Hypothesis

Open Access

The Severe Acute Respiratory Syndrome (SARS)-coronavirus 3a protein may function as a modulator of the trafficking properties of the spike protein

Yee-Joo Tan*

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

Email: Yee-Joo Tan* - mcbtanyj@imcb.a-star.edu.sg

* Corresponding author

Published: 10 February 2005

Virology Journal 2005, 2:5 doi:10.1186/1743-422X-2-5

This article is available from: http://www.virologyj.com/content/2/1/5

Received: 17 January 2005 Accepted: 10 February 2005

© 2005 Tan; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<u>http://creativecommons.org/licenses/by/2.0</u>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: A recent publication reported that a tyrosine-dependent sorting signal, present in cytoplasmic tail of the spike protein of most coronaviruses, mediates the intracellular retention of the spike protein. This motif is missing from the spike protein of the severe acute respiratory syndrome-coronavirus (SARS-CoV), resulting in high level of surface expression of the spike protein when it is expressed on its own *in vitro*.

Presentation of the hypothesis: It has been shown that the severe acute respiratory syndromecoronavirus genome contains open reading frames that encode for proteins with no homologue in other coronaviruses. One of them is the 3a protein, which is expressed during infection *in vitro* and *in vivo*. The 3a protein, which contains a tyrosine-dependent sorting signal in its cytoplasmic domain, is expressed on the cell surface and can undergo internalization. In addition, 3a can bind to the spike protein and through this interaction, it may be able to cause the spike protein to become internalized, resulting in a decrease in its surface expression.

Testing the hypothesis: The effects of 3a on the internalization of cell surface spike protein can be examined biochemically and the significance of the interplay between these two viral proteins during viral infection can be studied using reverse genetics methodology.

Implication of the hypothesis: If this hypothesis is proven, it will indicate that the severe acute respiratory syndrome-coronavirus modulates the surface expression of the spike protein via a different mechanism from other coronaviruses. The interaction between 3a and S, which are expressed from separate subgenomic RNA, would be important for controlling the trafficking properties of S. The cell surface expression of S in infected cells significantly impacts viral assembly, viral spread and viral pathogenesis. Modulation by this unique pathway could confer certain advantages during the replication of the severe acute respiratory syndrome-coronavirus.

Background

The recent severe acute respiratory syndrome (SARS) epidemic, which affected over 30 countries, resulted in more than 8000 cases of infection and more than 800 fatalities (World Health Organization, <u>http://www.who.int/csr/</u> <u>sars/country/en/</u>). A novel coronavirus was identified as the aetiological agent of SARS [1]. Analysis of the nucleotide sequence of this novel SARS coronavirus (SARS- CoV) showed that the viral genome is nearly 30 kb in length and contains 14 potential open reading frames (ORFs) [2-4]. These viral proteins can be broadly classified into 3 groups; (i) the replicase 1a/1b gene products which are important for viral replication, (ii) the structural proteins, spike (S), nucleocapsid (N), membrane (M) and envelope (E), which have homologues in all known coronaviruses, and are important for viral assembly, and (iii) the "accessory" proteins that are specifically encoded by SARS-CoV. Much progress have been made in characterizing these SARS-CoV proteins [5,6], but the molecular determinant for the severe clinical manifestations of SARS-CoV infection in contrast to the mild diseases caused by most coronaviruses, remains to be determined. In addition, the exact roles of "accessory" proteins of SARS-CoV are still poorly understood.

The subject of this hypothesis relate to the S protein and one of the "accessory" proteins, the SARS-CoV 3a protein. The S protein, which forms morphologically characteristic projections on the virion surface, mediates binding to cellular receptor and the fusion of viral and host membranes, both of these processes being critical for virus entry into host cells [7,8]. As such, S is known to be responsible for inducing host immune responses and virus neutralization by antibodies [9,10]. 3a (also termed ORF3 in [2] and [11], as X1 in [3], and as U274 in [12,13]) is the largest "accessory" protein of SARS-CoV, consisting of 274 amino acids and 3 putative transmembrane domains. Three groups independently reported the expression of 3a in SARS-CoV infected cells [13-15] and it was also detected in a SARS-CoV infected patient's lung specimen [14]. Antibodies against 3a were also found in convalescent patients [11,12,14].

This article hypotheses that the endocytotic properties of 3a allow it to modulate the surface expression of S and explores a functional significance for the interaction between S and 3a, which has been observed experimentally [13,15].

Presentation of the hypothesis

The cellular fate of the S protein has been well mapped [16,17]: S is cotranslationally glycosylated and oligomerized at the endoplasmic reticulum. Its N-linked high mannose side chains are trimmed, modified and become endoglycosidase H-resistant during the transportation to the Golgi apparatus. Only this fully-matured form of S can be assembled into virions and/or transported to the cell surface. The latter could cause cell-cell fusion and the formation of syncytia. Recently, Schwegmann-Wessels and co-worker reported that a novel sorting signal for intracellular localization is present in the S protein of most coronaviruses, but absent from SARS-CoV S [18]. Site-directed mutagenesis studies confirmed that a Yxx Φ motif (where x is any amino acid and Φ is an amino acid with a bulky hydrophobic side chain) retains the S protein of TGEV intracellularly when it is expressed alone. On the other hand, SARS-CoV S is transported efficiently to the cell surface unless such a motif is introduced into its cytoplasmic tail by mutagenesis.

The $Yxx\Phi$ motif has been implicated in directing protein localization to various intracellular compartments [19-21]. Furthermore, most $Yxx\Phi$ motifs are capable of mediating rapid internalization from the plasma membrane into the endosomes. Interaction between the adaptor protein complex 2 (AP-2) with the $Yxx\Phi$ motif present in the cytoplasmic domain of the internalizing protein concentrated the protein in clathrin-coated vesicle, which then budded from the plasma membrane resulting in internalization. However, it appears that the $Yxx\Phi$ motif can also bind other adaptor protein complexes, like AP-1, 3 and 4, and the differential binding to the different adaptors will determine the pathway of a cargo protein containing a particular $Yxx\Phi$ motif [21]. Coincidently, a $Yxx\Phi$ motif in the cytoplasmic domain of 3a has previously been identified [13]. Furthermore, the juxtaposition of the $Yxx\Phi$ motif and a ExD (diacidic) motif was found to be essential for the transport of 3a to the cell surface, consistent with the role of these motifs in the transportation of other proteins to the plasma membrane [22]. 3a on the cell surface can also undergo internalization [13].

Analyzing the experimental results present in these publications collectively, it is possible to postulate a functional role for the evolution of the SARS-CoV 3a protein. The SARS-CoV S protein lacks the YxxΦ motif but it can bind to the 3a protein which has internalization properties. In SARS-CoV infected cells, S is rapidly transported to the cell surface. But if 3a is expressed in the same cell, it is also transported to the cell surface where it can bind S. The interaction between 3a and S enables both proteins to become internalized, resulting in a decrease in the expression of S on the cell surface. Thus, this viral-viral interaction confers the functional role for the $Yxx\Phi$ motif found in other coronaviruses to the SARS-CoV S. This hypothesis also implies that the precise mechanisms used by TGEV and SARS-CoV to reduce the expression of S are different although in both cases, the $Yxx\Phi$ motifs will be crucial. In TGEV, the $Yxx\Phi$ motif in S caused it to be retained intracellularly, while in SARS-CoV, S that is transported to the cell surface becomes internalized again after it interacts with the 3a protein.

Testing the hypothesis

Using mammalian cell culture system and biochemical methods, it will be possible to determine the exact effects of 3a on the trafficking properties of S. Mutagenesis studies can be used to map the protein domains that are

Protein	Amino acid sequences in the cytoplasmic tail ^a
TGEV S ^b	TM-CLGSCCHSICSRRQFEN <u>YEPI</u> EKVHVH
PRCoV S ^b	TM-CLGSCCHSIFSRRQFEN <u>YEPI</u> EKVHVH
CCoV S ^b	TM-CLGSCCHSICSRGQFES <u>YEPI</u> EKVHVH
FCoV S ^b	TM-CLGSCCHSICSRRQFEN <u>YEPI</u> EKVHVH
PEDV S ^b	TM-CCGACFSGCCRGPRLQP <u>YEAF</u> EKVHVQ
HCoV-229E S ^b	TM-CFASSIRGCCESTKLP <u>YYDV</u> EKIHIQ
HCoV-NL63 S ^b	TM-CLTSSMRGCCDCGSTKLP <u>YYEF</u> EKVHVQ
BCoV Sc	TM-ICGGCCDD <u>YTGH</u> QELVIKTSHDD
HCoV-OC43 S ^c	TM-KCGGCCDD <u>YTGY</u> QELVIKTSHDD
HEV S ^c	TM-KCGGCCDD <u>YTGH</u> QEFVIKTSHDD
MHV S ^c	TM-KKCGNCCDECGGHQDSIVIHNISSHED
RtCoV S ^c	TM-KCGNCCDE <u>YGGR</u> QAGIVIHNISSHED
HCoV-HKUI S ^c	TM-KCHNCCDE <u>YGGH</u> HDFVIKTSHDD
SARS-CoV S ^c	TM-GACSCGSCCKFDEDDSEPVLKGVKLHYT
IBV S ^d	TM-KKSS <u>YYTT</u> FDNDVVTEQYRPKKSV
SARS-CoV 3a ^e	TM-38aa- <u>YNSV</u> TDTIVVTEGD-101aa
TfR ^e	19aa- <u>YTRF</u> SLARQVDGDNSHV-26aa-TM
LDLR (proximal) ^e	TM-17aa- <u>YQKT</u> TEDEVHICH-20aa
LDLR (distal) ^e	TM-34aa- <u>YSYP</u> SRQMVSLEDDVA
CD-M6PR ^e	TM-34aa- <u>YRGV</u> GDDGLGEESEERDDHLLPM
ASGPR ^e	MTKE <u>YQDL</u> QHLDNEES-24aa

Table 1: Amino acid sequences of the cytoplasmic tail of spike (S) proteins of coronaviruses are compared with the $Yxx\Phi$ (where x is any amino acid and Φ is an amino acid with a bulky hydrophobic side chain) motifs found in SARS-CoV 3a protein and other cellular proteins that are known to undergo endocytosis.

^aSequences were obtained from National Center for Biotechnology Information (NCBI). Yxxx tetrapeptides are underlined and abbreviations used are: TM, transmembrane domain, aa, amino acids.

^bS proteins of group I coronaviruses: TGEV, transmissible gastroenteritis virus (AJ271965); PRCoV, porcine respiratory coronavirus (Z24675); CCoV, canine coronavirus (D13096); FCoV, feline coronavirus (AY204704); PEDV, porcine epidemic diarrhea virus (AF353511); HCoV-229E, human coronavirus 229E (AF304460); HCoV-NL63, human coronavirus NL63(AY518894).

^cS proteins of group 2 coronaviruses: BCoV, bovine coronavirus (AF220295), HCoV-OC43, human coronavirus OC43 (AY585228), HEV, porcine hemagglutinating encephalomyelitis virus (AY078417), MHV, murine hepatitis virus (AF201929), RtCoV, rat coronavirus (AF207551), HCoV-HKUI, human coronavirus HKUI (AY597011), SARS-CoV, SARS coronavirus (AY283798).

^dS protein of group 3 coronavirus: IBV, infectious bronchitis virus (M95169).

eSARS-CoV 3a protein (AY283798) and other cellular proteins that are known to undergo endocytosis. Abbreviations: TfR, transferrin receptor (P02786), LDLR, low-density lipoprotein receptor (P01130); CD-M6PR, cation-dependent mannose 6-phosphate receptor (P24668); ASGPR, asialoglycoprotein receptor (P07306).

important for the interaction between 3a and S and for the defining the manner by which 3a contributes to the reduction of cell surface expression of S. Given that a full-length infectious clone of SARS-CoV has been assembled [23], the use of reverse genetics would certainly reveal more about the interplay between 3a and S during SARS-CoV infection.

Implication of the hypothesis

This hypothesis, if proven, will indicate that the interaction between SARS-CoV-unique 3a protein and S results in a reduction of S on the cell surface through the endocytotic properties of 3a [13]. During SARS-CoV infection, expression of S on the cell surface of an infected cell mediates fusion with un-infected neighboring cells, leading to syncytium formation. It follows that reducing the cell surface expression of S will delay this cell-damaging effect and prevent the premature release of unassembled viral RNA. It may also enhance virus packaging as it appears that the assembly of coronavirus occurs intracellularly, probably in the intermediate compartments between the endoplasmic reticulum and Golgi apparatus [24]. Clearly, this has certain advantages for the virus at certain stages of its life cycle. In addition, a reduction in the cell surface expression of S may also help the infected cell evade the host defense system and reduce the production of anti-S neutralizing antibodies. Conversely, host or viral factors that disrupt the interaction between S and 3a would favor the expression of S on the cell surface and enhance cellcell fusion, a process that is important for viral spreading.

Table 1 shows a comparison of the amino acid sequences of the cytoplasmic tails of the S protein of different coronaviruses, including SARS-CoV, which is distantly related to the established group 2 coronaviruses [25], as well as two recently identified novel human coronaviruses,

HCoV-NL63 [26] and HCoV-HKU1 [27]. The YxxΦ motifs are clearly present in all group 1 coronaviruses and also in IBV, which belongs to group 3. However, no $Yxx\Phi$ motif is present in SARS-CoV and MHV, both group 2 coronaviruses. In addition, there is a YGGR motif in the S protein of RtCoV and YxxH motifs in the S proteins of the other group 2 coronaviruses, BCoV, HEV and HCoV-HKU1. However, these motifs may not be able to function as signaling motifs because both R and H are not hydrophobic amino-acids. Therefore, HCoV-OC43 is the only one of these group 2 coronaviruses that encodes a S protein with a Yxx Φ motif. It is still unclear how the localization of S is modulated in those viruses that lack $Yxx\Phi$ motifs in the S proteins and further studies will be needed to understand the different signaling pathways that are important for regulating the trafficking properties of S. Indeed, the dilysine endoplasmic reticulum retrieval signal, which is a different type of sorting signal from the $Yxx\Phi$ motif, in the cytoplasmic tail of IBV was reported to be important for intracellular retention of S [28].

It therefore appears that the cell surface expression of S protein of SARS-CoV can be reduced like that for other coronaviruses, but the mechanism may be different. The trafficking of SARS-CoV S may be mediated through 2 separate viral proteins, expressed from separate subgenomic RNA, and regulated by numerous complex cellular processes including the efficiency of transcription and translation, post-translation modification and stability of the viral proteins, as well as their interactions with host factors. Indeed, it is crucial to determine how this unique pathway benefits replication of the SARS-CoV. It is also interesting to note that sequence comparison of isolates from different clusters of infection showed that both S and 3a showed a positive selection during virus evolution [29,30], implying that these proteins play important roles in the virus life cycle and/or disease development and is consistent with the proposal that 3a has evolved to modulate the trafficking properties of the spike protein.

Competing interests

The author(s) declare that they have no competing interests.

Author's contributions

Yee-Joo Tan is responsible for the entire manuscript.

Acknowledgements

This work was supported by grants from the Agency for Science, Technology and Research (A*STAR), Singapore.

References

- Drosten C, Preiser W, Gunther S, Schmitz H, Doerr HW: Severe acute respiratory syndrome: identification of the etiological agent. Trends Mol Med 2003, 9:325-327.
- 2. Marra MA, Jones SJ, Astell CR., Holt RA, Brooks-Wilson A, Butterfield YS, Khattra J, Asano JK, Barber SA, Chan SY, Cloutier A, Cough-

lin SM, Freeman D, Girn N, Griffith OL, Leach SR, Mayo M, McDonald H, Montgomery SB, Pandoh PK, Petrescu AS, Robertson AG, Schein JE, Siddiqui A, Smailus DE, Stott JM, Yang GS, Plummer F, Andonov A, Artsob H, Bastien N, Bernard K, Booth TF, Bowness D, Czub M, Drebot M, Fernando L, Flick R, Garbutt M, Gray M, Grolla A, Jones S, Feldmann H, Meyers A, Kabani A, Li Y, Normand S, Stroher U, Tipples GA, Tyler S, Vogrig R, Ward D, Watson B, Brunham RC, Krajden M, Petric M, Skowronski DM, Upton C, Roper RL: The Genome sequence of the SARS-associated coronavirus. *Science* 2003, 300:1399-1404.

- Rota PA, Oberste MS, Monroe SS, Nix WA, Campagnoli R, Icenogle JP, Penaranda S, Bankamp B, Maher K, Chen MH, Tong S, Tamin A, Lowe L, Frace M, DeRisi JL, Chen Q, Wang D, Erdman DD, Peret TC, Burns C, Ksiazek TG, Rollin PE, Sanchez A, Liffick S, Holloway B, Limor J, McCaustland K, Olsen-Rasmussen M, Fouchier R, Gunther S, Osterhaus AD, Drosten C, Pallansch MA, Anderson LJ, Bellini WJ: Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* 2003, 300:1394-1399.
- Thiel V, Ivanov KA, Putics A, Hertzig T, Schelle B, Bayer S, Weissbrich B, Snijder EJ, Rabenau H, Doerr HW, Gorbalenya AE, Ziebuhr J: Mechanisms and enzymes involved in SARS coronavirus genome expression. J Gen Virol 2003, 84:2305-2315.
- 5. Žiebuhr J: Molecular biology of severe acute respiratory syndrome coronavirus. *Curr Opin Microbiol* 2004, **7:**412-419.
- Tan Y-J, Lim SG, Hong W: Characterization of viral proteins encoded by the SARS-Coronavirus genome. Antiviral Research 2005, 65:69-78.
- Cavanagh D: The coronavirus surface glycoprotein protein. In The Coronaviridae Edited by: Siddell SG. New York: Plenum Press; 1995:73-113.
- 8. Gallagher TM, Buchmeier MJ: Coronavirus spike proteins in viral entry and pathogenesis. *Virology* 2001, **279**:371-374.
- 9. Holmes KV: SARS coronavirus: a new challenge for prevention and therapy. J Clin Invest 2003, 111:1605-1609.
- 10. Navas-Martin S, Weiss SR: SARS: lessons learned from other coronaviruses. Viral Immunol 2003, 16:461-474.
- Guo JP, Petric M, Campbell W, McGeer PL: SARS corona virus peptides recognized by antibodies in the sera of convalescent cases. Virology 2004, 324:251-256.
- Tan Y-J, Goh P-Y, Fielding BC, Shen S, Chou C-F, Fu J-L, Leong HN, Leo YS, Ooi EE, Ling AE, Lim SG, Hong W: Profile of antibody responses against SARS-Coronavirus recombinant proteins and their potential use as diagnostic markers. *Clin Diag Lab Immunol* 2004, 11:362-371.
- Tan Y-J., Teng E, Shen S, Tan THP, Goh P-Y, Fielding BC, Ooi E-E, Tan H-C, Lim SG, Hong W: A novel SARS coronavirus protein, U274, is transported to the cell surface and undergoes endocytosis. J Virol 2004, 78:6723-6734.
- 14. Yu C-J, Chen Y-C, Hsiao C-H, Kuo T-C, Chang SC, Lu C-Y, Wei W-C, Lee C-H, Huang L-M, Chang M-F, Ho H-N, Lee FJS: Identification of a novel protein 3a from severe acute respiratory syndrome coronavirus. FEBS Lett 2004, 565:111-116.
- 15. Zeng R, Yang RF, Shi MD, Jiang MR, Xie YH, Ruan HQ, Jiang XS, Shi L, Zhou H, Zhang L, Wu XD, Lin Y, Ji YY, Xiong L, Jin Y, Dai EH, Wang XY, Si BY, Wang J, Wang HX, Wang CE, Gan YH, Li YC, Cao JT, Zuo JP, Shan SF, Xie E, Chen SH, Jiang ZQ, Zhang X, Wang Y, Pei G, Sun B, Wu JR: Characterization of the 3a protein of SARS-associated coronavirus in infected vero E6 cells and SARS patients. *J Mol Biol* 2004, 341:271-279.
- Parker MD, Yoo D, Cox GJ, Babiuk LA: Primary structure of the S peplomer gene of bovine coronavirus and surface expression in insect cells. J Gen Virol 1990, 71:263-270.
- Vennema H, Heijnen L, Zijderveld A, Horzinek MC, Spaan WJ: Intracellular transport of recombinant coronavirus spike proteins: implications for virus assembly. J Virol 1990, 64:339-346.
- Schwegmann-Wessels C, Al-Falah M, Escors D, Wang Z, Zimmer G, Deng H, Enjuanes L, Naim HY, Herrler G: A novel sorting signal for intracellular localization is present in the S protein of a porcine coronavirus but absent from severe acute respiratory syndrome-associated coronavirus. J Biol Chem 2004, 279:43661-4366.
- Trowbridge IS, Collawn JF, Hopkins CR: Signal-dependent membrane protein trafficking in the endocytic pathway. Ann Rev Cell Biol 1993, 9:129-161.

- 20. Marks MS, Ohno H, Kirchhausen T, Bonifacino JS: **Protein sorting** by tyrosine- based signals: adapting to the Ys and wherefores. *Trends Cell Biol* 1997, 7:124-128.
- 21. Bonifacino JS, Traub LM: Signals for sorting of transmembrane proteins to endosomes and lysosomes. Annu Rev Biochem 2003, 72:395-447.
- 22. Bannykh SI, Nishimura N, Balch WE: Getting into the Golgi. Trends Cell Biol 1998, 8:21-25.
- Yount B, Curtis KM, Fritz EA, Hensley LE, Jahrling PB, Prentice E, Denison MR, Geisbert TW, Baric RS: Reverse genetics with a fulllength infectious cDNA of severe acute respiratory syndrome coronavirus. Proc Natl Acad Sci USA 2003, 100:12995-13000.
- Klumperman J, Locker JK, Meijer A, Horzinek MC, Geuze HJ, Rottier PJ: Coronavirus M proteins accumulate in the Golgi complex beyond the site of virion budding. J Virol 1994, 68:6523-6534.
 Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL,
- Snijder EJ, Bredenbeek PJ, Dobbe JC, Thiel V, Ziebuhr J, Poon LL, Guan Y, Rozanov M, Spaan WJ, Gorbalenya A: Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage. J Mol Biol 2003, 331:991-1004.
- van der Hoek L, Pyrc K, Jebbink MF, Vermeulen-Oost W, Berkhout RJ, Wolthers KC, Wertheim-van Dillen PM, Kaandorp J, Spaargaren J, Berkhout B: Identification of a new human coronavirus. Nat Med 2004, 10:368-373.
- Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, Huang Y, Wong BH, Poon RW, Cai JJ, Luk WK, Poon LL, Wong SS, Guan Y, Peiris JS, Yuen KY: Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. J Virol 2005, 79:884-895.
- Lontok E, Corse E, Machamer CE: Intracellular targeting signals contribute to localization of coronavirus spike proteins near the virus assembly site. J Virol 2004, 78:5913-5922.
- Guan Y, Peiris JS, Zheng B, Poon LL, Chan KH, Zeng FY, Chan CW, Chan MN, Chen JD, Chow KY, Hon CC, Hui KH, Li J, Li VY, Wang Y, Leung SW, Yuen KY, Leung FC: Molecular epidemiology of the novel coronavirus that causes severe acute respiratory syndrome. Lancet 2004, 363:99-104.
- Yeh SH, Wang HY, Tsai CY, Kao CL, Yang JY, Liu HW, Su IJ, Tsai SF, Chen DS, Chen PJ, National Taiwan University SARS Research Team: Characterization of severe acute respiratory syndrome coronavirus genomes in Taiwan: molecular epidemiology and genome evolution. Proc Natl Acad Sci USA 2004, 101:2542-2547.

