JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Characterization of the 3a Protein of SARS-associated Coronavirus in Infected Vero E6 Cells and SARS Patients

Rong Zeng[1]*†, Rui-Fu Yang[4]†, Mu-De Shi[2]†, Man-Rong Jiang[2]†
You-Hua Xie[3], Hong-Qiang Ruan[1], Xiao-Sheng Jiang[1], Lv Shi[1], Hu Zhou[1]
Lei Zhang[1], Xiao-Dong Wu[2], Ying Lin[2], Yong-Yong Ji[2], Lei Xiong[2]
Yan Jin[2], Er-Hei Dai[4], Xiao-Yi Wang[4], Bin-Ying Si[4], Jin Wang[4]
Hong-Xia Wang[4], Cui-E Wang[4], Yong-Hua Gan[4], Yu-Chuan Li[4]
Ju-Tian Cao[4], Jiang-Ping Zuo[5], Shi-Fang Shan[3], En Xie[6]
Song-Hua Chen[7], Zhi-Qin Jiang[2], Xi Zhang[8], Yuan Wang[3], Gang Pei[2]
Bing Sun[2]* and Jia-Rui Wu[2]*

[1]*Research Center for Proteome Analysis Key Lab of Proteomics, Shanghai Institutes for Biological Sciences Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031 China*

[2]*Laboratory of Molecular Cell Biology Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences 320 Yue-Yang Road, Shanghai 200031 China*

[3]*State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences 320 Yue-Yang Road, Shanghai 200031 China*

[4]*Institute of Microbiology and Epidemiology, Academy of Military Medical Sciences, Beijing 100071, China*

[5]*State Key Laboratory of Drug Research Shanghai Institute of Materia Medica Shanghai Institutes for Biological Sciences Chinese Academy of Sciences, 555 Zu-Chong-Zhi Road, Shanghai 201203 China*

[6]*Shanghai Laboratorial Animal Center Shanghai Institutes for Biological Sciences Chinese Academy of Sciences, Jiu-Tin Town, Shanghai 201615, China*

[7]*Cell Bank of Type Culture Collection Committee, Cell Resources Center Shanghai Institutes for Biological Sciences Chinese Academy of Sciences, 320 Yue-Yang Road, Shanghai 200031, China*

[8]*Shanghai Center for Disease Control and Prevention, 1360 Zhong-Shang Road Shanghai 200336, China*

Proteomics was used to identify a protein encoded by ORF 3a in a SARS-associated coronavirus (SARS-CoV). Immuno-blotting revealed that inter-chain disulfide bonds might be formed between this protein and the spike protein. ELISA indicated that sera from SARS patients have significant positive reactions with synthesized peptides derived from the 3a protein. These results are concordant with that of a spike protein-derived peptide. A tendency exists for co-mutation between the 3a protein and the spike protein of SARS-CoV isolates, suggesting that the function of the 3a protein correlates with the spike protein. Taken together, the 3a protein might be tightly correlated to the spike protein in the SARS-CoV functions. The 3a protein may serve as a new clinical marker or drug target for SARS treatment.

*Corresponding authors

Keywords: SARS; coronavirus; 3a protein; spike protein; proteome

Introduction

Coronaviruses are a large and diverse family of enveloped, positive-stranded RNA viruses. Members of the coronavirus family can be divided into three groups according to serologic properties.[1] Although an individual coronavirus naturally infects distinct animal hosts and causes particular diseases, the whole-genome sequencing of several coronaviruses such as avian infectious bronchitis virus (IBV), murine hepatitis virus (MHV), human coronavirus 229E (HCoV-229E), and porcine transmissible gastroenteritis virus (TGEV), revealed similar genomic structure among these different coronaviruses.[2−5]

A novel coronavirus has recently been identified, which causes the latest worldwide outbreak of severe acute respiratory syndrome (SARS). The analysis of the complete nucleotide sequence of SARS-associated coronavirus (SARS-CoV) shows that its genomic organization is similar to that of other known coronaviruses.[6,7] The genome of SARS-CoV is approximately 30 kb in size and has 14 open reading frames (ORFS).[6−8]

It looks easy to analyze the entire genome of coronaviruses, but their protein component identification has proven to be a difficult task. So far, only five types of structural proteins have been identified as members of the coronavirus family.[1] The spike (S) glycoprotein, together with small envelope (E) protein and membrane (M) glycoprotein, consist of the viral envelope, while the nucleocapsid (N) protein interacts with genomic RNA of the virus to form the viral nucleocapsid.[1−9] In addition, a hemagglutinin-esterase (HE) glycoprotein has been found in the virion surface of some coronaviruses such as turkey coronavirus.[1] The analysis of mRNAs of SARS-CoV revealed many transcription units covering unknown ORFs of the genome.[6,8,10] Although some researchers speculated that these unknown ORFs encode nonstructural proteins,[6] there is no evidence to support this assumption.

Analysis of the genomic organization of SARS-CoV showed that a gene locus containing ORF3a and 3b located between the S and E genes,[6−8] is frequently found in different members of the coronavirus family (see Figure 4A). It has been hypothesized that ORF 3/3a and ORF3-1/3b may be related to viral virulence and pathogenesis,[11−13] whereas some suggested that ORF3a of TGEV is not essential for enteric virulence.[14] However, these previous studies are not conclusive, since all data are based on the analysis of genomic sequences or genes. Here, we detected with proteomics a protein derived from ORF3a of the SARS-CoV genome. Immunoblotting revealed interchain disulfide linkages between the spike protein and the 3a protein. Mutation analysis of SARS-CoV isolates and ELISA of SARS patients' sera indicated that the ORF3a product might function together with the S protein *in vivo*.

Results

Proteomic identification of ORF3a product in SARS-CoV

The "shot-gun" strategy using two-dimensional liquid chromatography electrospray tandem mass spectrometry (2D-LC-MS/MS) is a newly developed proteomic approach in which a protein mixture of adenovirus can be directly analyzed with mass spectrometry without separation by two-dimensional gel electrophoresis.[15] In the present study, the cytosol of Vero E6 cells infected with SARS-CoV was collected and subjected to 2D-LC-MS/MS. The tryptic-digested peptides of S, M, N, E and other known proteins of SARS-CoV were identified (not shown). In addition, two novel peptides were detected, corresponding to the residues 180−193 and 182−193 of a predicted ORF3a protein, respectively (Table 1; Supplementary Material).

According to the genomic sequence information for SARS-CoV, ORF 3a encoded a putative 31 kDa protein containing 274 amino acid residues.[6,7] In order to confirm the result by shot-gun assay, the collected cytosol of SARS-CoV-infected Vero E6 cells was separated with SDS-PAGE, and then the gel slice around the 31 kDa region was excised and analyzed by capillary liquid chromatography electrospray tandem mass spectrometry (μLC-ESI-MS/MS). The MS results showed four peptides of this ORF 3a protein including one N-terminal peptide from position 7−19, one peptide from 180−193 and two identical C-terminal peptides from 236−274, encompassing 24.08% of the whole 3a protein sequence (Table 1; Supplementary Material). The results from two complementary proteomic studies indicate the existence of a 31 kDa protein encoded by ORF 3a of SARS-CoV.

The 3a protein sequence characterization

The predicted amino acid sequence of the 3a protein of SARS-CoV isolate BJ-01 is shown in Figure 1†. This information combined with DNASTAR analysis (DNASTAR Inc. USA) predicted the existence of three transmembrane regions of the 3a protein, which is in agreement with that of the Tor2 isolate.[7] The membrane topology of this protein might belong to type 3b: its N terminus (about 40 amino acid residues) is on the outer surface of the viral particle, which contains a putative signal peptide from residues 1 to 15 predicted by SignalP.[17,18] Its C-terminal domain of 160 amino acid residues is located in the interior of the viral particle, of which there is a potential cytoplasmic region of Ca ATPase (GI: 737940) from residues E209 to D264 (also see Marra *et al*.[7]).

---

† TMPRED[16] as well as ALOM available at http://psort.nibb.ac.jp/; TMHMM available at http://www.cbs.dtu.dk/services/TMHMM/

We noticed that a cysteine-rich region overlaps the third membrane-spanning domain and the cytoplasmic tail of the 3a protein (Figure 1). Moreover, three cysteine residues within this cysteine-rich region are arranged every three residues (CWLCWKC from residues C127 to C133, Figure 1). If these cysteine residues exist in one helix, they would be located in the same surface, suggesting that the 3a protein containing these cysteine residues is able to form interchain disulfide linkages with other viral structural proteins. Interestingly, comparison with S protein sequences of the coronavirus family also shows a cysteine-rich region overlapping the junction of the membrane-spanning domain and the cytoplasmic tail, which is a conserved motif of spike proteins in all analyzed coronaviruses (see Figure 3 of Marra *et al*.[7]). This feature is not seen in other known structural proteins of SARS-CoV. Therefore, we hypothesize that the protein encoded by ORF3a forms the interchain disulfide bonds with S protein of SARS-CoV though their "cysteine-rich" regions.

Formation of interchain disulfide linkages between the 3a protein and spike glycoprotein

In order to address whether the 3a protein forms interchain disulfide bonds with S protein, we generated a polyclonal antibody against a peptide (3a1) derived from the cytoplasmic region of the 3a protein of SARS-CoV (see Materials and Methods). Western blotting detected 3a proteins in the cytosol of infected Vero E6 cells (Figure 2A, lane 2), which agrees with the results of 2D-LC-MS/MS analysis. Moreover, the 3a protein was also detected in the crude virions of SARS-CoV (Figure 2A, lane 3), suggesting that the 3a protein might be a structural protein of SARS-CoV, although further validation is required employing more purified SARS-CoV virions.

If according to sequence analysis, the 3a protein forms interchain disulfide linkages with S protein, a protein complex would be formed of around 210 kDa in SARS-CoV virions. To test this hypothesis, the crude virions were analyzed by immuno-blotting under both reducing and non-reducing conditions. A polyclonal antibody against a peptide (S2) derived from the sequence of S protein detected two bands around 210 kDa and 180 kDa under non-reducing conditions (Figure 2B, lane 2 of left panel), whereas only one band corresponding to 180 kDa was detected under reducing conditions (Figure 2B, lane 1 of left panel). Furthermore, the immuno-blotting with the anti-3a1 antibody also showed a band at 210 kDa under non-reducing conditions (Figure 2B, lane 2 of right panel). In order to further confirm the observations, the antibodies were stripped out from these blotted PVDF membranes, and then the membranes were re-probed by the exchange of the antibodies (S2 for 3a1 and 3a1 for S2, respectively). The same patterns were detected again (not shown). These results suggest that the 3a protein binds to the spike protein with the interchain disulfide bonds on the interior side of the viral envelope of SARS-CoV. Since two bands corresponding to S protein co-existed in the gel under the non-reducing conditions (Figure 2B, lane 2 of left panel) and no band was detected by the anti-3a1 antibody at 31 kDa in the gel under the non-reducing conditions (not shown), it suggests that all molecules of the 3a protein are used for the formation of interchain disulfide linkages with the S proteins of SARS-CoV.

Functional correlation of the 3a protein with spike glycoprotein

If the 3a protein binds to S protein with disulfide bonds, it is possible that the 3a protein-specific antibodies may be detected in the sera of SARS patients. To address this question, two synthetic peptides, 3a1 and 3a2 corresponding to the cytoplasmic region and the extracellular region of the 3a protein, respectively, were used as antigens to probe for 3a protein-specific antibodies in the sera from clinically diagnosed SARS patients. ELISA results showed that the SARS patients could be divided into two groups corresponding to high responders and low responders based on the *A* values, whereas the control produced basal low

Table 1. Identified peptides of the 3a protein of SARS-CoV with ESI-MS/MS

Peptide sequence	Position	Calculated $[M + H]^+$	Observed $[M + H]^+$	Observed charge	Xcorr	Delta CN	Peptide (Hits)
LKEDYQIGGYSEDR[a]	180–193	1673.76	1673.71	3	4.20	0.78	27/52
EDYQIGGYSEDR[a]	182–193	1432.43	1432.12	2	2.44	0.64	13/22
FFTLGSITAQPVK	7–19	1409.66	1409.86	2	3.30	0.90	18/24
LKEDYQIGGYSEDR	180–193	1673.76	1674.28	2	4.93	0.84	22/26
LV KDPPN V QIHTIDGSSGV ANP AMDPIYDEPTTTTSVPL	236–274	4092.54	4092.52	3	4.20	0.84	26/152
LV KDPPNV QIHTIDGSSG V ANP AM[b]DPIYDEPTTTTSV PL	236–274	4108.53	4108.00	3	3.86	0.95	26/152

The MS/MS spectra of the peptides are given in the Supplementary Material.
[a] Indication of peptides identified with the shotgun strategy.
[b] Indication of oxidised Met.

:

```
MDLFMRFFTL G SITAQPVK I DNASPASTV H ATATIPLQAS LPFGWLVIGV AFLAVFQSA T KIIALNKRWQ 70
LALYK GFQFI CNLLLLFV TI YSHLLLVAAG M EAQFLYLYA LIYFLQCINA CRIIMRCWLC WKCKS Q NPLL  140
YDANYFVCWH THNYDYCIPY NSVTDTIVVT E GDGISTPKL KEDYQIGGYS ED R HSGVKDY VVVHGYFTEV  210
YYQLESTQIT T D TGIENATF FIFNKLVKDP PNVQIHTIDG SSGVANPAMD PIYDEPTTTT SVPL          274
```

Figure 1. Sequence analysis of the 3a protein of SARS-CoV. The ORF 3 gene encodes a putative protein of 274 amino acid residues. Predicted transmembrane regions are demarcated by arrows and the direction is from the outer surface to the interior of virion envelope. Cysteine residues are indicated with triangles and the cysteine-rich region is underlined. Amino acid residues in italics have mutations among the isolates of SARS-CoV, especially G marked with dots has two different mutations.

levels of antibodies against 3a1 and 3a2 (Figure 3A and B). As a negative control, a peptide derived from mice IL-12 β2 receptor was used and no peptide-specific antibodies were detected in all negative controls and SARS patients (no shown). The results suggest that when the 3a protein was exposed to patients during SARS-CoV infection, 3a protein-specific antibodies were produced.

Since S protein is a major structural protein of SARS-CoV, the peptide-specific antibodies against S protein in the patients' sera should also be detected by the same approach. The same sera-samples of SARS patients that were analyzed with 3a2 peptide, were measured again with S1 peptide derived from the extracellular region of S protein. As expected, the pattern of ELISA assay with S1 peptide was similar to that of 3a1 and 3a2 peptides (Figure 3C). In addition, statistical analysis of the correlation between the responder of S1 and 3a2 with SigmaPlot (SPSS Inc. USA) indicated that the response of the 3a2 was associated with that of S1 ($r = 0.60$, $P < 0.01$; Figure 3D). Our data strongly suggest that the 3a protein and S protein are equally exposed to the human immune system, and then the induced antibodies are expressed at comparable levels to each other in the same SARS patient.

Until now, there have been 60 complete or partial genomic sequences of SARS-CoV isolates from human and from four available Himalayan palm civets. According to these sequencing data, we analyzed the 3a protein mutations of these isolates. The results showed that the 3a protein sequences of 48 SARS-CoV isolates were identical, which is referred to as the wild-type, whereas the rest of the isolates showed different mutations on their 3a protein sequences, of which 15 isolates showed point-mutations, including six synonymous mutations, and one isolate showed a frame-shift mutation. Interestingly, the analysis of S protein sequences showed that the sequences of 44 isolates were identical, of which 43 isolates also belong to the "wild-type" of the 3a protein, whereas another 19 isolates had mutations on their spike sequences. Statistical analysis with SigmaPlot indicated that the mutations of the 3a protein were associated with that of S protein ($r = 0.79$, $P < 0.05$). We did not find similar correlation with other structural proteins of SARS-CoV isolates (not shown).

Further analysis indicated that the mutation patterns in the S protein sequences from Himalayan palm civets were identical only in three human SARS-CoV isolates (GD01, AY278489.2; GZ43, AY304490.1; GZ60, AY304491.1), whereas the same mutant site (19 t → a) of 3a protein sequences from Himalayan palm civets was also only found in those three human SARS-CoV isolates. This kind of mutation-correlation between S protein and 3a protein suggests that there is some relationship between the SARS-CoV isolates from Himalayan palm civet and those from human being, at least, GD01, GZ43 and GZ60. Taken together, these results suggest that the biological function of the 3a protein is tightly correlated with S protein of SARS-CoV.

Phylogenetic analyses of the sequence of SNE locus of coronaviruses

It has been shown that ORF 3a is located within the region between the S and E genes in the SARS-CoV genome.[6,7] By analyzing the genome organizations of other known coronaviruses, we found that the pattern of a potential gene locus between the S and E genes is conserved in all three groups of coronavirus genomes (Figure 4A). Therefore, we designate this locus as SNE (S neighbor E). In general, the SNE locus consists of one to three ORFs in different species of coronaviruses, which encodes putative proteins containing at least one predicted transmembrane region (Figure 4A).

Figure 4A shows that only the genomes of SARS-CoV and PEDV have one complete ORF in the SNE locus, whereas the putative genes corresponding to ORF 3 in SNEs of five other coronaviruses are truncated and short ORFs are produced due to a mutation. However, if the entire sequence of the SNE locus was used for phylogenetic analysis, the topology of the resulting phylograms of SNE would be remarkably similar to that of S proteins (Figure 4B).[6,7] In other words, the phylogenetic analysis of SNE sequences indicates that SARS-CoV is not closely related to any of the three known groups of coronaviruses, which is in agreement with the phylogenetic analyses of all other known structural proteins of SARS-CoV.[6,7]
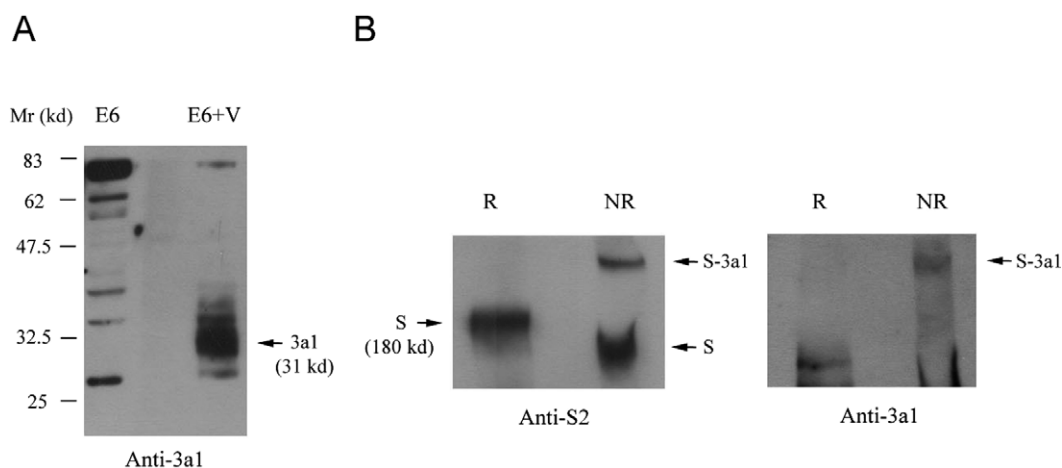
Figure 2. Formation of interchain disulfide bonds between the 3a protein and spike protein. A, Identification of the 3a protein in the cytosol of SARS-CoV-infected Vero E6 cells. Vero E6 cells (lane E6) and the cytosol of infected Vero E6 cells (lane E6 + V) were subjected to Western blotting with anti-3a1 antibody (against the 3a protein). An arrow indicates the band of the 3a protein. B, The 3a protein forms interchain disulfide linkages with S protein. The crude SARS-CoV virions were lysed in a sample loading buffer either with 100 mM DTT (reducing sample, lane R) or without DTT (non-reducing sample, lane NR), and then subjected to Western blotting with anti-S2 antibody (against S protein) (left panel) or anti-3a1 antibody (right panel).

Discussion

The 3a protein might be a minor structural protein

Though proteomics identified the 3a protein, an ORF3a gene product of SARS-CoV, it is a minor structural protein on the surface of SARS-CoV viral envelope. This is apparent because: (1) Western blotting with anti-3a1 antibody detected the 3a protein in the crude virions (Figure 2). (2) Since the conserved motif of the cysteine-rich region of the S protein is located between the end of the membrane-spanning region and the cytoplasmic tail, this motif may provide a frame for formation of interchain disulfide bonds (see Figure 3 of Rota *et al.*[6]), the spatial orientation of the 3a protein might determine the usage of its cysteine-rich region with the frame of the S protein in the virus envelope interior. (3) It is well known that the
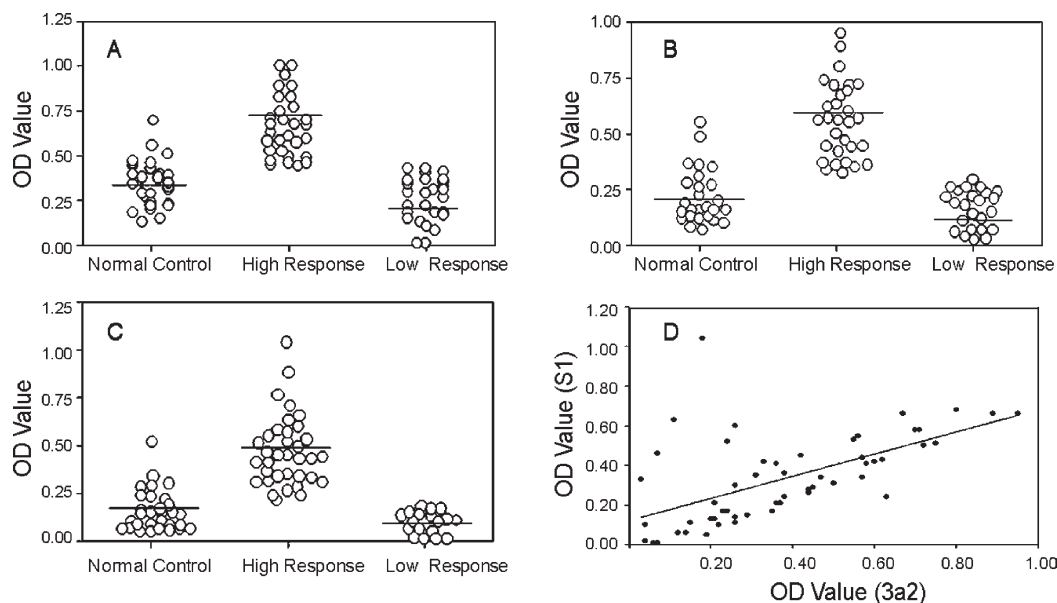

Figure 3. Detection of ORF3a peptide-specific IgG antibodies in SARS patients. The IgG antibodies in the sera of SARS patients were measured by ELISA with: A, 3a1 peptide; B, 3a2 peptide; C, S1 peptide. The low-response group consists of individual patients who had low *A* values, which were below the average of the normal control, whereas the high-response group had *A* values that were above the average of the normal control. D, The correlation of antibodies induced by the 3a protein and S protein in the same patient in response to 3a2 and S1-peptides with SigmaPlot software.

reducing environment of the cytosol in cells decreases the likelihood of disulfide bond formation. Therefore, the formation of the interchain disulfide bonds between the 3a protein and S protein must be in the interior of the viral envelope, in which there is a non-reducing environment. However, it might be necessary to further confirm this assumption with the more purified SARS-CoV virions, since the crude virions may be contaminated with cellular proteins and non-structural viral proteins.

Cysteine-rich regions may play an important role in SARS-CoV

Spike glycoprotein of coronaviruses contains many cysteine residues. In addition, there is a cysteine-rich region, overlapping the junction of the membrane-spanning region and the cytoplasmic region of S proteins, and this conserved structure of S protein exists in all three groups of coronaviruses.[6] The molecular function of the cysteine-rich domain in S protein is largely unknown. Recently, some evidence showed that the cysteine-rich domain of S protein was required for coronavirus-induced membrane fusion.[19,20] The sequence analysis showed that the 3a protein of SARS-CoV has a similar cysteine-rich region that is also located at the junction of the putative transmembrane and cytoplasmic region (Figure 1). (4) Here for the first time we show that the 3a protein might form interchain disulfide bonds with S proteins, probably through the cysteine-rich domains (Figure 2). Since no such kind of cysteine-rich domain exists in any other known SNE loci of coronaviruses, interchain disulfide bonds might not form between the proteins encoded by SNE loci and S protein in other coronaviruses.

Therefore, we postulate that the interaction between the 3a and S proteins occurs through their cysteine-rich regions, which is a special case in the coronavirus family, which may play an important role in affecting SARS-CoV virulence.

The 3a protein probably evolves from the SNE locus

An obvious question to ask is: what is the origin of the 3a protein of SARS-CoV during the evolutionary process of the coronaviridae family. The analysis on the SNE locus may provide some clues to answer this question, since the ORF gene encoding the 3a protein is located in this locus (Figure 4). First of all, the SNE locus exists widely in all three groups of coronavirus family besides SARS-CoV, which consists of one to three ORFs (Figure 4). It is likely that the SNE locus in the primitive coronavirus has only one ORF, and many short ORFs within the SNE locus may be produced in different strains of coronaviruses due to high mutation rates in this locus during the evolutionary process. Based on complete genome comparison, phylogenetic analysis of SARS-CoV showed that SARS-CoV is closer to group 1 of coronaviruses,[21] which is in agreement with our phylogenetic tree of the SNE locus (Figure 4B). Since it was proposed that PEDV and H229E diverged later than TEGV from their common ancestor in the group 1 according to the analysis of SNE locus,[10] we speculate that the 3a protein of SARS-CoV evolved from the SNE locus of PEDV or a PEDV-like ancestor.

It is not clear whether the SNE is necessary or dispensable for virulence of coronaviruses. It has been hypothesized that ORF 3/3a and ORF3-1/3b
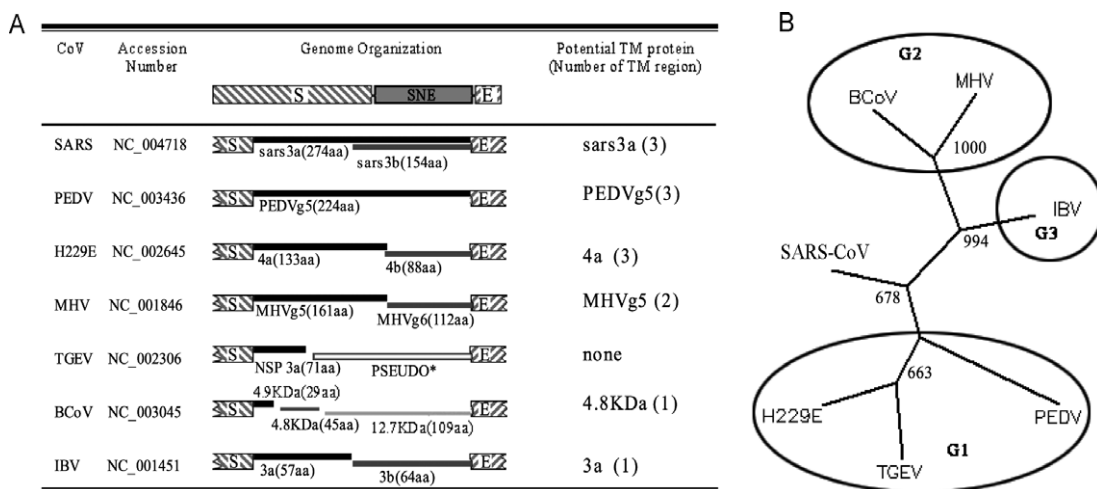

Figure 4. Analysis of SNE locus in the coronavirus family. A, Comparison of genome organization between S-gene and E-gene of coronaviruses. The predicted ORFs between the gene encoding the spike (S) protein and the gene encoding the small envelope (E) protein of seven species from three groups of coronaviruses were compared. The transmembrane (TM) region was also analyzed. B, Phylogenetic analysis of SNE locus. Unrooted phylogenetic tree of SNE region was generated by using CLUSTAL W 1.74 with the BLOSUM comparison matrix, and was drawn with the Phylip Drawtree program 3.6a3. Branch lengths indicate the number of substitutions per residue. Numbers indicate bootstrap replicates supporting each node.

may be related to viral virulence and pathogenesis,[11–13] whereas some suggested that ORF3a of TGEV was not essential for enteric virulence.[14] According to the data collected by ELISA (Figure 3), we postulate that the 3a protein plays an important role in the virulence of SARS-CoV, although it might not be the same case in other coronaviruses. Therefore, the 3a protein may serve as a new clinical marker or drug target in the treatment of SARS.

Materials and Methods

Cell culture and virus infection

African green monkey kidney cells (Vero E6, ATCC) were maintained in Dulbecco's modified Eagle's medium (DMEM, Gibco-BRL) supplemented with 10% (v/v) fetal bovine serum (FBS, Gibco-BRL) at 37 °C in a 5% (v/v) CO_2 atmosphere.

For virus infection, Vero E6 cells were treated for one hour with the DMEM (2% FBS) containing SARS-CoV virions (BJ-01 isolate, provided by Academy of Military Medical Sciences). The virus-medium was removed after the infection and the infected cells were cultured in DMEM with 2% FBS at 37 °C in 5% CO_2. After 48 hours post-infection, the supernatant medium of the infected cells was collected and centrifuged at 12,000 rpm for 30 minutes (Eppendorf 5415D) to eliminate the cell debris as crude SARS-CoV virions, of which $TCID_{50}$ (tissue culture infectious dose) was identified as 10^6 dilution. All the experiments using the virus were carried out in a Bio-safety Level 3 laboratory.

Preparation of cytosol of infected cells

Vero E6 cells were infected with SARS virus for 24 hours. The infected cells were lysed with a solution containing 40 mM Tris (pH 8.3) and 0.5% (v/v) Nonidet P-40 at room temperature for five minutes. The lysate was centrifuged at 8000 rpm for five minutes (Eppendorf 5415D), and then the supernatant was collected and heated at 100 °C for five minutes.

Peptide synthesis and raising polyclonal antibodies

The peptide sequences were determined based on the prediction of epitopes of the spike protein (gi:29836496) and the 3a protein (gi:30275670) as described by Rammensee *et al*.[22] combined with analysis of antigenicity (DNASTAR,†). The S1 peptide (amino acid residues 332–351, *TKFPSVYAWERKKISNCVAD*) and the S2 peptide (amino acid residues 758–780, *RNTREVFAQVKQMYKT PTLKYFG*) correspond to the spike protein; the 3a1 peptide (amino acid residues 176–199, *STPKLKEDYQIGGYSED RHSGVKD*) and the 3a2 peptide (amino acid residues 7–26, *FFTLGSITAQPVKIDNASPA*) correspond to the 3a protein. These peptides were synthesized using conventional solid-phase chemistry and purified by GL Biochem (Shanghai) Ltd. A peptide from the sequence of mice IL-12 2 receptor (amino acid residues *CNRLDLGINLSPD LAESPRFI*) was synthesized as a negative control. In addition, the nucleocapsid protein

was expressed in *Escherichia coli* and purified as an antigen for generating a polyclonal antibody.[23]

Rabbits were immunized with 200 μg of peptide-BSA or protein in 0.4 ml of emulsion mixed 1/1 (v/v) with complete Freund's adjuvant (CFA) supplemented with *Mycobacterium tuberculosis* to a final concentration of 1 mg/ml. The rabbits were boosted twice with freshly prepared emulsion of the conjugated peptide or protein and Freund's incomplete adjuvant at three week intervals. Blood was drawn from the rabbits at seven weeks following immunization, the blood was allowed to clot at 4 °C, and the antiserum was recovered by centrifugation. The antiserum was then purified with a BSA affinity column (CNBr-Sepharose 4B) according to the manufacture's instruction (Amersham Pharmacia).

ELISA analysis

The sera of 56 SARS patients, who were diagnosed by clinical symptoms and laboratory examination, were collected from different hospitals in Beijing. As normal control, the sera of 36 healthy people were collected. The sera dilution was 1 : 10 (v/v) for ELISA assay.

ELISA assay was done as described.[24] In brief, 96-well microtiter plates were coated with the peptides (S1, 3a1 and 3a2, respectively) in 0.1 M carbonate buffer (pH 9.6) at 4 °C overnight. After blocking with PBS containing 3% (w/v) gelatin and 0.1% (v/v) Tween 20, the plates were incubated with diluted sera from individual SARS patients or healthy controls at 37 °C for two hours. Bound antibodies were detected with horseradish peroxidase-coupled goat anti-human IgG (Bio-rad) and the absorbance values were measure by microplate autoreader (Bio-tek) at 450 nm.

Western blotting

The cellular fractions or viruses were lysed either in the non-reducing loading buffer (50 mM Tris (pH 6.8), 2% (w/v) SDS, 10% (v/v) glycerol, 0.1% (w/v) bromophenol blue) or reducing loading buffer containing 100 mM DTT. The mixtures were subjected to SDS-PAGE, and then transferred to Immobilon-P membrane (Millipore). Immunoblotting was carried out with the anti-3a1 antibody (against the 3a protein) or anti-S2 antibody (against S protein). The proteins were detected by enhanced chemiluminescence (ECL, Amersham Pharmacia Biotech).

"Shot-gun" approach

The cytosol fraction of infected Vero E6 cells was digested with trypsin and analyzed with 2D-LC-MS/MS system (ProteomeX, Thermo Finnigan) as described.[15] Protein identification was performed with BioWorks version 3.1(Thermo Finnigan) and the SEQUEST algorithm. Since the Vero-E6 cells were from monkey, both the human database and the SARS database from NCBI were merged. The MS results were searched against either the combined database or the SARS database alone. Protein results were further filtered with Xcorr $(1 + > = 1.5, 2 + > = 2.0, 3 + > = 2.5)$.

Mass spectrometry for identification of SDS-PAGE samples

The cytosol fractions of infected Vero E6 cells were subjected to PAGE and stained with Coomassie brilliant

blue. Gel slices were excised and in-gel digestion was carried out with trypsin as described.[25] The digested peptides were analyzed by mass spectrometry (LCQ Deca XP Plus, Thermo Finnigan). Data analysis was carried out as described for the shot-gun approach.

References

1. Lai, M. & Holmes, K. (2001). Coronaviridae: the viruses and their replication. In *Fields Virology* (Knipe, D. M. & Howley, P. M., eds), vol. 1, pp. 1163–1185, Lippincott-Willams and Wilkins, Philadelphia.
2. Boursnell, M. E., Brown, T. D., Foulds, I. J., Green, P. F., Tomley, F. M. & Binns, M. M. (1987). Completion of the sequence of the genome of the coronavirus avian infectious bronchitis virus. *J. Gen. Virol.* **68**, 57–77.
3. Lee, H. J., Shieh, C. K., Gorbalenya, A. E., Koonin, E. V., La Monica, N., Tuler, J. *et al.* (1991). The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology*, **180**, 567–582.
4. Herold, J., Raabe, T. & Siddell, S. (1993). Molecular analysis of the human coronavirus (strain 229E) genome. *Arch. Virol. Suppl.* **7**, 63–74.
5. Eleouet, J. F., Rasschaert, D., Lambert, P., Levy, L., Vende, P. & Laude, H. (1995). Complete genomic sequence of the transmissible gastroenteritis virus. *Advan Expt. Med. Biol.* **380**, 459–461.
6. Rota, P. A., Oberste, M. S., Monroe, S. S., Nix, W. A., Campagnoli, R., Icenogle, J. P. *et al.* (2003). Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science*, **300**, 1394–1399.
7. Marra, M. A., Jones, S. J., Astell, C. R., Holt, R. A., Brooks-Wilson, A., Butterfield, Y. S. *et al.* (2003). The genome sequence of the SARS-associated coronavirus. *Science*, **300**, 1399–1404.
8. Snijder, E. J., Bredenbeek, P. J., Dobbe, J. C., Thiel, V., Ziebuhr, J., Poon, L. *et al.* (2003). Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. *J. Mol. Biol.* **331**, 991–1004.
9. Holmes, K. V. & Enjuanes, L. (2003). Virology. The SARS coronavirus: a postgenomic era. *Science*, **300**, 1377–1378.
10. Duarte, M., Tobler, K., Bridgen, A., Rasschaert, D., Ackermann, M. & Laude, H. (1994). Sequence analysis of the porcine epidemic diarrhea virus genome between the nucleocapsid and spike protein genes reveals a polymorphic ORF. *Virology*, **198**, 466–476.
11. Paul, P. S., Vaughn, E. M. & Halbur, P. G. (1997). Pathogenicity and sequence analysis studies suggest potential role of gene 3 in virulence of swine enteric and respiratory coronaviruses. In *Mechanisms in the Pathogenesis of Enteric Diseases* (Paul, P. S., Francis, D. H. & Benfield, D. A., eds), pp. 317–321, Plenum Press, New York.
12. Vaughn, E. M., Halbur, P. G. & Paul, P. S. (1995). Sequence comparison of porcine respiratory coronavirus isolates reveals heterogeneity in the S, 3, and 3-1 genes. *J. Virol.* **69**, 3176–3184.
13. Wesley, R. D., Woods, R. D. & Cheung, A. K. (1991). Genetic analysis of porcine respiratory coronavirus, an attenuated variant of transmissible gastroenteritis virus. *J. Virol.* **65**, 3369–3373.
14. McGoldrick, A., Lowings, J. P. & Paton, D. J. (1999). Characterisation of a recent virulent transmissible gastroenteritis virus from Britain with a deleted ORF 3a. *Arch. Virol.* **144**, 763–770.
15. Chelius, D., Huhmer, A. F., Shieh, C. H., Lehmberg, E., Traina, J. A., Slattery, T. K. & Pungor, E., Jr (2002). Analysis of the adenovirus type 5 proteome by liquid chromatography and tandem mass spectrometry methods. *J. Proteome Res.* **1**, 501–513.
16. Stoffel, K. H. W. (1993). TMbase–a database of membrane spanning proteins segments. *Biol. Chem. Hoppe-Seyler*, **374**, 166.
17. Nielsen, H., Engelbrecht, J., Brunak, S. & von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**, 1–6.
18. Nielsen, H. & Krogh, A. (1998). Prediction of signal peptides and signal anchors by a hidden Markov model. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **6**, 122–130.
19. Chang, K. W., Sheng, Y. & Gombold, J. L. (2000). Coronavirus-induced membrane fusion requires the cysteine-rich domain in the spike protein. *Virology*, **269**, 212–224.
20. Gallagher, T. M. (1996). Murine coronavirus membrane fusion is blocked by modification of thiols buried within the spike protein. *J. Virol.* **70**, 4683–4690.
21. Qi, Z., Hu, Y., Li, W., Chen, Y. J., Zhang, Z. H., Sun, S. W. *et al.* (2003). Phylogeny of SARS-CoV as inferred from complete genome comparison. *Chinese Sci. Bull.* **48**, 1175–1178.
22. Rammensee, H., Bachmann, J., Emmerich, N. P., Bachor, O. A. & Stevanovic, S. (1999). SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*, **50**, 213–219.
23. Lin, Y., Shen, X., Yang, R. F., Li, Y. X., Ji, Y. Y., He, Y. Y. *et al.* (2003). Identification of an epitope of SARS-coronavirus nucleocapsid protein. *Cell Res.* **13**, 141–145.

24. Hou, W., Wu, Y., Sun, S., Shi, M., Sun, Y., Yang, C. *et al.* (2003). Pertussis toxin enhances Th1 responses by stimulation of dendritic cells. *J. Immunol.* **170**, 1728–1736.

25. Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T. & Mann, M. (1996). Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature*, **379**, 466–469.

Edited by M. Yaniv

SCIENCE *d* **DIRECT**®

www.sciencedirect.com

Supplementary Material comprising six Figures is available on Science Direct