Carboxyl Terminus of Severe Acute Respiratory Syndrome Coronavirus Nucleocapsid Protein: Self-Association Analysis and Nucleic Acid Binding Characterization†

Haibin Luo,‡ Jing Chen,‡ Kaixian Chen,‡ Xu Shen,*,‡,§ and Hualiang Jiang*,‡,§

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

Received May 11, 2006; Revised Manuscript Received August 10, 2006

ABSTRACT: Coronavirus nucleocapsid (N) protein envelops the genomic RNA to form long helical nucleocapsid during virion assembly. Since N protein oligomerization is usually a crucial step in this process, characterization of such an oligomerization will help in the understanding of the possible mechanisms for nucleocapsid formation. The N protein of severe acute respiratory syndrome coronavirus (SARS-CoV) was recently discovered to self-associate by its carboxyl terminus. In this study, to further address the detailed understanding of the association feature of this C-terminus, its oligomerization was systematically investigated by size exclusion chromatography and chemical cross-linking assays. Our results clearly indicated that the C-terminal domain of SARS-CoV N protein could form not only dimers but also trimers, tetramers, and hexamers. Further analyses against six deletion mutants showed that residues 343−402 were necessary and sufficient for this C-terminus oligomerization. Although this segment contains many charged residues, differences in ionic strength have no effects on its oligomerization, indicating the absence of electrostatic force in SARS-CoV N protein C-terminus self-association. Gel shift assay results revealed that the SARS-CoV N protein C-terminus is also able to associate with nucleic acids and residues 363−382 are the responsible interaction partner, demonstrating that this fragment might involve genomic RNA binding sites. The fact that nucleic acid binding could promote the SARS-CoV N protein C-terminus to form high-order oligomers implies that the oligomeric SARS-CoV N protein probably combines with the viral genomic RNA in triggering long nucleocapsid formation.

Severe acute respiratory syndrome (SARS) is an emerging infectious disease caused by SARS coronavirus (SARS-CoV)[1] (*1*−*4*). SARS-CoV is a novel type of *Coronaviridae* (*5*) and is moderately related to the other known coronaviruses (*6*). It contains several structural proteins, namely, spike protein (S), small envelope protein (E), membrane protein (M), and nucleocapsid protein (N) (*6*, *7*). SARS-CoV N protein is one of the most abundant structural proteins and performs many biological functions. It may bind to the structure M protein, forming a virion (*8*, *9*), and interact with the heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1), playing an important role in the process of mRNA synthesis (*10*, *11*). SARS-CoV N protein inhibits the activity of the cyclin−CDK complex and blocks S phase progression in mammalian cells (*12*). Recently, it was even found to be able to activate the expression of cyclooxygenase-2, resulting in inflammation (*13*).

It is known that coronavirus N protein is able to encapsidate RNA into ribonucleoprotein (RNP) or a long helical nucleocapsid structure (*14*, *15*), and SARS-CoV N protein is suggested to play an essential role in viral RNA packaging (*16*). Self-association of N protein is an important step within virus particle assembly for many viruses (*15*, *17*−*19*). The detailed oligomerization study has revealed that the full-length SARS-CoV N protein is able to form high-order oligomers but exists predominantly as dimers (*20*−*24*). The self-association domain in SARS-CoV N was first mapped out as the C-terminal 209 amino acids using the yeast two-hybrid (*23*). Later this domain was shortened to the C-terminal 138 (residues 285−422) (*20*) and 140 (residues 283−422) amino acids (*25*) by two independent groups. Yu et al. reported that residues 285−422 of SARS-CoV N formed only dimers (*20*), while our previous work revealed that residues 283−422 formed not only dimers but also multimers (*25*). This illegibility needs more exploration to clarify. In addition, NMR structural analysis and the crystal structure revealed that parts of the C-terminus, residues 248−365 and residues 270−370, are able to form dimers (*26*, *27*); however, the relevant oligomeric property of the whole C-terminus is still poorly characterized, and the necessary sequence for this oligomerization remains unknown. In this

† This work was supported by the State Key Program of Basic Research of China (Grant 2004CB58905), the National Natural Science Foundation of China (Grants 30525024 and 20372069), and the Sino-European Project on SARS Diagnostics and Antivirals (Proposal/Contract No. 003831).

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

‡ Chinese Academy of Sciences.

§ East China University of Science and Technology.

[1] Abbreviations: SARS-CoV, severe acute respiratory syndrome coronavirus; SARS-CoV N, the nucleocapsid protein of SARS-CoV; aa, amino acid.

10.1021/bi0609319 CCC: $33.50
Published on Web 09/12/2006

work, the oligomerization of the SARS-CoV N carboxyl terminus was systematically investigated by various biophysical and biochemical analyses. Size exclusion chromatography and chemical cross-linking assays revealed that the C-terminal 140 amino acids (residues 283−422, $N_{283-422}$) were able to form dimers, trimers, tetramers, and hexamers depending on the protein's concentration. Analysis of SARS-CoV N protein deletion mutants further identified that residues 343−402 are important for $N_{283-422}$ oligomerization. Although this fragment contains many charged residues, changes in ionic strength could not affect its oligomerization, which thus indicates that the self-association of the C-terminal domain of SARS-CoV N protein is not related to electrostatic forces. Additionally, DNA band shift assays revealed that $N_{283-422}$ is able to interact with nucleic acids, and the nucleic acid binding region was mapped as residues 363−382. The fact that ssDNA binding could promote $N_{283-4222}$ to form high-order oligomers implies that it is possible that the oligomeric SARS-CoV N protein might combine with viral genomic RNA to generate higher order oligomers, which in turn triggers the formation of the long nucleocapsid structure.

EXPERIMENTAL PROCEDURES

Plasmid Construction. For construction of the related truncated SARS-CoV N proteins, the coding sequences for the C-terminal 283−422, 303−422, 283−402, 283−382, 263−362, 243−342, and 343−422 residues of SARS-CoV N proteins were amplified from the full-length SARS-CoV N gene with the following primers: $N_{283-422}$ sense 5′-AATT*GGATCC*ACCCAAGGAAATTTCGGGGACCAA-3′, $N_{283-422}$ antisense 5′-ACGG*GTCGAC*TTATGCCTGAGTTGAATCAGCA-3′, $N_{303-422}$ sense 5′-AATT*GGATCC*CCGCAAATTGCACAATTTG-3′, $N_{303-422}$ antisense 5′-GGC*CAAGCT*TTTATGCCTGAGTTGAATC-3′, $N_{283-402}$ sense 5′-ACGT*CATATG*ACCCAAGGAAATTTCGGGGACCA-3′, $N_{283-402}$ antisense 5′-TCAA*GGATCC*TTAATCCATGTCAGCCGCAGGAA-3′, $N_{283-382}$ sense 5′-ACGT*CATATG*ACCCAAGGAAATTTCGGGGACCA-3′, $N_{283-382}$ antisense 5′-AATT*GGATCC*TTAAGGCTGAGCTTCATCAGTCT-3′, $N_{263-362}$ sense 5′-CCTT*CATATG*CGTACTGCCACAAAAC-3′, $N_{263-362}$ antisense 5′-CCGT*GGATCC*TTATTTGTATGCGTCAA-3′, $N_{243-342}$ sense 5′-AATT*GGATCC*CAAGGCCAAACTGTCAC-3′, $N_{243-342}$ antisense 5′-GGCC*GAATTC*TCAGTCATCCAATTTAATG-3′, $N_{343-422}$ sense 5′-ACCG*GGATCC*AAAGATCCACAATTCAAAG-3′, and $N_{343-422}$ antisense 5′-GGCC*GAATTC*TTATGCCTGAGTTGAATC-3′. The PCR products for $N_{283-422}$ were digested by *Bam*HI and *Sal*I and ligated into pQE30 expression vectors (Qiagen). The PCR products for $N_{303-422}$ were digested by *Bam*HI and *Hin*dIII and ligated into pET28a expression vectors (Novagen). The PCR products for $N_{283-402}$, $N_{283-382}$, and $N_{263-362}$ were digested by *Nde*I and *Bam*HI and ligated into pET15b expression vectors (Novagen). The PCR products for $N_{243-342}$ and $N_{343-422}$ were digested by *Bam*HI and *Eco*RI and ligated into pGEX4T-1 expression vectors (Amersham Parmacia Biotech). The constructs were identified by PCR reactions with the corresponding primers and digested with the restriction enzymes. The sequences of all the constructs were further confirmed by DNA sequencing. All of the constructs were transformed into DH5α competent cells and kept in 20% glycerol at −80 °C.

Expression and Purification of His-Tagged Proteins. The C-terminal domain ($N_{283-422}$) of SARS-CoV N protein was cloned into a pQE30 vector with an N-terminal His_6-tag and expressed under the control of the T5 promoter. The constructed pQE30−$N_{283-422}$ recombinant plasmid was transformed into *Escherichia coli* strain M15 and inoculated into 10 mL of LB medium containing 100 μg/mL ampicillin and 50 μg/mL kanamycin at 37 °C. The 10 mL culture was transferred into 1 L of LB medium with the same antibiotics 10 h later, and the culture was grown at 37 °C until the OD_{600} (optical density at 600 nm) reached 0.8. IPTG (isopropyl β-D-thiogalactopyranoside) was then added to a final concentration of 1.0 mM, and the induction was allowed to proceed at 37 °C for 10 h. The *E. coli* cells were harvested by centrifugation at 8000*g* for 10 min and stored at −80 °C.

The recombinant plasmids pET28a−$N_{303-422}$, pET15b−$N_{283-402}$, pET15b−$N_{283-382}$, and pET15b−$N_{263-362}$ were transformed into *E. coli* strain BL21(DE3)pLysS. A single colony was inoculated into 10 mL of LB medium supplemented with ampicillin (final concentration 100 μg/mL). The 10 mL culture was transferred into 1 L of LB medium with the same antibiotics 10 h later and grown at 37 °C with vigorous shaking. When the OD_{600} reached 0.7, the expression of His_6-tag fusion protein was induced by the addition of IPTG to a final concentration of 0.5 mM, followed by 5 h of shaking at 25 °C. The *E. coli* cells were harvested by centrifugation at 8000*g* for 10 min and stored at −80 °C.

All the His_6-tag fusion proteins were purified by Ni−NTA affinity chromatography according to similar protocols. In brief, cell pellets from 0.5 L of culture were suspended in 20 mL of sonication buffer (20 mM Tris−HCl buffer at pH 8.0, 500 mM NaCl, 10 mM imidazole, 20 μM PMSF (phenylmethylsulfonyl fluoride)) and disrupted by sonication for 30 min in an ice bath. Soluble and insoluble fractions were separated by centrifugation for 45 min at 14000 rpm and 4 °C. The supernatant was incubated with preequilibrated Ni−NTA resin (Qiagen) in sonication buffer by rotating constantly (15 rpm) at 4 °C for 3 h. The unbound proteins were removed by washing with 10 mL of sonication buffer. The column was washed with 30 mL each of washing buffers I (20 mM Tris−HCl buffer at pH 8.0, 500 mM NaCl, 60 mM imidazole) and II (20 mM Tris−HCl buffer at pH 8.0, 500 mM NaCl, 100 mM imidazole). The protein of interest was then eluted with elution buffer (20 mM Tris−HCl buffer at pH 8.0, 500 mM NaCl, 500 mM imidazole). Fractions were collected, and the recombinant protein in the eluted fractions was confirmed by 10% (w/v, the mass concentration of the acrylamide/bisacrylamide (29:1) mixture in the separating gel) SDS−PAGE analysis.

Expression and Purification of GST Fusion Proteins. The expression and purification procedures for GST-tag fusion proteins were carried out according to instructions from Amersham Pharmacia Biotech with some modifications. The constructs of pGEX4T-1−$N_{243-342}$ and pGEX4T-1−$N_{343-422}$ were transformed into BL21(DE3), and the transformants were grown at 37 °C overnight in LB medium in the presence of ampicillin (100 μg/mL). Bacterial cultures were diluted with fresh LB medium (plus 100 μg/mL ampicillin) and incubated for 4 h at 37 °C. For 7 h with the induction of IPTG (final concentration at 0.8 mM) at 30 °C, the bacterial cultures were collected by centrifugation, resuspended in PBS

buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4) containing 20 μM PMSF, and lysed by sonication for 30 min in an ice bath. The cell lysates were clarified by centrifugation at 14000 rpm at 4 °C for 60 min. The supernatant was mixed with 3 mL of PBS preequilibrated glutathione−Sepharose beads (Amersham Pharmacia Biotech) on a rotating bed at 4 °C for 3 h. A 100 mL sample of prechilled PBS buffer was used to wash the beads to eliminate the loose-binding contaminants. The bound proteins were then cleaved on-column to get rid of GST-tag by 5 units of thrombin (Amersham Pharmacia Biotech) in 5 mL of PBS buffer at 16 °C for 12 h. The digested products were then further purified by size exclusion chromatography (Hiload 16/60 Superdex 75). Fractions were collected, and the presence of recombinant protein in the eluted fractions was confirmed by 10% SDS−PAGE analysis.

Size Exclusion Chromatography Based Assay. The oligomeric features of the purified fragments of SARS-CoV N were analyzed by size exclusion chromatography using HiLoad 16/60 Superdex 75 (separation range 3−70 kDa) and Hiload 16/60 Superdex 200 (separation range 10−600 kDa) on an AKTA FPLC platform (Amersham Pharmacia Biotech). The column was equilibrated at a flow rate of 1 mL/min with PBS buffer at room temperature. The protein sample (500 μL) was injected at a given concentration and detected by 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 Pharmacia Biotech. The $N_{283-422}$ oligomerization was studied according to the procedure modified from that of HIV Vpr (*28*).

Chemical Cross-Linking Assays. All the tested mutants were purified by gel filtration column and digested by DNAase and RNAase to get rid of the possible binding nucleic acids. Glutaraldehyde (25%, Sigma) was diluted to a series of concentrations (1%, 2%, 4%, 6%, 8%, 10%) by distilled water. The proteins in PBS buffer were reacted with glutaraldehyde on an ice bath for 20 min. Dimethyl pimelimidate dihydrochloride (DMP) was resolved in distilled water at 300 mM. Samples were exchanged into PBS buffer and treated with 0−30 mM DMP for 60 min at room temperature. The reactions were quenched by an equal volume of SDS−PAGE loading buffer, followed by heating at 99 °C for 5 min. To evaluate the effects of ionic strength differences on $N_{283-422}$ oligomerization, the purified protein was prepared by different sets in 10 mM HEPES buffer at pH 7.4 containing different salts (NaCl, KCl, and $CaCl_2$).

Native PAGE Assays. To investigate the molecular mass of $N_{283-422}$ in the native state, native polyacrylamide gel electrophoresis (native PAGE) was performed in 5% stacking gel (pH 9.5) and 10% separating gel (pH 9.5). The purified protein sample was electrophoresed at 120 V for 3 h. After electrophoresis, the gel was stained with Coomassie Brilliant Blue R-250 and then destained. The molecular mass of $N_{283-422}$ was further estimated according to the following protein standards: bovine serum albumin (67 kDa), ovalbumin (45 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa).

Gel Shift Assays. To test whether the purified proteins contain nucleic acids, the samples were electrophoresed on 0.8% agarose gel in TAE buffer (40 mM Tris−acetate buffer at pH 8.0, 1 mM EDTA) with the presence of ethidium bromide (0.5 μg/mL final concentration) at low voltage (3 V/cm). The electrophoretic result was directly visualized using a short-wavelength UV transilluminator for nucleic acid detection, and the same gel was then stained with Coomassie Blue R-250 for protein visualization. In the investigation of the nucleic acid binding activity for $N_{283-422}$, the purified proteins were digested with DNAase and RNAase to get rid of the possible binding nucleic acids. ssDNA (single-strand DNA, 300−400-mer) was purchased from Sigma and resolved in doubly distilled water (2 mg/mL) as a stock solution. The RNA was extracted from SD rat liver cells with a Trizole kit (Sangon). The purified proteins without nucleic acids (estimated by OD_{260}/OD_{280}) were prepared at 1 mg/mL in PBS buffer and then mixed with ssDNA or RNA at a molar ratio from 2:1 to 1:3 (protein:nucleic acid). These mixtures were incubated at 4 °C for 1 h before analysis by 0.8% agarose gel.

RESULTS AND DISCUSSION

Oligomerization Characterization of the SARS-CoV N Protein C-Terminus ($N_{283-422}$). The oligomerization of the full-length SARS-CoV N protein has been studied widely (*20*−*24*), and its carboxyl terminal was proved to be necessary for its self-association (*20*, *23*). It was discovered previously that the C-terminal domain is able to form multimers (*25*), but it was not clear how high the formed oligomers could be, and the primary sequence for the oligomerization remained unclear. In this work, to investigate the oligomeric properties of the C-terminal domain (residues 283−422, $N_{283-422}$), the gel filtration (Figure 1), chemical cross-linking, and native gel electrophoresis technology based assays (Figure 2) were applied.

In the gel filtration assay, $N_{283-422}$ samples at concentrations ranging from 8 to 469 μM were applied to a Hiload Superdex 200 column and eluted by 180 mL of PBS buffer. A subset of these elution profiles is shown in Figure 1A. It is noticed that, for each case, a high peak (around 46 mL) always appears close to the void volume. On the basis of their large Stokes radii and invariant behaviors, we assumed that this peak probably corresponds to a large, inert aggregate of $N_{283-422}$. As also indicated in Figure 1A, at low concentration (8 μM), $N_{283-422}$ protein was eluted in two separated peaks. One is close to the void volume, and the other is at 83 mL, which is attributed to a dimer on the basis of the standard curve (Figure 1B). Interestingly, with an increase of the protein concentration to 15, 31, and 44 μM, the corresponding retention volume for the second peak decreased (Figure 1) corresponding to a trimer as determined from the standard curve in Figure 1B. At higher concentrations ($\geq$55 μM), such a peak shifted to a tetrameric status (Figure 1). However, at even higher concentrations ($\geq$75 μM), the protein sample started to produce a third peak around 66 mL apart from retaining the second peak characteristic of the tetramers. By comparison with the standard curve (Figure 1B), this third peak was attributed to the molecular size of a hexamer. As the concentration continuously increased (from 166 to 469 μM), the peak characteristic of a hexamer increased significantly in concentration, while the tetrameric peak increased slowly (Figure 1), which implied that one tetramer associated with two

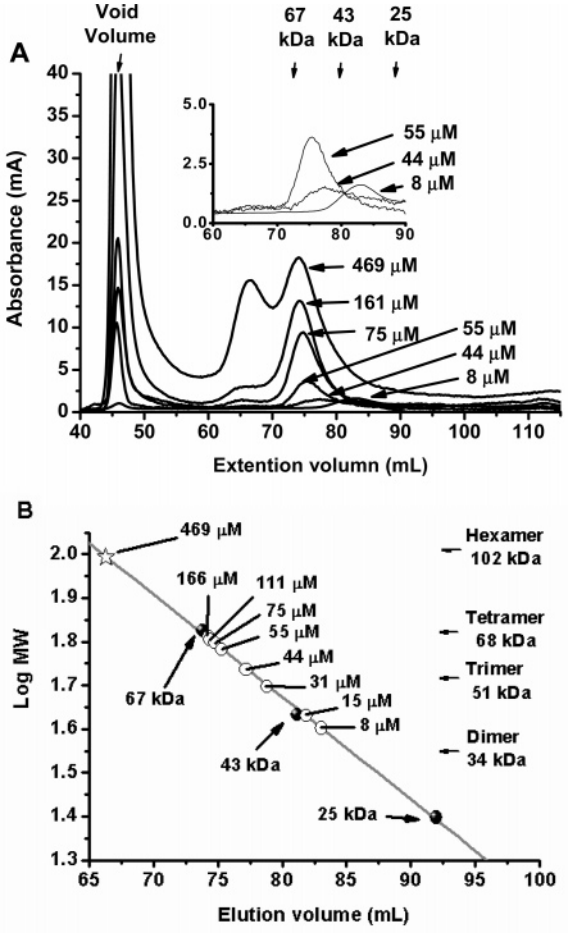

FIGURE 1: Size exclusion chromatography based analysis of $N_{283-422}$. (A) Elution profiles of pure $N_{283-422}$ on Superdex 200 at the indicated concentrations. (B) Standard curve generated by protein markers showing the molecular masses of $N_{283-422}$ at the indicated concentrations. The molecular mass standards are indicated as follows: bovine serum albumin, 67 kDa; ovalbumin, 45 kDa; chymotrypsinogen A, 25 kDa; ribonuclease A, 13.7 kDa.

monomers or one dimer into a hexamer, and the tetramer and hexamer might be the two concomitant components during the virion assembly.

In the chemical cross-linking assay, if two proteins physically interact with each other, they can be covalently cross-linked by using bifunctional reagents containing reactive end groups that react with functional groups (primary amines or sulfhydryls). The formation of cross-links could be analyzed by the common SDS−PAGE, and it is direct and convincing evidence of the existence of oligomers. Here the cross-linking experiments were performed using glutaraldehyde, a short self-polymerizing reagent mostly reacting with the ϵ-amino groups of lysine, and DMP, a more specific reagent which cross-links lysine residues at shorter distances. As shown in Figure 2A,B, the recombinant $N_{283-422}$ migrated as a single band (lane 2) with a relative molecular mass around 17 kDa. The cross-linking experiments by increasing the concentration of glutaraldehyde and DMP showed that $N_{283-422}$ was able to form dimers, trimers, and tetramers (Figure 2A,B). Furthermore, the native electrophoresis assay also detects the presence of trimeric and tetrameric $N_{283-422}$ (Figure 2C). Accordingly, these data indicated that $N_{283-422}$ could form dimers, trimers, and tetramers in solution in a concentration-dependent manner.

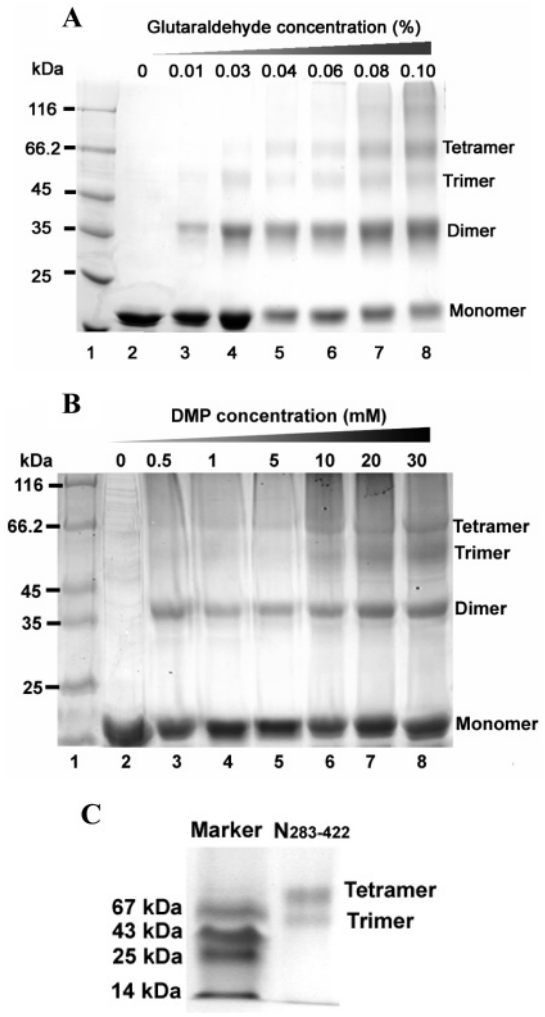

FIGURE 2: Chemical cross-linking and native gel electrophoresis analysis of $N_{283-422}$. (A) Cross-linking results of $N_{283-422}$ with glutaraldehyde. $N_{283-422}$ was incubated with the indicated concentrations of glutaraldehyde for 20 min on an ice bath. (B) Cross-linking results of $N_{283-422}$ with DMP. $N_{283-422}$ was incubated with the indicated concentrations of DMP for 20 min at room temperature. The cross-linked samples were run on 8% SDS−PAGE and visualized by Coomassie Blue staining. Cross-linking intermediates representing monomer, dimer, trimer, and tetramer forms of $N_{283-422}$ are indicated. The values on the left are molecular masses of the protein markers. (C) Analysis of the oligomeric state of $N_{283-422}$ by native PAGE assay. A 40 μg sample of protein was electrophoresed and stained with Coomassie Blue R-250. Protein standards used for molecular mass evaluation were run on the same gel.

Mapping the Oligomeric Domain of the SARS-CoV N Protein Carboxyl Terminus. To determine the primary sequence requirements for $N_{283-422}$ self-association, six truncated fragments of SARS-CoV N protein (Figure 3A) were expressed in an *E. coli* strain and purified by affinity chromatography. $N_{303-422}$, $N_{283-402}$, $N_{283-382}$, and $N_{263-362}$ were purified by a His_6-tag constructed at the N-terminus. $N_{243-342}$ and $N_{343-422}$ were engineered as GST-tag fusion proteins to aid expression and purification. The GST-tag was removed by thrombin (protease) digestion before further analysis. The oligomeric status of each fragment was probed at first by size exclusion chromatography (Figure 3B). Every tested sample was prepared in PBS buffer at 150 μM followed by analysis on an FPLC platform using a HiLoad Superdex 75 column. $N_{303-422}$ and $N_{283-402}$ are able to form trimers and tetramers (Figure 3B), indicating that residues

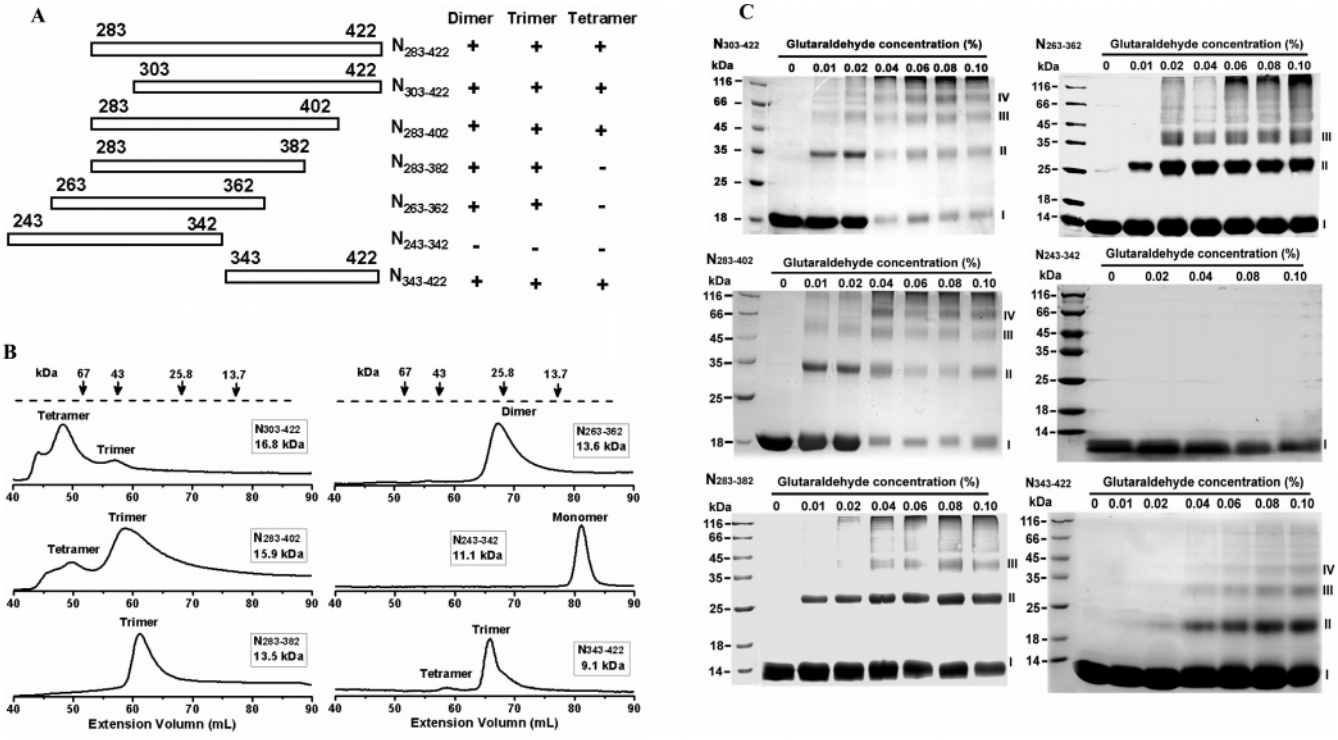

FIGURE 3: Mapping of the oligomeric domain in SARS-CoV N protein carboxyl terminus. (A) Schematic description of the truncated mutants and summary of their oligomeric activities determined by chemical cross-linking assay. (B) Size exclusion chromatography based analyses for the six truncated mutants. All retention volume graphs were obtained on a Superdex 75 column. Pure proteins (2 mL) at 150 μM protein were loaded and eluted by PBS buffer at room temperature. (C) Chemical cross-linking results of the truncated mutants. The proteins were reacted with glutaraldehyde on an ice bath for 20 min. After the cross-linking reaction was quenched, the cross-linked samples were analyzed by 10% SDS−PAGE. The letters on the right indicate the following: I, monomer; II, dimer; III, trimer; IV, tetramer.

283−302 and 403−422 might not be involved in the oligomerization of the SARS-CoV N protein C-terminus. However, the removal of a further 20 amino acids from the C-terminus apparently decreased the ability of the protein to oligomerize, for $N_{283-382}$ was eluted only as a single peak corresponding to a trimer. To further orientate the oligomeric domain, it is necessary to generate and analyze a shorter fragment. However, shorter constructs are too small to express and purify. Thus, two other truncated fragments, $N_{263-362}$ and $N_{243-342}$ (Figure 3A), were prepared by N-terminal extension of 20 and 40 amino acids, respectively. As shown in Figure 3B, $N_{263-362}$ formed only dimers as reported (*22*, *26*). Remarkably, a further C-terminal 20-residue truncation caused the protein to completely lose its oligomeric ability by showing that the $N_{243-342}$ fragment formed only monomers. Hereunto, gel filtration results suggested that the fragment of residues 343−402 is likely the oligomeric domain. To further confirm this, a fusion protein containing this domain, $N_{343-422}$ (Figure 3A), was prepared, and its oligomeric status was also studied by gel filtration assay. The fact that $N_{343-422}$ was able to form trimers and tetramers indicated that residues 343−402 are necessary and sufficient for the oligomerization of the SARS-CoV N protein C-terminus.

Being a widely used approach in studying protein oligomerization, chemical cross-linking was also used to investigate the above six truncated fragments. During the assay, the purified proteins were incubated with increasing concentrations of glutaraldehyde for 20 min in an ice bath. Cross-linked samples were run on 10% SDS−PAGE and visualized by staining with Coomassie Blue. As shown in Figure 3A,C, the fragments $N_{303-422}$ and $N_{283-402}$ existed as monomers, dimers, trimers, and tetramers, $N_{283-382}$ and $N_{263-362}$ were able to form dimers and trimers, $N_{243-342}$ existed as only monomers, and $N_{343-422}$ had the ability to form dimers, trimers, and tetramers. These results indicated that residues 343−402 are crucial and sufficient for the oligomerization of the SARS-CoV N protein C-terminus, in agreement with the results revealed by the gel filtration assay. Furthermore, these results have suggested that residues 343−362 and 383−402 are responsible for the formation of dimers and trimers, and tetramers, respectively. It is noticed that the gel filtration results showed $N_{263-362}$ forms only dimers, while the chemical cross-linking results revealed it forms both dimers and trimers. The trimers' presence in the cross-linking assay might be due to the high reactive activity of glutaraldehyde as a linker.

A Change in Ionic Strength Has No Effect on $N_{283-422}$ Oligomerization. As has been reported (*7*), the C-terminus of SARS-CoV N protein ($N_{283-422}$) contains a number of charged residues, which leads us to assume that SARS-CoV N protein probably self-associates by electrostatic forces. If so, the differences in ionic strength will affect the protein's oligomeric activity. To evaluate this assumption, the oligomeric behaviors of $N_{283-422}$ at different ionic strengths were studied by chemical cross-linking assay, in which the purified $N_{283-422}$ proteins were cross-linked in HEPES buffer containing monovalent and divalent chloride salts. As shown in Figure 4, $N_{283-422}$ showed almost the same oligomeric activity at different ionic strengths, implying the absence of electrostatic forces in the oligomerization of SARS-CoV N protein.

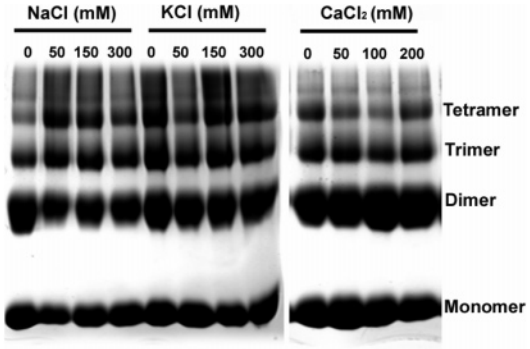

FIGURE 4: Effect of ionic strength on the oligomeric activity of $N_{283-422}$. The pure $N_{283-422}$ was prepared in different sets of 10 mM HEPES buffer at pH 7.4 each containing different salts (NaCl, KCl, and $CaCl_2$). These samples reacted with 0.03% glutaraldehyde on an ice bath for 30 min. After the cross-linking reaction was quenched, the cross-linked samples were analyzed by 10% SDS−PAGE.

$N_{283-422}$ Has the Ability to Associate with Nucleic Acids. As discussed above, a high peak was always observed close to the void volume when pure $N_{283-422}$ was analyzed by gel filtration assay (Figure 1A). On the basis of their large Stokes radii and invariant behavior, we think these complexes probably represent large aggregates of $N_{283-422}$. The absorbance in the ultraviolet light range indicated that the OD_{280}/OD_{260} value of this peak was 0.86, indicating the presence of a high concentration of nucleic acids. To confirm the presence of nucleic acids, the sample was further subjected to gel shift assays, similar to the established methods for analysis of RNA binding to protein (*20*, *22*, *28*). During the assay, the samples were electrophoresed on an 0.8% agarose gel in the presence of ethidium bromide. The nucleic acids in the gel were visualized by ultraviolet excitation. As shown in Figure 5A, the sample apparently contained a large amount of nucleic acids. The same gel was then stained with Coomassie Blue for checking protein migration. The p*I* (isoelectric point) of $N_{283-422}$ is 9.5, and this protein could not run into the gel due to its positively charged state at pH 8.0. In our gel shift assay, however, the proteins and nucleic acids were found at the same position on the gel, indicating that the interactions with negatively charged nucleic acids help the protein run into the gel. These results thus suggested the possibility that $N_{283-422}$ had the ability to associate with nucleic acids, which further promotes $N_{283-422}$ to form high-order oligomers.

Recently, it was reported that the recombinant SARS-CoV N protein binds to nonspecific DNA and RNA (*20*, *22*). To investigate whether the N protein C-terminus is involved in the association with nucleic acids, the possible interactions between $N_{283-422}$ protein and DNA as well as RNA were studied by gel shift assays. During the experiments, a constant amount of $N_{283-422}$ protein was mixed with nucleic acids at different molar ratios. As shown in Figure 5B, the ssDNA and RNA itself migrated farther away from the loading well, while the ssDNA/protein and RNA/protein mixtures migrated a much shorter distance. Meanwhile, such a migration distance (Figure 5B) increased with the amount of ssDNA and RNA, which indicated that this association with ssDNA or RNA allowed $N_{283-422}$ to run further on the gel. However, this effect was not seen in BSA control assays. Accordingly, our results clearly indicated that the C-terminus ($N_{283-422}$) of SARS-CoV N protein also has the ability to

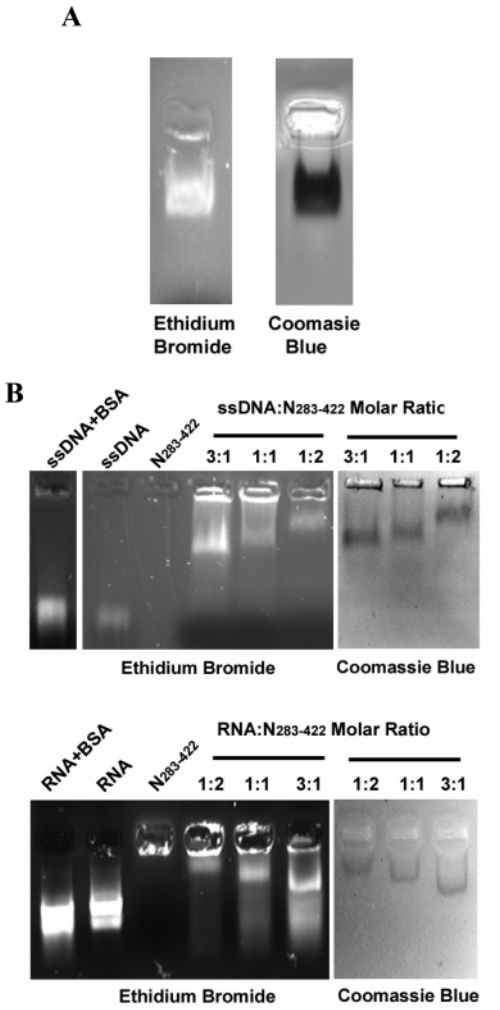

FIGURE 5: $N_{283-422}$ is able to interact with nucleic acids. (A) The peak at the void volume contains protein and a high concentration of nucleic acids. (B) The $N_{283-422}$ interaction with ssDNA and RNA was studied by gel shift assay.

associate with nucleic acids without sequence specificity, apart from the reported fact that the SARS-CoV N protein N-terminus (residues 45−181) could interact with RNA (*30*). Meanwhile, our result has supported the recent report that both N-terminal and C-terminal regions of SARS-CoV N protein have binding activities to the viral RNA (*16*).

ssDNA's Association Promotes $N_{283-422}$ to Form High-Order Oligomers. To further confirm whether the association with nucleic acids can promote $N_{283-422}$ to form high-order oligomers, the above ssDNA/protein complexes (DNA is more stable than RNA) were analyzed by chemical cross-linking (Figure 6A) and gel filtration assays (Figure 6B). The corresponding samples without cross-linking were also tested by SDS−PAGE as references. In the absence of ssDNA as shown in Figure 6A, the cross-linked $N_{283-422}$ existed as monomers, dimers, trimers, and tetramers. As the ssDNA concentration increased, the concentrations of monomers, dimers, trimers, and tetramers decreased, which suggested that the binding with ssDNA might promote $N_{283-422}$ to form high-order oligomers. Such a result could also be confirmed by gel filtration assays as indicated in Figure 6B. The pure $N_{283-422}$ was eluted as a single peak corresponding to tetramers. When ssDNA was mixed with $N_{283-422}$ at different molar ratios, the peak close to the void

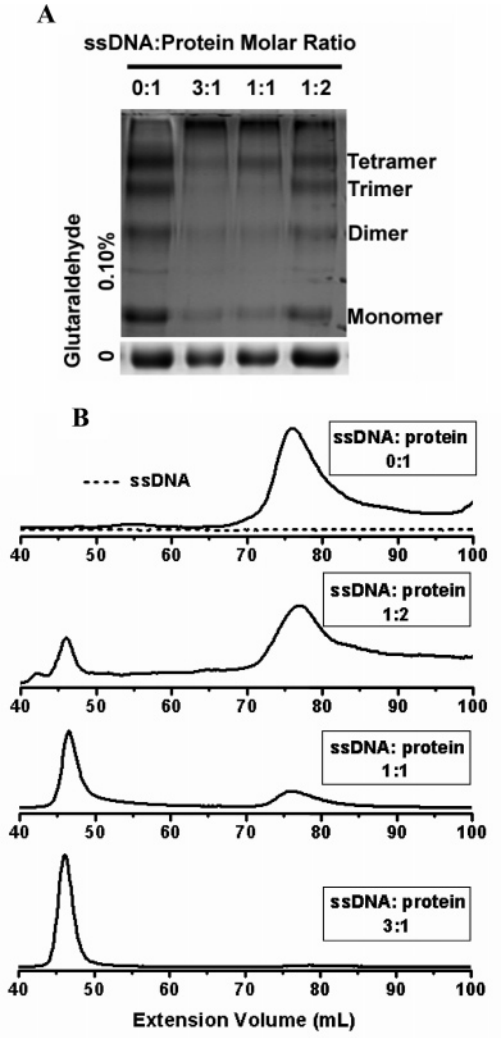

FIGURE 6: ssDNA binding helps $N_{283-422}$ form high-order oligomers. (A) Chemical cross-linking results of ssDNA/$N_{283-422}$ complexes. Those complexes without glutaraldehyde cross-linking were also tested by SDS−PAGE and used as controls. (B) Gel filtration analysis of ssDNA/$N_{283-422}$ complexes.

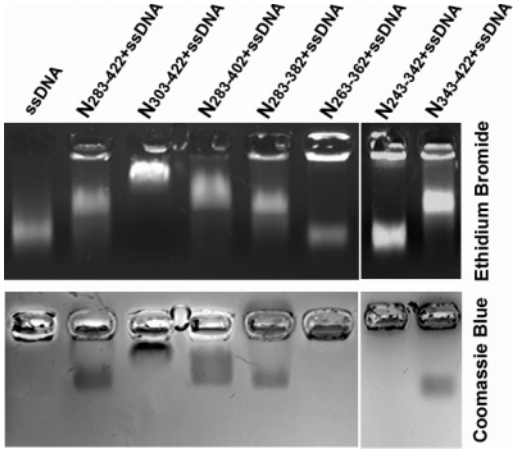

FIGURE 7: Mapping of the nucleic acid binding domain by DNA gel shift assays. The protein sample was mixed with ssDNA at a molar ratio of 1:3. The samples were electrophoresed on 0.8% agarose gel in TAE buffer with low voltage. The electrophoretic result was directly visualized using a UV transilluminator for nucleic acid detection, and the same gel was then stained with Coomassie Blue R-250 for protein visualization.

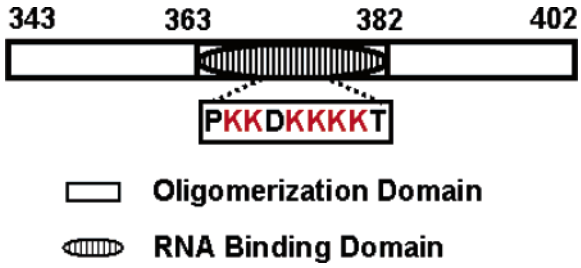

FIGURE 8: Summary of the oligomeric domains and nucleic acid binding domain in the carboxyl terminus of SARS-CoV N protein.

volume appeared and increased with the ssDNA concentration. In contrast, the tetrameric peak decreased with an increase of ssDNA and almost disappeared at a 3:1 (ssDNA: $N_{283-422}$) molar ratio. Hence, these results indicated that ssDNA binding could indeed promote the $N_{283-422}$ to form high-order oligomers. Actually, many viral proteins were reported to interact with nucleic acid to form high-order complexes, for instance, nucleocapsid protein of chandipura virus (*31*), nucleocapsid protein of vesicular stomatitis virus (*32*), the VP 40 of Marburg virus (*33*), and the nucleocapsid protein of sindbis virus (*29*).

Residues 363−382 of the SARS-CoV N Protein Carboxyl Terminus are Required for Nucleic Acid Binding. To address which regions in the carboxyl terminus of SARS-CoV N protein were required for nucleic acid binding, six deletion mutants were assayed by DNA band shift assays. The p*I* values for $N_{303-422}$, $N_{283-402}$, $N_{283-382}$, $N_{263-362}$, $N_{243-342}$, and $N_{343-422}$ are predicted to be 9.9, 9.7, 9.5, 9.8, 10.3, and 9.8, respectively. As shown in Figure 7, $N_{303-422}$, $N_{283-402}$, and $N_{283-382}$ could associate with ssDNA and migrate into the gel; however, the removal of a further 20 and 40 amino acids, generating $N_{263-362}$ and $N_{243-343}$ (Figure 3A), caused the proteins to lose the capability of binding to ssDNA. The ability of interacting with ssDNA was restored for $N_{343-422}$ (Figure 3A), as $N_{283-422}$. These results thus indicated that residues 363−382 are responsible for nucleic acid binding of the SARS-CoV N protein carboxyl terminus. As demonstrated in Figure 8, such a region consists of six lysines and is likely to be involved in nonspecific interactions with the RNA through charge neutralization.

In conclusion, in this work we have completely discovered that the C-terminus of SARS-CoV N protein (residues 283−422) could form dimers, trimers, tetramers, and hexamers in a concentration-dependent manner. The crucial sequence for such oligomerization is determined to be residues 343−402 (Figure 8). Although the C-terminus of SARS-CoV N protein contains many charged residues, electrostatic force was not involved in its oligomerization. Furthermore, it is found that the SARS-CoV N protein C-terminus was able to associate with nucleic acids, and the interaction domain was mapped out as the residues 363−382, in which the six lysines are supposed to be responsible for its RNA binding (Figure 8). Moreover, ssDNA could incorporate $N_{283-422}$ tetramers into high-order oligomers. Taken together, these data suggested that the overlapping oligomerization and nucleic acid binding domain, i.e., residues 343−402, might be the most critical sequence for nucleocapsid structure formation. Since much of the driving force for nucleocapsid assembly comes from protein−protein interactions and

protein−RNA interactions involving the C-terminus of SARS-CoV N protein, a hypothesis was proposed for the nucleocapsid assembly. Self-association of SARS-CoV N protein to form oligomers, especially tetramers, is the first step of the assembly process. Then these oligomers continuously bind to the genomic RNA in triggering the formation of a long nucleocapsid structure.

REFERENCES

1. 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.
2. 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.
3. 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., Burguiere, A. M., Cinatl, J., Eickmann, M., Escriou, 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.
4. 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.
5. 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.
6. Marra, M. A., Jones, S. J., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S., Khattra, J., Asano, J. K., Barber, S. A., Chan, S. Y., Cloutier, A., Coughlin, S. M., Freeman, D., Girn, 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.
7. 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, L., 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.
8. Fang, X., Ye, L., Timani, K. A., Li, S., Zen, Y., Zhao, M., Zheng, H., and Wu, Z. (2005) Peptide domain involved in the interaction between membrane protein and nucleocapsid protein of SARS-associated coronavirus, *J. Biochem. Mol. Biol. 38*, 381−385.
9. Luo, H., Wu, D., Shen, C., Chen, K., Shen, X., and Jiang, H. (2006) Severe acute respiratory syndrome coronavirus membrane protein interacts with nucleocapsid protein mostly through their carboxyl termini by electrostatic attraction, *Int. J. Biochem. Cell Biol. 38*, 589−599.
10. 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.
11. Jiang, X. S., Tang, L. Y., Dai, J., Zhou, H., Li, S. J., Xia, Q. C., Wu, J. R., and Zeng, R. (2005) Quantitative analysis of severe acute respiratory syndrome (SARS)-associated coronavirus-infected cells using proteomic approaches: implications for cellular responses to virus infection, *Mol. Cell Proteomics 4*, 902−913.
12. Surjit, M., Liu, B., Chow, V. T., and Lal, S. K. (2006) The nucleocapsid protein of sars-coronavirus inhibits the activity of cyclin-CDK complex and blocks S phase progression in mammalian cells, *J. Biol. Chem. 281*, 10669−10681.
13. Yan, X., Hao, Q., Mu, Y., Timani, K. A., Ye, L., Zhu, Y., and Wu, J. (2006) Nucleocapsid protein of SARS-CoV activates the expression of cyclooxygenase-2 by binding directly to regulatory elements for nuclear factor-kappa B and CCAAT/enhancer binding protein, *Int. J. Biochem. Cell Biol.* (in press).
14. Lai, M. M. and Cavanagh, D. (1997) The molecular biology of coronaviruses, *Adv. Virus Res. 48*, 1−100.
15. Narayanan, K., Kim, K. H., and Makino, S. (2003) Characterization of N protein self-association in coronavirus ribonucleoprotein complexes, *Virus Res. 98*, 131−140.
16. Hsieh, P. K., Chang, S. C., Huang, C. C., Lee, T. T., Hsiao, C. W., Kou, Y. H., Chen, I. Y., Chang, C. K., Huang, T. H., and Chang, M. F. (2005) Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent, *J. Virol. 79*, 13848−13855.
17. 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.
18. Doan, D. N., and Dokland, T. (2003) Structure of the nucleocapsid protein of porcine reproductive and respiratory syndrome virus, *Structure (London) 11*, 1445−1451.
19. 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.
20. Yu, I. M., Gustafson, C. L., Diao, J., Burgner, J. W., Li, Z., Zhang, J., and 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.
21. Luo, H. B., Ye, F., Sun, T., Yue, L. D., Peng, S. Y., Chen, J., Li, G., Du, Y., Xie, Y. H., Yang, Y. M., Shen, J. K., Wang, Y., Shen, X., and Jiang, H. L. (2004) In vitro biochemical and thermodynamic characterization of nucleocapsid protein of SARS, *Biophys. Chem. 112*, 15−25.
22. 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.
23. 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.
24. 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.
25. Luo, H. B., Ye, F., Chen, K. X., Shen, X., and Jiang, H. L. (2005) SR-rich motif plays a pivotal role in recombinant SARS coronavirus nucleocapsid protein multimerization, *Biochemistry 44*, 15351−15358.
26. Chang, C. K., Sue, S. C., Yu, T. H., Hsieh, C. M., Tsai, C. K., Chiang, Y. C., Lee, S. J., Hsiao, H. H., Wu, W. J., Chang, C. F., and Huang, T. H. (2005) The dimer interface of the SARS coronavirus nucleocapsid protein adapts a porcine respiratory and reproductive syndrome virus-like structure, *FEBS Lett. 579*, 5663−5668.
27. Yu, I. M, Oldham, M. L, Zhang, J., Chen, J. (2006) Crystal Structure of the Severe Acute Respiratory Syndrome (SARS) Coronavirus Nucleocapsid Protein Dimerization Domain Reveals Evolutionary Linkage between Corona- and Arteriviridae, *J. Biol. Chem. 281*, 17134−17139.

28. Jenkins, Y., Pornillos, O., Rich, R. L., Myszka, D. G., Sundquist, W. I., and Malim, M. H. (2001) Biochemical analyses of the interactions between human immunodeficiency virus type 1 Vpr and p6(Gag). *J. Virol. 75*, 10537−10542.
29. Tellinghuisen, T. L., Hamburger, A. E., Fisher, B. R., Ostendorp, R., and Kuhn, R. J. (1999) In vitro assembly of alphavirus cores by using nucleocapsid protein expressed in *Escherichia coli*, *J. Virol. 73*, 5309−5319.
30. Huang, Q., Yu, L., Petros, A. M., Gunasekera, A., Liu, Z., Xu, N., Hajduk, P., Mack, J., Fesik, S. W., and Olejniczak, E. T. (2004) Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein, *Biochemistry 43*, 6059−6063.
31. Bhattacharya, R., Basak, S., Chattopadhyay, D. J. (2006) Initiation of encapsidation as evidenced by deoxycholate-treated Nucleocapsid protein in the Chandipura virus life cycle, *Virology 349*, 197−211.
32. Green, T. J., Macpherson, S., Qiu, S., Lebowitz, J., Wertz, G. W., Luo, M. (2000) Study of the assembly of vesicular stomatitis virus N protein: role of the P protein, *J. Virol. 74*, 9515−9524.
33. Timmins, J., Schoehn, G., Kohlhaas, C., Klenk, H. D., Ruigrok, R. W., Weissenhorn, W. (2003) Oligomerization and polymerization of the filovirus matrix protein VP40, *Virology 312*, 359−368.

BI0609319