

Proteomic analysis on structural proteins of Severe Acute Respiratory Syndrome coronavirus

Wantao Ying^{1*}, Yunwei Hao^{1*}, Yangjun Zhang^{1*}, Wenming Peng^{2*}, Ede Qin^{2*}, Yun Cai¹, Kaihua Wei³, Jie Wang³, Guohui Chang², Wei Sun¹, Shujia Dai¹, Xiaohai Li¹, Yunping Zhu¹, Jianqi Li¹, Songfeng Wu¹, Lihai Guo¹, Jingquan Dai¹, Jinglan Wang¹, Ping Wan¹, Tinggui Chen¹, Chunjuan Du¹, Dong Li¹, Jia Wan¹, Xuezhong Kuai¹, Weihua Li³, Rong Shi¹, Handong Wei¹, Cheng Cao⁴, Man Yu², Hong Liu², Fangting Dong³, Donggen Wang¹, Xuemin Zhang³, Xiaohong Qian¹, Qingyu Zhu², Fuchu He¹

¹Beijing Institute of Radiation Medicine, Beijing, China

²Beijing Institute of Microbiology and Epidemiology, Beijing, China

³National Center of Biomedical Analysis, Beijing, China

⁴Beijing Institute of BioTechnology, Beijing, China

Recently, a new coronavirus was isolated from the lung tissue of autopsy sample and nasal/throat swabs of the patients with Severe Acute Respiratory Syndrome (SARS) and the causative association with SARS was determined. To reveal further the characteristics of the virus and to provide insight about the molecular mechanism of SARS etiology, a proteomic strategy was utilized to identify the structural proteins of SARS coronavirus (SARS-CoV) isolated from Vero E6 cells infected with the BJ-01 strain of the virus. At first, Western blotting with the convalescent sera from SARS patients demonstrated that there were various structural proteins of SARS-CoV in the cultured supernatant of virus infected-Vero E6 cells and that nucleocapsid (N) protein had a prominent immunogenicity to the convalescent sera from the patients with SARS, while the immune response of spike (S) protein probably binding with membrane (M) glycoprotein was much weaker. Then, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to separate the complex protein constituents, and the strategy of continuous slicing from loading well to the bottom of the gels was utilized to search thoroughly the structural proteins of the virus. The proteins in sliced slots were trypsinized in-gel and identified by mass spectrometry. Three structural proteins named S, N and M proteins of SARS-CoV were uncovered with the sequence coverage of 38.9, 93.1 and 28.1% respectively. Glycosylation modification in S protein was also analyzed and four glycosylation sites were discovered by comparing the mass spectra before and after deglycosylation of the peptides with PNGase F digestion. Matrix-assisted laser desorption/ionization-mass spectrometry determination showed that relative molecular weight of intact N protein is 45 929 Da, which is very close to its theoretically calculated molecular weight 45 935 Da based on the amino acid sequence deduced from the genome with the first amino acid methionine at the N-terminus depleted and second, serine, acetylated, indicating that phosphorylation does not happen at all in the predicted phosphorylation sites within infected cells nor in virus particles. Intriguingly, a series of shorter isoforms of N protein was observed by SDS-PAGE and identified by mass spectrometry characterization. For further confirmation of this phenomenon and its related mechanism, recombinant N protein of SARS-CoV was cleaved *in vitro* by caspase-3 and -6 respectively. The results demonstrated that these shorter isoforms could be the products from cleavage of caspase-3 rather than that of caspase-6. Further, the relationship between the caspase cleavage and the viral infection to the host cell is discussed.

Keywords: Proteome / Severe Acute Respiratory Syndrome

PRO 0676

1 Introduction

Severe Acute Respiratory Syndrome (SARS), as a newly infectious disease, has seriously threatened the health of people worldwide. There were 8402 probable SARS

cases with 772 deaths having been reported from 29 countries up to June 4 2003 (<http://www.who.int.crs/sars/country/en>). An overall estimate of case fatality reached 14–15% as reported by WHO [1] and the mortality rate in people older than 60 years could even be as high as 43–55% [2].

Correspondence: Dr. Fuchu He, Beijing Institute of Radiation Medicine, 27 Taiping Road, Beijing 100850, China

E-mail: hefc@nic.bmi.ac.cn

Fax: +86-10-6817-7208

Abbreviations: CoV, coronavirus; FA, formic acid; IBV, infectious bronchitis virus; M, membrane glycoprotein; MCP, microchannel plate; MHV, murine hepatitis virus; N protein, nucleocapsid protein; SARS, severe acute respiratory syndrome; S protein, spike protein

A number of laboratories worldwide have undertaken research on the identification of the causative agent of the SARS. An unknown virus that causes SARS was first isolated and announced on March 22 2003 [3].

* These authors contributed equally to this work.

Then coronavirus as a possible causal agent of SARS was identified from patients by using serological and RT-PCR methods and the possible route of transmission of the virus was analyzed [4]. Further research in different laboratories indicated that a new coronavirus was associated with SARS by characterization of cytopathological features and ultrastructural features, and this new coronavirus is only distantly related to known coronavirus by genetic characterization [5, 6]. The genomes of SARS-associated coronavirus (SARS-CoV) from different strains were sequenced and declared successively [7–10].

A new strain of SARS-CoV was isolated from the lung tissue of autopsy sample and nasal/throat swabs of the patients with SARS and identified by morphology, serology, animal experiments, RT-PCR and partial sequence analysis by Qin *et al.* [11]. The causative association between the isolate and SARS was also determined. Complete genome sequencing and comparative analysis of the isolate (BJ01 GenBank accession number AY278488) indicated that the genome size is 29.725 Kb and has 11 ORFs [9]. The whole genome is composed of a stable region encoding an RNA-dependent RNA polymerase and a variable region representing four coding sequences for viral structural proteins (the S, E, M, N proteins) and five putative uncharacterized proteins. Its gene order is identical to that of other known coronaviruses.

Although the genome sequencing and comparative analysis provided abundant information to realize the characteristics of SARS-CoV and various predictions about the structures and functions of the proteins composing the virus particles, the information on natural proteins with post-translational modifications and possible isoforms or cleavage products is difficult to obtain from the genome sequence. But this information is very important to understand the functions of these proteins and further to reveal the properties of the virus. Thus, a systematic proteomics research is necessary to identify these proteins at their natural forms and to probe their processing and modification directly. For this purpose, a mass spectrometric characterization of proteins from the SARS virus was recently reported [12]. Two antigenic proteins with molecular masses ~ 46 and ~ 139 kDa, were characterized respectively. The glycosylation modification of the spike protein was determined.

In this study, the structural proteins of SARS-CoV isolate BJ01 were investigated by proteomics strategies. Three out of four structural proteins with antigenicity against the convalescent sera from patients with SARS were characterized by SDS-PAGE and/or RP-HPLC combined with mass spectrometry. The peptide sequences with

some modification and cleavage that existed in these proteins were also analyzed. The relationship between these modifications and cleavages with their probable functions are discussed.

2 Materials and methods

2.1 Sample source

The BJ01 strain of SARS-CoV was separated from lung tissue of deceased patients and cultured in Vero E6 cells. When the cytopathic effect was observed in more than 75% cells infected with BJ01 strain of SARS-CoV, the cultured supernatant containing virus and infected cells was harvested. The cells were frozen and thawed repeatedly in the medium to completely release the virus particles. After centrifugation at 6000 rpm for 10 min (JA-25.50, Beckmann, Fullerton, CA, USA), the lysates were dialyzed and subsequently measured at 595 nm for protein concentration according to Lowry methods [13] and then lyophilized. As for controls, the noninfected Vero E6 cells were cultured and processed in the same way as the infected cells.

2.2 Chemicals and reagents

Electrophoresis reagents including acrylamide, *N,N*-methylenebisacrylamide (Bis), TEMED, Tris base, glycine, DTT, Low Molecular Marker were purchased from Amersham Biosciences (Uppsala, Sweden). Iodoacetamide and TFA were from Acros (New Jersey, USA). Trypsin (sequencing grade) and DTT were obtained from Promega (Madison, WI, USA). Endoproteinase Glu-C (sequencing grade), PNGase F, ammonium bicarbonate and ammonium acetic acid were purchased from Sigma (St. Louis, MO, USA). Caspase-3 and caspase-6 were from BD Pharmingen (San Diego, CA, USA). Acetonitrile (HPLC grade) was purchased from J. T. Baker (Phillipsburg, NJ, USA); formic acid (FA) was obtained from Beijing Chemicals (Beijing, China).

2.3 SDS-PAGE

The lyophilized samples of Vero E6 cells lysates infected by SARS-CoV (S) and control (V) were suspended in loading buffer (50 mM/L Tris-HCl pH 6.8, 100 mM/L DDT, 2% SDS, 0.1% bromophenol blue, 10% glycerol), respectively. The samples were run on SDS-PAGE (T = 13%) in Tris-glycine running buffer with 100 μ g protein *per* lane, and stained with Coomassie blue R250.

2.4 Western blotting

Human specific anti-SARS-CoV sera were obtained from 10 clinical cases of convalescent SARS patients (14–28 d after being diagnosed as SARS) and one patient (21 d after onset of SARS), and the control normal human serum was collected from 6 uninfected donors with their permission.

One group (A) of 100 μ g and two groups (B and C) of 50 μ g lyophilized samples of SARS-CoV infected Vero E6 cells lysates (S) and Vero E6 cells lysates (V) were set up for Western blot experiments respectively. All three groups were electrophoresed in the same conditions with 13% SDS-PAGE. The separated proteins were transferred to Hybond-P PVDF membrane (Amersham Biosciences) at 20°C for 3 h and the remaining gel was stained with Coomassie blue for protein identification [6]. After overnight incubation at 4°C in blocking buffer (20 mM Tris-HCl pH 7.5, 140 mM NaCl, 0.05% Tween-20, 5% nonfat dried milk), the membranes of group A were probed by the addition of the antisera (1:1000, v/v in PBST, 5% nonfat dried milk) from one clinical case of convalescent SARS and incubated for 2 h at room temperature. Group B membranes was hybridized with pooled antisera from 10 convalescent patients, which were all qualified to be positive for antibodies to the SARS-CoV by indirect immunofluorescence assay (IFA). Group C was testified by adding sera from uninfected donors (negative in IFA). After washing in PBST 3 \times 10 min each, the membrane was incubated for 1 h with horseradish peroxidase-conjugated second antibody (Amersham Biosciences, 1:10 000, v/v in PBST, 5% nonfat dried milk), and then washed in PBST three times. Finally the blots were developed with ECL Western blot kit (Santa Cruz Biotechnology, Santa Cruz, CA, USA) and reactive bands were detected by exposure to Kodak X-Omat K film for 3 min at ambient temperature. Bands that showed an apparent reaction with antisera were cut out and stored at 4°C until analyzed by MS.

2.5 In-gel digestion

After SDS-PAGE, the gel was sliced into 30 \times 2 mm strips *per* lane manually from the loading well to the bottom. The gel slices were destained with 50% ACN/25 mM NH_4HCO_3 , reduced with 10 mM DTT at 56°C and alkylated in the dark with 50 mM iodoacetamide at room temperature for 1 h. Then the gel plugs were lyophilized and immersed in 15 μ L of 10 ng/ μ L trypsin solution in 25 mM NH_4HCO_3 . The digestion was kept at 37°C for 15 h. Trypsin peptide mixtures were first extracted with 100 μ L 5% TFA and then with same volume of 2.5% TFA/50% ACN. The extracted solutions were blended, lyophilized and

used for further identification by MS. Furthermore, Glu-C in PBS (pH 7.8), which hydrolyzes peptide bonds at the carboxyl side of glutamyl and aspartyl residues, was also used for in-gel protein digestion and peptide extraction to improve sequence coverage of the nucleocapsid protein. For the identification of the glycosylated spike protein, the dried gel particles were suspended in 15 μ L PNGase F solution (500 U/mL) and incubated at 4°C for at least 40 min, then at 37°C overnight. During the incubation, 10–25 μ L of water was added to ensure the gel plugs were covered with liquid at all times.

2.6 RP-HPLC

RP-HPLC of the protein mixture was performed on a prepacked column (50 mm \times 4.6 mm id, Hypersil C18, 5 μ m spherical particles with pore diameter 300 Å; Elite, Dalian, China). The flow rate was 1.0 mL/min and detection wavelength was set at 280 nm. Mobile phase A consisted of water/ACN (95/5, v/v) with 0.1% TFA. Mobile phase B consisted of water/ACN (5/95, v/v) with 0.1% TFA. The separation was performed by running a non-linear gradient: 10–90% B, for 60 min, 90–100% B, for 5 min, retaining 100% B for 5 min, then coming to 100% A for 5 min and keeping the system in 100% A for 10 min for another run. The lyophilized protein mixture was dissolved in 8 M urea and 25–50 μ L sample was injected by a Rheodyne injection valve (Rheodyne, Rohnert Park, CA, USA) in multiloading mode. The chromatographic fractions were collected and lyophilized, followed by trypsin digestion and MS identification. For measurement of the M_r of the nucleocapsid (N) protein, the relative fraction was lyophilized for MALDI-MS analysis.

2.7 Capillary RP-HPLC

Capillary RP-HPLC of the peptide mixture was carried out on a Micromass CapLC liquid chromatography system including three pumps A, B and C (Micromass, Manchester, UK). Fused silica tubing (150 mm \times 75 μ m id) packed with PepMap C18, 3 μ m spherical particles with pore diameter 100 Å (LC Packings, Amsterdam, Netherlands) was used. The flow rate was set at 2.0 μ L/min and split into *ca.* 0.15 μ L/min prior to the precolumn and analytical column. Samples were injected at a flow rate of 30 μ L/min with pump C by the autosampler and salts were removed on the precolumn of 320 μ m \times 5 mm PepMap C18, 3 μ m spherical particles with pore diameter 100 Å (LC Packings). The precolumn was connected in the 10-port switching valve, and switched to the analytical column after the sample was desalted. Mobile phase A consisted of water/ACN (95/5, v/v) with 0.1% FA. Mobile phase B consisted of water/TFA (5/95, v/v) with 0.1% FA. The

separation was performed by running a nonlinear gradient: 4% B, in 0.1–3.5 min for injection; 4–50% B, in 3.5–63.5 min; 50–100% B, in 63.5–73.5 min; 100% B in 73.5–80 min; 100–4% B, in 80–85 min. After 15 min equilibration in 100% A, another analysis could be run. The CapLC is coupled on-line with a Q-TOF Micro mass spectrometer (Micromass) for detection and protein identification.

2.8 MALDI-TOF MS

2.8.1 Molecular weight determination

Molecular weight measurements of proteins or peptides were carried out on a Reflex III MALDI mass spectrometer (Bruker Daltonics, Bremen, Germany), equipped with a flight tube (linear mode, 1.6 m long), laser (N_2 , 337 nm) and scout 384 target system. Accelerating voltage was 20 kV and Microchannel plate (MCP) detector working at 1.6 kV. Mass spectra were acquired in positive mode and 300 shots were summed for each spectrum. One μ L sample dissolved in 1% TFA was mixed with 1 μ L matrix solution (sinapic acid; Sigma) and centrifuged, 1 μ L of supernatant was spotted on the target. One pmol of BSA was used to calibrate the instrument.

2.8.2 Peptide mass fingerprinting

Mass spectra were recorded with a MALDI-R MALDI mass spectrometer (Micromass). The instrument was calibrated with a tryptic digest mixture of alcohol dehydrogenase. Positive ion mass spectra were recorded in reflectron mode with α -cyano-4-hydroxycinnamic acid as the matrix. Samples dissolved in 0.5–1 μ L of water were crystallized with 0.5 μ L of a saturated solution of the matrix in ACN on the target. Reflection spectra were acquired using the delayed extraction technique in positive ion mode with an acceleration voltage of 1.5 kV. About 100 laser shots were summed to acquire the spectra and MassLynx software (Micromass) was used to process the data. Database searching was manually performed using the MASCOT (<http://www.matrixscience.com/>), or PeptIdent (<http://www.expasy.ch/tools/peptident.html>) programs available on the web.

2.9 LC-ESI MS/MS

2.9.1 Nanospray ESI MS/MS

All MS/MS measurements were carried out on hybrid quadrupole-time of flight mass spectrometer (Q-TOF2; Micromass) with a nanospray needle sample introducing

system at an applied spray voltage of 3000 V, MCP detector with 2250 V of working voltage, energy adjustable collision cell filled with pure argon gas. Typically, a 2 μ L sample was loaded in the Nanoflow Probe Tip (Micromass), the sample cone working on 25–40 V. The instrument was controlled in MassLynx 3.5 and sequences were manually read out in BioLynx. Generally, spectra were generated from 100–500 MS/MS scans. The accuracy of external calibration of Glu-Fib was 3 ppm. A local protein search engine Global Server 1.1 beta was setup with local NCBI database for automatic protein identification (using peaklist files) and local BLAST software with the same protein database for sequence alignment.

2.9.2 Nanoflow ESI MS/MS

For analysis of peptide mixture by LC-ESI MS/MS, lyophilized peptide mixtures were dissolved with 5.5 μ L of 0.1% FA in 2% ACN and injected by autosampler onto a 0.3 \times 1 mm trapping column (PepMap C18; LC Packings) using a CapLC system. Peptides were directly eluted into a Q-TOF mass spectrometer (Q-TOF Micro; Micromass) at 200 nL/min on a C18 column (75 μ m \times 15 cm; LC Packings). MS/MS data were processed using MassLynx 3.5 and searched against NCBI protein sequence databases via internet available MS/MS ion searching program MASCOT (<http://www.matrixscience.com>).

2.10 *In vitro* cleavage of N protein by caspase-3 and -6

In order to probe the cleavage mechanism of the N protein in Vero E6 cells infected with SARS-CoV, recombinant N protein was used as a substrate to test the possibility of the protein cleavage by cysteine proteases, which play a central role in cell apoptosis. Caspase-3 and caspase-6 were selected and added into the reaction system respectively. The reactions were carried out in caspase reaction buffer (20 mM/L piperazine-*N,N'*-bis(2 ethanesulfonic acid)-NaOH (pH 7.2), 100 mM/L NaCl, 2% sucrose, 0.2 mM/L EDTA, 10 mM/L DTT) and incubated at 37°C for 15 h. The reactant was analyzed by 13% SDS-PAGE.

2.11 Experimental procedures

A flowchart of the experimental procedure for the identification of structural proteins of SARS-CoV is shown in Fig. 1. Because it was difficult to get a plentiful amount of virus particles for the study, the original sample obtained for analysis was a complicated mixture of SARS-CoV par-

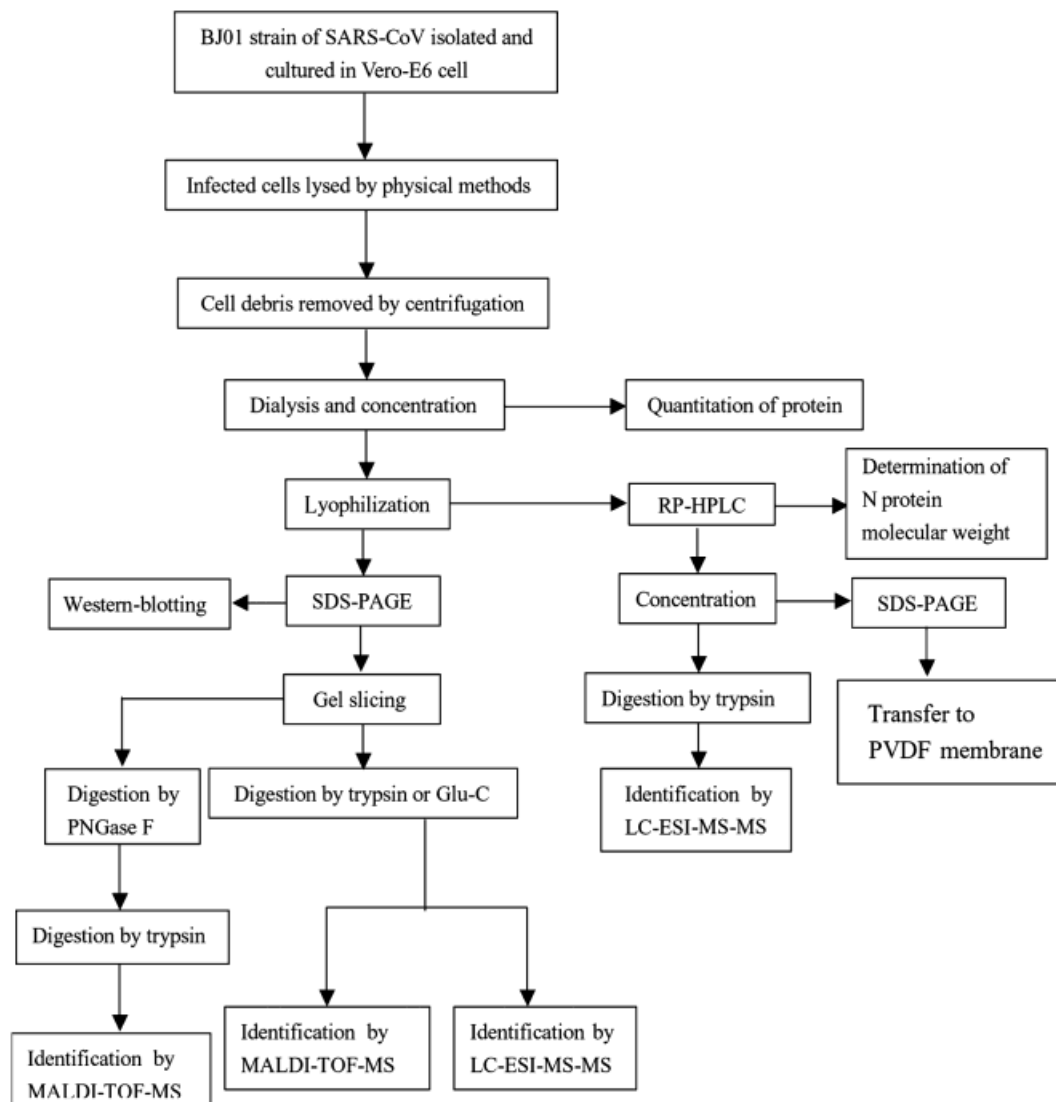


Figure 1. Flowchart of experimental procedures for proteome research on structural proteins of SARS-CoV.

ticles, Vero E6 cells, and culture media. A proteomics strategy based on SDS-PAGE was taken to separate the complex mixture at first, and then the total bands on the gel were sliced, in-gel digested and characterized with peptide sequencing by LC-ESI MS/MS. To acquire accurate molecular weight information as well as the *N*-terminus of the protein, RP-HPLC was performed to separate the protein of interest, MALDI-MS was employed to identify the protein with peptide mass fingerprinting. RP-HPLC was also used for characterization of the virus protein by prefractionation of the sample mixture to decrease the complexity of the samples, and MALDI-MS was used to characterize the peptide mixture to increase the sequence coverage of the proteins.

3 Results and discussion

3.1 Protein identification of SARS-CoV from infected cells and antigenicity analysis of viral structural proteins

The antisera from SARS patients and convalescent patients were utilized for antigenicity analysis of viral proteins. The results demonstrate that the antisera from a single patient and 10 convalescent patients notably reacted with SARS-CoV related proteins (Figs. 2A and 2B) with the apparent mass range approximately 21–200 kDa, which contains two very strong hybridized bands (4 and 5) with an apparent mass ~46 kDa, and three much weaker

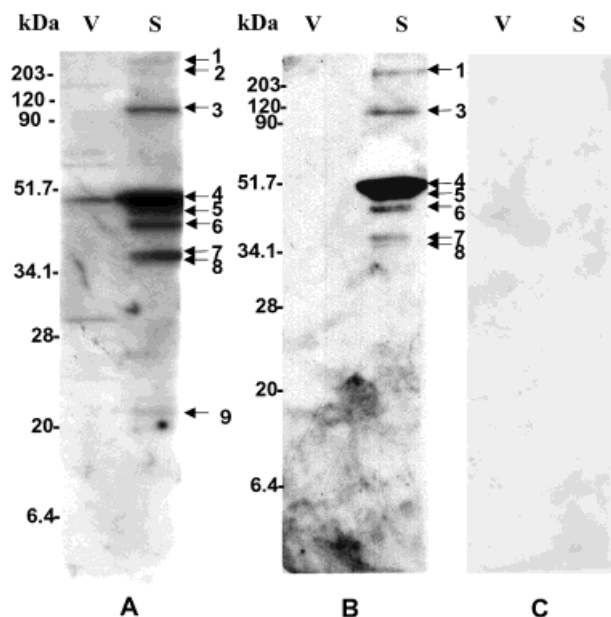


Figure 2. Western blot analysis of proteins from SARS-CoV in Vero E6 cell lysates. SARS-CoV in Vero E6 cell lysates (S) and control Vero E6 cell lysates (V) were separated by 13% SDS-PAGE and analyzed by Western blot. The transferred PVDF membranes were probed respectively with the antisera from one (A) and 10 convalescent patients (B) or a control pooled sera from 6 noninfected donors (C). The putative S glycoprotein and putative protein M from SARS-CoV were identified in the bands cut out at position 1, 2 and 3, and the putative N protein from SARS-CoV in bands 4–9.

protein bands (1–3) with an apparent mass range of more than 100 kDa. The reacted proteins in the gel bands were further identified by LC-ESI MS/MS. The putative S glycoprotein and putative M protein were found in bands 1–3 (Fig. 2A), 1 and 3 (Fig. 2B), and the putative N protein in bands 4–9 (Fig. 2A), 4–8 (Fig. 2B). Although different antisera demonstrated slight differences on antigenicity, the main antigenic proteins were identified as N protein. Intriguingly, a strong reaction with the antisera of SARS patients was observed when M protein (theoretical mass 25 060 Da) comigrated with S glycoprotein in the gel. However, when separated in the gel these two proteins did not show any characteristics of antigenicity. These results imply that the antigenicity of M and S proteins might depend on their interaction or physical binding.

3.2 Characterization of structural proteins of SARS-CoV

Figure 3 shows the results of SDS-PAGE separation and MS characterization of the total proteins from the lyophilized samples of Vero E6 cells infected by SARS-CoV, in which spike protein (S), nucleocapsid protein (N) and membrane glycoprotein (M) were identified respectively. Tables 1–3 show the calculated and measured mass values of peptides found in tryptic and Glu-C digests of S, N and M proteins. The sequence coverage was 38.9, 93.1 and 28.1% respectively. It should be noted that the sequence coverage of S and N proteins is close to that

Table 1. Calculated and measured mass values of peptides found in tryptic digest of spike (S) = protein

Position	Calc.	Meas.	Error (Da)	Peptide sequence	MALDI	ESI-QTOF
39–48	1258.6102	1258.6166	0.0064	GVYYPDEIFR	+	+
85–94	1114.5415	1114.5382	0.0033	DGIYFAATEK	+	+
189–198	1246.6446	1245.6900	0.0454	NKDGFLYVYK		+
191–198	1004.5087	1004.5148	0.0061	DGFLYVYK	+	
199–207	1046.5629	1046.5533	0.0096	GYQPIDVVR	+	
208–221	1576.8733	1576.8434	0.0299	DLPSGFNTLKPIFK	+	+
222–232	1257.7313	1257.6700	0.0613	LPLGINITNFR		+
298–306	1085.5374	1085.5242	0.0132	GIYQTSNFR	+	
298–315	1994.0454	1994.0020	0.0434	GIYQTSNFRVWVPSGDVVR	+	
334–342	1154.5629	1154.5600	0.0029	FPSVYAWER	+	+
374–390	1990.9578	1990.9167	0.0411	LNDLC ^a FSNVYADSFVVK	+	+
396–411	1737.8806	1737.8700	0.0106	QIAPGQTGVIADYNYK	+	+
412–426	1737.8087	1737.8363	0.0276	LPDDFMGCVLAWNTR	+	+
427–439	1460.6652	1460.7098	0.0446	NIDATSTGNYNYK	+	+
448–453	817.4679	817.4553	0.0126	LRFPER	+	
496–514	2014.1041	2014.0075	0.0966	VVLSFELLNAPATVC*GPK	+	
544–553	1310.6752	1310.6885	0.0133	RFQPFQQFGR	+	+
545–553	1154.5741	1154.5764	0.0023	FQPFQQFGR	+	+
554–563	1140.5167	1140.5516	0.0349	DVSDFTDSVR	+	+

Table 1. Continued

Position	Calc.	Meas.	Error (Da)	Peptide sequence	MALDI	ESI-QTOF
554–566	1480.6914	1480.7450	0.0536	DVSDFTDSVRDPK	+	+
748–758	1130.5800	1130.5587	0.0213	ALSGIAAEQDR	+	+
798–807	1225.6463	1225.6444	0.0019	SFIEDLLFNK	+	+
808–817	1052.5445	1052.5700	0.0255	VTLADAGFMK		+
818–829	1395.6321	1395.6429	0.0108	QYGEC ^a LGDINAR	+	+
888–903	1823.9286	1823.8540	0.0746	FNGIGVTQNVLYENQK	+	+
912–929	1848.9913	1848.9657	0.0256	AISQIQESLTTTSTALGK	+	+
930–946	1868.0236	1867.9619	0.0617	LQDVVNQNAQALNTLVK	+	+
947–965	2021.0662	2021.0233	0.0429	QLSSNFGAISSVLNDILSR	+	
966–977	1414.7536	1414.7466	0.0070	LDKVEAEVQIDR	+	
983–996	1690.9486	1690.9485	0.0001	LQSLQTYVTQQLIR	+	+
1028–1055	3167.5883	3167.5330	0.0553	GYHLMSFPQAAPHGVVFLH V TYVPSQER	+	
1164–1173	1186.6426	1186.7100	0.0674	EIDRLNEVAK		+
1238–1248	1293.5845	1293.6500	0.0655	FDEDDSEPVLK		+
1238–1251	1577.7693	1577.8517	0.0824	FDEDDSEPVLKGVK	+	
1249–1255	817.4566	817.4588	0.0022	GVKLHYT	+	

a) These cysteines were modified with iodoacetamide

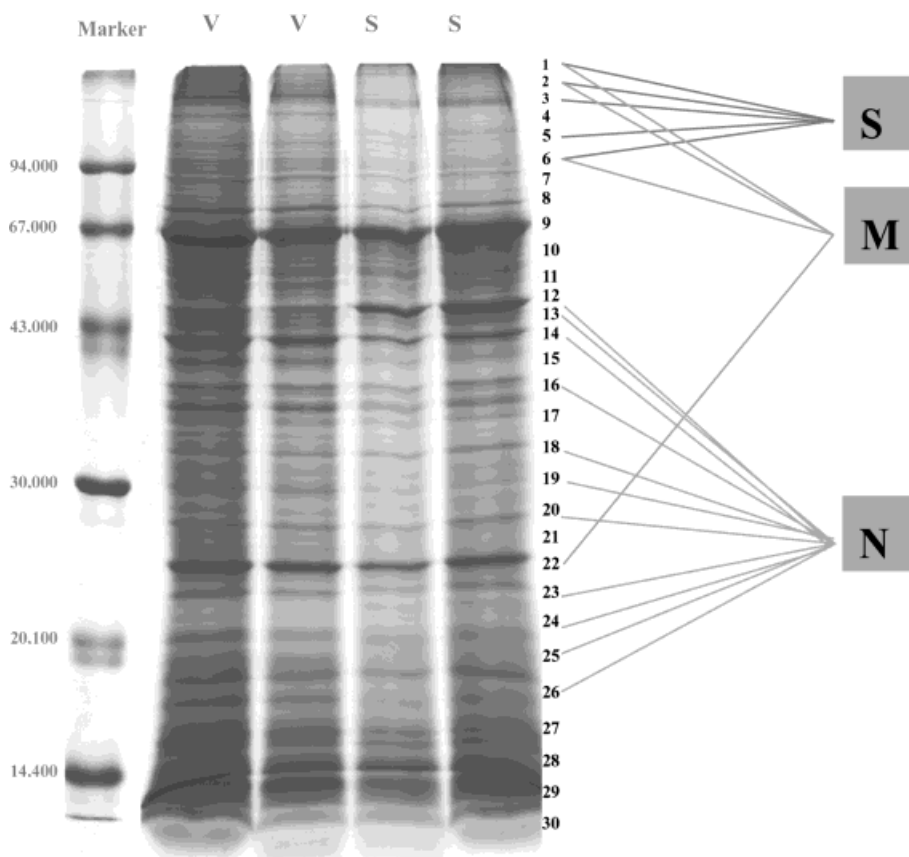


Figure 3. Identification of viral structural proteins of SARS-CoV from lysates of attacked Vero E6 cells by SDS-PAGE MS. The gel was stained with Coomassie Brilliant Blue R250. Thirty slices were cut from load well to the bottom of the gel. The bands in which SARS-CoV proteins were identified were denoted as S (spike protein), N (nucleocapsid protein) and M (membrane glycoprotein) respectively. The results showed that S protein mainly existed as a highly modified protein, thus

appeared separately at about 200 kDa (slice 3). N protein existed mainly as an integrity molecule at about 45 kDa (slice 12), but unlike the report [12], there were few fragmentation bands of this protein. M protein was detected not only at its theoretical molecular weight position (slice 22), but also at a very high position with S protein.

Table 2. Calculated and measured mass values of peptides found in tryptic and Glu-C digest of nucleocapsid protein

Position	Calc.	Meas.	Error (Da)	Peptide sequence	MALDI	ESI
1–10	1144.4990	1144.5079	0.0089	Ac-SDNGPQSNQR	+	
11–32	2262.0493	2262.0700	0.0207	SAPRITFGGPTDSTDNNQNGGR		+
15–32	1850.8263	1850.8726	0.0463	ITFGGPTDSTDNNQNGGR	+	+
33–40	926.5278	926.5300	0.0022	NGARPKQR		+
41–61	2324.1894	2324.1445	0.0449	RPQGLPNNTASWFTALTQHGK	+	+
41–65	2851.4597	2851.4800	0.0203	RPQGLPNNTASWFTALTQHGKEELR		+
62–68	946.5000	946.5200	0.0200	EELRFPR	+	+
63–81 ^{a)}	2095.0400	2095.0300	0.0100	ELRFPRGQGVPINTNSGPD	+	+
69–88	2151.0101	2151.0105	0.0004	GQGVPINTNSGPDDQIGYYR	+	+
94–102	947.5091	947.5657	0.0566	VRGGDGKMK	+	
104–118 ^{a)}	1848.8955	1848.9233	0.0278	LSPRWYFYLLGTGPE	+	+
108–127	2297.0913	2297.0747	0.0166	WYFYLLGTGPEASLPYGANK	+	+
119–128 ^{a)}	1049.5262	1049.5714	0.0452	ASLPYGANKE	+	+
128–143	1684.8904	1684.9500	0.0596	EGIVWVATEGALNTPK		+
137–174 ^{a)}	4019.0733	4019.0828	0.0095	GALNTPKDHIGTRNPNNNAATVLQLPQ GTTLPKGFYAE	+	+
150–169	2091.1193	2091.0713	0.0480	NPNNNAATVLQLPQGTTLPLK	+	+
170–177	886.4053	886.4592	0.0539	GFYAEGR	+	
178–189	1166.5508	1166.5620	0.0112	GGSQASSRSSSR	+	
210–226	1687.9047	1687.9534	0.0487	MASGGGETALALLLLDR	+	+
217–231 ^{a)}	1696.0003	1696.0493	0.0490	TALALLLLDRLNQL	+	+
232–253 ^{a)}	2275.2000	2275.2200	0.0200	SKVSGKGGQQGQTVTKKSA AE	+	+
238–249	1372.7100	1372.7900	0.0800	GQQQQGQTVTKK		+
254–280 ^{a)}	3103.6984	3103.7602	0.0618	ASKKPRQKRTATKQYNVTQA FGRRGPE	+	+
267–276	1183.5854	1183.6693	0.0839	QYNVTQAFGR	+	+
277–293	1930.9365	1930.9364	0.0001	RGPEQTQGNFGDQDLIR	+	+
278–293	1774.8354	1774.8784	0.0430	GPEQTQGNFGDQDLIR	+	+
294–319	2928.3886	2928.4100	0.0214	QGTDYKHWPQIAQFAPSASAFFGMSR		+
300–319	2236.0756	2236.0295	0.0461	HWPQIAQFAPSASAFFGMSR	+	+
320–338	2061.0473	2061.0054	0.0419	IGMEVTPSGTWLTYHGAIK	+	+
339–355	2015.0807	2015.0345	0.0462	LDDKDPQFKDNVILLNK	+	+
348–361	1655.9115	1655.8900	0.0215	DNVILLNKHIDAYK		+
349–358 ^{a)}	1178.6891	1178.7213	0.0322	NVILLNKHID	+	+
356–369	1685.8500	1685.9100	0.0600	HIDAYKTFPTEPK		+
359–371 ^{a)}	1521.7947	1521.8481	0.0534	AYKTFPTEPKKD	+	+
375–385	1282.6750	1282.6672	0.0078	KTDEAQPLPQR	+	+
376–385	1154.5800	1154.5985	0.0185	TDEAQPLPQR	+	
379–399 ^{a)}	2300.3084	2300.2952	0.0132	AQPLPQRQKKQPTVTLLPAAD	+	+
388–405	2005.0059	2005.0208	0.0149	KQPTVTLLPAADMDDFSR	+	+
389–405	1876.9109	1876.9183	0.0074	QPTVTLLPAADMDDFSR	+	+
406–421	1594.6900	1594.7800	0.0900	QLQNSMSGASADSTQA		+

a) Peptides from Glu-C digest

of a Canadian report (96% and 42%). Notably, this is the first time that M protein has been identified in its natural form [12].

3.2.1 Spike protein analysis

Coronavirus Spike Protein is a large, type I membrane glycoprotein that contains distinct functional domains near the amino (S1) and carboxy (S2) termini. These

spikes function to define viral tropism by their receptor specificity and perhaps also by their membrane fusion activity during virus entry into cells. For most coronaviruses, spike proteins were post-translationally cleaved into two subunits after synthesis, S1 and S2. The peripheral S1 portion can independently bind cellular receptors while the integral membrane S2 portion is required to mediate fusion of viral and cellular membranes. The extraordinary variations in host range and tissue tropism

Table 3. Calculated and measured mass values of peptides found in tryptic digest of membrane glycoprotein

Position	Peptide sequence	Calc.	Meas.	Error (Da)
107–124	SMWSFNPETNILLNVPLR	2130.0900	2130.1300	0.0400
205–221	LNTDHAGSNDNIALLVQ	1793.8900	1793.9200	0.0300
171–184	TLSYYKLGASQR	1385.7300	1385.7600	0.0300
186–197	VGTDSGFAAYNR	1256.5858	1256.6400	0.0542

among coronaviruses are in large part attributable to variations in the spike glycoprotein [14]. As a surface glycoprotein, spike proteins may offer an attractive target for new drugs. Illumination of the structures featured including glycosylation may lead to important therapeutic applications.

Here, spike protein was detected in five bands of the gel (Fig. 3, slices 1, 2, 3, 5, 6). The M_r of slices 1, 2 and 3 were higher than the theoretical one (139 kDa), which indicated the existence of a large quantity of modifications after the translation. Surprisingly, slices 5 and 6 were found at a position significantly lower than the theoretical M_r of intact S protein, which implicated the possible cleavage of S protein. In addition, besides S protein, M protein was also identified in bands 1, 2, 5 and 6, where the M_r s are

significantly higher than its theoretic M_r , implying that there was a strong interaction or physical binding between the two proteins.

To investigate glycosylations in the spike protein, the protein was first deglycosylated with PNGase F and then treated with trypsin in gels as described above in Section 2.5. After deglycosylation, asparagines residues were converted to aspartic acids, which specified the corresponding deglycosylated peptides through the observation of their mass difference of 0.984 Da *per* deglycosylated site from the values calculated from the predicted sequence. As a result, four glycosylated peptides were identified by comparing the mass spectra before and after deglycosylation of the peptides (Table 4). Figure 4 shows the mass spectra of glycosylated peptide T1074–1089.

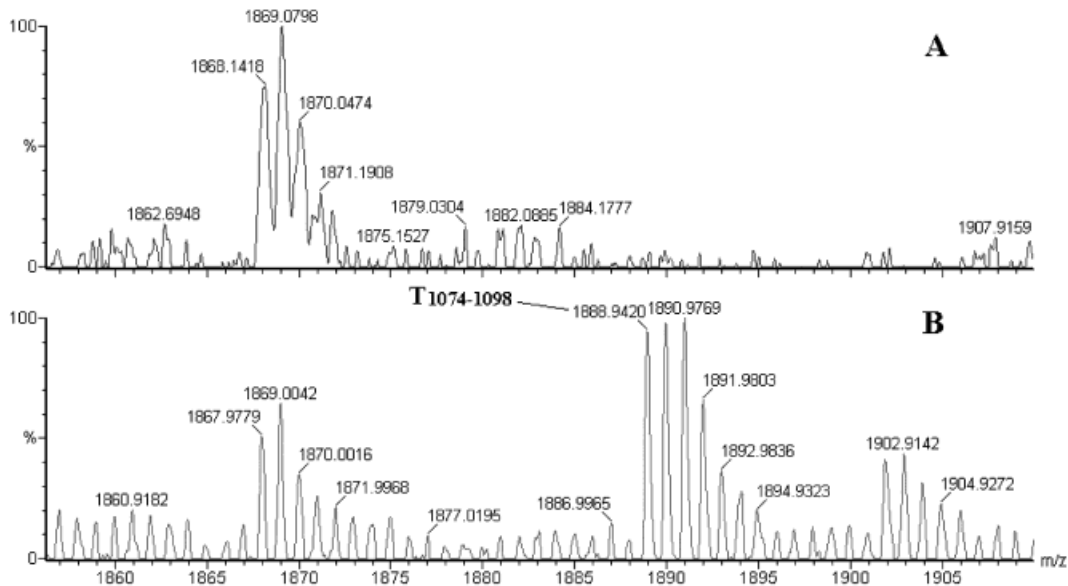


Figure 4. Comparison of the same section of mass spectra of tryptic digestion of spike protein. (A) without deglycosylation with PNGase F and (B) after deglycosylation with PNGase F by MALDI-MS. T1074–1098 indicates the residue numbers corresponding to the intact protein. The peak corresponding to T1074–1098 is absent in A, but it is present in B after deglycosylation. The measured m/z value of 1888.9420 for $[M+H]^+$ ions corresponds to a ~ 0.984 Da difference from the predicted value for nonglycosylated T1074–1098. The measured and predicted masses for the identified peaks are shown in Table 4.

Table 4. Deglycosylated peptides found in tryptic digest of PNGase F-treated SARS spike protein

Position	Peptide sequence	Calculated	Detected	ΔM
T1137–1163	YFKNHTSPDVLGDISGINA SVVNIQK	2931.4846	2932.5305	1.0469
T778–796	YFGGFNFSQILPDPLKPTK	2169.1378	2170.1252	0.9874
T1074–1089	EGVFVFNSTSWFITQR	1887.9388	1888.9672	0.9284
T111–126	SQSVIIINNSTNVVIR	1756.9915	1758.0128	1.0213

All the glycosylated peptides we found showed clearly in their mass spectra that each had only one glycosylated site. This result was different from that published by the Canadian group, in which two peptides with two glycosylation sites were characterized [12]. We also determined that peptide T222–232, which contained a potential glycosylation site, displayed no glycosylation at all.

3.2.2 Nucleocapsid protein analysis

N protein was identified by both SDS-PAGE/mass spectrometry and RP-HPLC/mass spectrometry. The fractions separated by RP-HPLC were collected, concentrated and digested. The analytical results by LC-ESI MS/MS showed that fraction 24 in the HPLC chromatogram con-

tained intact N protein (Fig. 5, peak 2), so MALDI-MS was used to measure the M_r , as shown in Fig. 5, which was defined as 45 929 Da. This result is in agreement with the data published [12] that the first amino acid methionine in the *N*-terminus of the protein was depleted and the second serine was acetylated. The theoretical molecular weight calculated based on the predicted amino acid sequence is 45 935 Da. Comparing the calculated molecular weight with the measured one, the relative error was less than 0.13/1000. To confirm the amino acid sequence in the *N*-terminus of N protein, Edman degradation was performed on a PVDF membrane blotted from SDS-PAGE of the N protein by protein sequencer following the instrument's manual (Procise Sequencer; Applied Biosystems, Foster City, CA, USA). Only after deacetylation with TFA according to [15] was the first amino acid serine identified (data not shown).

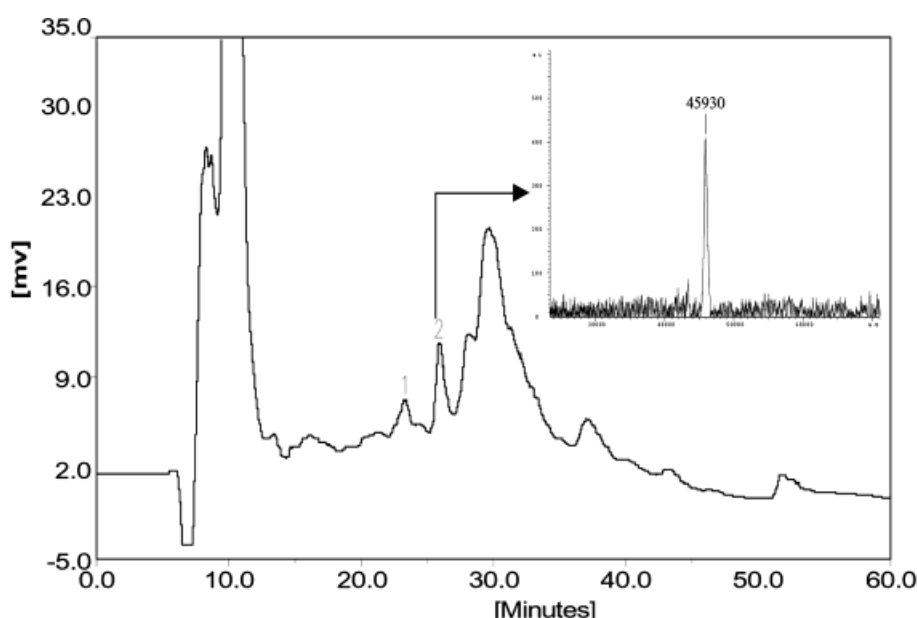


Figure 5. Chromatogram of Vero E6 cells lysates infected by SARS-CoV and mass spectrometry spectrum of intact N protein. Chromatographic condition: column, 50 × 4.6 mm id Hypersil C18, 5 μ m, 300 Å; mobile phase A: water/ACN (95/5, v/v) with 0.1% TFA, mobile phase B: water/ACN (5/95, v/v) with 0.1% TFA; flow rate: 1.0 mL/min; nonlinear gra-

dient: 10–90% B, in 60 min, 90–100% B, in 5 min and retaining 100% B for 5 min, and then coming to 100% A in 5 min. 1, cleaved N protein; 2, intact N protein, fraction 24. The arrow indicates the molecular weight spectrum of intact N protein in fraction 24 corresponding to peak 2.

The phosphorylation of N protein is a well documented phenomenon in many coronaviruses such as murine hepatitis virus (MHV), infectious bronchitis virus (IBV), bovine coronavirus (BCV) and porcine epidemic diarrhoea virus (PEDV) [16–20], in which phosphorylation were determined by metabolic labeling methods for recombinant forms rather than natural forms. The functional significance of N proteins' phosphorylation is still under investigation. In the case of MHV, the widely investigated coronavirus, phosphorylated N protein was considered as having a higher RNA binding capacity than the unphosphorylated one, and its dephosphorylation was found to be in connection with initiation of the infection [21, 22]. As phosphorylation of the N protein may be a general phenomenon in coronavirus, several methods were tried to confirm the hypothesis but all failed (data not shown). These included Western blot analysis for phosphoprotein from viral proteins separated by SDS-PAGE using an anti-phosphoprotein antibody, IMAC (immobilized metal affinity chromatography) combined with LC-MS/MS to looking for phosphopeptides from tryptic digest of N protein, and phosphatase digestion to find the new peptides after dephosphorylated. Although we did not find any evidence that the N protein is a phosphoprotein, its real status is still unknown due to the sensitivity of the methods used. Metabolic labeling using ^{32}P should be a more sensitive method and will provide the direct evidence for the phosphorylation status of SARS N protein. However, as a strong indirect evidence, we should indicate the result of molecular weight resolution of intact natural N protein, 45 929 Da (real M_r) compared with the theoretical one of 45 935 Da of almost naked protein (only with N-terminal acetylation, without any phosphorylation). Their difference, like the relative error, was less than 0.13/1000, indicating that there is no phosphorylation in natural N protein at all.

From SDS-PAGE, N protein was found existing in different bands distributed between 46 kDa and 20 kDa (Fig. 3). This phenomenon was also observed with various coronaviruses such as transmissible gastroenteritis virus (TGEV), MHV, feline coronavirus (FIPV), BCV, avian IBV and turkey coronavirus (TCV) late in infection in cell culture, which resulted from caspases cleavage in the host cell [23]. To testify the role of caspase in N protein cleavage, caspase-3 and -6 were selected to cleave recombinant N protein of SARS CoV. The experimental result showed that recombinant N protein of SARS CoV could be cleaved by caspase-3 rather than caspase-6 and the pattern of peptide distribution was similar to that in infected cells (Fig. 6), which may imply that caspase-dependent apoptosis occurs and that caspase-3 cleaves the N protein *in vivo* during the late phase of virus infection. This phenomenon was not reported by the Canadian group [12].

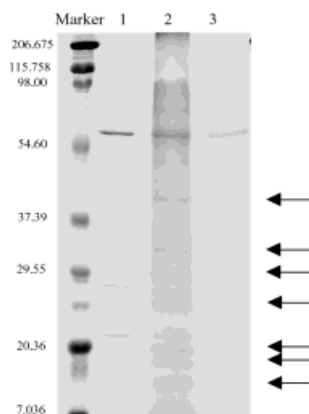


Figure 6. Cleavage of recombinant N protein of SARS CoV by caspase-3 and -6 *in vitro*. 1, recombinant N protein; 2, recombinant N protein incubated with caspase-3; 3, recombinant N protein incubated with caspase-6.

Apoptosis is an important process in the development and cell defense, which usually causes morphological and biochemical changes. Caspases, special kinds of proteases participating in such a reaction, are activated in a cascade triggered by apoptosis signals. Virus-infected cells undergo apoptosis by the attack of cytotoxic cells, including cytotoxic T cells and natural killer cells [24–26]. As for coronavirus, it is reported that infection with mouse hepatitis virus strain 3 (MHV-3) results in lethal fulminant hepatic necrosis [27]. In IBV research, it was found that replication of IBV in Vero cells caused extensive cytopathic effects, leading to destruction of the entire monolayer and the death of infected cells [28]. In some further research on apoptosis about coronavirus, the E protein of MHV has been confirmed as an apoptosis inducer in 17Cl-1 cells [29]. In the TEGV-infected cells, the nucleocapsid protein can be cleaved by caspase-6, and caspase-7 *in vitro* rather than caspase-3 [23]. Although apoptosis in the Vero E6 cell infected by SARS-CoV has not been reported, the result in our experiment may give a clue to apoptosis in the virus-infected cell and the N protein of SARS-CoV may be a substrate of caspase-3 *in vivo*.

3.2.3 Membrane glycoprotein analysis

M protein (the membrane glycoprotein or matrix protein, E1 membrane glycoprotein), a transmembrane protein, is the most abundant glycoprotein in infected cells as well as in the virus particle of known coronaviruses [30]. It has three domains: a short N-terminal ectodomain, a triple-spanning transmembrane domain in the N-terminal half of the protein, and a C-terminal endodomain. TMHMM (<http://www.cbs.dtu.dk/services/TMHMM> and Tmpred (<http://www.ch.embnet.org/software/TMPRED.html>) anal-

ysis indicated that the three transmembrane helices were approximately located at residues 15–37, 50–72 and 77–99, with the 121-amino acid hydrophilic domain on the inside of the virus particle [7]. In our experiment, however, only the sequence from 107–221 amino acids was covered (Table 3), which may be because of its high hydrophobicity, especially in its *N*-terminus. Furthermore, our experimental results showed that besides the 25 kDa position (slice 22 in Fig. 3), M protein also appeared at higher molecular weight position but only together with the spike protein in SDS-PAGE. Whether this is because of a possible interaction between M and S protein or the hydrophobic character of the M protein, which kept it at its higher molecule weight position, still needs further research.

3.2.4 Envelop protein analysis

The small envelope (E) protein has been recognized as a structural component of the coronaviruses such as MHV, IBV, TGEV [29, 31–32]. Recent research shows E protein has two major biological functions. MHV E protein can induce apoptosis in cells expressing E protein and overexpression of Bcl-2 oncoprotein suppresses MHV E protein-induced apoptosis indicating that initiation of the apoptotic pathway begins upstream of Bcl-2 [33]. Furthermore, coexpression of the genes encoding the MHV-A59 and E protein results in the production of virus-like particles, and E protein membrane vesicles can be released only from E protein expressing cells as well as MHV infected cells [34, 35]. These results indicate that the E protein plays a pivotal role in virus envelope formation.

In our experiments, two strategies were used to identify the E protein. First, the total proteins from Vero E6 cells infected by SARS-CoV were separated by 15% SDS-PAGE, all bands were sliced, in-gel digested and MS/MS analyzed. Three structural proteins (S, N and M) were identified, but E protein was not found. Then we used 2-D LC-MS/MS, the total proteins from the infected cells were digested directly and 2-D capillary HPLC separated the digestion with a strong cation ion exchange chromatography as the first dimension and reverse phase chromatography as the second dimension, and the protein was still not identified. It is therefore likely, as previously reported for other coronaviruses, that the E protein is present only in minute amounts in infected cells [31] and in the viral envelope. Its strong hydrophobicity, *i.e.* the *N*-terminal two-thirds region is highly hydrophobic [36, 37], also results in difficulties to discover this protein. However, we can not conclude so far that the E protein is not expressed in Vero E6 cells infected by SARS-CoV or are not present in the SARS-CoV envelope, because the E protein is not only present in envelopes, but also plays an essential role in the assembly of known coronaviruses.

4 Concluding remarks

Proteomic analysis was utilized to uncover the natural structural protein constitutes of BJ01 strain of SARS-CoV. Three structural proteins, *i.e.*, S, N and M proteins, were identified and characterized from the cultured Vero E6 cells infected with the virus and their antigenicity displayed with patient sera, of which the M protein was identified for the first time. Glycosylation modification of S protein was analyzed, and four glycosylation sites were characterized. Cleavage of S protein into two subunits was also suggested. Molecular weight determination of intact N protein showed that a post-translational modification only happened in its *N*-terminus as acetylation and no phosphorylation modification was detected within the entire N protein. Antigenicity analysis indicated that the N protein has a prominent immunogenicity to the convalescent sera from patients with SARS. The immune response of S protein probably depends on the strong interaction with the M protein. Cleavage of recombinant N protein with caspase-3 and -6 *in vitro* demonstrated that the series of shorter isoforms of N protein observed in SDS-PAGE might be the products of caspase-3 cleavage rather than caspase-6 and might have a relationship with the apoptosis of Vero E6 cells induced by the infection of SARS-CoV.

The experimental results indicate that proteomics strategy is a very powerful method to discover and identify the proteins from a complex system. The molecular biological information of natural proteins from SARS-CoV, especially the processing and modification of structural viral proteins could be a complement of the genomic information and provide direct molecular basis for further research on diagnosis, prevention and treatment of SARS.

This work was supported partly by State High Technology R&D Plan (No. 2001AA233031, 2002-BA711A11) and China National Key Program for Basic Research (No. 2001CB510201).

Received July 11, 2003

Revised October 21, 2003

Accepted October 27, 2003

5 References

- [1] WHO SARS case fatality ratio, incubation period, http://www.who.int/csr/sarsarchive/2003_05_07a/en/, May 7, 2003.
- [2] Donnelly, C. A., Ghani, A. C., Leung, G. M., Hedley, A. J. *et al.*, *Lancet* 2003, 361, 1761–1766.
- [3] WHO SARS virus isolated, new diagnostic test producing reliable results, http://www.who.int/entity/crs/sarsarchive/2003_03_22/en, March 22, 2003.
- [4] Peiris, J. S., Lai, S. T., Poon, L. L., Guan, Y. *et al.*, *Lancet* 2003, 361, 1319–1325.

- [5] Drosten, C., Gunther, S., Preiser, W., van der Werf, S. *et al.*, *N. Engl. J. Med.* 2003, **348**, 1967–1976.
- [6] Ksiazek, T. G., Erdman, D., Goldsmith, C., Zaki, S. R. *et al.*, *N. Engl. J. Med.* 2003, **348**, 1953–1966.
- [7] Marra, M. A., Jones, S. J., Astell, C. R., Holt, R. A., *et al.*, *Science* 2003, **300**, 1399–1404.
- [8] Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A. *et al.*, *Science* 2003, **300**, 1394–1399.
- [9] Qin, E. D., Zhu, Q. Y., Wang, J., Li, W. *et al.*, *Chin. Sci. Bull.* 2003, **48**, 941–948.
- [10] Ruan, Y. J., Wei, C. L., Ee, A. L., Vega, V. B. *et al.*, *Lancet* 2003, **361**, 1779–1785.
- [11] Zhu, Q. Y., Qin, E. D., Wang, C. E., Yu, M. *et al.*, *J. Chin. Biotechnol.* 2003, **23**, 106–112.
- [12] Krokhin, O., Li, Y., Andonov, A., Feldman, H. *et al.*, *Mol. Cell. Proteomics* 2003, **2**, 346–356.
- [13] Lowry, O. H., Rosebrough, N. J., Farr, A. L., Randall, R. J., *J. Biol. Chem.* 1951, **193**, 265–275.
- [14] Gallagher, T. M., Buchmeier, M. J., *Virology* 2001, **279**, 371–374.
- [15] Wellner, D., Panneerselvam, C., Horecker, B. L., *Proc. Natl. Acad. Sci. USA* 1990, **87**, 1947–1949.
- [16] Stohlman, S. A., Lai, M. M., *J. Virol.* 1979, **32**, 672–675.
- [17] Wilbur, S. M., Nelson, G. W., Lai, M. M., McMillan, M., Stohlman, S. A., *Biochem. Biophys. Res. Commun.* 1986, **141**, 7–12.
- [18] Lomniczi, B., Morser, J., *J. Gen. Virol.* 1981, **55**, 155–164.
- [19] King, B., Brian, D. A., *J. Virol.* 1982, **42**, 700–707.
- [20] Utiger, A., Tobler, K., Bridgen, A., Suter, M. *et al.*, *Adv. Exp. Med. Biol.* 1995, **380**, 287–290.
- [21] Wilson, G. A., Mohandas, D. V., Dales, S., *Adv. Exp. Med. Biol.* 1990, **276**, 261–266.
- [22] Mohandas, D. V., Dales, S., *FEBS Lett.* 1991, **282**, 419–424.
- [23] Eleouet, J. F., Slee, E. A., Saurini, F., Castagne, N. *et al.*, *J. Virol.* 2000, **74**, 3975–3983.
- [24] Shibata, S., Kyuwa, S., Lee, S. K., Toyoda, Y., Goto, N., *J. Virol.* 1994, **68**, 7540–7545.
- [25] Smyth, M. J., Trapani, J. A., *J. Virol.* 1998, **72**, 1–9.
- [26] Wege, H., Siddell, S., ter Meulen, V., *Curr. Top. Microbiol. Immunol.* 1982, **99**, 165–200.
- [27] Belyavskiy, M., Belyavskaya, E., Levy, G. A., Leibowitz, J. L., *Virology* 1998, **250**, 41–49.
- [28] Liu, C., Xu, H. Y., Liu, D. X., *J. Virol.* 2001, **75**, 6402–6409.
- [29] An, S., Chen, C. J., Yu, X., Leibowitz, J. L., Makino, S., *J. Virol.* 1999, **73**, 7853–7859.
- [30] Rottier, P. J. M., in: Siddell, S. G. (Eds.), *The Coronaviridae*, Plenum Press, New York 1995, pp. 115–139.
- [31] Liu, D. X., Inglis, S. C., *Virology* 1991, **185**, 911–917.
- [32] Godet, M., L'Haridon, R., Vautherot, J. F., Laude, H., *Virology* 1992, **188**, 666–675.
- [33] Maeda, J., Maeda, A., Makino, S., *Virology* 1999, **263**, 265–272.
- [34] Vennema, H., Godeke, G. J., Rossen, J. W., Voorhout, W. F. *et al.*, *EMBO J.* 1996, **15**, 2020–2028.
- [35] Yu, X., Bi, W., Weiss, S. R., Leibowitz, J. L., *Virology* 1994, **202**, 1018–1023.
- [36] Raamsman, M. J., Locker, J. K., de Hooge, A., de Vries, A. A. *et al.*, *J. Virol.* 2000, **74**, 2333–2342.
- [37] Maeda, J., Repass, J. F., Maeda, A., Makino, S., *Virology* 2001, **281**, 163–169.