacks the SR-rich motif and the central region. Is it possible that the interaction between the SR-rich motif and the central region might affect SARS-CoV N multimerization?

*SR-Rich Motif Perturbs N10 (Amino Acids 211–422) Multimerization.* To explore the above-mentioned possibility, the influence of the SR-rich motif on N10 (including the central region, Scheme 1) multimerization was investigated. Within the assays, 1.25 mg/mL N10 protein was incubated with different concentrations of the SR peptide for 3 h at room temperature before the mixture was applied to the chemical cross-linking assay. To confirm the assay specificity, the control peptide (AGGAGAGGEA) was used as a reference. As indicated in Figure 6A, with the increase of the SR peptide, the multi- and dimeric forms of N10 decrease, while the monomeric form increases (lanes 3–5), but the control peptide refers no influences on N10 multimerization (lane 2). This result thus showed the possibility that the SR-rich motif might perturb N10 multimerization. To further verify this, a gel-filtration assay was also performed. As clearly shown in Figure 6B, the SR peptide could inhibit N10 multimers formation. In addition, the SR peptide could also perturb the dimerization of the full-length SARS-CoV N protein as demonstrated in Figure 6C.



**FIGURE 6:** Inhibition of the SR peptide against N10 multimerization and SARS-CoV N dimerization. (A) Lane 1, the purified N10 protein (1.25 mg/mL); lanes 3–5, 1.25 mg/mL N10 cultured with different concentrations of the SR peptide (labeled above in molar ratio); lane 2, 1.25 mg/mL N10 cultured with the control peptide (AGGAGAGGEEA) was used as a reference. The final concentration of glutaraldehyde in each sample was 0.05%. (B) Retention volume graphs of the N10 protein (1.25 mg/mL) cultured with different concentrations of the SR peptide. The retention volume graph of the SR peptide was used as a control. (C) Retention volume graphs of SARS-CoV N (5 mg/mL) cultured with 1500  $\mu$ M SR peptide. The retention volume graph of the SR peptide and SARS-CoV N was used as a control.

## DISCUSSION

Self-association of the viral nucleocapsid protein to form dimers or multimers has been well-documented and suggested to be crucial for viral nucleocapsid core formation in viral particle assembly and maturation within a variety of viral systems (9, 12–15). The SARS-CoV N protein was recently reported to self-associate in forming homotypic dimers or multimers by different independent methods (13, 16–18). It is believed that self-association is also the

fundamental function for the SARS-CoV N protein in coating the viral genomic RNA.

Recently, Surjit et al. found that C-terminal 209 amino acids exhibited the capability of interacting with the full-length SARS-CoV N protein (18), whereas He et al. could not detect any interactions between this region and the full-length SARS-CoV N but found that the SR-rich motif was crucial for SARS-CoV N protein oligomerization (13). These results are obviously incompatible and make the SARS-CoV N multimerization study quite confusing. In our present work, we have analyzed the interactions among the different truncated mutants of the SARS-CoV N protein. From yeast two-hybrid assays, the interaction between residues 1–210 ( $\Delta N_N$ ) and the central region (amino acids 211–290) was found, while the  $\Delta N_N/\Delta N_N$  interaction was not detected (Figure 1B). Therefore, on the basis of the yeast two-hybrid assay results and SPR investigation (Figure 2A), it could be proposed that the SR-rich motif has a binding affinity against the central region instead of the SR-rich motif itself, different from the published result (13). Moreover, our results also indicated that the C-terminal fragment 140 residues (amino acids 283–422) of the SARS-CoV N protein are capable of self-association. Such a segment is shorter in sequence length than the reported 209-residues fragment by Surjit et al. (18), which is supposed to be a possible target for mutagenic studies on the disruption of SARS-CoV N protein multimerization.

Furthermore, the chemical cross-linking and gel-filtration results in our work evidently implied that the C-terminal fragment N11 (amino acids 283–422, Scheme 1) of the SARS-CoV N protein could form multimers (Figures 4B and 5B). However, Tang et al. reported that residues 279–370 of SARS-CoV N might self-associate to form dimers (17). This thereby suggests that the segment from residues 371–422 of SARS-CoV N is vital for SARS-CoV N protein multimerization. Moreover, three discrete forms existed in N11, including dimers, trimers, and multimers as shown in Figure 4B, which might have possibly reflected the functional multimeric contacts within SARS-CoV N. Additionally, it is noticed that SARS-CoV N is likely serine-phosphorylated, and phosphorylation was proven to be related to its location in the cytoplasm (20), although the relevance of phosphorylation in the N–N interaction is unclear. The bacterially expressed proteins are presumably not phosphorylated, and proteins expressed in the context of the yeast two-hybrid system may be phosphorylated. In the current work, the good agreement of the yeast two-hybrid system-based experimental analyses with the *E. coli* system-based assays suggests that the phosphorylation issue could be ignored while discussing.

As indicated in parts A and B of Figure 6, the chemical cross-linking and gel-filtration assays clearly showed that the SR peptide could drive the N10 multimers into monomers and make the full-length SARS-CoV N dimers into monomers. These results might explain the fact that N10 (amino acids 211–422) forms multimers for the absence of the SR-rich domain in sequence and lacks the SR-rich domain interaction with the central domain (amino acids 211–290). Accordingly, it is also the SR-rich motif contained in the full-length SARS-CoV N protein that makes the full-length SARS-CoV N prefer to form dimers instead of multimers as reported elsewhere (16–18, 21).

When taken together, we tentatively interpret the multimerization mechanism of the SARS-CoV N protein. It is supposed that during the nucleocapsid assembly process, the SR-rich motif of the SARS-CoV N protein could not interact with the central region (amino acids 211–290), and at this stage, the SARS-CoV N protein might possess the capability of forming multimers through its C-terminal 140-residues (amino acids 283–422) self-association. During the disassembly course, the central region (amino acids 211–290) of the SARS-CoV N protein probably interacts with the SR-rich motif to form dimers. Simultaneously, the C-terminal 140 amino acids could not self-associate to form multimers. Therefore, the SR-rich motif seems to play a pivotal role in SARS-CoV N protein multimerization.

In addition, despite the striking heterogeneity of the SARS-CoV N protein compared with other coronavirus nucleoproteins, there is still a SR-rich motif, a common feature for most coronavirus N proteins (22–26). This motif has been mapped as the RNA-binding domain of the MHV (murine hepatitis virus) N protein (23, 27). As has been discovered, the SARS-CoV N protein contains six SRXX motifs (6, 7). Although there has been no published investigation, the interaction of the SR-rich motif with the genomic RNA could be accepted as a conserved function for the SR-rich motif. On the basis of our current investigation, it is supposed that the SR-rich motif might alternatively associate with the central domain of SARS-CoV N and RNA and probably function in the regulation of SARS-CoV N protein multimerization, even though further studies are necessary to display the exact mechanism for nucleocapsid formation.

In conclusion, in this work, using related biochemical and biophysical techniques, we have discovered that SARS-CoV N might self-associate through the SR-rich motif interaction with the central region and its C-terminal 140-residues itself is able to form multimers. The SARS-CoV N protein prefers to form dimers rather than multimers because of the interaction between the contained SR-rich motif and the central domain. A possible mechanism was hypothesized for illumination of SARS-CoV N protein multimerization, in which the SR-rich motif is proposed to play an important role in SARS-CoV N protein multimerization during its binding to the SARS-CoV N central region or not. This current paper hopefully provided some new ideas in studying SARS-CoV N multimerization.

## REFERENCES

- Peiris, J. S., Chu, C. M., Cheng, V. C., Chan, K. S., Hung, I. F., Poon, L. L., Law, K. I., Tang, B. S., Hon, T. Y., Chan, C. S., Chan, K. H., Ng, J. S., Zheng, B. J., Ng, W. L., Lai, R. W., Guan, Y., and Yuen, K. Y. (2003) Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: A prospective study, *Lancet* 361, 1767–1772.
- Peiris, J. S., Lai, S. T., Poon, L. L., Guan, Y., Yam, L. Y., Lim, W., Nicholls, J., Yee, W. K., Yan, W. W., Cheung, M. T., Cheng, V. C., Chan, K. H., Tsang, D. N., Yung, R. W., Ng, T. K., and Yuen, K. Y. (2003) Coronavirus as a possible cause of severe acute respiratory syndrome, *Lancet* 361, 1319–1325.
- Drosten, C., Gunther, S., Preiser, W., Van der Werf, S., Brodt, H. R., Becker, S., Rabenau, H., Panning, M., Kolesnikova, L., Fouchier, R. A., Berger, A., Burguier, A. M., Cinatl, J., Eickmann, M., Escriviou, N., Grywna, K., Kramme, S., Manuguerra, J. C., Muller, S., Rickerts, V., Sturmer, M., Vieth, S., Osterhaus, A. D., Schmitz, H., and Doerr, H. W. (2003) Identification of a novel coronavirus in patients with severe acute respiratory syndrome, *N. Engl. J. Med.* 348, 1967–1976.
- Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J. A., Lim, W., Rollin, P. E., Dowell, S. F., Ling, A. E., Humphrey, C. D., Shieh, W. J., Guarner, J., Paddock, C. D., Rota, P., Fields, B., DeRisi, J., Yang, J. Y., Cox, N., and Hughes, J. M. (2003) A novel coronavirus associated with severe acute respiratory syndrome, *N. Engl. J. Med.* 348, 1953–1966.
- Ruan, Y. J., Wei, C. L., Ee, A. L., Vega, V. B., Thoreau, H., Su, S. T., Chia, J. M., Ng, P., Chiu, K. P., Lim, L., Zhang, T., Peng, C. K., Lin, E. O., Lee, N. M., Yee, S. L., Ng, L. F., Chee, R. E., Stanton, L. W., Long, P. M., and Liu, E. T. (2003) Comparative full-length genome sequence analysis of 14 SARS coronavirus isolates and common mutations associated with putative origins of infection, *Lancet* 361, 1779–1785.
- Marra, M. A., Jones, S. J., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S., Khattri, J., Asano, J. K., Barber, S. A., Chan, S. Y., Cloutier, A., Coughlin, S. M., Freeman, D., Girm, N., Griffith, O. L., Leach, S. R., Mayo, M., McDonald, H., Montgomery, S. B., Pandoh, P. K., Petrescu, A. S., Robertson, A. G., Schein, J. E., Siddiqui, A., Smailus, D. E., Stott, J. M., Yang, G. S., Plummer, F., Andonov, A., Artsob, H., Bastien, N., Bernard, K., Booth, T. F., Bowness, D., Czub, M., Drebot, M., Fernando, L., Flick, R., Garbutt, M., Gray, M., Grolla, A., Jones, S., Feldmann, H., Meyers, A., Kabani, A., Li, Y., Normand, S., Stroher, U., Tipples, G. A., Tyler, S., Vogrig, R., Ward, D., Watson, B., Brunham, R. C., Krajden, M., Petric, M., Skowronski, D. M., Upton, C., and Roper, R. L. (2003) The Genome sequence of the SARS-associated coronavirus, *Science* 300, 1399–1404.
- Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P., Penaranda, S., Bankamp, B., Maher, K., Chen, M. H., Tong, S., Tamin, A., Lowe, A., Frace, M., DeRisi, J. L., Chen, Q., Wang, D., Erdman, D. D., Peret, T. C., Burns, C., Ksiazek, T. G., Rollin, P. E., Sanchez, A., Liffick, S., Holloway, B., Limor, J., McCaustland, K., Olsen-Rasmussen, M., Fouchier, R., Gunther, S., Osterhaus, A. D., Drosten, C., Pallansch, M. A., Anderson, L. J., and Bellini, W. J. (2003) Characterization of a novel coronavirus associated with severe acute respiratory syndrome, *Science* 300, 1394–1399.
- Lai, M. M., and Cavanagh, D. (1997) The molecular biology of coronaviruses, *Adv. Virus Res.* 48, 1–100.
- Narayanan, K., Kim, K. H., and Makino, S. (2003) Characterization of N protein self-association in coronavirus ribonucleoprotein complexes, *Virus Res.* 98, 131–140.
- Luo, C., Luo, H., Zheng, S., Gui, C., Yue, L., Yu, C., Sun, T., He, P., Chen, J., Shen, J., Luo, X., Li, Y., Liu, H., Bai, D., Shen, J., Yang, Y., Li, F., Zuo, J., Hilgenfeld, R., Pei, G., Chen, K., Shen, X., and Jiang, H. (2004) Nucleocapsid protein of SARS coronavirus tightly binds to human cyclophilin A, *Biochem. Biophys. Res. Commun.* 321, 557–565.
- Luo, H., Chen, Q., Chen, J., Chen, K., Shen, X., and Jiang, H. (2005) The nucleocapsid protein of SARS coronavirus has a high binding affinity to the human cellular heterogeneous nuclear ribonucleoprotein A1, *FEBS Lett.* 579, 2623–2628.
- Kaukinen, P., Koistinen, V., Vapalahti, O., Vaheri, A., and Plyusnin, A. (2001) Interaction between molecules of hantavirus nucleocapsid protein, *J. Gen. Virol.* 82, 1845–1853.
- He, R., Dobie, F., Ballantine, M., Leeson, A., Li, Y., Bastien, N., Cutts, T., Andonov, A., Cao, J., Booth, T. F., Plummer, F. A., Tyler, S., Baker, L., and Li, X. (2004) Analysis of multimerization of the SARS coronavirus nucleocapsid protein, *Biochem. Biophys. Res. Commun.* 316, 476–483.
- Doan, D. N., and Dokland, T. (2003) Structure of the nucleocapsid protein of porcine reproductive and respiratory syndrome virus, *Structure* 11, 1445–1451.
- Wootton, S. K., and Yoo, D. (2003) Homo-oligomerization of the porcine reproductive and respiratory syndrome virus nucleocapsid protein and the role of disulfide linkages, *J. Virol.* 77, 4546–4557.
- Luo, H., Ye, F., Sun, T., Yue, L., Peng, S., Chen, J., Li, G., Du, Y., Xie, Y., Yang, Y., Shen, J., Wang, Y., Shen, X., and Jiang, H. (2004) *In vitro* biochemical and thermodynamic characterization of nucleocapsid protein of SARS, *Biophys. Chem.* 112, 15–25.
- Tang, T. K., Wu, M. P., Chen, S. T., Hou, M. H., Hong, M. H., Pan, F. M., Yu, H. M., Chen, J. H., Yao, C. W., and Wang, A. H. (2005) Biochemical and immunological studies of nucleocapsid proteins of severe acute respiratory syndrome and 229E human coronaviruses, *Proteomics* 5, 925–937.

18. Surjit, M., Liu, B., Kumar, P., Chow, V. T., and Lal, S. K. (2004) The nucleocapsid protein of the SARS coronavirus is capable of self-association through a C-terminal 209 amino acid interaction domain, *Biochem. Biophys. Res. Commun.* 317, 1030–1036.
19. He, R., Leeson, A., Ballantine, M., Andonov, A., Baker, L., Dobie, F., Li, Y., Bastien, N., Feldmann, H., Strocher, U., Theriault, S., Cutts, T., Cao, J., Booth, T. F., Plummer, F. A., Tyler, S., and Li, X. (2004) Characterization of protein–protein interactions between the nucleocapsid protein and membrane protein of the SARS coronavirus, *Virus Res.* 105, 121–125.
20. Surjit, M., Kumar, R., Mishra, R. N., Reddy, M. K., Chow, V. T., Lal, S. K. (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.
21. Yu, I. M., Gustafson, C. L., Diao, J., Burgner, J. W., II, Li, Z., Zhang, J., Chen J. (2005) Recombinant severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein forms a dimer through its C-terminal domain, *J. Biol. Chem.* 280, 23280–23286.
22. Krokhin, O., Li, Y., Andonov, A., Feldmann, H., Flick, R., Jones, S., Strocher, U., Bastien, N., Dasuri, K. V., Cheng, K., Simonsen, J. N., Perreault, H., Wilkins, J., Ens, W., Plummer, F., and Standing, K. G. (2003) Mass spectrometric characterization of proteins from the SARS virus: A preliminary report, *Mol. Cell Proteomics* 2, 346–356.
23. Nelson, G. W., and Stohlman, S. A. (1993) Localization of the RNA-binding domain of mouse hepatitis virus nucleocapsid protein, *J. Gen. Virol.* 74, 1975–1979.
24. Britton, P., Carmenes, R. S., Page, K. W., Garwes, D. J., and Parra, F. (1988) Sequence of the nucleoprotein gene from a virulent British field isolate of transmissible gastroenteritis virus and its expression in *Saccharomyces cerevisiae*, *Mol. Microbiol.* 2, 89–99.
25. Homberger, F. R. (1995) Sequence analysis of the nucleoprotein genes of three enterotropic strains of murine coronavirus, *Arch. Virol.* 140, 571–579.
26. Parker, M. M., and Masters, P. S. (1990) Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein, *Virology* 179, 463–468.
27. Nelson, G. W., Stohlman, S. A., and Tahara, S. M. (2000) High affinity interaction between nucleocapsid protein and leader/ intergenic sequence of mouse hepatitis virus RNA, *J. Gen. Virol.* 81, 181–188.

BI051122C