

# Cloning, sequencing, expression, and purification of SARS-associated coronavirus nucleocapsid protein for serodiagnosis of SARS

Khalid Amine Timani<sup>a</sup>, Li Ye<sup>a</sup>, Linbai Ye<sup>a,\*</sup>, Ying Zhu<sup>a</sup>, Zhenghui Wu<sup>a</sup>, Zuojiang Gong<sup>b</sup>

<sup>a</sup> Key Laboratory of Institute of Virology, College of Life Science, Wuhan University, Wuhan, Hubei 430072, China

<sup>b</sup> Department of Infectious Disease, Renmin Hospital of Wuhan University, Wuhan, Hubei 430060, China

Received in revised form 10 January 2004; accepted 11 January 2004

## Abstract

A novel coronavirus has been associated with a worldwide outbreak of atypical pneumonia referred to as Severe Acute Respiratory Syndrome (SARS-CoV). SARS-CoV nucleocapsid (N) protein has been cloned sequenced and expressed in *Escherichia coli* strain. Purified N protein was used to measure the SARS-CoV specific IgG antibodies from 16 SARS-CoV infected patients' sera and from 131 control subjects using ELISA assay. Specific antibody responses to the purified recombinant N protein after 10, 20, and 30 days of disease onset were observed in 13 of 16 (81.3%), 16 of 16 (100%) and 16 of 16 (100%) SARS patients sera, respectively. Comparison of detection results with a commercially available diagnostic kit coated with a mixture of SARS-CoV viral proteins showed 9 of 16 (56.3%), 13 of 16 (81.3%), and 15 of 16 (93.7%) positive responses, respectively. None of 131 control sera gave positive reaction in either assay. This data suggests that the N protein of SARS-CoV is immunodominant and this ELISA based test assay for detecting the SARS-CoV N antigen may hold a significant value for SARS diagnosis. © 2004 Elsevier B.V. All rights reserved.

**Keywords:** SARS-CoV; Nucleocapsid protein; Expression and purification; ELISA

## 1. Introduction

Severe Acute Respiratory Syndrome (SARS) is a recently recognized febrile respiratory illness that first appeared in southern China in November 2002, and has since spread to several other countries. The disease was caused by a novel coronavirus, now provisionally named "SARS-associated coronavirus" (SARS-CoV) (Drosten et al., 2003; Ksiazek et al., 2003; Peiris et al., 2003). The severity of this disease is such that the mortality rate appears to be 9.6% and diagnostic methods for SARS are urgently needed (Lipsitch et al., 2003; Riley et al., 2003). The SARS-CoV is an enveloped virus of a positive single-stranded RNA genome ~29.7 kb in length and has 11 open reading frames. The genome organization is similar to that of other coronaviruses, where phylogenetic analyses and sequence comparisons showed that the SARS-CoV is not closely related to any of previously characterized coronaviruses (Marra et al., 2003; Rota et al., 2003). The structural protein genes of SARS-CoV contain

four major open reading frames; including the surface spike (S) glycoprotein, the small membrane (M) protein, the envelope (E) glycoprotein, the nucleocapsid (N) protein, and a set of accessory proteins whose number and sequence vary among different coronaviruses (Siddell, 1995).

The N proteins have been shown to be strong immunogens in several coronaviruses, such as murine coronavirus (Wage et al., 1993), Turkey coronavirus (Akin et al., 2001), porcine reproductive and respiratory syndrome coronavirus (Casal et al., 1998). It has been reported that N protein accumulates intra-cellularly even before it is packed in the mature virus (Collisson et al., 2000) and is the most abundant virus derived-protein throughout the infection, probably because its template mRNA is the most abundant subgenomic RNA (Hiscox et al., 1995; Keck et al., 1988). These features make it a suitable candidate for raising antibodies for diagnostic applications.

In this study, SARS-CoV N protein gene was cloned, sequenced, expressed in *Escherichia coli* strain and purified to homogeneity. Using the purified N protein as an antigen for SARS diagnosis and comparing it with a commercially available detection kit coated with a mixture of SARS viral proteins, we demonstrated that the antibody to N protein has high sensitivity and specificity for SARS diagnosis.

\* Corresponding author. Tel.: +86-27-8768-2372;

fax: +86-27-8768-4763.

E-mail addresses: linbaiye@whu.edu.cn, linbaiye@hotmail.com (L. Ye).

## 2. Materials and methods

### 2.1. Clinical samples

Serum samples from 16 SARS-CoV infected patients (the clinical diagnostic for SARS followed the clinical description of SARS released by WHO), 3 from Hubei, 1 from Beijing, and 12 from Henan were collected from 10, 20, and 30 days of disease onset. One hundred and thirty-one serum samples of symptom free direct contacts to the above patients (physicians, nurses, and their family members) or non-SARS patients who showed symptoms of fever  $>38^{\circ}\text{C}$ , cough and sore throat, were collected as controls by Dr. Zuojiang Gong from Renmin Hospital of Wuhan University, Wuhan, China.

### 2.2. Cells and virus infection

VeroE6 cells were maintained in minimal essential medium (GIBCO) with 10% fetal bovine serum (Hyclone), 200  $\mu\text{M}$  L-glutamine, penicillin, streptomycin, at  $37^{\circ}\text{C}$  in humidified atmosphere. Virus isolation was performed on nasopharyngeal swab specimen of a fatal SARS case belonging to original case cluster from Wuhan in Hubei province, China, which was inoculated into VeroE6 cells. Viral particles were purified by centrifugation and its genetic material RNA was extracted by Trizol reagent (Invitrogen) according to the manufacturer's instructions.

### 2.3. RT-PCR

The cDNA was made by adding 5  $\mu\text{l}$  of extracted viral RNA to 20  $\mu\text{l}$  of total reaction mixture using AMV-reverse transcriptase kit (TAKARA Biomedical, Japan), containing 50  $\mu\text{M}$  random primer. The Viral cDNA was used as a template for PCR by adding 3  $\mu\text{l}$  to 50  $\mu\text{l}$  of total reaction mixture containing  $10\times$  transcription buffer, 20 mM  $\text{MgCl}_2$ , 1 mM of each dNTP, 20 pmol of each primer, and 2.5 U of LATAq<sup>TM</sup> DNA polymerase (TAKARA Biomedical, Japan). The primers were synthesized according to HKU39849 accession number AY278491, i.e. forward (AGCTGGATC-CATGTCTGATAATGGACCCCAATCAAAC) and reverse (AGCTGAATTCCATCATGAGTGTATTATGCCTGAGT). Products were amplified using the following conditions,  $94^{\circ}\text{C}$  for 4 min, then 31 cycles of  $94^{\circ}\text{C}$  for 30 s,  $55^{\circ}\text{C}$  for 45 s, and  $72^{\circ}\text{C}$  for 1 min followed by one cycle of  $72^{\circ}\text{C}$  for 10 min, then examined in 0.7% agarose gel electrophoresis.

### 2.4. Plasmid construction

The N protein gene cDNA was cloned into pGEM<sup>®</sup>-T Easy vector (Promega, UK) by standard techniques then transformed into *E. coli* strain DH5 $\alpha$ . The constructed pGEMT-NP was digested with BamHI and EcoRI (MBI Fermentas, Germany), sub-cloned into pETHis expression vector thereby introducing a small hexahistidine-T7 Tag

at the N-terminus of the protein to facilitate subsequent purification. The recombinant plasmid was transformed into DH5 $\alpha$  strain. Seven positive colonies were analyzed, confirmed by PCR and sequenced.

### 2.5. Expression and purification of SARS-CoV nucleocapsid protein

The pETHis-N plasmid was transformed into *E. coli* strain BL21 (DE3). Protein expression was induced by addition of 1 mmol/l isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation and the pellet was suspended in binding buffer (20 mM Tris-HCl pH 7.9, 500 mM NaCl, 5 mM Imidazole, 1 mM NaF, and 1 mM PMSF), sonicated and centrifuged at  $12,000 \times g$  at  $4^{\circ}\text{C}$  for 30 min. The clear supernatant was applied in to Ni-NTA affinity column (Qiagen) then washed with gradient washing buffers (20 mM Tris-HCl pH 7.9, 500 mM NaCl, 5, 20, 40, and 60 mM Imidazole). Purified N protein was eluted with elution buffer (20 mM Tris-HCl pH 7.9, 500 mM NaCl, 100 mM Imidazole) and concentrated by dialysis against polyethyleneglycol 6000.

### 2.6. ELISA

Fifty nanogram purified N protein, dissolved in coating buffer (0.016 M  $\text{Na}_2\text{CO}_3$ , 0.034 M  $\text{NaHCO}_3$  pH 9.6), were added to a 96-well microplate and incubated at  $4^{\circ}\text{C}$  overnight. Wells were then blocked with 4% non-fat milk, 0.5% BSA in phosphate buffer saline (PBS), incubated for 1 h at RT, washed with PBS and dried at  $37^{\circ}\text{C}$  for 1 h. For detection by chemiluminescent method, sera were diluted in sample buffer (1:25) then incubated at  $37^{\circ}\text{C}$  for 30 min. Each well was washed then incubated with peroxidase-conjugated rabbit anti-human antibody at  $37^{\circ}\text{C}$  for 30 min. The wells were washed again with PBS containing 0.5% Tween-20. The peroxidase reaction was visualized by using *O*-phenylenediamine as a substrate, after 10 min at  $37^{\circ}\text{C}$  the reaction was stopped by adding 50  $\mu\text{l}$ /well of 1N  $\text{H}_2\text{SO}_4$  and absorbance was read at 490 nm. The serum samples were run in triplicate. The cutoff value was defined as the mean OD of control samples plus two standard deviations (S.D.).

The results of the N protein ELISA were compared with results from a commercially available ELISA kit based on viral S, N, and E proteins prepared by the Beijing Genomics Institute, Chinese Academy of Sciences, Beijing, China.

## 3. Result

The nucleocapsid protein gene sequence of SARS-CoVHB was performed and the sequence (1269 nt) was deposited into the gene bank accession number AY365036. The full length of N protein (422 amino acids) with an amino terminal hexahistidine-tag was expressed in BL21 (DE3) and purified by Ni-NTA affinity column. After 2 h induction

Table 1

Results of the SARS-CoV recombinant N protein ELISA in comparison to the commercially available ELISA for the sero-diagnosis of SARS

Coated antigen	Number positive/number tested (% positive) SARS infected patients sera			Control sera <sup>a</sup>
	10 days	20 days	30 days	
Nucleocapsid protein	13/16 (81.3%)	16/16 (100%)	16/16 (100%)	0/131 (0%)
Mixture of antigens	9/16 (56.3%)	13/16 (81.3%)	15/16 (93.7%)	0/131 (0%)

<sup>a</sup> Serum samples of both the symptom free direct contacts to the above patients (physicians, nurses, and their family members) and non-SARS patients who showed symptoms of fever >38°C, cough and sore throat.

with IPTG, high level expression of N protein was observed (Fig. 1, lane 1) compared to uninduced cells lysate (Fig. 1, lane 2). The His-tagged protein was purified from the soluble fraction resulting in a single band of the expected mass of 47 kDa (Fig. 1, lane 3). This data demonstrates that the SARS-CoV N protein was successfully expressed in *E. coli* and purified to homogeneity. The N protein was identified by western blot using human SARS-CoV infected serum (data not shown).

To investigate the immunoreactivity of the nucleocapsid protein with the 16 SARS-CoV infected patient's sera collected at 10, 20, and 30 days after disease onset, an ELISA assay was established using the purified recombinant N protein. Interestingly, we found that 13 of 16 (81.3%), 16 of 16 (100%), and 16 of 16 (100%) SARS patients had a significant positive response to the N protein after 10, 20, and 30 days of disease onset, respectively. When comparing this result with a commercially available detection ELISA kit coated with a mixture of viral proteins (S, N, and E), prepared by Beijing Genomics Institute, Chinese Academy of Science, Beijing, China, we observed that the immunoreactivity with the SARS patient's antibodies were 9 of 16 (56.3%), 13 of

16 (81.3%), and 15 of 16 (93.7%) at 10, 20, and 30 days of disease onset, respectively, while none of the 131 control sera from symptom free and non-SARS patients gave a positive reaction in either assay (Table 1). This result showed that the IgG antibodies to N protein is commonly present in all 16 SARS patients sera and could be detected as early as 10 days of disease onset. The mean OD  $\pm$  S.D. obtained using 131 control sera were  $0.079 \pm 0.052$  and  $0.0688 \pm 0.034$  for N protein coated ELISA and commercially available detection kit, respectively.

#### 4. Discussion

Despite the initial rapid progress in the discovery of the causative agent and the early development of diagnostic tests, there is a critical clinical need for further laboratory tests that can confirm a diagnosis of SARS-CoV infection in its early stages. A real time RT-PCR with optimized RNA extraction methods could detect 40 of 50 (80%) nasopharyngeal aspirate samples from confirmed SARS patients at the first 3 days of disease onset (Poon et al., 2003). However, comparable improvements in serological methods are needed.

We used purified N protein as an antigen in an ELISA assay and studied its sensitivity and specificity for sero-diagnosis of SARS. The results indicate that the ELISA based on the SARS-CoV N protein has high sensitivity and high specificity. Eighty-one percent of the SARS patients had a positive result 10 days after disease onset while 100% had positive results at 20 and 30 days post-infection. When comparing this result with the commercially available diagnosis kit using a mixture of viral proteins, only 56.3% of the sera were positive 10 days after disease onset while 81.3% and 93.7% were positive after 20 and 30 days, respectively. No reactivity was observed in the control samples in either assay. The difference in the positive response between these two assays may be due to the quantity of coated N protein or to the expression and purification procedures that has been used for the recombinant viral proteins. Our work suggests that the high specificity and sensitivity of nucleocapsid protein ELISA indicates that it has potential to be an useful option for the sero-diagnosis of SARS, and will be particularly in laboratories that lack BSL 3 facilities. However, the cross reactivity of this assay in patients

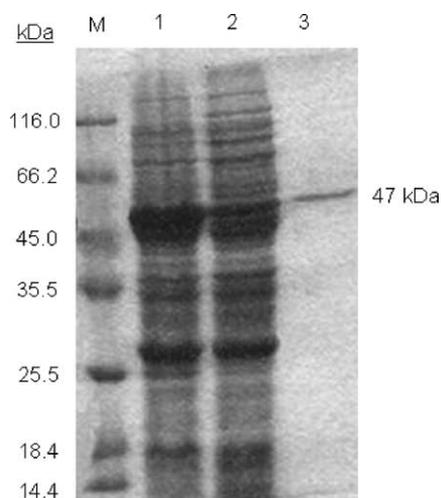


Fig. 1. SDS-PAGE analysis of the expression of the nucleocapsid protein in *E. coli*. Lanes 1 and 2: crude material from bacterial cultures containing the pETHis expression vectors with cDNA insert encoding the N protein, before (lane 2) and after (lane 1) the 2h induction. Lane 3: purified recombinant protein by Ni-NTA affinity column. M: markers with their corresponding molecular masses.

with recent infection with human coronaviruses 229E and OC43 require to be studied more closely. Therefore, until further experience is accumulated with this assay, it is suggested that the nucleocapsid protein ELISA can be useful as a screening test, with positive results being confirmed by immunofluorescence or virus neutralization tests.

### Acknowledgements

We thank Dr. Zhenhui Gong (Hubei Centre of Disease Control, Wuhan, China) for providing the commercially available ELISA detection kit. Center of disease control (CDC, Henan, China) for providing the 12 SARS patients sera and performing the ELISA assays at their laboratory. The authors are grateful to Dr. Stephen J. Polyak (Department of Laboratory Medicine, University of Washington, Seattle, USA) for his support and proof-reading the manuscript.

### References

- Akin A, Lin TL, Wu CC, Bryan TA, Hooper T, Schrader D. Nucleocapsid protein gene sequence analysis reveals close genomic relationship between Turkey coronavirus and avian infectious bronchitis virus. *Acta Virol* 2001;45:31–8.
- Casal JI, Rodriguez MJ, Sarraseca J, Garcia J, Plana-Duran J, Sanz A. Identification of a common antigenic site in the nucleocapsid protein of European and North American isolates of porcine reproductive and respiratory syndrome virus. *Adv Exp Med Biol* 1998;440:469–77.
- Collisson EW, Pei J, Dzielawa J, Seo SH. Cytotoxic T lymphocytes are critical in the control of infectious bronchitis virus in poultry. *Dev Comput Immunol* 2000;24:187–200.
- Drosten C, Gunther S, Preiser W, Van der Werf S, Brodt HR, Becker S, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *New Engl J Med* 2003;348:1967–76.
- Hiscox JA, Caavanagh D, Britton P. Quantification of individual subgenomic mRNA species during replication of the coronavirus transmissible gastroenteritis. *Virus Res* 1995;36:119–30.
- Keck JG, Hogue BG, Brian DA, Lai MMC. Temporal regulation of bovine coronavirus RNA synthesis. *Virus Res* 1988;9:343–56.
- Ksiazek TG, Erdman D, Goldsmith CS, Zaki SR, Peret T, Emery S, et al. A novel coronavirus associated with severe acute respiratory syndrome. *New Engl J Med* 2003;348:1953–66.
- Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, et al. Transmission dynamics and control of severe acute respiratory syndrome. *Science* 2003;300:1966–70.
- Marra MA, Jones SJM, Caroline R, Astell CR, Holt RA, et al. The genome sequence of the SARS-associated coronavirus. *Science* 2003;300:1399–404.
- Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, Lim W, et al. Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 2003;361:1319–25.
- Poon LL, Wong OK, Chan KH, Luk W, Yuen KY, Peiris JS, et al. Rapid diagnosis of coronavirus associated with severe acute respiratory syndrome (SARS). *Clin Chem* 2003;49:953–5.
- Riley S, Fraser C, Donnelly CA, Ghani AC, Abu-Raddad LJ, Hedley AJ, et al. Transmission dynamics of the etiological agent of severe acute respiratory syndrome (SARS) in Hong Kong: the impact of public health investigations. *Science* 2003;300:1961–6.
- Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* 2003;300:1394–9.
- Siddell SG. *The coronaviridae*. New York: Plenum Press; 1995.
- Wage H, Schliephake A, Korner H, Flory E, Wage H. An immunodominant CD4+ T cells site on the nucleocapsid protein of murine coronavirus contributes to protection against encephalomyelitis. *J Gen Virol* 1993;74:1287–94.