Understanding the accessory viral proteins unique to the severe acute respiratory syndrome (SARS) coronavirus

Yee-Joo Tan*, Seng Gee Lim, Wanjin Hong

Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673, Singapore

Received 24 March 2006; accepted 15 May 2006

Abstract

A novel coronavirus, termed the severe acute respiratory syndrome coronavirus (SARS-CoV), infected humans in Guangdong, China, in November 2002 and the subsequent efficient human-to-human transmissions of this virus caused profound disturbances in over 30 countries worldwide in 2003. Eventually, this epidemic was controlled by isolation and there has been no human infection reported since January 2004. However, research on different aspects of the SARS-CoV is not waning, as it is not known if this virus will re-emerge, especially since its origins and potential reservoir(s) are unresolved. The SARS-CoV genome is nearly 30 kb in length and contains 14 potential open reading frames (ORFs). Some of these ORFs encode for genes that are homologous to proteins found in all known coronaviruses, namely the replicase genes (ORFs 1a and 1b) and the four structural proteins: nucleocapsid, spike, membrane and envelope, and these proteins are expected to be essential for the replication of the virus. The remaining eight ORFs encodes for accessory proteins, varying in length from 39 to 274 amino acids, which are unique to SARS-CoV. This review will summarize the expeditious research on these accessory viral proteins in three major areas: (i) the detection of antibodies against accessory proteins in the serum of infected patients, (ii) the expression, processing and cellular localization of the accessory proteins, and (iii) the effects of the accessory proteins on cellular functions. These in-depth molecular and biochemical characterizations of the SARS-CoV accessory proteins, which have no homologues in other coronaviruses, may offer clues as to why the SARS-CoV causes such a severe and rapid attack in humans, while other coronaviruses that infect humans seem to be more forgiving.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Severe acute respiratory syndrome (SARS); Coronavirus; Accessory viral proteins; Virus–virus interactions; Virus–host interactions

Contents

1. Introduction ............................................................................................................. 79
2. Detection of antibodies in the serum of infected patients .............................................. 80
3. Expression, processing and cellular localization ............................................................ 80
   3.1. ORFs 3a and 3b .................................................................................................. 81
   3.2. ORF 6 ........................................................................................................... 82
   3.3. ORFs 7a and 7b ............................................................................................... 82
   3.4. ORFs 8a and 8b ............................................................................................... 82
   3.5. ORF 9b ........................................................................................................... 83
4. Effects on cellular functions ......................................................................................... 83
   4.1. ORFs 3a and 3b ............................................................................................... 83
   4.2. ORF 6 ........................................................................................................... 84
   4.3. ORF 7a ........................................................................................................... 84
5. Future directions ...................................................................................................... 85
Acknowledgements ........................................................................................................ 85
References .................................................................................................................. 85

* Corresponding author. Tel.: +65 65869625; fax: +65 67791117.
E-mail address: mcbbtanyj@imcb.a-star.edu.sg (Y.-J. Tan).

0166-3542/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.antiviral.2006.05.010
1. Introduction

As illustrated by the outbreaks of novel viral diseases in recent years, it is clear that novel viruses can emerge in any part of the world at anytime. Since 2004, there have been reports of bird-to-human, and probable human-to-human, transmissions of the H5N1 avian influenza virus and this raised the fear that the virus may be mutating to a form that can cause a flu pandemic with the magnitude of the Spanish flu pandemic of 1918 (Lipatov et al., 2004). Previous to this, a new respiratory viral disease, termed the severe acute respiratory syndrome (SARS), caused profound disturbances worldwide and by the end of the SARS epidemic, more than 8000 cases were reported with 774 fatalities (World Health Organization, http://www.who.int/csr/sars/country/en/).

The concerted efforts of scientists all over the world resulted in the rapid identification and characterization of a novel coronavirus (termed as SARS coronavirus, SARS-CoV) as the etiological agent of SARS (see reviews by Peiris et al., 2003; Osterhaus et al., 2004). Coronaviruses are a family of enveloped, single-stranded, positive-strand RNA viruses with very large genomic size of \( \sim 30 \text{ kb} \) and have been known to infect many animal species as well as humans (Siddell, 1995).

Since the outbreak of SARS, much research efforts have been directed to understanding the SARS-CoV including characterization of the evolution of the virus in the human population, isolation of the virus from wild animals, the development of vaccines and the establishment of animal models suitable for evaluation of pathogenesis and immunity (see recent reviews by Berger et al., 2004; Christian et al., 2004; Donnelly et al., 2004; Peiris et al., 2004; Poon et al., 2004; Osterhaus et al., 2004; Ziebuhr, 2004; Cinatl et al., 2005; Groneberg et al., 2005; Jiang et al., 2005; Tan et al., 2005b; Weiss and Navas-Martin, 2005; Tsunetsugu-Yokota et al., 2006). After the SARS epidemic, numerous novel coronaviruses have been isolated from different animals as well as from patients with pneumonia (Guan et al., 2003; Fouchier et al., 2004; van der Hoek et al., 2004; Jonassen et al., 2005; Kan et al., 2005; Lau et al., 2005; Li et al., 2005; Poon et al., 2005a; Woo et al., 2005; Wise et al., 2006). Some of these showed very high sequence homology to the SARS-CoV isolated from humans (Guan et al., 2003; Kan et al., 2005; Lau et al., 2005; Li et al., 2005), suggesting that they may be reservoirs for the virus. Epidemiological studies also revealed that it is likely that the SARS-CoV was transmitted from an animal source to the human population and this has happened at least twice, once in Guangdong (China) in November 2002, leading to the outbreak in the human population, and a second time in Guangzhou (China) in December 2003 to January 2004 (Chinese SARS Consortium, 2004; Song et al., 2005; Wang et al., 2005). The latter refers to the infection of four people who did not have any contact history with previously documented SARS cases, and no subsequent human-to-human transmission was reported (Liang et al., 2004). Sequence analysis of viruses isolated from these patients showed that they were not derived from the preceding epidemic in 2003 but rather suggested that these cases represented new zoonotic transmissions (Song et al., 2005; Wang et al., 2005). These findings indicate that the SARS-CoV is of animal origin and that its precursor is still present in animal populations, and underscore the potential for the re-emergence of SARS and the need for continuing research on the virus.

This review summarizes present knowledge on a group of SARS-CoV proteins that show little amino acids (aa) homology with proteins of other coronaviruses, i.e. the so-called group-specific accessory proteins. The SARS-CoV genome has an organization, which is typical of other members of the Coronaviridae family (Fig. 1) (Marra et al., 2003; Rota et al., 2003). The first 2/3 of the SARS-CoV genome encodes the replicase genes (ORFs 1a and 1b), which translates into two large

---

**Table 1**

<table>
<thead>
<tr>
<th>Viral gene</th>
<th>Alternate names</th>
<th>Genome position (bp)</th>
<th>No. of amino acids</th>
<th>Cellular localization of viral protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>3a</td>
<td>ORF 3, X, U274</td>
<td>25252–26076</td>
<td>274</td>
<td>Golgi apparatus and cell surface (Tan et al., 2004c; Ito et al., 2005)</td>
</tr>
<tr>
<td>3b</td>
<td>ORF 4, X2</td>
<td>25673–26137</td>
<td>154</td>
<td>Nucleus (Yuan et al., 2005c)</td>
</tr>
<tr>
<td>6</td>
<td>ORF 7, X3</td>
<td>27058–27249</td>
<td>63</td>
<td>Throughout cytoplasm and somewhat concentrated in the endoplasmic reticulum and Golgi apparatus (Geng et al., 2005; Pewe et al., 2005)</td>
</tr>
<tr>
<td>7a</td>
<td>ORF 8, X4, U122</td>
<td>27257–27625</td>
<td>122</td>
<td>Intermediate compartments between the endoplasmic reticulum and Golgi apparatus (Fielding et al., 2004; Nelson et al., 2005)</td>
</tr>
<tr>
<td>7b</td>
<td>ORF 9</td>
<td>27622–27756</td>
<td>44</td>
<td>N.D.*</td>
</tr>
<tr>
<td>7a</td>
<td>ORF 10</td>
<td>27763–27882</td>
<td>39</td>
<td>Punctuate vesicle-like structures throughout the cytoplasm (unpublished data)</td>
</tr>
<tr>
<td>8b</td>
<td>ORF 11, X5</td>
<td>27848–28102</td>
<td>84</td>
<td>Punctuate vesicle-like structures throughout the cytoplasm (unpublished data)</td>
</tr>
<tr>
<td>9b</td>
<td>ORF 13</td>
<td>28114–28410</td>
<td>98</td>
<td>N.D.**</td>
</tr>
</tbody>
</table>

* This article adopts the nomenclatures used by Snijder et al. (2003) and Thiel et al. (2003).

** Alternative names that have been used in published articles.

---

---

---

---
polyproteins, pp1a (486 kDa) and pp1ab (790 kDa) (Thiel et al., 2003). Proteolytic processing of these polyproteins, which are mediated by viral cysteine proteinases, produce non-structural proteins, some of which are responsible for replicating the viral genome and/or generating a nested set of subgenomic mRNAs to express all the other ORFs in the genome (Ziebuhr, 2004). The ORFs for the structural proteins, spike (S), envelope (E), membrane (M) and nucleocapsid (N), are encoded in the remaining portion of the genome and interspaced between these are the ORFs for a number of putative accessory proteins. Each coronavirus encodes different number of accessory proteins and the predicted sequences of these proteins do not share high level of homology although there may be some degree of conservation within each subgroup of coronaviruses (Brown and Brierly, 1995). The SARS-CoV genome encodes eight putative accessory proteins (i.e. ORFs 3a, 3b, 6, 7a, 7b, 8a, 8b and 9b) varying in length from 39 to 274aa (Table 1). While the SARS-CoV replicate genes and structural proteins share some degree of sequence homology with those of other coronaviruses, the accessory proteins do not show significant homology, at the amino acids level, to viral proteins of known coronaviruses. However, emerging evidence suggests that some of SARS-CoV accessory proteins, like 3a, 3b and 9b, may share some structural and biological features with other coronaviral proteins (see Section 3).

2. Detection of antibodies in the serum of infected patients

During viral infection, the host immune system is stimulated to provide host defence mechanisms against the invading virus, and these responses, which are of paramount importance for the outcome of the infection, can be classified into innate immunity, cell-mediated immunity and humoral responses (Hilleman, 2004). Our current understanding on the host immune responses during SARS-CoV infection and their roles in disease progression was recently reviewed (Lo et al., 2006; Tsunetsugu-Yokota et al., 2006) and will not be discussed in detail. Antibodies against all of the SARS-CoV accessory proteins were detected in the sera of SARS patients, suggesting that these proteins were expressed during infection in vivo (Guan et al., 2004a,b; Guo et al., 2004; Tan et al., 2004b; Yu et al., 2004; Zeng et al., 2004; Chan et al., 2005; Qiu et al., 2005; Chow et al., 2006). However, for some of them, antibodies were found in only a few cases, suggesting that they may be expressed only at low levels or that they are not highly immunogenic.

So far, the antibodies against 3a have been found in the most cohorts of SARS patients, suggesting that it is one of the most abundant or immunogenic accessory protein. Another reason could be that 3a is a novel coronavirus structural protein (Ito et al., 2005; Shen et al., 2005), which is also expressed on the cell surface of infected cells (Tan et al., 2004c; Ito et al., 2005). The topology of 3a on the cell surface was determined experimentally: its first 34aa, i.e. before the first transmembrane domain, is facing the extracellular matrix and its C-terminus after the third transmembrane domain (i.e. 134–274aa) is facing the cytoplasm (Tan et al., 2004c). Accordingly, the N-terminal ectodomain would be expected to protrude out of the virion. Interestingly, in two separate cohorts of SARS patients, one from Taiwan (Liu et al., 2004) and one from Hong Kong (Zhong et al., 2005), B cells recognizing the N-terminal region of 3a were isolated from patients. In another study, it was reported that 48.8% of patients who recovered from SARS had antibodies against the N terminus of 3a while only 7.4% of the diseased patients has such antibodies (Zhong et al., 2006). Thus far, anti-3a antibody has not been demonstrated to inhibit virus infection. However, Zhong et al. demonstrated that the 3a N-terminal ectodomain elicits strong antibody response and that anti-3a antibodies could bind cells expressing 3a and induce the elimination of these cells in the presence of the human complement system (Zhong et al., 2006). Taken together, these data show that 3a is presented to the host immune system during infection and that it may stimulate protective humoral responses in infected patients.

3. Expression, processing and cellular localization

Like other coronaviruses, SARS-CoV uses a transcription attenuation mechanism to synthesize both full-length and subgenomic-length negative-strand RNAs which then function as templates for synthesis of full-length genomic mRNA and subgenomic mRNAs (sgRNAs) (Snijder et al., 2003; Thiel et al., 2003; Yount et al., 2003). The first ORF is translated from the full-length genomic mRNA, while the remaining ORFs are translated from eight sgRNAs synthesized as a nested set of 3′ coterminal RNA species in which the leader RNA sequences on the 5′ end of the genome are joined to the body sequences at distinct transcription regulatory sequences containing a highly conserved consensus sequence. Consequently, a total of nine mRNA species of different length were detected in SARS-CoV-infected cells and the SARS-CoV accessory proteins are encoded by sgRNAs 3, 6, 7, 8 and 9. While sgRNA 6 is monocristonic, sgRNA 3, 7, 8 and 9 are bicristonic with the second ORFs (i.e. 3b, 7b, 8b
and 9b) being expressed via an internal ribosomal entry mechanism or by a leaky ribosomal scanning mode of translation. In order to determine if these accessory proteins are expressed during SARS-CoV infection, various laboratories have expressed recombinant forms of the individual proteins and raised specific antibodies that can be used to study the expression, processing and cellular localization of the accessory proteins. The results are summarized in this section.

3.1. ORFs 3a and 3b

The second largest subgenomic RNA of SARS-CoV (sgRNA3) contains two ORFs, 3a and 3b. The 3a protein has been characterized to a great extent and has been shown to be expressed in infected cells both in vitro and in vivo (Tan et al., 2004c; Yu et al., 2004; Zeng et al., 2004; Chan et al., 2005; Ito et al., 2005; Law et al., 2005). It is localized in the Golgi apparatus as well as on the cell surface and has been shown to be an integral membrane protein (Tan et al., 2004c; Yuan et al., 2005a; Ito et al., 2005). The 3a protein is efficiently transported to the cell surface (Tan et al., 2004c). In the cytoplasmic domain of 3a, there is a region that consists of the juxtaposition of two sorting motifs, a YxxΦ (where x is any aa and Φ is an aa with a bulky hydrophobic sidechain) and a ExD (diacidic) motif. When this domain is deleted, the 3a protein is retained intracellularly and cannot go to the cell surface (Tan et al., 2004c). The 3a protein on the cell surface can undergo internalization and this may be important for modulating expression of S on the cell surface (Tan, 2005).

3a is a novel coronavirus structural protein as it is associated with virions purified from SARS-CoV infected cells and it is incorporated into viral-like particles when co-expressed with M and E in the baculovirus system (Ito et al., 2005; Shen et al., 2005). Interestingly, multiple forms of 3a have been detected in both transfected and infected cells as well as in the purified virion (Tan et al., 2004c; Ito et al., 2005; Huang et al., 2006; Oostra et al., 2006). Some of these arose from nucleotide insertions at a 6T stretch near to the initiation codon of the 3a gene while others appeared to be the result of post-translational modifications. At least two forms of 3a proteins can be easily detected in infected cells or cells transfected with a cDNA expressing 3a. There is a major form of ∼37 kDa (termed as 3a-1) and a minor form of ∼31 kDa (termed as 3a-2) and both forms can be secreted out of the cells into the culture supernatant (Huang et al., 2006). In infected cells, the buoyant density of 3a-1 is close to that of the SARS virion, suggesting that it has been incorporated into virions. In contrast, the buoyant density of the 3a-2 is lower indicating that it is released from the cells as extracellular membranous structures. The deletion of the sorting motifs from the cytoplasmic domain of 3a, which eliminates its ability to go to the cell surface, also prevented 3a from being secreted. The nature of these structures and the biological significance for the secretion of 3a remain to be determined. Interestingly, when SARS-CoV N protein is expressed alone in mammalian cells, it can also be transported out of the cells (Goh et al., 2004). Several groups also reported that N protein can be detected in patients’ sera, indicating that the extracellular export of N protein may occur in vivo (Che et al., 2004; Lau et al., 2004).

In a detailed study, Oostra et al. showed that the unglycosylated 3a protein has a molecular weight of ∼31 kDa and is likely to correspond to the 3a-2 protein described above (Oostra et al., 2006). This 31 kDa form is then O-linked glycosylated to give a higher molecular weight form that is likely to correspond to the 3a-1 protein described above. They also noted the striking similarity between 3a and coronavirus M proteins; they are triple-spanning membrane glycosylated proteins with short N-terminal ectodomains and large C-terminal endodomains and they are found mainly in the Golgi compartment. Despite little sequence homology, the hydrophobicity profiles of 3a and M are also highly similar, suggesting they may share some structural and biological features. Interestingly, the accessory proteins ORF 3 of another human coronavirus, HCoV-NL63, and ORF 3c of feline coronavirus (FeCoV) are also triple membrane proteins and they also have similar hydrophobicity profiles as 3a and M (Oostra et al., 2006).

Interestingly, the 3a protein interacts with the S protein intracellularly (Tan et al., 2004c) and can detected in disulfide-linked complexes with S in the medium of SARS-CoV infected cells, indicating that 3a could be secreted together with S, possibly through the formation of virus particles (Zeng et al., 2004). By comparing the sequences of the SARS-CoV isolated from different infected individuals and/or palm civets, it was found that the rates of non-synonymous changes are greater than the rates of synonymous changes in the S and 3a genes, which indicated that both S and 3a showed positive selections during virus evolution (Yeh et al., 2004; Song et al., 2005). Hence, changes in S and 3a proteins allowed the virus to improve adaptation to its human host, implying that these proteins play important roles in the virus life cycle and/or disease development. Schwegmann-Wessels et al. (2004) reported that a novel sorting signal for intracellular localization is present in the S protein of most coronaviruses, like TGEV, but absent from SARS-CoV S. Site-directed mutagenesis studies confirmed that a YxxΦ motif retains the S protein of TGEV intracellularly when it is expressed alone, while the SARS-CoV S is transported efficiently to the cell surface unless such a motif is introduced into its cytoplasmic tail by mutagenesis. Given that the 3a protein expressed on the cell surface can undergo internalization (Tan et al., 2004c), it may be postulated that 3a can modulate the expression of S on the cell surface via its interaction with the latter (Tan, 2005). However, Yount et al. (2005) did not observe a significant change in the surface expression of S when the 3a gene was deleted from a full-length infectious clone. Alternatively, a novel dibasic signal present in group 1 coronavirus and SARS-CoV S proteins may be sufficient to retain S intracellularly (Lontok et al., 2004). Further studies are needed to determine if there is any cross-talking between S and 3a. The 3a protein has also been shown to interact with the other structural proteins, M and E (Tan et al., 2004c), but whether these interactions are important for the function of 3a or for the incorporation of 3a into virions have not been determined.

Mutations in the 3a gene in a Singapore isolate were observed after continuous passaging of the virus in Vero E6 cells and as
a result, several different forms of 3a protein were expressed in infected cells (Tan et al., 2004c). When viral RNA isolated directly from clinical samples were examined, similar mutations were found suggesting that the mutation in 3a is not due to cell culture adaptation, but indicate the presence of quasi-species (Tan et al., 2005a). A recent study showed that if the 3a gene contained frame-shift mutations, full-length 3a protein can still be expressed but at much lower level (Wang et al., 2006). As some of the clinical samples only contained viruses with frame-shift mutations in the 3a gene, it may predict that little or no full-length 3a protein will be expressed and this may explain why some patients do not have anti-3a antibodies (Guan et al., 2004a,b; Tan et al., 2004c; Zeng et al., 2004; Zhong et al., 2006).

The 3b protein has been detected in SARS-CoV infected cells (Chan et al., 2005), and anti-3b antibody has been detected in a SARS patient’s serum (Guo et al., 2004). The 3b protein contains two predicted nuclear localization signals and has been shown to localize to the nucleolus of transfected cells in the absence of any other SARS-CoV proteins (Yuan et al., 2005c). The 3b protein of another coronavirus, infectious bronchitis virus (IBV), is also localized in the nucleus and it was shown that the continuous passage of IBV in Vero E6 cells resulted in a truncation of the 3b gene and, consequently, there is increase in a growth advantage of IBV in Vero E6 cells. It was shown that the continuous passage of IBV in Vero E6 cells resulted in a truncation of the 3b gene, and, consequently, there is increase in a growth advantage of IBV in Vero E6 cells (Wang et al., 2006). As some of the clinical samples only contained viruses with frame-shift mutations in the 3a gene, it may predict that little or no full-length 3a protein will be expressed and this may explain why some patients do not have anti-3a antibodies (Guan et al., 2004a,b; Tan et al., 2004c; Zeng et al., 2004; Zhong et al., 2006).

The 3b protein has been detected in SARS-CoV infected cells (Chan et al., 2005), and anti-3b antibody has been detected in a SARS patient’s serum (Guo et al., 2004). The 3b protein contains two predicted nuclear localization signals and has been shown to localize to the nucleolus of transfected cells in the absence of any other SARS-CoV proteins (Yuan et al., 2005c). The 3b protein of another coronavirus, infectious bronchitis virus (IBV), is also localized in the nucleus and it was shown that the continuous passage of IBV in Vero E6 cells resulted in a truncation of the 3b gene and, consequently, there is increase in a growth advantage of IBV in Vero E6 cells (Wang et al., 2006). As some of the clinical samples only contained viruses with frame-shift mutations in the 3a gene, it may predict that little or no full-length 3a protein will be expressed and this may explain why some patients do not have anti-3a antibodies (Guan et al., 2004a,b; Tan et al., 2004c; Zeng et al., 2004; Zhong et al., 2006).

3.2. ORF 6

ORF 6 was detected in SARS-CoV infected cells and was found throughout cytoplasm and somewhat concentrated in the endoplasmic reticulum and Golgi apparatus (Geng et al., 2005; Pewe et al., 2005). The stability of this protein, when it is expressed from a DNA construct via cationic liposomes transient transfection method, seems to be low (Geng et al., 2005), while it is easily detected when it is expressed via a recombinant virus method (Pewe et al., 2005). This suggests that some cellular factors up-regulated during viral infection may be needed to prevent the rapid degradation of ORF 6 in the cell. Further studies to identify such factors will contribute to an understanding on the potential role of ORF 6 in viral replication and pathogenesis (see next section).

3.3. ORFs 7a and 7b

The SARS-CoV 7a protein is encoded by the first ORF of sgRNA7 (Snijder et al., 2003; Thiel et al., 2003; Yount et al., 2003). 7a is expressed in SARS-CoV infected cells, and is found in the intermediate compartments between the endoplasmic reticulum and Golgi apparatus (Fielding et al., 2004; Chen et al., 2005; Nelson et al., 2005), where coronaviruses are known to assemble and bud (Klumperman et al., 1994). It has also been detected in the lung tissues from SARS-CoV infected patients, especially in the bronchial epithelial cells, peripheral blood erythrocytes and leukocytes (Chen et al., 2005).

The crystal and NMR solution structures of the bacterially expressed soluble domain of 7a (i.e. 16–97aa and not including the signal peptide, transmembrane and cytoplasmic domains), have been determined (Nelson et al., 2005; Hanel et al., 2005) and the three-dimensional structures obtained by both methods were highly similar. Interestingly, 16–85aa of 7a forms a compact structure with an Ig-like fold that is distinctive from the typical Ig-superfamily members and may be classified as a subtype of the family (Hanel et al., 2005). Based on the structural comparison to other proteins, Hanel et al. found that the structure of 7a resembles the D1 domain of ICAM-1 or ICAM-2, which are cell adhesion molecules expressed on the surface of cells (Dustin and Springer, 1999). The 85–97aa (also called the “stalk” in Nelson et al., 2005) is highly flexible in the crystalline state as well as in solution, although it is likely that this domain is more structured in the full-length protein as Nelson et al. found that this domain contains some signals that contribute partly for the cellular localization of 7a. Thus far, the 7a protein is the only SARS-CoV accessory protein whose three-dimensional structure has been determined. Because of the structural similarity between 7a and ICAM-1 (see above), it was predicted that 7a may be able to bind the integrin I domain subunit I-domain of LFA-1, but this has not yet been determined experimentally (Hanel et al., 2005).

Thus far, the expression of the accessory protein 7b in infected cells has not been demonstrated although Guo et al. detected antibody against 7b in a SARS patient (Guo et al., 2004). One study reported an in frame deletion of 45 nucleotides in ORF 7b in the Frankfurt isolate after three passages in tissue culture (Thiel et al., 2003), suggesting that 7b may not be important for viral replication in cell-culture. Indeed, Yount et al. showed by reverse genetics that deletions of ORF 7a and 7b did not have a significant effect on viral replication in cell culture or in the murine model (Yount et al., 2005). However, it remains to be determined if 7a or 7b plays a role during SARS-CoV infection of the natural host.

3.4. ORFs 8a and 8b

The expression of 8a and 8b proteins in SARS-CoV-infected Vero E6 cells has been observed (unpublished data). 8a and 8b, which are encoded in subgenomic RNA 8 of SARS-CoV isolated from most of the human isolates, contain 39 and 84 amino acids, respectively. Interestingly, the SARS-CoV isolated from animals contains an extra 29-nucleotide in this region such that the 8a and 8b proteins are fused to become a single protein, 8ab, of 122 amino acids, whose N-terminus is identical to 8a and C-terminus is identical to 8b (Guan et al., 2003). Another extensive study of 63 SARS-CoV isolates obtained from early, middle and late phases of the SARS epidemic in China also showed that there are major deletions in this region of the viral genome (Chinese SARS Consortium, 2004). Interestingly, the clustering of patients with different patterns of deletion was cor-
related with the different phases of the epidemic. Although these mutations in the ORF 8 region do not appear to have any adverse effect on the survival of the virus, it is conceivable that the 8a, 8b and 8ab proteins may have different stabilities and/or functions, and hence would contribute differently to viral replication and/or pathogenesis in vivo. As SARS-CoV appeared to have only recently crossed the barrier species from animal to human and it has been observed that the profiles of these mutations correlated with clusters of infection, it is plausible that the properties of these accessory proteins are changing as the virus undergoes adaptive evolution. However, it is also possible that the unusually high mutation in the ORF 8 region is a result of genomic instability and that the viral proteins encoded by ORF 8 are dispensable for viral replication.

Indeed, it has been demonstrated that the palm civets are equally susceptible to the human SARS-CoV isolate BJ01 from the middle phase (without the extra 29 nucleotides) and the isolate GZ01 (with the extra 29 nucleotides like in the animal SARS-CoV isolates) from the early phase (Wu et al., 2005). Using reverse genetic methods, Yount et al. also reported similar findings in the murine model (Yount et al., 2005). A frame-shift mutation in the 8a gene was also reported after the virus was isolated in Vero E6 or FRhK4 cells but not in the original clinical samples, suggesting that this mutation was acquired as a result of cell culture adaptation (Poon et al., 2005b).

3.5. ORF 9b

The expression of the 9b protein in infected cells has been reported (Chan et al., 2005) and anti-9b antibodies were detected in the sera of different cohorts of SARS patients (Qiu et al., 2005; Chow et al., 2006). Interestingly, 9b is encoded by an internal ORF within the N gene and this has also been reported previously for other coronaviruses, like the mouse hepatitis virus and bovine coronavirus (Fischer et al., 1997; Senanayake and Brian, 1997).

4. Effects on cellular functions

From the studies of other coronaviruses, accessory proteins are usually dispensable for viral replication in cell culture systems but may be important for virus–host interactions and thus contribute to viral stability and/or pathogenesis in vivo. For example, although the 7b gene of feline coronavirus is easily lost upon virus adaptation to cell culture, it is strictly maintained in naturally occurring strains and its loss was correlated with reduced virulence (Herrewegh et al., 1995). Recent studies also showed that some of these group-specific genes are not essential for viral replication in cell culture, but their deletion, by reverse genetics, attenuates the virus in the natural host (de Haan et al., 2002; Ortego et al., 2003; Hajjema et al., 2004).

A study on the role of the accessory proteins for viral replication in cell culture and murine model was recently published (Yount et al., 2005). In this study, reverse genetics experiments showed that the five of the accessory proteins, 3a, 3b, 6, 7a and 7b, are non-essential for SARS-CoV replication because when these ORFs were systematically deleted, the recombinant virus still replicates with similar kinetics as the wild-type. However, SARS-CoV replication in mice is self-limiting and the animals do not show the clinical symptoms and pathology observed in SARS patients (Subbarao et al., 2004; Wentworth et al., 2004). Hence it is possible that the effects of these accessory proteins on viral pathogenesis in the natural host are not presented in the mouse model. Some insights on the contribution of the SARS-CoV accessory proteins to viral pathogenesis may be gained from studying their abilities to interfere with cellular machineries or interact with cellular proteins. Thus far, the effects of the over-expression of 3a, 3b, 6 and 7a on cellular homeostasis have been described and are summarized in this section (Table 2).

4.1. ORFs 3a and 3b

When the gene expression profile of A549 (lung epithelial) cells that stably expressed 3a was analyzed, it was found that the mRNA levels of all three subunits, Aα, Bβ, and γ, of fibrinogen were highly up-regulated (Tan et al., 2005c). Consequently, the intracellular levels as well as the secretion of fibrinogen were increased. The level of fibrinogen was also increased in infected Vero E6 cells, which is consistent with a recent report that fibrinogen mRNA was up-regulated in SARS-CoV infected human peripheral blood mononuclear cells (PBMCs) (Ng et al., 2004), indicating that the excessive production of fibrinogen during

<table>
<thead>
<tr>
<th>Proteins</th>
<th>Characteristics of protein and its effects on cellular functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>3a</td>
<td>Undergoes efficient endocytosis (Tan et al., 2004c)</td>
</tr>
<tr>
<td></td>
<td>Interacts with structural proteins, spike, envelope and membrane, and accessory protein 7a (Tan et al., 2004c; Zeng et al., 2004)</td>
</tr>
<tr>
<td></td>
<td>Is a structural protein (Shen et al., 2005; Ito et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>Is secreted out of the cell (Huang et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>Is O-linked glycosylated (Oostra et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>Up-regulates the expression of fibrinogen in lung cells (Tan et al., 2005b)</td>
</tr>
<tr>
<td></td>
<td>Induces apoptosis via a caspase-8 dependent pathway (Law et al., 2005)</td>
</tr>
<tr>
<td>3b</td>
<td>Induce cell cycle arrest at the G0/G1 phase and apoptosis (Yuan et al., 2005b)</td>
</tr>
<tr>
<td>6</td>
<td>Enhance the virulence of an attenuated murine coronavirus (Pewe et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>Enhanced DNA synthesis (Geng et al., 2005)</td>
</tr>
<tr>
<td>7a</td>
<td>Has an Ig-like fold and showed structural similarity to D1 domain of ICAM-1 (Nelson et al., 2005; Hanel et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>Induces apoptosis (Tan et al., 2004a; Kopecky-Bromberg et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>Inhibits cellular protein synthesis (Kopecky-Bromberg et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>Induces the phosphorylation and activation of p38 mitogen-activated protein kinase (Kopecky-Bromberg et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>Blocks cell cycle progression at G0/G1 phase (Yuan et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>Interacts with host protein SGT and viral structural proteins, envelope and membrane (Fielding et al., 2006)</td>
</tr>
</tbody>
</table>
a SARS-CoV infection may have important consequences for pathogenesis.

During an acute-phase response to infection, injury, or neoplasia, the production of fibrinogen by the liver increases to restore homeostasis (Bini et al., 2000). Increased production of fibrinogen at ex-hepatic tissues, like the lung (Rybarczyk et al., 2003; Lawrence and Simpson-Haidaris, 2004), also helps in this process. However, the excessive production of fibrinogen and formation of fibrin at the site of injury may enhance cytokine production or imbalance procoagulant and/or fibrinolytic activities (Ware and Matthy, 2000; Idell, 2002). Post-mortem examinations of patients who died of SARS-CoV infection revealed extensive lung damages that are typical of acute respiratory distress syndrome (Franks et al., 2003; Nicholls et al., 2003; Hwang et al., 2005). In addition, most SARS patients have abnormal hematological manifestations, like thrombocytopenia, elevated D-dimers and prolonged activated partial-thromboplastin time and this suggests the dysregulation of coagulation and fibrin polymerization pathways during SARS-CoV infection (for reviews, see Peiris et al., 2003; Hui et al., 2004; Vijayanand et al., 2004). Taken together, the laboratory features, lung pathological findings and in vitro demonstration of fibrinogen up-regulation in both PBMC and Vero E6 cells by SARS-CoV infection, indicate that increased fibrinogen expression and fibrin degradation products may contribute to disease progression in SARS patients. The expression of 3a alone can up-regulate the expression of fibrinogen in lung cells suggesting that 3a may contribute to SARS pathogenesis.

The over-expression of 3a in Vero E6 cells can also trigger apoptosis via a caspase-8 dependent pathway (Law et al., 2005). A similar phenomenon was observed in the Drosophila if 3a is mis-expressed (Wong et al., 2005). In addition, the 3a transgenic Drosophila showed a dominant rough eye phenotype that can be suppressed by genes that mediate clathrin-dependent endocytosis. Hence, the effect of 3a on the eye in Drosophila appeared to be related to the endocytosis properties of 3a (Tan et al., 2004c). In future studies, it will be crucial to determine how the different properties of 3a allow it to influence viral replication or pathogenesis. In the 3a transgenic Drosophila as well as in the A549 cell-line stably expressing 3a (see above), a number of cellular genes have been significantly affected by the over-expression of 3a and future investigations will reveal what are the impact of these cellular changes on viral replication and/or pathogenesis.

Much less information is available for 3b, which is the second ORF in subgenomic RNA3, but it has been demonstrated that the over-expression of 3b can induce cell cycle arrest at the G0/G1 phase and apoptosis (Yuan et al., 2005b). However, it was not determined if apoptosis was occurring via a caspase-dependent pathway and if it is a consequence of the cell cycle arrest. The 3b protein is likely to be expressed via an internal ribosomal entry mechanism since its translation initiation codon is not the first AUG in sgRNA3 (Snijder et al., 2003). Interestingly, Hussen et al. reported that the 3b protein can also be expressed from an independent subgenomic RNA but this is a rare event (Hussen et al., 2005). In future studies, it will be crucial to determine the temporal expression of 3b during the viral replication cycle.

### 4.2. ORF 6

Interestingly, a recent study showed that the SARS-CoV ORF 6 protein can enhance the virulence of an attenuated murine coronavirus (MHV) (Pewe et al., 2005). The insertion of the ORF 6 gene into the MHV genome enhanced the virus growth in cell culture and increased the lethality of infection in mice. Although the importance of ORF 6 in the context of SARS-CoV remains to be determined, these latest findings suggested that ORF 6 may have a direct role in enhancing virus replication or assembly. In addition, the over-expression of ORF 6 also enhanced cellular DNA synthesis (Geng et al., 2005).

### 4.3. ORF 7a

Over-expression of 7a induces apoptosis via a caspase-dependent pathway, and in cell-lines derived from different organs, including lung, kidney and liver (Tan et al., 2004a; Kopecky-Bromberg et al., 2006). However, several lines of evidence showed that 7a is not the only apoptosis-inducing factor during SARS-CoV infection in Vero E6 cells. Firstly, the level of apoptosis in Vero E6 cells transfected with 7a cDNA was significantly lower than that for SARS-CoV-infected Vero E6 cells (Tan et al., 2004a). Secondly, other SARS-CoV proteins have been shown to induce apoptosis (Surjit et al., 2004; Chow et al., 2005; Law et al., 2005; Yang et al., 2005; Yuan et al., 2005b). Thirdly, deletion of 7a and 7b by reverse genetic methods did not abolish apoptosis in infected cells (Yount et al., 2005). More studies are needed to delineate the precise contributions of the different viral proteins to the cytopathic effects of SARS-CoV infection.

In addition, the 7a protein can also block cell cycle progression at the G0/G1 phase by reducing the expression levels of cyclin D3 and phosphorylation of retinoblastoma protein (Yuan et al., 2006). It was also shown that the over-expression of 7a inhibits cellular protein synthesis and induces the phosphorylation and activation of p38 mitogen-activated protein kinase (MAPK) (Kopecky-Bromberg et al., 2006). However, inhibition of p38 phosphorylation did not prevent the inhibition of translation or the induction of apoptosis by 7a. Different laboratories independently reported that the over-expression of 7a protein decreases cell proliferation, but there is currently no consensus as to the mechanism (Tan et al., 2004a; Chen et al., 2005; Kopecky-Bromberg et al., 2006; Yuan et al., 2006). As these studies were performed in different laboratories, it is unclear if these events, i.e. apoptosis, cell-cycle arrest, p38 MAPK activation, translation inhibition, occur independently, simultaneously or sequentially. Moreover, as these effects of 7a on cellular machineries were observed when 7a was over-expressed, their contributions during SARS-CoV infection remain unclear as there are currently no detailed studies on the expression levels of 7a and its temporal expression during the viral replication cycle. In addition, it has been observed that 7a can interact with 3a, which can interact with M and E, hence suggesting that these viral proteins may form complexes in SARS-CoV infected cells (Tan et al., 2004c).

A recent study also reported that 7a interacts with a cellular protein called the small glutamine-rich tetratricopeptide repeat
(TPR)-containing protein (SGT) (Fielding et al., 2006). TPR motif-containing proteins have been shown to be involved in many cellular processes including cell cycle control, transcription and splicing events, protein transport and protein folding (Blatch and Lassle, 1999). The SGT protein was first identified as a binding partner for the non-structural protein, NS1, of Blatch, G.L., Lassle, M., 1999. The tetratricopeptide repeat: a structural motif-containing proteins have been shown to be involved in intricate interplay between the SARS-CoV and its human host and may contribute to the development of antiviral therapeutics and design of efficient strategies for disease control.

5. Future directions

Reverse genetics experiments showed that the accessory proteins, 3a, 3b, 6, 7a and 7b, are non-essential for SARS-CoV replication in cell culture and in the murine model (Yount et al., 2005). This is not surprising because for many animal coronaviruses, the group-specific accessory proteins have been shown to be dispensable for viral replication, but they can contribute to viral stability and pathogenesis in the natural hosts (see recent review by Weiss and Navas-Martin, 2005). However, some subtle effects were observed suggesting that the accessory proteins may have modulating effects on viral replication (Yount et al., 2005). For example, deletion of 3a resulted in a 0.5–1 log reduction in virus yield in cell culture, while the deletion of 3a, 3b and ORF 6 together resulted in a 1–1.5 log decrease.

In order to determine if the in vitro properties of the accessory proteins that have been characterized so far are important for SARS-CoV replication and/or pathogenesis, it will be crucial to carry out more studies using non-human primate models, where the animals could develop a disease comparable to that of SARS in patients (Fouchier et al., 2003). Further investigations on the interaction between accessory proteins and other viral proteins as well as cellular factors will lead to a better understanding of the intricate interplay between the SARS-CoV and its human host and may contribute to the development of antiviral therapeutics and design of efficient strategies for disease control.

Acknowledgements

We apologize to any investigators whose work has been inadvertently omitted. This study was supported by grants from the Agency for Science, Technology and Research (A*STAR), Singapore.

References


