# The Molecular Biology of SARS Coronavirus

NAMITA SATIJA AND SUNIL K. LAL

*Virology Group, International Centre for Genetic Engineering & Biotechnology, New Delhi, India*

**ABSTRACT: Severe acute respiratory syndrome (SARS) is the first emerging infectious disease of the 21st century that has been highly transmissible and fatal and was caused by a previously unknown coronavirus (SARS-CoV). The SARS epidemic in 2003 resulted in more than 8400 SARS cases and approximately 800 deaths. Existing in non-identified animal reservoirs, SARS-CoV continues to represent a threat to humans although more than four years have passed since a large outbreak of SARS, and no new cases have been reported. However, we cannot exclude the possibility of reemergence of SARS. It is hence necessary to understand the biology of the SARS-CoV to deal adequately with the next outbreak, whenever it happens. The SARS-CoV is a novel coronavirus with a large (∼30 thousand nucleotides) positive-sense, single-stranded RNA containing 14 functional open reading frames (ORFs) of which 2 large ORFs constitute the replicase gene which encodes proteins required for viral RNA syntheses. The remaining 12 ORFs encode the 4 structural proteins: spike, membrane, nucleocapsid and envelope; and eight accessory proteins. The viral genome and its expression within the host cell undergoes extensive translational and enzymatic processing to form the 4 structural, 8 accessory and 16 nonstructural proteins. In an effort to understand the molecular mechanisms or capsid assembly and viral pathogenesis, laboratories around the world have adopted a variety of approaches to answering these trivial questions. It has been our effort to consolidate all information known to date about the molecular mechanisms of the SARS-CoV into this chapter to update our readership on the current status of research.**

**KEYWORDS: coronavirus; structural proteins; accessory proteins; viral capsid; assembly; proteolytic cleavage**

Address for correspondence: Sunil K. Lal, Virology Group, International Centre for Genetic Engineering & Biotechnology, P.O. Box 10504, Aruna Asaf Ali Road, New Delhi 110067, India. Voice: 91-11-26177357; fax: 91-11-26162316.
sunillal@icgeb.res.in

**Ann. N.Y. Acad. Sci. 1102: 26–38 (2007). © 2007 New York Academy of Sciences.**
**doi: 10.1196/annals.1408.002**

## INTRODUCTION

Severe acute respiratory syndrome (SARS) is a newly emerged infectious disease that appeared in Guangdong Province, mainland China, in November 2002. By March 2003, the disease had spread globally and by July there were 8,447 probable SARS cases including 811 deaths reported from 32 countries or regions worldwide to the WHO.

The etiological agent was identified as a coronavirus (CoV)[1] and the complete genome sequence of SARS-CoV Tor2 (Toronto) isolate was published in April 2003, which established it as a novel member of the family.[2,3]

The genome of SARS-CoV is a positive sense single-stranded ribonucleic acid (RNA) consisting of 29,751 bases. Phylogenetic analysis revealed that this CoV was only moderately related to other known CoVs, including two human CoVs, HCoV-OC43 and HCoV-229E, known to cause common colds and lower respiratory tract infections and diarrhea in humans.

The clinical symptoms of the syndrome include fever, chills, rigors, cough, and headache and the pathological aspects include lymphopenia, thrombocytopenia, elevated lactate dehydrogenase, and creatine kinase levels.

Although currently the spread of the virus seems to be confined to rigorous and timely quarantine measures (http://www.who.int/csr/sars/country/table2003 0923/en/), it may still be circulating in the animal reservoir and it is impossible to predict when it will return. Because this scare of recurrence of SARS looms over mankind, better monitoring of SARS outbreaks through accurate diagnostic tests and the development of effective antiviral therapies are urgently required. These in turn depend directly on better molecular understanding of the SARS-CoV infection cycle. Research toward the understanding of the SARS virus at a molecular level is therefore of vital importance to combat SARS-CoV infection if another SARS pandemic was to occur in the future.

## TISSUE TROPISM

SARS mainly shows pneumonia-like symptoms and the lung is pathologically the most affected organ. Detailed investigations, however, suggest that SARS is a systemic disease with widespread extrapulmonary dissemination, resulting in viral shedding in respiratory secretions, stools, urine, and possibly even in sweat.[4,5] Immunohistochemistry and *in situ* hybridization of organs from four SARS patients who died revealed that the virus was found not only in the lung and intestine, but also in the liver, distal convoluted renal tubules, sweat glands, parathyroid, pituitary, pancreas, adrenal gland, and cerebrum.[4] SARS-CoV RNA was detected in the lung, small and large bowels, lymph nodes, spleen, liver, heart, kidney, and skeletal muscle, in descending order of viral load per gram of tissue using reverse-transcriptase polymerase chain reaction.[5]

## CELLULAR RECEPTOR

The virus seems to gain entry into the target cells by direct membrane fusion at the target cell surface. Apart from this it has been postulated that the SARS-CoV may gain cell entry via pH-dependent endocytosis. The spike protein of the virus is the key player in this regard.[6] A metallopeptidase, angiotensin-converting enzyme 2 (ACE2), the C-type lectin CD209L (also known as L-SIGN), and DC-SIGN bind SARS-CoV, but ACE2 appears to be the key functional receptor for the virus.

The ACE2 protein is reportedly present in type 1 and type 2 pneumocytes, enterocytes of all parts of the small intestine, the brush border of the proximal tubular cells of the kidney as well as the endothelial cells of small and large arteries and veins of all tissues studied, and arterial smooth muscle cells.[7] This localization of ACE2 explains the tissue tropism of SARS-CoV for the lung, small intestine, and kidney; however, notable discrepancies include virus replication in colonic epithelium, which has no ACE2, and no virus infection in endothelial cells, which have ACE2. Other receptors or co-receptors such as L-SIGN may explain such discrepancies.

## GENOME ORGANIZATION AND PROTEIN PROCESSING

The SARS-CoV genome has been predicted to contain 14 functional open reading frames (ORFs). The comparison of different SARS-CoV ORFs with those of other CoVs revealed a similar pattern of structural gene arrangement with the replicase and protease genes (gene 1a-1b) and the spike (S), envelope (E), membrane (M), and nucleocapsid (N) genes in a 5′ to 3′ order of appearance. Interspersed between these well-characterized genes is a series of ORFs, many of whose functions are yet unknown. There are two ORFs situated between the S and the E genes and three to five ORFs between the M and N genes (FIG.1). Such gene organization most closely resembles that of group III CoVs. Also, the SARS-CoV genomic sequence does not contain a gene for the hemagglutinin-esterase (HE) protein, which is present in most group II CoVs.

Two-thirds of the SARS-CoV RNA is organized in the genes 1a-1b. The sequence of these genes is highly conserved among all CoVs with the exception that SARS-CoV lacks the sequence coding for papain-like protease (PL1pro), one of the two papain-like proteinases operating on cleavage sites at the N terminus of the polyproteins. ORFs 1a and 1b encode two polyproteins, pp1a and pp1ab, the latter is predicted to translate through a -1 ribosomal frameshifting mechanism. These two polyproteins are processed by virus-encoded proteinases, to yield 16 individual nonstructural proteins (nsp). Many of the postulated nsp proteins to date have undesignated functions but it is suggested that they participate primarily in viral RNA replication. Most

potential gene 1a-1b products are fairly well conserved between SARS-CoV and other CoVs. PL2pro of the SARS virus is responsible for the cleavage of all the N-terminal proteins of gene 1a. The main proteinase (Mpro), also called the 3C-like protease (3CLpro), is responsible for the cleavage of all the remaining proteins encoded by gene 1a-1b.

### *The SARS-CoV 5′ and 3′ UTR*

The 5′ UTR (untranslated region) of the SARS-CoV genome was initially characterized by 5′ rapid amplification of the cDNA ends and Northern blot assays. These procedures elucidated the leader sequence and the transcription regulatory sequence (TRS). The leader sequences found in the viral sg-mRNA (subgenomic-mRNA) transcripts were at least 72 nucleotides long. After aligning the leader sequence at the 5′ end of the eight sg-mRNAs, a minimal consensus TRS was found with the sequence 5′-ACGAAC-3′,[8] which participated in the discontinuous synthesis of sg-mRNAs as a signaling sequence. TRS showed no clear relationship with the abundance of the sg-mRNAs. A highly conserved s2m motif with 32 nucleotides was also identified in the 3′ region of the genome, which has also been discovered in the avian infectious bronchitis virus.

### *The Replicase Gene*

The replicase gene of the SARS-CoV encodes for two proteins as a consequence of the proteolytic processing of the large polyproteins (pp1a and pp1ab). The translation of segment 1b of such a polyprotein is interrupted by the -1 ribosomal frameshifting by a putative “slippery” sequence and a corresponding putative pseudoknot structure, which is required for stalling the translating ribosome at the slippery site to undergo -1 translational frameshifting. Two functional domains, the papain-like cysteine proteinase (PL2pro) and the 3C-like cysteine proteinase (3CLpro), were identified experimentally and are considered responsible for the proteolytic processing of the polyprotein into 16 subunits. A 375 amino acid SARS-CoV unique domain was identified upstream of the PL2pro domain, not found in any other known CoVs. In addition, seven more putative regions encoding RNA processing enzymes were identified, namely, RNA-dependent RNA polymerase (RDRP), RNA helicase (HEL), poly (U)-specific endoribonuclease (NendoU), 3′-5′ exonuclease (ExoN), S-adenosylmethionine-dependent ribose 2′-O-methyltransferase (2′-O-MT), adenosine diphosphate-ribose 1′-phosphatase (ADRP), and a cyclic phosphodiesterase (CPD). The translation of two polyproteins from the ORF 1a and 1b starts the genome expression (FIG. 1). The polyproteins are then proteolytically processed by PL2pro and 3CLpro into 16 units. PL2pro is responsible for the N-proximal cleavage and 3CLpro is responsible for the

C-proximal cleavage. The helicase is then released. ATPase activity and DNA duplex-unwinding activity were demonstrated by purified helicase, indicating that the protein has RNA polymerase activity.[9]

### *The Spike Protein*

Together with the M protein, the spike (S) protein is believed to be incorporated into the viral envelope before the mature virion is released. Initial analysis of the 1255a.a. peplomer protein of the virus reveals the possible existence of a signal peptide that would most likely be cleaved between residues 13 and 14.[11,12] It forms typical petal-shaped spikes on the surface of the virion and consists of three domains, the external N-terminal domain with its conserved S1 and S2 subdomains, a transmembrane domain, and a short cytoplasmic domain at the C terminus. S1 is the receptor-binding unit in the N terminus. Molecular modeling of the S1 and S2 subunits of the S glycoprotein suggest that the former unit consists mainly of antiparallel β-sheets with dispersed α and β regions, in addition to the three domains identified in the S2 unit. Entry of CoVs into target cells is initiated by binding of the viral S protein to receptor molecules. The S1 domain of SARS-CoV is significantly different from that of other human CoVs, implying that these viruses are using different receptors for host cell entry. The cellular receptor of several group I CoVs are aminopeptidase N and zinc metalloprotease, whereas mouse hepatitis virus (MHV) (group II) uses carcinoembryonic antigen-related cell adhesion molecules as a cellular receptor. In the case of SARS-CoV, ACE2 was demonstrated to be a functional receptor for the SARS-CoV *in-vitro*. Syncytia were observed in cell culture expressing ACE2 and the SARS-CoV S1 domain, which could be inhibited using anti-ACE2 antibody. Fine mapping on the N-terminal unit of the S protein indicates that the minimal receptor-binding domain is located between the residues 318–510.[10] Antibodies specific for this S1 subunit of the SARS-CoV S protein were shown to neutralize SARS-CoV infection. The C-type lectin CD209L (also called L-SIGN), a human cellular glycoprotein has been reported to serve as an alternative receptor for SARS-CoV. The S protein of SARS-CoV contains 23 putative N-glycosylation sites, among which 12 have been described to be effectively glycosylated.[11,12] The S protein, when expressed alone, acquires EndoH-resistant complex N-glycans in the Golgi within 30 min following expression. Both high-mannose and complex glycan N-glycoforms have been detected on S trimers within ER and Golgi, respectively, suggesting that trimers form in the ER and pass the quality control to move toward the Golgi to acquire complex N-glycans. It has also been shown that the S glycoprotein is present all along the secretory pathway from the ER to the plasma membrane.[13]

Recently, structure of SARS spike protein's receptor-binding domain in complex with its receptor, ACE2 has been determined using X-ray crystallography. The structure though determined at 2.9A, shows potential for usage in designing vaccines against the virus.[14]

### *The Nucleocapsid Protein*

Nucleocapsid or the N protein of SARS-CoV is a 46-kDa phosphoprotein, which shows homology with other members of CoV family. Similar to other CoVs, it gets abundantly expressed during infection and antibodies against N have been detected in SARS patients.[15–18] N protein of most CoVs contains multiple epitopic sites lying most commonly at their C terminus. The epitope site for TGEV is located around codons 360–382,[19] for MHV it is at 381–405[20] and for infectious bronchitis virus (IBV) it is located at 360–409,[21] all at the C terminus. Likewise SARS-CoV, N protein has been found to contain its most antigenic site at the C terminus from codons 371–407.[22] This makes N an excellent candidate for diagnostic purposes.

Being a multipurpose protein, it plays a vital role during virion assembly. The process of virion assembly requires efficient packaging of the viral RNA into the virion. For this, N protein associates with the viral genomic RNA, and together they form the ribonucleoprotein. In MHV it has been shown that N protein does possess RNA-binding activity.[16] The structure of the N terminal (49–178 aa) region of the N protein has been solved by a group at the Abbott Laboratories, Illinois, USA. The structure reveals an RNA-binding domain similar to U1A RNA-binding protein.[23] Studies have also revealed that packaging of SARS RNA into virus-like particles (VLPs) is dependent on N protein. Packaging signal within a domain from nucleotide 19715–20294 of genomic RNA and the N terminal structure of N protein are essential for packaging and RNA-binding activity, respectively.[24] Purified SARS-CoV has been shown to have RNA chaperone activity.[51]

Self-association of N protein governs the formation of viral capsid to a large extent and has been observed in many viruses. For SARS-CoV, two groups have shown that N protein does dimerize. The interaction domain responsible for dimerization was first shown to be residing in the C terminal 209 amino acid residues[25] and was further narrowed down to 138 amino acid residues (between 285 and 422) at the C terminus of the protein.[26] N's ability to form oligomers (trimers, tetramers, and hexamers) has been demonstrated to be residing in residues 343–402.[49] N also interacts with the membrane protein. Amino acids 211–254, especially tetrad glutamines ($^{240}$QQQQ$^{243}$), are essential for this interaction.[50]

Apart from interactions among viral proteins, interactions with the proteins of the host play a major role in the establishment of viral infection. Hence mapping these viral host protein interactions provides vital hints about how infection actually takes place and progresses. It has been shown that N binds tightly to human cyclophilin A (cypA).[27] CypA is a family of proteins that has been extensively studied but their function remains unknown. Recently, cyclophilins have been shown to play an important role in HIV infection. CypA specifically has been shown to play a vital role in viral core disassembly.[28] Hence CypA interaction with N protein may provide new hints about how SARS CoV infects host cells.

N protein has also been reported to bind to human cellular heterogenous nuclear ribonucleoprotein A1 (hnRNP A1) with high affinity.[29] The regions responsible for this protein–protein interaction have been shown to be residing in the region from amino acid 161–210 of SARS CoV and C terminus amino acid 203–320 of human hnRNP A1. Studies have indicated that hnRNP A1 acts as a global regulator of alternative pre-mRNA splicing and its activities result in inactivation of alternative distal 5′ splicing sites and skipping of optional exons.[30] Also a recent publication has demonstrated that the nucleocapsid protein of MHV could interact with hnRNP A1.[31] Hence it may be postulated that interaction of SARS-CoV, N protein with hnRNP A1 may play a role in transcription and replication of viral genome.

Another aspect important for establishing an infection is modulation of host pathways by viral proteins. A report by He *et al.* in 2003 has shown that N protein activates AP1 signal transduction pathway. However, the mechanism underlying the observation is not known yet. Neither is it known whether N interacts with the components of AP1 directly or through an intermediate. AP1 or activator protein 1 pathway is a regulator of a wide variety of cellular processes. These include cell proliferation, differentiation, and apoptosis. This makes this pathway an important target for modification by the viral proteins to gain control over the host cell cycle.[32] Moreover, binding of N protein to cyclin-dependent kinase (CDK) complex leads to downregulation of S phase gene products, which in turn leads to inhibition of S phase progression.[33]

SARS-CoV N protein has also been shown to induce apoptosis in Cos 1 cells. Cells transfected with N protein show significant cell death under conditions of stress, induced by serum starvation. N protein has been found to downregulate extracellular signal-regulated kinase (ERK), but upregulate c-jun-N-terminal kinase (JNK) and p38 mitogen-activated map kinase (MAPK). Further, p38 MAPK activation was also found to induce actin reorganization.[34]

Nucleocapsid protein of several CoVs exhibits cytoplasmic as well as nucleolar localization at some point of time during the infection cycle. In the case of SARS-CoV, the N protein had been predicted to contain as many as eight nuclear localization signal (NLS) motifs, both monopartite and bipartite. A report by Zhu *et al.*[35] showed that the wild-type N was distributed mainly in the cytoplasm. The N-terminal of N, which contains the NLS1 (aa38–44), was localized to the nucleus. The C terminus of N, which contains both NLS2 (aa257–265) and NLS3 (aa369–390), was localized to the cytoplasm and the nucleolus. Also the region containing amino acids 226–289 was able to mediate nucleolar localization. Furthermore, deletion of the leucine-rich region (220-LALLLLDRLNRL) resulted in the accumulation of N to the cytoplasm and nucleolus, and when fusing this peptide to enhanced green fluorescence protein (EGFP) localization was cytoplasmic, suggesting that N may act as a shuttle protein.[35] A report by Surjit *et al.* (2005) provides evidence that N is indeed a nucleocytoplasmic-shuttling protein.[36] N proteins of many CoVs have been shown to be phosphorylated.[37] How-

ever, the mechanism by which this happens and the functional significance of this process remains largely unknown. N protein of SARS-CoV has been shown to be a phosphoprotein, which is stable and localizes in the cytoplasm. The phosphorylation takes place at the serine residues and that too by multiple kinases. It has been shown to be a substrate for CDK, glycogen synthase kinase, MAPK, and casein kinase II. A possible mechanism for phosphorylation-dependent nucleocytoplasmic shuttling of the N protein has also been elucidated. It has been shown that phosphorylated N is translocated to cytoplasm with the help of 14-3-3 (tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein), a protein family that functions as adapters for modulating interactions between components of various cellular signaling and cell-cycle regulatory pathways through phosphorylation-dependent protein–protein interactions.[36]

It can be postulated that N phosphorylation may be serving an array of purposes ranging from acting as a deciding factor for localization of protein to aiding in creating a more favorable environment for replication of the virus.[36]

### *The Matrix Protein*

CoV matrix (M) protein is the most abundant structural protein at the surface of virus particles and contains three putative transmembrane domains, a short N-terminal ectodomain, and a long C-terminal endodomain.

Interaction of the M protein with N has been described for many CoVs. For SARS-CoV, the M–N interaction was first reported by Fang *et al.* (2005) using glutathione-*S*-transferase pulldowns. Laboratory experiments have shown that the 12 amino acids domain (194–205) in the M protein is responsible for binding to the N protein.[39]

Linear epitope mapping of the M protein using synthetic peptides revealed that amino acid residues 2,137,158 interacted with SARS patient sera by enzyme-linked immunosorbent assay, implying the potential capability of the M protein to induce an immune response. The M protein resides on the viral envelope and in the internal core. In the cells, the M protein is anchored in the Golgi complex, thus dictating the site of virus assembly to the ER-Golgi complex.

It has been shown that SARS-CoV M protein is N-glycosylated at Asn 4 residue in mammalian cells.[40] Also C-terminally tagged M glycoprotein strongly co-localizes with Golgi markers and partially with the ERGIC-53 protein, a lectin that cycles between ER, ERGIC, and cis-Golgi.[41]

### *The Envelope Protein*

The envelope (E) protein of the SARS-CoV is, as the name suggests, the main component of the virus envelope. It is an integral membrane protein,

76 amino acids long, and is highly hydrophobic in nature. In mammalian cells it is found to be localized to endoplasmic reticulum (ER), Golgi apparatus, and lipid rafts of cell membrane. The transmembrane domain of the E protein has been found to alter the membrane permeability.[42]

### *The ORF 3a or X1*

ORF 3a is also called SARS X1 and was initially named U274. The 3a locus encodes an ORF of unknown function and shows no homology to any known protein. The 3a mRNA and 3a protein have been detected in SARS-CoV-infected monkey kidney Vero E6 cells. In addition, the 3a protein has been detected in the lung tissue of a SARS patient. A signal sequence and three transmembrane domains have been predicted when 3a sequence was searched against SMART server.[43] It is a novel SARS-CoV protein that has been shown to be transported efficiently to the cell surface of infected and transfected cells. Its cytoplasmic domain contains sorting signals important for this process. The 3a protein expressed on the cell surface can also undergo endocytosis. It has been shown to interact with M, E, and S proteins suggesting that it may play a role in viral assembly and/or release of virus from infected cells.[44] Also it has been shown to form ion channels in Xenopus oocytes. Formation of a pore in the infected cells may aid in virus release. The modulation of virus release may also be one of the important functions of 3a as reduction in 3a protein expression in SARS-infected FRhK4 cells led to a decrease in virus release.[48]

3a has also been shown to upregulate expression of fibrinogen in lung epithelial cells. mRNA levels of all three subunits of fibrinogen Aα, Bβ, and γ were found to be upregulated. As a result of this intracellular levels as well as secretion of fibrinogen were found to increase.[45]

In addition, 3a protein has also been shown to be expressed in the lungs and intestinal tissues of SARS patients and that the protein localized to the ER in 3a-transfected monkey kidney Vero E6 cells. *In vitro* experiments of chromatin condensation and DNA fragmentation suggested that the 3a protein may trigger apoptosis through a caspase-8-dependent pathway in Vero E6 cells.[46]

### *The ORF 7a or X4*

Like ORF 3a, sequence homology search yielded no significant result for any existing proteins, but the existence of a cleavage site (between residues 15 and 16) and a transmembrane helix were predicted for the 7a protein. It has recently been shown that overexpression of 7a induces apoptosis via a caspase-dependent pathway in different cell lines derived from the lung, kidney, and liver.[47]

## CONCLUSIONS

Four years since the discovery of SARS, a great deal of scientific coordination across the globe has led to unfolding of various facts about this virus. But much needs to be learned about pathogenesis of the disease and biology of the virus. These aspects are not fully understood as yet and consorted efforts and collaboration among the scientific community are required to unravel the mysteries surrounding the first pandemic of the 21st century. Reemergence of SARS looms as an uncertain possibility.

The challenge now is to maintain interest in efforts to understand the biology of the virus and pathogenesis of the disease caused by it so as to minimize the risk of being surprised by its sudden reappearance. The lessons learned from SARS would certainly be helpful to counter any reemergence of SARS itself or emergence of any other pandemic.

## REFERENCES

1. PEIRIS, J.S., C.M. CHU, V.C. CHENG, *et al.* HKU/UCH SARS Study Group. 2003. Coronavirus as a possible cause of severe acute respiratory syndrome. Lancet **361:** 1319–1325.
2. ROTA, P.A., M.S. OBERSTE, S.S. MONROE, *et al.* 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science **300:** 1394–1399.
3. MARRA, M.A., S.J. JONES, C.R. ASTELL, *et al.* 2003. The genome sequence of the SARS-associated coronavirus. Science **300:** 1399–1404.
4. DING, Y., L. HE, Q. ZHANG, *et al.* 2004. Organ distribution of severe acute respiratory syndrome (SARS) associated coronavirus (SARS-CoV) in SARS patients: implications for pathogenesis and virus transmission pathways. J. Pathol. **203:** 622–630.
5. FARCAS, G.A., S.M. POUTANEN, T. MAZZULLI, *et al.* 2005. Fatal severe acute respiratory syndrome is associated with multiorgan involvement by coronavirus. J. Infect. Dis. **191:** 193–197.
6. YANG, Z.Y., Y. HUANG, L. GANESH, *et al.* 2004. pH-dependent entry of severe acute respiratory syndrome coronavirus is mediated by the spike glycoprotein and enhanced by dendritic cell transfer through DC-SIGN. J. Virol. **78:** 5642–5650.
7. HAMMING, I., W. TIMENS, M.L.C. BULTHUIS, *et al.* 2004. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. J. Pathol. **203:** 631–637.
8. THEIL, V., K.A. IVANOV, A. PUTICS, *et al.* 2003. Mechanisms and enzymes involved in SARS coronavirus genome expression. J. Gen. Virol. **84:** 2305–2315.
9. ZEIBUHR, J. 2004. Molecular biology of severe acute respiratory syndrome coronavirus. Curr. Opin. Microbiol. **7:** 412–419.
10. WONG, S.K., L. WENHUI, M.J. MOORE, *et al.* 2004. A 193 amino acid fragment of SARS coronavirus S protein efficiently binds angiotensin converting enzyme 2. J. Biol. Chem. **279:** 3197–3201.

11. Krokhin, O., Y. Li, A. Andonov, *et al*. 2003. Mass spectrometric characterization of proteins from the SARS virus: a preliminary report. Mol. Cell Prot. **2:** 346–356.
12. Ying, W., Y. Hao, Y. Zhang, *et al*. 2004. Proteomic analysis on structural proteins of severe acute respiratory syndrome coronavirus. Proteomics **4:** 492–504.
13. Nal, B., C. Chan, F. Kien, *et al*. 2005. Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. J. Gen. Virol. **86:** 1423–1434.
14. Li, F., W. Li, M. Farzan & S.C. Harrison. 2005. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science **309:** 1864–1868.
15. Liu, X., Y. Shi, P. Li, *et al*. 2004. Profile of antibodies to the nucleocapsid protein of the severe acute respiratory syndrome (SARS)-associated coronavirus in probable SARS patients. Clin. Diagn. Lab. Immunol. **11:** 227–228.
16. Shi, Y., Y. Yi, P. Li, *et al*. 2003. Diagnosis of severe acute respiratory syndrome (SARS) by detection of SARS coronavirus nucleocapsid antibodies in an antigen-capturing enzyme-linked immunosorbent assay. J. Clin. Microbiol. **41:** 5781–5782.
17. Wang, J., J. Wen, J. Li, *et al*. 2003. Assessment of immunoreactive synthetic peptides from the structural proteins of severe acute respiratory syndrome coronavirus. Clin. Chem. **49:** 1989–1996.
18. Yuxian, H., Y. Zhou, H. Wu, *et al*. 2004. Mapping of antigenic sites on the nucleocapsid protein of the severe acute respiratory syndrome coronavirus. J. Clin. Microbiol. **42:** 5309–5314.
19. Martin, A.J.M. 1992. Antigenic structure of transmissible gastroenteritis virus nucleoprotein. Virology **188:** 168–174.
20. Stohlman, S.A. 1994. Location of antibody epitopes within the mouse hepatitis virus nucleocapsid protein. Virology **202:** 146–153.
21. Seah, J.N. 2000. Localization of linear B-cell epitopes on infectious bronchitis virus nucleocapsid protein. Vet. Microbiol. **75:** 11–16.
22. Li, S., L. Lin, H. Wang, *et al*. 2003. The epitope study on the SARS-CoV nucleocapsid protein. Genomics Proteomics Bioinformatics **1:** 198–206.
23. Huang, Q., L. Yu, A.M. Petros, *et al*. 2004. Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. Biochemistry **43:** 6059–6063.
24. Hsieh, P.K., S.C. Chang, C.C. Huang, *et al*. 2005. Assembly of severe acute respiratory syndrome coronavirus RNA packaging signal into virus-like particles is nucleocapsid dependent. J. Virol. **79:** 13848–13855.
25. Surjit, M., B. Liu, P. Kumar, *et al*. 2004. The nucleocapsid protein of the SARS coronavirus is capable of self-association through a C-terminal 209 amino acid interaction domain. Biochem. Biophys. Res. Commun. **317:** 1030–1036.
26. Yu, I.M., C.L. Gustafson, J. Diao, *et al*. 2005. Recombinant severe acute respiratory syndrome (SARS) coronavirus nucleocapsid protein forms a dimer through its C-terminal domain. J. Biol. Chem. **280:** 23280–23286.
27. Luo, C., H. Luo, S. Zheng, *et al*. 2004. Nucleocapsid protein of SARS coronavirus tightly binds to human cyclophilin A. Biochem. Biophys. Res. Commun. **321:** 557–565.
28. Gamble, T.R., F.F. Vajdos, S. Yoo, *et al*. 1996. Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. Cell **87:** 1285–1294.

29. Luo, H., Q. Chen, J. Chen, *et al*. 2005. The nucleocapsid protein of SARS coronavirus has a high binding affinity to the human cellular heterogeneous nuclear ribonucleoprotein A1. FEBS Lett. **579:** 2623–2628.
30. Xu, R.M., L. Jokhan, X. Cheng, *et al*. 1997. Crystal structure of human UP1, the domain of hnRNP A1that contains two RNA-recognition motifs. Structure **5:** 559–570.
31. Wang, Y. & X. Zhang. 1999. The nucleocapsid protein of coronavirus mouse hepatitis virus interacts with the cellular heterogeneous nuclear ribonucleoprotein A1 *in vitro* and *in vivo*. Virology **265:** 96–109.
32. He, R., A. Leeson, A. Andonov, *et al*. 2003. Activation of AP-1 signal transduction pathway by SARS coronavirus nucleocapsid protein. Biochem. Biophys. Res. Commun. **311:** 870–876.
33. Surjit, M., B. Liu, V.T. Chow & S.K. Lal. 2006. The nucleocapsid protein of sars-coronavirus inhibits the activity of cyclin-CDK complex and blocks S phase progression in mammalian cells. J. Biol. Chem. **281:** 10669–10681.
34. Surjit, M., B. Liu, S. Jameel, *et al*. 2004. The SARS coronavirus nucleocapsid protein induces actin reorganization and apoptosis in COS-1 cells in the absence of growth factors. Biochem. J. **383:** 1–6.
35. Timani, K.A., Q. Liao, L. Ye, *et al*. 2005. Nuclear/nucleolar localization properties of C-terminal nucleocapsid protein of SARS coronavirus. Virus Res. **114:** 23–34.
36. Surjit, M., R. Kumar, R.N. Mishra, *et al*. 2005. The severe acute respiratory syndrome coronavirus nucleocapsid protein is phosphorylated and localizes in the cytoplasm by 14-3-3-mediated translocation. J. Virol. **79:** 11476–11486.
37. Wootton, S.K., R.R. Rowland & D. Yoo. 2002. Phosphorylation of the porcine reproductive and respiratory syndrome virus nucleocapsid protein. J. Virol. **76:** 10569–10576.
38. Rowland, R.R., V. Chauhan, Y. Fang, *et al*. 2005. Intracellular localization of the severe acute respiratory syndrome coronavirus nucleocapsid protein: absence of nucleolar accumulation during infection and after expression as a recombinant protein in vero cells. J. Virol. **79:** 11507–11512.
39. Fang, X., L. Ye, K.A. Timani, *et al*. 2005. Peptide domain involved in the interaction between membrane protein and nucleocapsid protein of SARS-associated coronavirus. J. Biochem. Mol. Biol. **38:** 381–385.
40. Voss, D., A. Kern, E. Traggiai, *et al*. 2006. Characterization of severe acute respiratory syndrome coronavirus membrane protein. FEBS Lett. **580:** 968–973.
41. Nal, B., C. Chan, F. Kien, *et al*. 2005. Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. J. Gen. Virol. **86:** 1423–1434.
42. Liao, Y., Q. Yuan, J. Torres, *et al*. 2006. Biochemical and functional characterization of the membrane association and membrane permeabilizing activity of the severe acute respiratory syndrome coronavirus envelope protein. Virology **349:** 264–275.
43. Law, P.T., C.H. Wong, T.C. Au, *et al*. 2005. The 3a protein of severe acute respiratory syndrome-associated coronavirus induces apoptosis in Vero E6 cells. J. Gen. Virol. **86:** 1921–1930.
44. Tan, Y.-J. 2005. The severe acute respiratory syndrome (SARS) coronavirus 3a protein may function as a modulator of the trafficking properties of the spike protein. Virol. J. **2:** 5.

45. TAN, Y.-J., P.-Y. THAM, D.Z.L. CHAN, *et al*. 2005. The severe acute respiratory syndrome coronavirus 3a protein up-regulates expression of fibrinogen in lung epithelial cells. J. Virol. **79:** 10083–10087.
46. LAW, P.T.W., C.-H. WONG, T.C.C. AU, *et al*. 2005. The 3a protein of severe acute respiratory syndrome-associated coronavirus induces apoptosis in Vero E6 cells. J. Gen. Virol. **86:** 1921–1930.
47. TAN, Y.-J., B.C. FIELDING, P.-Y. GOH, *et al*. 2004. Overexpression of 7a, a protein specifically encoded by severe acute respiratory syndrome coronavirus, induces apoptosis via a caspase-dependent pathway. J. Virol. **78:** 14043–14047.
48. LU, W., B.-J. ZHENG, K. XU, *et al*. 2006. Severe acute respiratory syndrome-associated coronavirus 3a protein forms an ion channel and modulates virus release. Proc. Natl. Acad. Sci. USA **103:** 12540–12545.
49. LUO, H., J. CHEN, K. CHEN, *et al*. 2006. Carboxyl terminus of severe acute respiratory coronavirus nucleocapsid protein: self-association analysis and nucleic acid binding characteristics. Biochemistry **45:** 11827–11835.
50. FANG, X., L.-B. YE, Y. ZHANG, *et al*. 2006. Nucleocapsid amino acids 211 to 254, in particular, tetrad glutamines, are essential for the interaction between the nucleocapsid and membrane proteins of SARS-associated coronavirus. J. Microbiol. **44:** 577–580.
51. ZUNIGA, S., I. SOLA, J.L. MORENO, *et al*. 2007. Coronavirus nucleocapsid protein is an RNA chaperone. Virology **357:** 215–227.