REGULAR ARTICLE

Biochemical and immunological studies of nucleocapsid proteins of severe acute respiratory syndrome and 229E human coronaviruses

Tswen-Kei Tang[1]*, Mark P.-J. Wu[1]*, Shui-Tsung Chen[1], Ming-Hon Hou[1], Ming-Hsiang Hong[1], Fu-Ming Pan[1], Hui-Ming Yu[1], Jenn-Han Chen[2, 3], Chen-Wen Yao[3] and Andrew H.-J. Wang[1]

[1] Institute of Biological Chemistry, Academia Sinica
[2] School of Dentistry, National Defense Medical Center
[3] Biochip R&D Center, Department of Pathology, Tri-Service General Hospital, National Defense University, Taipei, Taiwan

Severe acute respiratory syndrome (SARS) is a serious health threat and its early diagnosis is important for infection control and potential treatment of the disease. Diagnostic tools require rapid and accurate methods, of which a capture ELISA method may be useful. Toward this goal, we have prepared and characterized soluble full-length nucleocapsid proteins (N protein) from SARS and 229E human coronaviruses. N proteins form oligomers, mostly as dimers at low concentration. These two N proteins degrade rapidly upon storage and the major degraded N protein is the *C*-terminal fragment of amino acid (aa) 169–422. Taken together with other data, we suggest that N protein is a two-domain protein, with the *N*-terminal aa 50–150 as the RNA-binding domain and the *C*-terminal aa 169–422 as the dimerization domain. Polyclonal antibodies against the SARS N protein have been produced and the strong binding sites of the anti-nucleocapsid protein (NP) antibodies produced were mapped to aa 1–20, aa 150–170 and aa 390–410. These sites are generally consistent with those mapped by sera obtained from SARS patients. The SARS anti-NP antibody was able to clearly detect SARS virus grown in Vero E6 cells and did not cross-react with the NP from the human coronavirus 229E. We have predicted several antigenic sites (15–20 amino acids) of S, M and N proteins and produced antibodies against those peptides, some of which could be recognized by sera obtained from SARS patients. Antibodies against the NP peptides could detect the cognate N protein clearly. Further refinement of these antibodies, particularly large-scale production of monoclonal antibodies, could lead to the development of useful diagnostic kits for diseases associated with SARS and other human coronaviruses.

Correspondence: Dr. Andrew H.-J. Wang, Institute of Biological Chemistry, Academia Sinica, 128 Academia Road, Taipei, 115, Taiwan
E-mail: ahjwang@gate.sinica.edu.tw
Fax: +886-2-2788-2043

Abbreviations: aa, amino acid; **CD**, circular dichroism; **HCoV**, human coronavirus; **Ig**, immunoglobulin; **MAP**, multiple antigen peptide; **NP**, nucleocapsid protein; **SARS**, severe acute respiratory syndrome

1 Introduction

The severe acute respiratory syndrome (SARS) epidemic in Asia and North America during 2002–2003 caused a serious worldwide health concern. SARS is a type of viral pneumonia, with symptoms including fever, a dry cough, dyspnea, headache and hypoxemia. Death may result

---

* These two authors contributed equally.

from progressive respiratory failure due to alveolar damage [1–3]. A new type of human coronavirus has now been conclusively shown to be the single most probable cause of SARS [4, 5]. The SARS virus can be grown in Vero cells.

The complete sequence of ~29 727 nucleotides of the SARS virus genome from several isolates has been determined [6] (see also: NCBI accession no. NC_004718). The plus-strand RNA genome of SARS human coronavirus (HCoV) has a characteristic, strictly conserved organization with the essential genes occurring in the order 5'-polymerase(pol)-S-E-M-N-3' (Fig. 1A). Sequence comparison of the SARS genes shows that the SARS virus is a brand new type of coronavirus [7]. The availability of the complete sequence now affords us the opportunity to design new experiments to develop effective diagnostic kits, vaccine or therapeutic agents.

There are currently no effective antiviral drugs in treating SARS or any coronavirus infection, nor any proven vaccine against SARS. Diagnostic tests for coronavirus infection fall into two types: ELISA to detect a virus-induced antibody in patients (which is slow) [8–14] and RT-PCR (which may result in a false negative) [15–18]. This problem can be complemented and overcome by using a capture ELISA method in viral diagnosis. Typically, the capture ELISA is achieved with a polyclonal antibody raised in rabbit and a mouse monoclonal antibody, which detect at least two spatially separated epitopes on the antigens [19]. Therefore the availability of suitable antibodies is essential for the development of a immunological diagnostic tool.

Nucleocapsid N protein (NP) is the most abundant structural protein produced during SARS viral infection. The anti-NP antibody has been found to be detectable early in the sera of SARS patients [13]; therefore N protein may be suitable as a candidate for early detection of SARS viral infection. Here we report the preparation and the characterization of soluble full-length N proteins from SARS and 229E human coronaviruses. Specific and high affinity polyclonal antibodies against SARS virus are prepared for possible development of diagnostic kits. The antigenic amino acid sequences of the SARS N protein are delineated using these antibodies. The information may be relevant to help the design of therapeutic proteins.

2   Materials and methods

2.1   cDNA cloning

The cDNA clones of the SARS virus were obtained from the College of Medicine, National Taiwan University [20] and were used to clone the structural proteins. The complete sequence of the Taiwan isolate of SARS (NCBI accession no. AY291451) is available at http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucleotide&val=30698326. The ex-

pression construct, pET21b-N, carrying the full-length NP gene behind T7 promoter, was made by the vector pET21b (Novagen & EMD Biosciences, San Diego, CA, USA). The full-length NP gene was obtained as a PCR fragment amplified from the SARS virus genome cDNA clone. By using oligonucleotides: NP (SARS) F-primer: 5'-CTTCGGCCA TATGTCTGATAATGGACCCCAATCA-3' and NP (SARS) R-primer: 5'-AAACGGCCGCTGCCTGAGTTGAATCAGCA GAAGC-3', the *Nde*I restriction site at the 5' end and the *Not*I site at the 3' end were introduced. The resulting PCR fragment was subsequently digested with *Nde*I and *Not*I and cloned into the digested pET21b vector carrying the same enzyme cutting sites, leading to plasmid pET21b-N. The sequence of the insert was confirmed and then used for generating the recombinant SARS N protein with the $(His)_6$-tag (AAALEHHHHHH) attached at the *C*-terminal end. The N protein gene of 229E coronavirus was similarly cloned. The *C*-terminal partial protein ND4 (amino acid (aa) 279–370) was cloned from the full-length SARS N protein clone using similar methods described above.

2.2   Expression and purification of the full-length N proteins

For generating the recombinant SARS N protein, we transformed *Escherichia coli* strain BL21(DE3) carrying with the plasmid pET21b-N. Induction of the expression was initiated by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to 0.5 mM final concentration and then incubated at 30°C for 4 h. After harvesting the bacteria by centrifugation (using JLA-8.1000 rotor (Beckman Coulter, Fullerton, CA, USA), 6000 rpm, 30 min, 4°C), we lysed the bacterial pellet with the lysis buffer (20 mM Tris-HCl, 150 mM NaCl, 20 mM imidazole, pH 8.0) under protease inhibitors (Complete cocktail EDTA-free; Roche, Penzberg, Germany) protection. Soluble proteins were obtained from the supernatant by centrifugation (15 000 rpm, 30 min at 4°C) to remove the precipitates. All purification steps were at low temperature conditions with a final concentration of 0.1 mM PMSF in the buffers. The SARS N protein was purified using a nickel(II)-nitrilotriacetic acid (Ni-NTA) column (Amersham Biosciences, Piscataway, NJ, USA) with an elution gradient from 0–300 mM imidazole in the buffer solution (20 mM Tris-HCl, 150 mM NaCl, pH 8.0) and the pure fractions were collected and dialyzed against a low salt buffer (20 mM Tris, 50 mM NaCl, pH 7.5). Since the N protein is positively-charged (p*I* 10.11), the protein was further purified by sulfopropyl (SP) cation exchange column Amersham Biosciences using a gradient from 50–1500 mM NaCl buffer (20 mM Tris, pH 7.5). High purity N protein, judged by SDS-PAGE analysis, was obtained and ready for next characterization and immunological assays. ND4 and 229E N proteins were purified using a similar procedure as that for the SARS N protein, except the ion exchange chromatography step was omitted in the purification of ND4 protein.

A

ORF1(Rep1b)

ORF1(Rep1a)

E   N

S   M

B

```
                    10         20         30         40         50
              . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
SARS   1  - - - - - - - - - M S D N G P Q S N Q R S A P R I T F G G P T D S T D N N Q N G G R N G A R P K Q R R   42
229E   1  - - - - - - - - - - - - - - - - - M A T V K W A D A S E - - - - - P Q R C R Q G R - - - - - -   19
OC43   1  M S F T P G K Q S S S R A S S G N R S G N G I L K W A D Q S D Q F R N V Q T R G R R - A Q P K Q T A   49

                    60         70         80         90        100
              . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
SARS  43  P Q G L P N N - - - - - T A S W F T A L T Q H G K - E E L R F P R G Q G V P I N T N S G P D D Q I G   86
229E  19  - - - I P - - - - - - - Y S L Y S P L L V D S E - Q P W K V I P R N L V P I N K K - D K N K L I G   56
OC43  50  T S Q Q P S G G N V V P Y Y S W F S G I T Q F Q K G K E F E F V E G Q G V P I A P G V P A T E A K G   99

                   110        120        130        140        150
              . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
SARS  87  Y Y R R A T R R - A R G G D G K M K E L S P R W Y F Y Y L G T G P E A S L P Y G A N K E G I V W V A   135
229E  57  Y W N V Q K R - - F R T R K G K R V D L S P K L H F Y Y L G T G P H K D A K F R E R V E G V V W V A   104
OC43 100  Y W Y R H N R R S F K T A D G N Q R Q L L P R W Y F Y Y L G T G P H A K D Q Y G T D I D G V Y W V A   149

                   160        170        180        190        200
              . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
SARS 136  T E G A L N T P K D H I G T R N P N N N - - A A T V L Q L P Q G T T L P K G F Y A E G S R G G S Q A   183
229E 105  V D G A K T E P T G Y G V R R K N S E P E I P H F N Q K L P N G V T V E E P D S R A P - S R S Q S   153
OC43 150  S N Q A D V N T P A D I V D R D P S S D - - E A I P T R F P P G T V L P Q G Y Y I E G S - G R S A P   196

                   210        220        230        240        250
              . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
SARS 184  S S R S S S R S R G N S R N S T P G S - S R G N S P A R M A S G G G E T A L A L L L L D R L N R L E   232
229E 154  R S Q S R G R G E S K P Q S R N P S S - D R N H N S Q D D I M K A V A A A L K S L G F D K P Q E K D   202
OC43 197  N S R S T S R T S - - S R A S S A G S R S R A N S G N R T P T S G V T P D M A D Q I A S L V L A K L   244

                   260        270        280        290        300
              . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
SARS 233  S K V S G K G Q Q Q - - - - - - - - - - - - - - - - - - Q G Q T V T K K S A A E A S K K P R Q K R T   264
229E 203  K K S A K T G T P K P S R N Q S P A S S Q T S A K S L A R S Q S S E T K E Q K H E M Q K P R W K R Q   252
OC43 245  G K D A T K - P Q Q - - - - - - - - - - - - - - - - - - V T K H T A K E V R Q K I L N K P R Q K R S   275

                   310        320        330        340        350
              . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
SARS 265  A T K - - Q Y N V T Q A F G R R G P E Q T Q G N F G D Q D L I R Q G T D Y K H W P Q I A Q F A P S A   312
229E 253  P N D D V T S N V T Q C F G P R D L D H - - - N F G S A G V V A N G V K A K G Y P Q F A E L V P S T   299
OC43 276  P N K - - Q C T V Q Q C F G K R G P N Q - - - N F G G G E M L K L G T S D P Q F P I L A E L A P T A   320

                   360        370        380        390        400
              . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
SARS 313  S A F F G M S R I G M E V T P S - - G - - T - - - - - - W L T Y H G A I K L D D K D P Q F K D N V   351
229E 300  A A M L F D S H I V S K E S G N - - - - - - - - - - T V V L T F T T R V T V P K D H P H L G - - -   335
OC43 321  G A F F F G S R L E L A K V Q N L S G N P D E P Q K D V Y E L R Y N G A I R F D S T L S G F E T I M   370

                   410        420        430        440        450
              . . . | . . . . | . . . . | . . . . | . . . . | . . . . |
SARS 352  I L L N K H I D A Y K T - - - - F - - P P T E P K K D K K K K T D E A Q P L P Q R Q K K Q P T V T L   395
229E 335  - K F L E E L N A F T R - - - - - - - - - E M Q Q H P L L N P S A L E F N P S Q T S P - - - A T A   371
OC43 371  K V L N E N L N A Y Q Q Q D G M M N M S P K P Q R Q R G H K N G Q G E N D N I S V A V P K S R V Q Q   420

                   460        470
              . . . | . . . . | . . . . | . . . . | . . . -
SARS 396  L P A A D M D D F S R Q L Q N S M S G A S A D S T Q A -   422
229E 372  E P V R D E V S I E T D I I D E V N - - - - - - - - - -   389
OC43 421  N K S R E L T A E D I S L L K K M D E P Y T E D T S E I   448
```

Figure 1. A, Map of the ORFs in the SARS coronavirus. All coronaviruses share similar ORF arrangement and have nearly the same length of genome. Usually, the total length of complete RNA genome is about 30 Kb. The relative locations of structure proteins, which are encoded by the SARS genome, has been determined by sequencing. N and other structural proteins are near the 3' end of the genome. B, Amino acid multiple sequence alignment of three human infectable coronaviruses: SARS, 229E and OC43 N proteins. Residues that are conserved are boxed and similar residues are colored. There are two conserved regions among the three N proteins, one is from aa 50–170 and the other is from aa 250–360. These conserved regions are not found in any other protein except in the coronavirus family by BLAST. It means that they are highly unique and conserved. These regions also correspond to the ordered regions obtained from the PONDR analysis (see Fig. 3B).

2.3 Circular dichroism spectroscopy

Circular dichroism (CD) spectra were obtained in a JASCO-720 CD spectropolarimeter (Jasco, Tokyo, Japan). Temperature was controlled by water circulation at 4°C in the cell jacket. The concentration of the proteins in each sample was 0.4 mg/mL in 20 mM Tris-buffered solution, pH 8.3 with 150 mM NaCl. The CD spectra were collected between 250–190 nm with bandwidth at 1 nm intervals. All spectra were the average of five runs.

2.4 RNA and DNA band shift assay

For the RNA band shift assay, purified SARS and 229E N proteins were adjusted to 0.1 µg/µL in saline buffer. Each reaction was added with 1, 2, 4, 8, 16 µL N protein and incubated with 3 µL single strand RNA 5'-CGCAAUUGCGCG-CAAUUGCG (100 ng/lane). Double-distilled H_2O was added to the solution to make a final total volume of 20 µL (10 mM Tris, 300 mM NaCl, pH 7.5) and incubated for 15 min at room temperature. Two microliters of 50% glycerol was added and mixed well. A total of 22 µL solution of each lane was loaded onto a 8% nondenaturing polyacrylamide gel and the gel was run at 150 V for 45 min in prechilled TBE buffer (Tris-borate-EDTA). The gel was then stained with SYBR Green II stain (Sigma-Aldrich, St. Louis, MO, USA) at 1:5000 dilution for 30 min at room temperature, then washed twice with ddH_2O for 15 s to remove any excess stain that might interfere with image analysis. The image was recorded using a gel photo system (Vilber Lourmat, Torey, France) and scanned for analysis. All solutions and buffers were treated with the RNase inhibitor diethyl pyrocarbonate (Merck, Whitehouse Station, NJ, USA) and the buffer tank and other accessories were cleaned with RNaseZAP (Sigma-Aldrich) to inhibit the RNase activity. Two complementary ssDNA, 5'-GATCC AGCTATACTTGGTCAGGGCGAATTCTAACTA and 5'-TA GTTAGAATTCGCCCTGACCAAGTATAGCTGGATC, were similarly employed for DNA band shift assay by utilizing 1.5% agarose gel and stained with ethidium bromide.

2.5 Chemical cross-link assay

To investigate the polymerization features of the SARS and 229E N proteins, a chemical cross-linking experiment was performed. A series of protein solutions containing the same amount of SARS N protein with the concentration of 1.0, 0.5, 0.25 and 0.125 mg/mL, respectively, was added to make a 0.25% v/v final concentration of glutaraldehyde and reacted at room temperature for 10 min. The reaction was stopped by adding overabundant 1 M Tris buffer (0.5% v/v) and then set on ice. The sample of each reaction was concentrated to 10 µL and used for SDS-PAGE analysis. For 229E N protein, which is dissolved in buffer with detergent (0.5% Triton X-100), the reaction condition was modified with the reaction time increased to 15 min and the glutaraldehyde final concentration increased to 0.5% v/v.

2.6 Determination of N protein oligomerization by gel filtration

Size-exclusion chromatography assays were carried out on a XK 16/70 column with Sephacyl S-100 HR media on an AKTA FPLC system (Amersham Biosciences, Piscataway, NJ, USA). The column was equilibrated and run with the balance buffer (10 mM Tris, 150 mM NaCl, pH 7.5) at 4°C. Both SARS and 229E N proteins were adjusted to 1 mg/mL in the balance buffer. The loading volume of the protein sample was 2 mL with a flow rate of 0.4 mL/min with a detection of 280 nm absorbance. Three proteins, phosphorylase B (97 kDa), ovalbumin (45 kDa) and chymotrypsinogen A (25 kDa) were used as size markers. ND4 protein was similarly employed for gel filtration assay.

2.7 Sequence alignment and order/disorder analysis of N proteins

The ClustalX program, V1.8 [21] was used to align the sequences of SARS (422 amino acid), 229E (389 amino acid) and OC43 (448 amino acid) N proteins. The resulting file was transferred to Bioedit V5.06 (Isis Pharmaceutical, Carlsbad, CA, USA) to prepare for graphic figures. The PONDR program (Molecular Kinetics, Indianapolis, IN, USA; http://www.pondr.com/) with the VL3-BA neuronetwork feedback predictor was used to predict the order/disorder regions of all three N proteins.

2.8 Synthesis of peptides

Twenty-eight peptides, each 20 amino acids long, derived from the N protein sequence of SARS-CoV were synthesized by a stepwise FastMoc protocol [33] and used without further purification. The peptide was synthesized by solid-phase peptide synthesis using 433A peptide synthesizer (Applied Biosystems, Foster City, CA, USA). Starting with 0.10 mmol (0.101 g) of HMP (p-hydroxymethyl phenoxymethyl polystyrene) resin (1.01 mmol/g). The amino acids were introduced using the manufacturer's prepacked cartridges (1 mmol each). After synthesis, 0.1 mmol peptide resin was placed in a round-bottom flask containing a microstirring bar. The cool mixture containing 0.75 g crystalline phenol, 0.25 mL EDT (1,2-ethandithol), 0.5 mL thioanisole, 0.5 mL water, and 10 mL TFA was put into the flask and stirred for 1–1.5 h at room temperature. In general their purity is greater than 80%. Mass spectra were determined using a Finnigan LCQ mass spectrometer (Thermofinnigan, San Jose, CA, USA) with an electron spray ion source. The synthesis of an octa-branched matrix core with peptide antigen attached was accomplished manually by the same synthesis. After cleavage, the octa-branched core matrix containing eight functional amino groups were determined by MS.

2.9 Immunization of rabbits and assays

Soluble recombinant N protein prepared as described in Section 2.2 was used as antigens for immunization and immunoassays. Octameric multiple antigen peptide (MAP) synthetic peptides (predicted to have high antigenicity and low hydrophobicity) were used as antigens without purification. Rabbits (New Zealand White strain), weighing 3–3.5 kg, were immunized by intrasplenic injection with the recombinant SARS N proteins or octameric MAP peptides (15–20 aa long) derived from SARS structural proteins at 250 μg or 500 μg, respectively, *per* immunization. The antigen was administered together with an equal amount of Gold TiterMax adjuvant (CytRx, Norcross, GA, USA). The rabbit antisera were used for most of the subsequent experiments without purification. We analyzed the titer of rabbit sera using Western blot assay for N protein antigen and dot blot assay for MAP synthetic peptide. In general, we could obtain high titer polyclonal antibodies in 6–8 weeks. If necessary, additional booster immunizations were administered in order to obtain good titer of the antisera.

2.10 Protein array fabrication and assay

A protein array was designed to detect the antibodies against SARS N protein. Quantities of 1 nL of full-length and various peptides of SARS N protein with the concentration of near 400 mg/mL were spotted onto the aldehyde-coated glass slides (CEL Associates, Pearland, TX, USA) simultaneously using a PixSys 5000 robot arrayer (Cartesian Technologies, Irvine, CA, USA). After blocking with PBS with 3% nonfat milk and 0.5% Tween-20, the complete protein array was incubated with sera generated from rabbits or sera from patients with SARS and followed by incubation of goat antihuman immunoglobulin (Ig)G and IgM or antirabbit IgG antibodies conjugated with Cy3 or Cy5 fluorescent dye (Jackson Immuno Research, West Grove, PA, USA), respectively, at room temperature for 30 min. Subsequently, protein arrays were spun dry and scanned using a GenePix 4000B scanner (Axon Laboratories, Foster City, CA, USA).

2.11 Immunostain assay

Immunofluoresence stain analysis was performed in laminar-flow safety cabinets in a BSL-3 (Biological Safety Level 3) laboratory. The SARS coronavirus was propagated in Vero E6 cells at 37°C until cytopathogenic effects were seen in 75% of the cell monolayer, after which the cells were harvested, spotted onto 24 well plates, and fixed with 1:1 of cold acetone/methanol. Uninfected Vero E6 cells were used as controls for this experiment. Anti-N protein or anti-M protein antibodies were tested at a 1:2000 dilution applied to 24 well plates which were precoated with SARS HcoV infected Vero E6 cells and washed with

$1 \times$ PBS after being incubated for 60 min at 37°C. After rinsing with PBS buffer, followed by incubation with a fluorescein isothiocyanate-conjugated goat anti-rabbit IgG (Jackson Immuno Research) for 30 min at 37°C, the 24 well plates were subjected to another washing cycle before being monitored for specific fluorescence under an immunofluorescence microscope. Immunostained images were visualized and recorded using a Zeiss imaging microscope (Carl Zeiss, Oberkochen, Germany).

3 Results

3.1 Characterizations of soluble N proteins from SARS and 229E HCoV

For detailed biochemical, biophysical and immunological studies of the structural proteins of coronavirsues, it is desirable to have soluble (thus likely to have native conformation) full-length proteins. Here we have successfully produced soluble full-length nucleocapsid N proteins of SARS HCoV and HCoV-229E in large (milligram) quantity. Figure 2 shows the SDS-PAGE of the expression and purification steps of SARS (lanes 2–5) and 229E (lanes 7–10) HCoV N proteins. It can be seen that a significant amount of N proteins are found in the total soluble fractions (lanes 4 and 9). The yields were in the range of 10–15 mg/L culture. They were purified using Ni-NTA affinity column chromatography as a single band of molecular mass ∼50 kDa (lanes 5 and 10). Their M_rs were determined to be 47303.3 and 44745.1 for SARS and 229E N proteins, respectively, by MS. These are in agreement with the calculated values of N protein (46025 for SARS and 43466.7 for 229E) plus the *C*-terminal AAALEHHHHHH (1296.3).

We have noticed that the N proteins are labile and are degraded rapidly into several bands with lower M_r during storage, even at 4°C (Fig. 2, lanes 12 and 13). Over time, only a major band of ∼30 kDa remained. We have identified this band as being aa 169–422 protein fragment by *N*-terminal end sequencing (determined to be 169-PKGFYA) and LC-ESI-TOF MS (confirming four peptides of aa 179–190, aa 211–249, aa 264–349 and aa 376–406, all located within the *C*-terminal region, by MASCOT analysis). The band could be detected clearly by Western blot analysis using anti-His-tag antibody, in agreement with the fact that (His)$_6$ tag is fused at the *C*-terminal end of the N protein.

Thus far, no 3-D structural information of full-length N protein from any coronavirus is available. Circular dichroism (CD) spectra of the two N proteins were obtained (Fig. 3A). The analysis showed that no significant signature of α-helix exists. Most compositions are turns and coils, with some beta sheets present. Analysis by the program PONDR [22] indeed suggested that a significant part of the N protein molecule is disordered (Fig. 3B), but 229E seems to be slightly more ordered than SARS in the *N*-terminal region. It is interesting
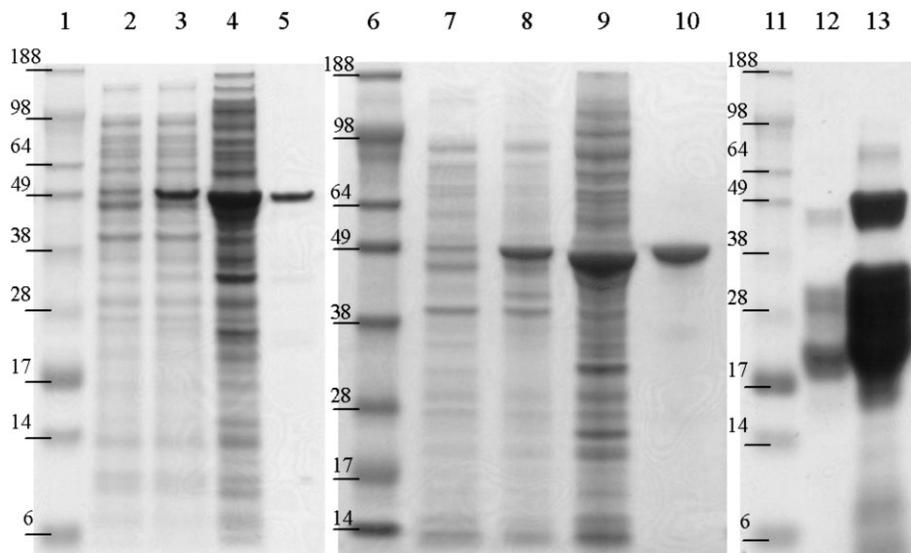
Figure 2. Expression and purification of SARS and 229E N proteins. SDS-PAGE result of SARS (lanes 2–5) and HCoV 229E (lanes 7–10) N proteins. Lanes 1, 6 and 11 are markers. Proteins obtained from lysates of non-induction bacteria (lanes 2, 7), after induction (lanes 3, 8), soluble total proteins (lanes 4, 9), and purified N proteins (lanes 5, 10) are shown. SARS N protein is unstable without protease inhibitor protection after storage at 4°C for 2 weeks (lane 13) and one month (lane 12).
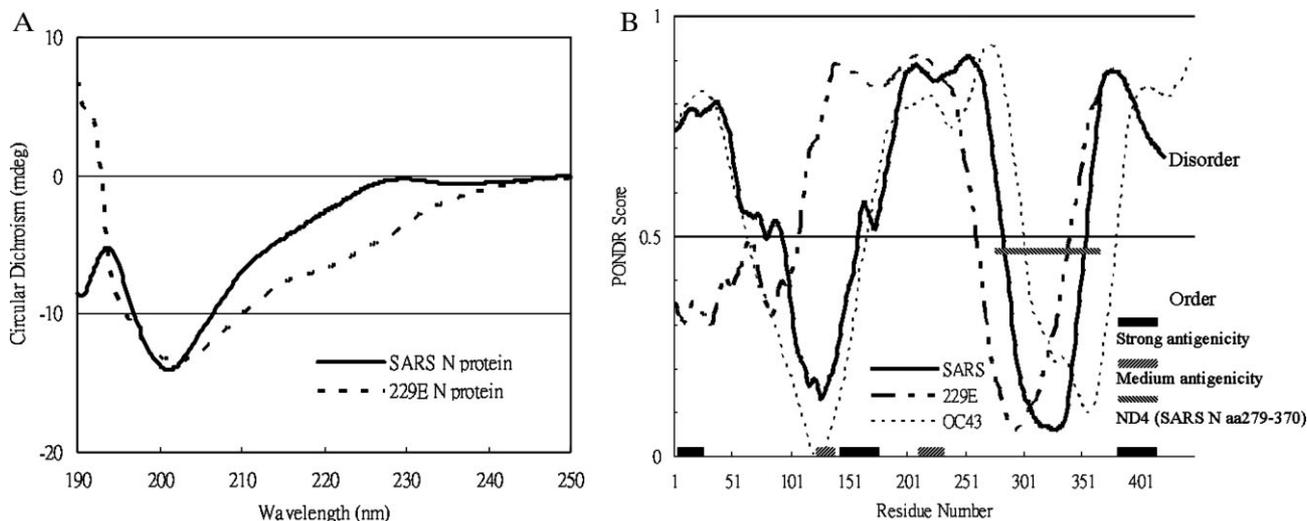
Figure 3. A, CD spectra of SARS and 229E N proteins. Both N proteins have no significant secondary structure with or without DNA/RNA binding, but the result shows that 229E N protein has slightly more secondary structures than SARS N protein. This is consistent with the predicted result of the PONDR analysis. B, Order/disorder analysis of SARS, 229E and OC43 N protein by the VL3-BA predictor of the PONDR program. SARS N protein is the most disordered protein of the three. OC43 N protein pattern is much like SARS N protein. NP of 229E is the most ordered one, particularly in the *N*-terminal region. The amino acid sequences have high antigeneicity and medium antigeneicity are marked by bars at the bottom. The mapping of partial *C*-terminal region of SARS N protein, ND4 protein, is also shown.

to note that in both N proteins there appears to be two distinct ordered domains located around aa 50–170 and aa 250–360 of SARS N protein, and aa 20–140 and aa 240–340 of 229E N protein, respectively. This observation is supported by the sequence alignment of the three N proteins (Fig. 1B), which shows that the conserved regions match the two predicted ordered regions by PONDR analysis.

We also measured the CD spectra of the mixture of N protein and RNA. No apparent reorganization of the N protein structure could be detected as seen from the resulting CD spectra (data not shown).

3.2 Chemical cross-linking studies and size exclusion assay of N proteins

N protein had been reported capable of self-association [23, 24]. We further characterized the SARS HCoV N protein by chemical cross-linking experiments. The protein with different concentrations are cross-linked with glutaraldehyde (0.25%) for 10 min. Interestingly, we noted that at low concentration (~0.1 mg/mL), the N protein appeared as a major band at a mass of ~90 kDa, suggesting that the N protein exists predominantly as a dimer (Fig. 4A, lane 2). At higher
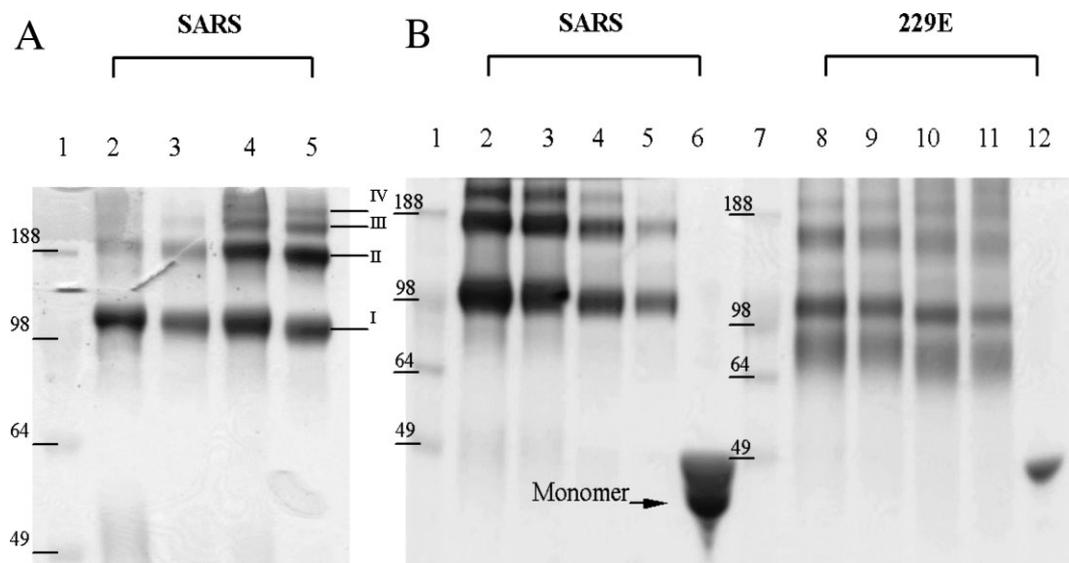
Figure 4. A, Chemical cross-linking assay of SARS N protein. Markers are in lane 1 and different SARS N protein concentrations of 0.125, 0.25, 0.5, and 1 mg/mL for the cross-linking reaction is in lanes 2, 3, 4, and 5, respectively. The result shows that at low concentration of N protein the dimer (I) and trimer (II) are the major forms. At high concentration, not only the dimer and trimer forms, but also the tetramer (III) and pentamer (IV) forms exist. B, Chemical cross-linking of SARS and 229E N proteins in the presence of detergent. Lanes 1 and 7 are markers (SeeBlue® Plus2; Invitrogen), lanes 2–5 are SARS N protein and lanes 8–11 are 229E N protein in 1, 0.5, 0.25, 0.125 mg/mL concentration for chemical cross-linking reactions, respectively. Lanes 6 and 12 are SARS and 229E N proteins without cross-linking reaction, respectively. The cross-linking results show that the dimer and trimer are the major forms of all tested concentrations in the presence of detergent.

concentrations, minor bands with M_r approximately that of a trimer, tetramer and pentamer appeared (Fig. 4A, lanes 3–5). Interestingly, in the presence of detergent Triton X-100 (0.5%), dimers and trimers are the major forms at all concentrations (Fig. 4B, lanes 2–5 for SARS and lanes 8–9 for 229E). It is possible that oligomerization of N protein is facilitated in the membrane environment.

We also performed gel filtration experiments to further clarify the oligomerization phenomenon of N proteins (Fig. 5). The molecular masses of the standard proteins phosphorylase B (97 kDa) and ovalbumin (45 kDa) correspond to those of the dimer and monomer form of full-length N proteins. Chymotrypsinogen A (25 kDa) corresponds to the dimer form of ND4 (24 kDa) protein (Fig. 5A). SARS and 229E N proteins show similar results. The dimer form is the major form of both N proteins at 1 mg/mL concentration. The monomer form is presenting a small fraction. The small peak at about 30 mL retention volume is out of the resolution of Sephacyl S-100 HR media, which represents the higher oligomer forms (Fig. 5B, 5C). The ND4 protein gel filtration experiment demonstrated that SARS N protein *C*-terminal region has self-association ability. The data indicate that the dimer form is the major form of ND4 protein at 1 mg/mL concentration, suggesting that the *C*-terminal region is responsible for the oligomerization of full-length N protein (Fig. 5D).

3.3 Nucleic acid binding of N protein

Nucleocapsid N protein is an RNA binding protein in coronaviruses. The p*I* values of the three N proteins are predicted to be 10.11, 9.72 and 9.65 for SARS, 229E and OC43 HCoVs, respectively. We used the band shift method to study the interactions between SARS N protein and nucleic acids (Fig. 6). Different ratios of protein/RNA (Fig. 6A) and protein/DNA (Fig. 6B) are used and the complexes are run on 8% nondenaturing polyacrylamide gel and 1.5% agarose gel, respectively. Each gel has been visualized by SYBR Green II and ethidium bromide staining under UV light.

It can be seen that at 1:1 (w/w) SARS NP/RNA ratio (Fig. 6A, lane 3), the RNA band begins to have a visible shift for both SARS and 229E N proteins. At 16:1 ratio (Fig. 6A, lane 7), the SARS NP/RNA complex is completely retarded. The band patterns for the 229E NP/RNA complex are similar to those of the SARS N protein/RNA complex. The bands of protein/RNA complexes remain sharp in the gel, suggesting that the complexes have uniform size and the RNase activity has been inhibited. The interactions between SARS N protein with both ssDNA and dsDNA appeared to be less specific (Fig. 6B). The protein/DNA bands became more smeared with increasing ratio of protein to DNA. Such observations are likely due to the fact that N protein is an RNA binding protein and specific interactions between ssRNA and N protein exist.
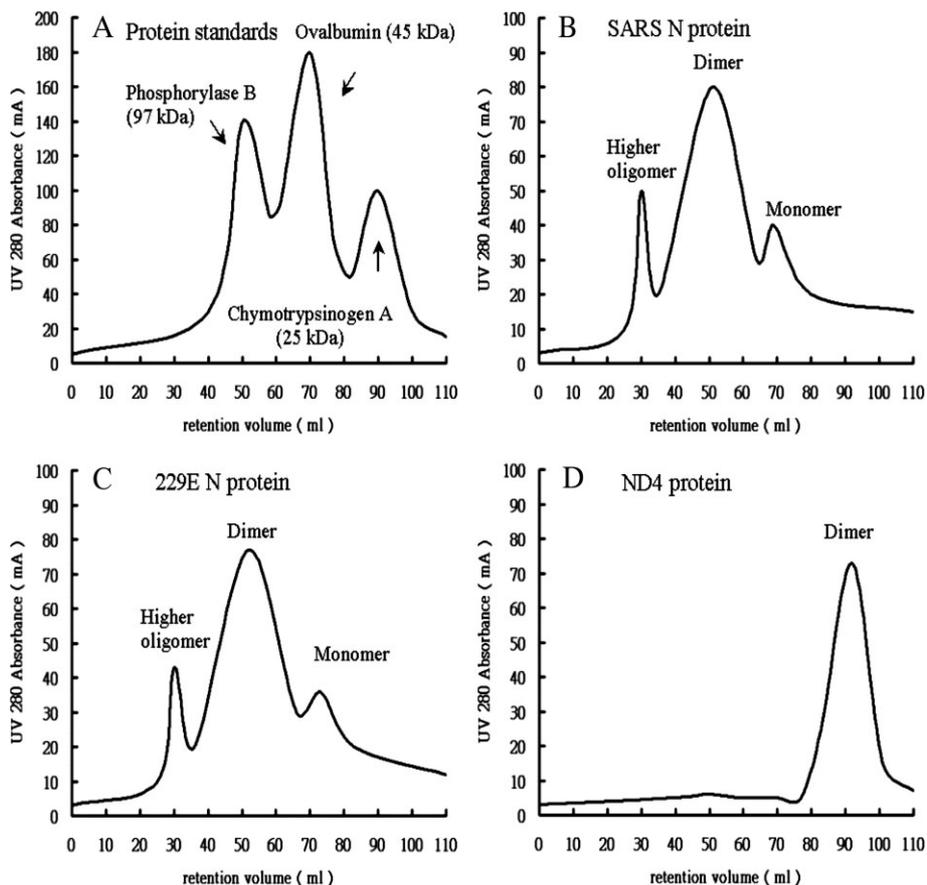
Figure 5. Protein gel filtration assay. A, The retention volume of standard proteins. Three proteins, phosphorylase B (97 kDa), ovalbumin (45 kDa) and chymotrypsinogen A (25 kDa) whose masses correspond to the sizes of dimer, monomer of the full-length N protein (47 kDa) and the dimer form of the ND4 protein (12 kDa) respectively. B, The retention volume graph of the SARS N protein. The sample concentration is 1 mg/mL. The major peak (retention volume ~50 mL) represents the dimer form. A small amount of the monomer still remains (retention volume ~70 mL). The small peak at ~30 mL retention volume, beyond the column resolution, represents all other higher oligomer forms. C, The retention volume graph of the 229E N protein, showing a similar result to that of the SARS N protein. The broader dimer peaks in both B and C may imply the coexistence of dimer isoforms. D, The retention volume graph of the ND4 protein. At 1 mg/mL concentration, only one peak (retention volume ~90 mL) representing the dimer form exists.
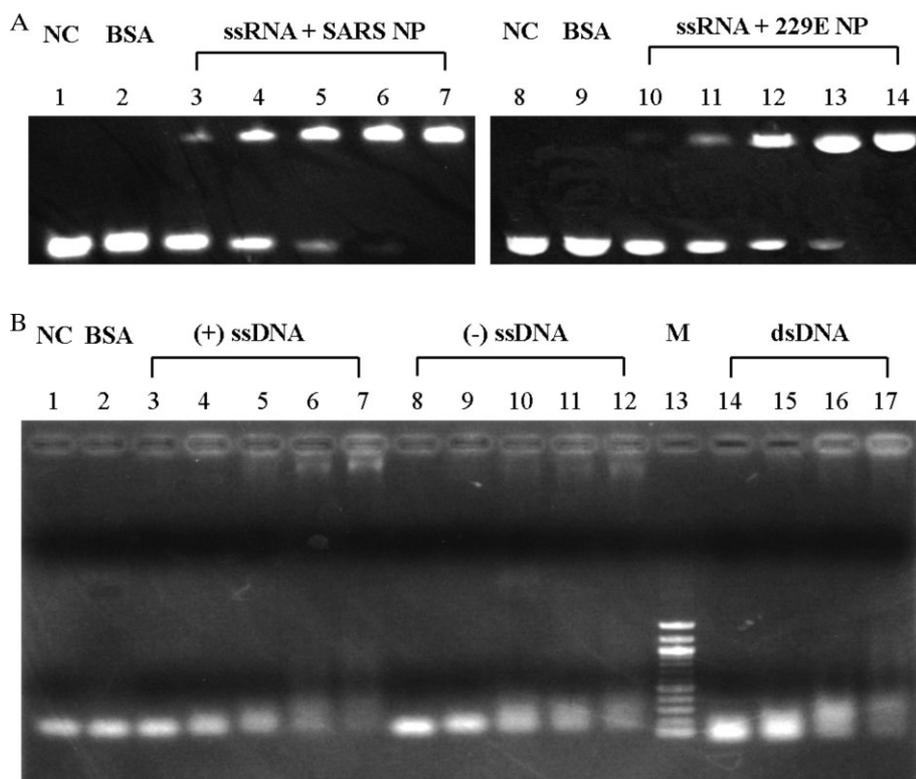
Figure 6. A, RNA binding band shift assay with SARS and 229E N proteins. Negative control mock (lanes 1, 8) and BSA binding with ssRNA (lanes 2, 9), ssRNA added with 1, 2, 4, 8, 16 μL SARS N protein (lanes 3, 4, 5, 6, 7) and ssRNA added with 1, 2, 4, 8, 16 μL 229E N protein (lanes 10, 11, 12, 13, 14), respectively. B, DNA binding band shift assay with SARS N protein. Lane 1 is DNA alone as negative control, lane 2 is BSA binding with ssDNA, ssDNA added with 1, 2, 3, 4, 5 μL SARS N protein (lanes 3, 4, 5, 6, 7), complementary ssDNA added with 1, 2, 3, 4, 5 μL SARS N protein (lanes 8, 9, 10, 11, 12), respectively. Marker (lane 13), dsDNA added with 2, 3, 4, 5 μL SARS N protein (lanes 14, 15, 16, 17), respectively. The band shift result shows that SARS N protein can also bind DNA due to the positive charges of DNA.

3.4 Preparation and analysis of polyclonal anti-N antibodies

The availability of soluble recombinant SARS full-length N protein afforded us the opportunity to produce high quality anti-NP antibody. We have used an intraspleen immunization method to prepare the antibodies. Our Western blot analysis showed that we could detect 0.3 µg of SARS N protein easily using antisera after 320 000-fold dilution (Fig. 7, lane 2). For comparison, the (His)$_6$ tag at the *C*-terminal end of the recombinant N protein could be detected using commercial anti-His tag mAb after 2000-fold dilution (equivalent of 10 µg of antibody) (Fig. 7, lane 1). Two minor bands, one slightly below the major band and the other at ~30 kDa, could be seen. In lane 2 of Fig. 7, the N protein was detected clearly using the rabbit antisera at 320 000-fold dilution. Here six minor bands are seen, two of which are the same with those seen in lane 1. Four minor bands are not detected by the anti-His-tag antibody indicating that those protein fragments have lost the *C*-terminal ends.

We have also prepared rabbit polyclonal antibodies using peptide antigens derived from N protein. Five peptides were chosen based on antigeneicity/hydrophobicity analysis: aa 65–79, aa 107–121, aa 245–259, aa 339–353 and aa 366–380. Three of them (aa 65–79, aa 107–121, and aa 339–353) were able to elicit immune response to produce antibodies with good titer when blotting against full-length SARS N protein (Fig. 7, lanes 3–5). We have also purified a large quantity of soluble 229E N protein (Fig. 2, lane 10). We have found that the rate of degradation of 229E N protein is faster than that of SARS N protein. We noted that a significant band of protein ~30 kDa can be detected by anti-His-tag antibody (Fig. 7, lane 6). We could see only the ~30 kDa protein after the full-length protein is stored for only a few days at 4°C (data not shown).

The excellent specificity of the SARS anti-NP polyclonal antibody is also reflected in the observation that it did not cross-react with the purified HCoV-229E N protein (Fig. 7, lane 7), despite the fact that the two N proteins share 23% sequence identity (see Fig. 1B). In fact we could detect 0.3 µg of SARS N protein using the rabbit anti-NP (SARS) antisera even after 1 200 000-fold dilution (data not shown). This superior quality of the antibody will be very useful for diagnostic purposes, since potential cross-reaction among different coronaviruses could be avoided. Conversely, the availability of antibodies against N proteins of 229E and OC43 CoV will allow us to detect which infection the patient has acquired so that proper clinical care can be applied.

3.5 Mapping of epitopes of SARS N protein

To map the epitopes of the SARS N protein that are recognized by the anti-NP (SARS) antibody, 28 synthetic peptides of 20 aa each with five amino acids overlap in each sequence, encompassing the entire N protein sequence, were synthesized and used as antigens for protein array assay as described in Section 2.10. As seen in Fig. 8A, BSA showed essentially background signal, whereas the full-length SARS N protein reacted strongly with the antibody. Among the 28 peptides, peptide 1 (aa 1–20), peptide 11 (aa 150–170) and peptide 27 (aa 390–410) showed very strong responses, close to that of full-length N protein. Four other regions showed medium responses, including aa 60–95, aa 120–155, aa 210–230 and aa 270–305 (Fig. 3B). Our results are consistent with those obtained using sera from SARS patients in which peptides of aa 51–71, aa 134–208 and aa 349–422 were identified as having strong antigeneicity [25, 26].
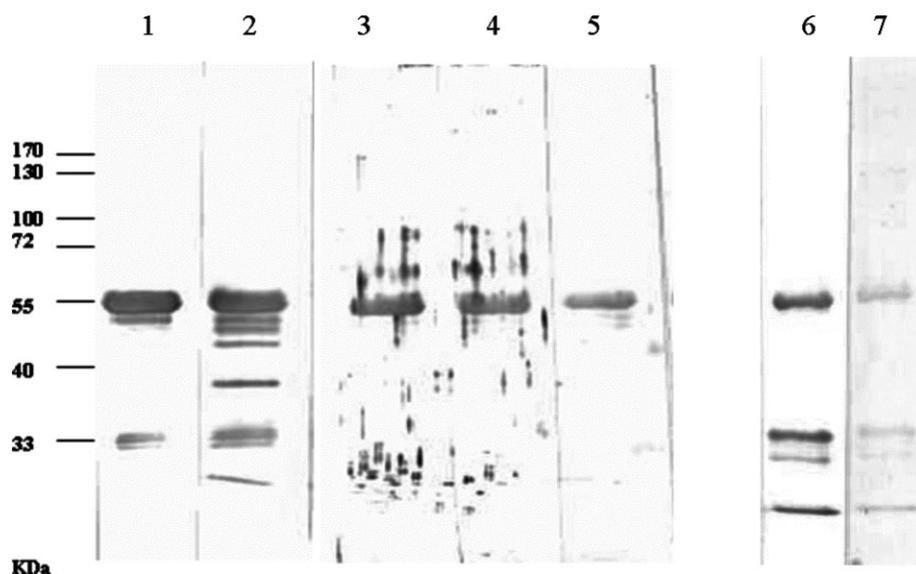
Figure 7. Analysis of various polyclonal antibodies against full-length SARS and 229E N proteins using Western blot assays. The full-length SARS and 229E N proteins (0.3 µg each lane) are used as antigen in lanes 1–5 and 6–7, respectively. Antibodies are as follows: lanes 1 and 6, anti-his tag antibody (0.5 µg); lanes 2 and 7, anti-N protein serum 320 000X dilution; lane 3, anti-N (aa 65–79) peptide serum 1000X dilution; lane 4, anti-N (aa 107–121) peptide serum 1000X dilution; lane 5, purified anti-N (aa 339–353) peptide (12.8 µg) antibody.

The results here suggest that the immunological responses induced by N protein in human and rabbit are generally similar. It is likely that antibodies produced by rabbit or mouse will recognize the N protein effectively, and are therefore suitable for diagnostic purposes. The three strong epitopes mapped by our polyclonal antibody can be used to produce antibodies that have high specificity towards those sequences.

3.6 Design of peptide antigens derived from SARS structural proteins and their detection by sera of SARS patients

When the nucleotide sequence of the SARS genome became available early in 2003 [6], we chose to raise antibodies against viral S, M, N, and E structural proteins. There are two strategies to produce these antibodies. The first includes cloning, expression and purification of these proteins in bacterial and mammalian systems, preferably using epitope tagging, followed by antibody production. However, this approach is time-consuming and it was uncertain whether suitable proteins could be purified.

Another approach is the direct production of antibodies using the synthetic peptides derived from these protein sequences. We designed several peptide sequences derived from the S, M, N, and E proteins as immunogens prepared based on (predicted) good antigeneicity and low hydrophobicity criteria. These included five peptides from N protein (aa 65–84, aa 107–121, aa 245–259, aa 339–353, aa 366–380), three peptides from M protein (aa 149–163, aa 173–187, aa 203–217), three peptides from S protein (aa 172–186, aa 385–399, aa 434–448) and one peptide from E protein (aa 47–76). MAP forms of these peptides were synthesized (except E peptide which is directly used) against which polyclonal antibodies in rabbits have been raised. Specific peptide-based antibodies have the advantage that peptides can be readily prepared and the antibodies can avoid possible cross-reactions with other proteins of host cells. Those polyclonal antibodies may act like mAbs due to the well-defined antigens.

Figure 8B shows the result of protein array assays designed for evaluating the antigenicity of synthesized peptides by human sera from two patients with SARS (#217 and #1045) and from one healthy person as control. BSA (lane 1) and recombinant SARS N protein (lane 4) were used as negative and positive controls. Recombinant N protein showed a strong response toward the IgG from those two patients, but a weak response toward the IgM. The result is consistent with an earlier study by Wu *et al.* [13]. Interestingly, among those 11 peptides, most of the N peptides showed varying degrees of response, especially in patient #1045. Surprisingly all S peptides showed very weak responses to both IgG and IgM in both patients. Peptide M173 (aa 173–187 of SARS M protein) showed medium to strong response to both IgG and IgM in both patients, although a weak response was also found in the healthy control.

Our results here, when compared with the epitopes mapped by polyclonal anti-NP antibodies or SARS patients' sera (Fig. 8A, 8B), show that many of the peptides from N, S and M structural proteins are not good antigens. Nevertheless, antibodies against three NP peptides (aa 65–79, aa 107–121, and aa 339–353) were successfully prepared and they were still able to detect full-length N protein by Western blot analysis as shown above (Fig. 7, lanes 3–5).

3.7 Immunological detection of SARS proteins expressed in Vero E6 cells infected by SARS virus

To assess whether the antibodies could detect the structural proteins in the virus-infected cells, we performed immunofluorescence assays. When the polyclonal anti-NP antibodies were used, we could see that the entire cytosol, but not the nucleus, of the SARS virus-infected Vero cells is uniformly stained (Fig. 8C). The immunofluorescent signal is very strong in this assay, suggesting that a high level of N proteins, not packaged into intact virus particles, exists in the virus-infected cells. Interestingly, the immunofluorescence assay using M149 antibody (raised against peptide (aa 149–163) of M protein) shows that the immunofluorescent signal occurs at the periphery of the virus-infected cells (Fig. 8D). The result is consistent with the fact that M protein is a membrane protein.

4   Discussion

A major goal of this study is to prepare specific and high affinity antibodies for the early detection of SARS viral infection. However, due to the high risk in handling the SARS virus, we chose to raise antibodies against viral structural proteins. Earlier immunogenicity studies of coronaviruses suggested that the spike (S1) surface glycoprotein is the major antigen of the virus [27–29]. However, in another study [30], it was reported that the immune response to S protein was barely detectable in naturally infected dogs, whereas anti-M and anti-N antibodies were detected with a very strong reaction and for a long time after infection [30]. The study of antibodies produced in SARS patients suggested that different patients have distinct immunological responses, although almost all patients showed responses against N protein either in the appearance of IgA or IgG [25]. Therefore it is important that proteins other than the S protein should also be considered for antibody production. N protein is a suitable candidate.

We have optimized the expression and protein purification procedures so that we could obtain significant amounts of soluble full-length N proteins from SARS and 229E CoVs for further biochemical and immunological studies. CD spectra of the soluble full-length N proteins from SARS and 229E CoVs indicate that the N protein is relatively disordered and they have nearly 50% turn/disorder in nonsecondary structures. We have performed order/disorder analyses of
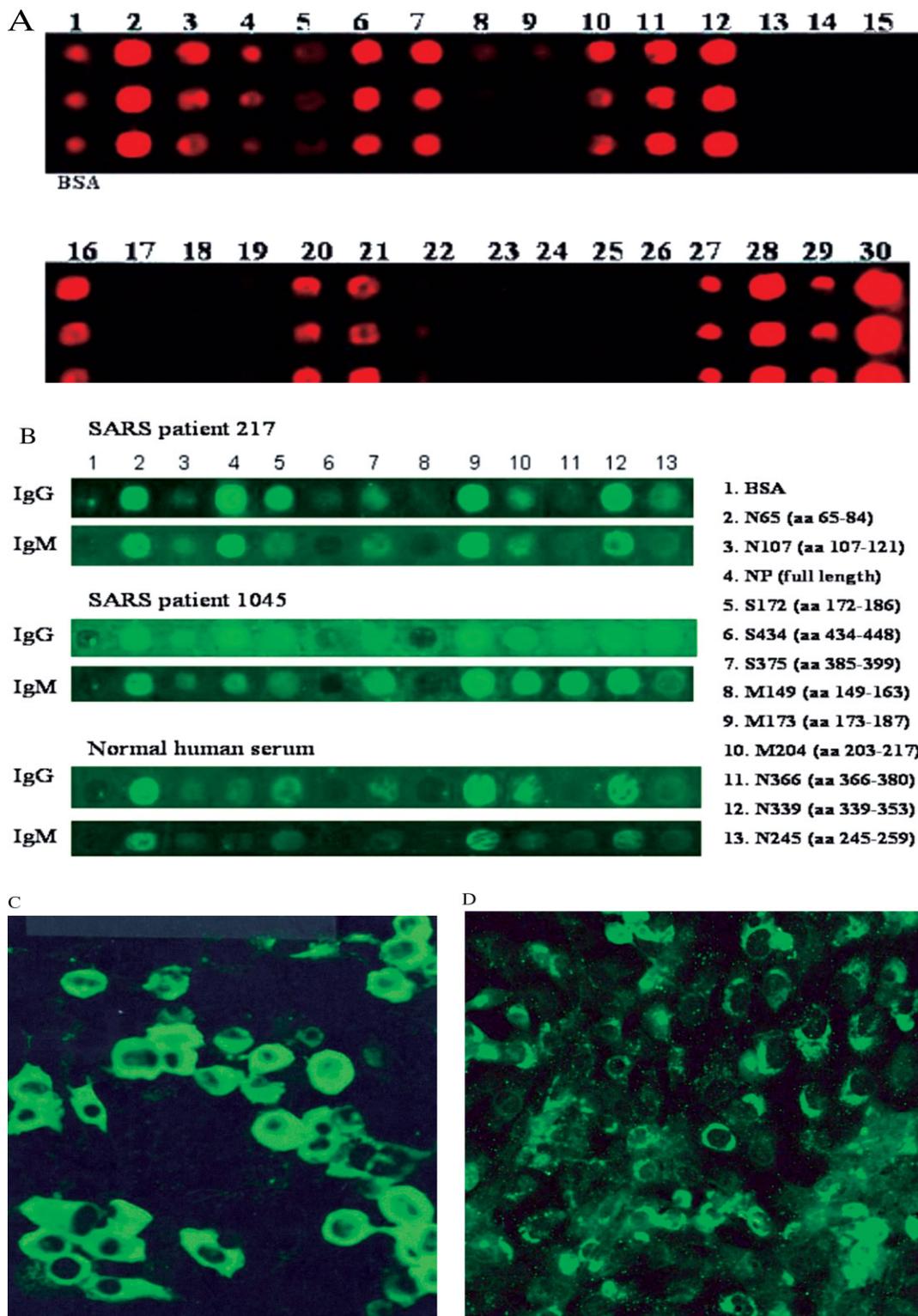
Figure 8. A, Twenty-eight synthetic peptides are reacted with rabbit anti-N protein antibody in triplicate. Column 1 is BSA and column 30 is the SARS full-length N protein as controls. Peptide 1 (aa 1–20), peptide 11 (aa 150–170), peptide 27 (aa 390–410) showed strong responses. B, Immunofluorenscence assays of peptides derived from SARS structural proteins N, S and M using IgG and IgM antibodies purified from sera of SARS patients. C and D, Immunological detection of SARS proteins expressed in Vero E6 cells infected by SARS virus, detected with anti-SARS N protein antibody (C) and detected with anti-M149 peptide antibody (D).

the three N proteins from SARS, 229E and OC43 CoVs. The results suggest that the N protein has two ordered regions, covering aa 90–160 in the *N*-terminal region and aa 270–360 in the *C*-terminal region for SARS, and aa 1–101 in the *N*-terminal region and aa 250–350 in the *C*-terminal region for 229E, respectively (Fig. 3B). The sequence alignment of the three coronaviruses reveals that two conserved regions are found between aa 50–170 and aa 250–360 (SARS numbering) (Fig. 1B). The solution structure of the *N*-terminal domain of aa 45–181 has been determined by NMR, which revealed a mostly β-sheet structure [31]. The protein fold appears to resemble a RNA-binding domain of RNP RNA-binding proteins. It is likely that the conserved *N*-terminal domain of coronaviruses is responsible for RNA binding.

Many viral nucleocapsid proteins are known to form a dimer as the start building motif for the formation of the higher-order structure of nucleocapsid. The full-length N proteins form oligomers, predominantly as a dimer, as judged from our cross-linking and gel filtration experiments. The gel filtration result of the SARS N protein agreed with that of Luo *et al.* [32]. Full-length N proteins are labile and easily degraded into lower M_r fragments. Upon storage, the end product appears to be the SARS N protein *C*-terminal domain of ~30 kDa based on the LC-ESI-TOF MS analysis. The *N*-terminal amino acid has been identified to be around aa 169. Therefore we suggest that the *C*-terminal 30 kDa, which contains the *C*-terminal conserved region, is an oligomerization domain. This is also supported by our ND4 protein gel filtration assay result. Both results of He *et al.* [23] and Surjit *et al.* [24] based on the sequence truncation and two hybrid system studies have suggested different dimerization regions. Taken together, we suggest that there may be more than one sequence responsible for SARS N protein dimerization within the *C*-terminal 30 kDa region. The somewhat broader dimer peaks of the full-length N proteins of the gel filtration diagram compared to that of monomer protein (97 kDa) peak also imply this hypothesis. There could be N protein dimer isoforms coexisting. For example, dimers could be made through the association of the SR sequence [23] or the ND4 domain.

Taking the results so far, we propose that the N protein of coronavirus is a two-domain protein, with the RNA-binding domain (near aa 50–170) in the *N*-terminal region and the dimerization domain (near aa 170–360) in the *C*-terminal domain. Presumably, the N protein binds to the RNA genome of coronavirus as a dimer first. Further interactions between the N protein dimers result in a higher order structure of the NP/RNA complex, which interacts with other structural proteins such as S, M and E proteins to form the complete virion.

The availability of the soluble full-length SARS N protein allowed us to produce useful high binding-affinity polyclonal antibodies. We have used these antibodies to map out three strong antigenic sites of N protein, consisting of aa 1–20, aa 150–170 and aa 390–410, which match those identified using sera of SARS patients. It is interesting to compare the anti-

genic sites, both strong and medium, on SARS N protein with order/disorder regions in the amino acid sequence. There are five sequence regions (aa 1–35, aa 60–95, aa 120–170, aa 210–230, aa 270–305 and aa 375–422) possessing medium/strong antigeneicity and all these sequences do not fall in the ordered regions predicted by PONDR program. In fact, two strong antigenic regions are at the *N*-terminal and *C*-terminal ends, while the other three lie at the boundaries between order and disorder regions. It seems that this information could be useful in the prediction of antigeneicity of proteins.

5 Concluding remarks

Finally, these antibodies were tested against SARS virus infected Vero E6 cells and they clearly detected the N protein in the Vero cells. Therefore we believe that the antibodies produced by rabbits are physiologically relevant and can be used as detection tools. We have also prepared mouse monoclonal antibodies against SARS NP, which are being characterized in details (data not shown). These mAb reagents will be valuable tools to design diagnostic kits later.

6 References

[1] Peiris, J. S., Yuen, K. Y., Osterhaus, A. D., Stohr, K., *N. Engl. J. Med.* 2003, *349*, 2431–2441.

[2] Berger, A., Drosten, C., Doerr, H. W., Sturmer, M., Preiser, W., *J. Clin. Virol.* 2004, *29*, 13–22.

[3] Tsang, K. W., Mok, T. Y., Wong, P. C., Ooi, G. C., *Respirology* 2003, *8*, 259–265.

[4] Falsey, A. R., Walsh, E. E. *Lancet* 2003, *361*, 1312–1313.

[5] Ksiazek, T. G., Erdman, D., Goldsmith, C., *N. Engl. J. Med.* 2003, *348*, 1953–1966.

[6] Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A. *et al.*, *Science* 2003, *300*, 1394–1399.

[7] Ruan, Y. J., Wei, C. L., Ee, A. L., Vega, V. B. *et al.*, *Lancet* 2003, *361*, 1779–1785.

[8] Huang, L. R., Chiu, C. M., Yeh, S. H., Huang, W. H. *et al.*, *J. Med. Virol.* 2004, *73*, 338–346.

[9] Ksiazek, T. G., West, C. P., Rollin, P. E., Jahrling, P. B., Peters, C. J., *J. Infect. Dis.* 1999, *179 (Suppl 1)*, S192–S198.

[10] He, Q., Chong, K. H., Chng, H. H., Leung, B. *et al.*, *Clin. Diagn. Lab. Immunol.* 2004, *11*, 417–422.

[11] Tan, Y. J., Goh, P. Y., Fielding, B. C., Shen, S. *et al.*, *Clin. Diagn. Lab. Immunol.* 2004, *11*, 362–371.

[12] Chang, M. S., Lu, Y. T., Ho, S. T., Wu, C. C. *et al.*, *Biochem. Biophys. Res. Commun.* 2004, *314*, 931–936.

[13] Wu, H. S., Hsieh, Y. C., Su, I. J., Lin, T. H. *et al.*, *J. Biomed. Sci.* 2004, *11*, 117–126.

[14] Liu, X., Shi, Y., Li, P., Li, L. *et al.*, *Clin. Diagn. Lab. Immunol.* 2004, *11*, 227–228.

[15] Lau, L. T., Fung, Y. W., Wong, F. P., Lin, S. S. *et al.*, *Biochem. Biophys. Res. Commun.* 2003, *312*, 1290–1296.

[16] Bressler, A. M., Nolte, F. S., *J. Clin. Microbiol.* 2004, *42*, 987–991.

[17] Poon, L. L., Chan, K. H., Wong, O. K., Cheung, T. K. *et al.*, *Clin. Chem.* 2004, *50*, 67–72.

[18] Ng, E. K., Hui, D. S., Chan, K. C., Hung, E. C. *et al.*, *Clin. Chem.* 2003, *49*, 1976–1980.

[19] Shi, Y., Yi, Y., Li, P., Kuang, T. *et al.*, *J. Clin. Microbiol.* 2003, *41*, 5781–5782.

[20] Yeh, S. H., Wang, H. Y., Tsai, C. Y., Kao, C. L. *et al.*, *Proc. Natl. Acad. Sci. USA* 2004, *101*, 2542–2547.

[21] Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., Higgins, D. G., *Nucleic Acids Res.* 1997, *24*, 4876–4882.

[22] Dunker, A. K., Brown, C. J., Lawson, J. D., Iakoucheva, L. M., Obradovic, Z., *Biochemistry* 2002, *41*, 6573–6582.

[23] He, R., Dobie, F., Ballantine, M., Leeson, A. *et al.*, *Biochem. Biophys. Res. Commun.* 2004, *316*, 476–483.

[24] Surjit, M., Liu, B., Kumar, P., Chow, V. T., Lal, S. K., *Biochem. Biophys. Res. Commun.* 2004, *317*, 1030–1036.

[25] Lin, Y., Shen, X., Yang, R. F., Li, Y. X. *et al.*, *Cell Res.* 2003, *13*, 141–145.

[26] Chen, Z., Pei, D., Jiang, L., Song, Y. *et al.*, *Clin. Chem.* 2004, *50*, 988–995.

[27] Tsai, J. C., Zelus, B. D., Holmes, K. V., Weiss, S. R., *J. Virol.* 2003, *77*, 841–850.

[28] Lewicki, D. N., Gallagher, T. M., *J. Biol. Chem.* 2002, *277*, 19727–19734.

[29] Bonavia, A., Zelus, B. D., Wentworth, D. E., Talbot, P. J., Holmes, K. V., *J. Virol.* 2003, *77*, 2530–2538.

[30] Elia, G., Decaro, N., Tinelli, A., Martella, V. *et al.*, *New Microbiol.* 2002, *25*, 275–280.

[31] Huang, Q., Yu, L., Petros, A. M., Gunasekera, A. *et al.*, *Biochemistry* 2004, *43*, 6059–6063.

[32] Luo, H., Ye, F., Sun, T., Yue, L. *et al.*, *Biophys. Chem.* 2004, *112*, 15–25.

[33] Fields, C. G., Lloyd, D. H., Macdonald, R. L., Otteson, K. M., Noble, R. L., *Peptide Res.* 1991, *4*, 95–101.