*Chinese Science Bulletin* 2003 Vol. 48 No.11 1115—1117

# Spike protein homology between the SARS-associated virus and murine hepatitis virus implies existence of a putative receptor-binding region

LU Yun & CHEN Yinghua

Laboratory of Immunology, Department of Biology, Tsinghua University; Protein Science Laboratory of Ministry of Education, Beijing 100084, China
Correspondence should be addressed to Chen Yinghua (e-mail: chenyh@mail.tsinghua.edu.cn)

**Abstract** **Coronavirus has been determined to be the cause of the recent outbreak of severe acute respiratory syndrome (SARS). Human coronavirus 229E had been studied well and its receptor-binding domain was restricted to aa417—547 of S protein. However, this region has no homology with the newly separated SARS-associated virus (Hong Kong isolate CUHK-W1). Then we analyzed the phylogenesis of S1 subunit of the coronavirus spike protein (SARS-associated virus, Hong Kong isolate CUHK-W1). Interestingly, the highest homology between murine hepatitis virus (MHV) and SARS-associated coronavirus was found. And the important sites (aa62—65 and aa214—216) on the spike protein of MHV with receptor-binding capacity were highly conservative in comparison with the newly separated SARS-associated virus (the corresponding sites are aa51—54 and aa195—197). These results from bioinformatics analysis might help us to study the receptor-binding sites of SARS-associated virus and the mechanism of the virus entry into the target cell, and design antiviral drugs and potent vaccines.**

**Keywords: SARS-associated virus, receptor-binding sites, human coronavirus, murine hepatitis virus.**

**DOI: 10.1360/03wc0197**

The envelop protein 2 (E2 protein) of coronavirus, which is also called spike (S) protein, contains two subunits S1 and S2. S1 can bind to the cell receptor and S2 mediate membrane fusion. Like many other virus such as HIV and influenza, the receptor-binding subunit has most of antigenic sites and induces neutralizing antibodies, which block the virus attachment[1]. Then this part of envelop protein often becomes the ideal target of antiviral drugs and potent vaccine component. Also the specificity of virus-receptor interactions clearly affects the species specificity of virus infection, and in some instances may be an important determinant of viral tissue tropism[2]. Unfortunately, the exact receptor-binding sites of many coronaviruses are still unknown, especially these viruses newly separated. Now, it is urgent to offer some information of the receptor-binding site of SARS-associated coronavirus for antiviral research, as the disease had spread over 30 countries and cause the mortality of about 7.2% within several months[1)].

Recently, a receptor-binding domain of the S protein of human coronavirus 229E (human CoV-229E) has been identified. A few of truncated proteins that contain different lengths of N-terminal amino acids of the protein were obtained and their activities of binding to 3T3 mouse cells that express human aminopeptidase N (hAPN) were tested. The region from amino acid 417 to 547 was identified as receptor-binding domain on human CoV-229E[3]. We compared the domain (aa417—547) with sequences of other viruses in the genebank and found some homologous regions in other species of coronavirus (Fig. 1). Unfortunately, it was not found any homology between S proteins of SARS-associated virus (strain CUHK-W1) and human CoV-229E in this region. Up to now, however, there is no evidence that the spike protein of the SARS-associated virus has no receptor-binding sites in the S1 subunit.

In order to get more information of the receptor-binding sites on the SARS-associated virus, we do the phylogenetic analysis of the N-terminal 840 amino acids (SN840) of most coronavirus species (Fig. 2). Because the conservative cleavage site RRSRR was not found through all the sequence of strain CUHK-W1[4], it is now difficult to divide the spike protein into two subunits. Since the two subunits usually have similar size[1] and most of the cleavage sites are before the amino acid 770 (Table 1), we reckon that SN840 could overlap the S1 subunit and also the receptor-binding domain. From the tree (Fig. 2), we could see that the SARS virus, murine hepatitis virus, bovine coronavirus and human coronavirus OC43 were in one group. This result is similar to the recent report by a Canadian research group[5]. It is cheerful that receptor-binding site of murine hepatitis virus had been studied. A partial protein containing 330 amino acid residues of the spike protein and larger N-terminal truncated proteins could bind to its receptor MHVR (also referred to as Bgp1a or C-CAM), which indicated that the N terminus 330 amino acid residues were responsible for the receptor-binding[6—8]. Furthermore, Site-directed mutagenesis analysis and soluble receptor-resistant mutants suggested that positions aa62—65 and aa214—216 are important for receptor-binding capacity[9—11]. And these two sites was highly conservative among the murine coronaviruses, especially aa62—65[10]. Interestingly, the important sites

1) http://www.who.int/csr/sarscontury/2003_05_10/en

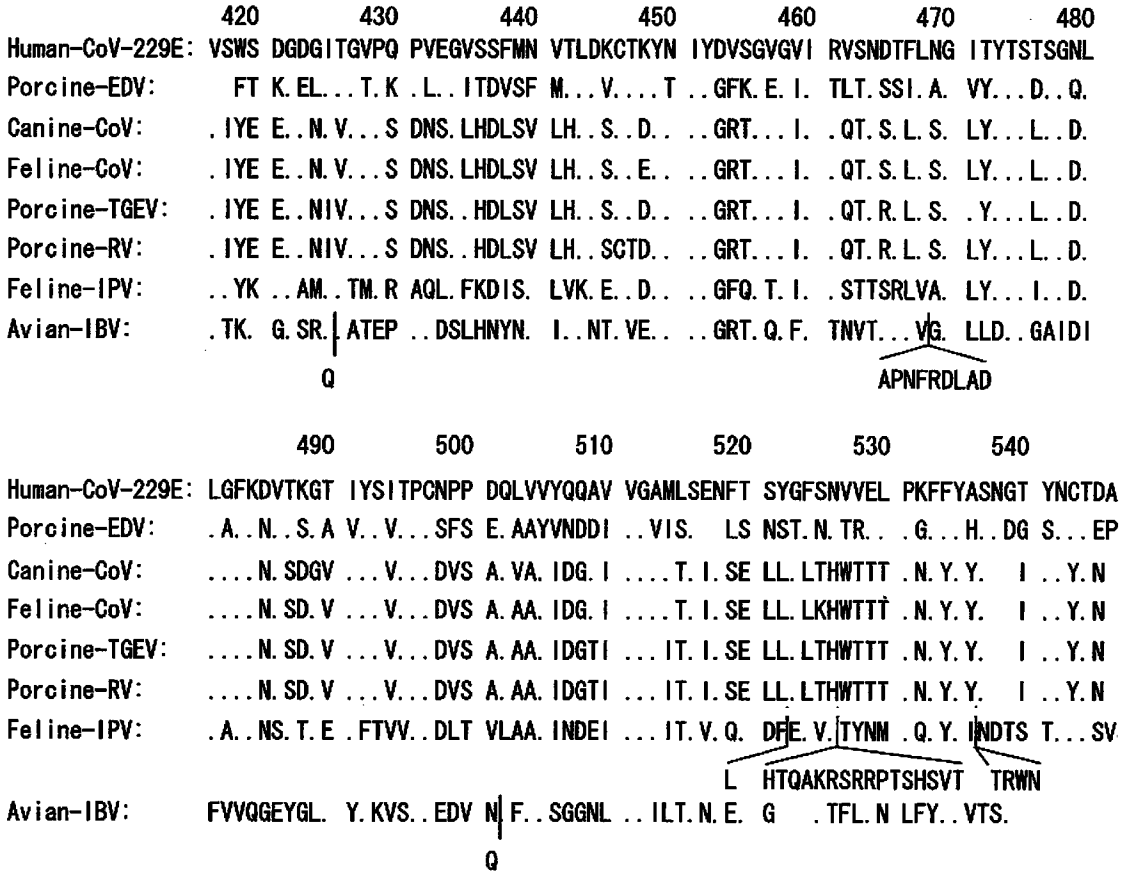


Fig. 1. Sequences homologous to the receptor-binding domain of human-CoV-229E. The first line is the alignment of the receptor-binding domain of human CoV-229E, from amino acid 417 to 547. Homologous sequences to this region were searched in the genebank and was aligned from the second to eighth line. Human-CoV-229E donates human coronavirus (strain 229E), porcine-EDV porcine epidemic diarrhea virus, canine-CoV canine coronavirus, feline-CoV feline coronavirus, porcine-TGEV porcine transmissible gastroenteritis virus, porcine-RV porcine respiratory virus, feline-IPV feline infectious peritonitis virus, avian-IBV avian infectious bronchitis virus. A dot indicates that the position is identical to the reference sequence (human-CoV-299E).

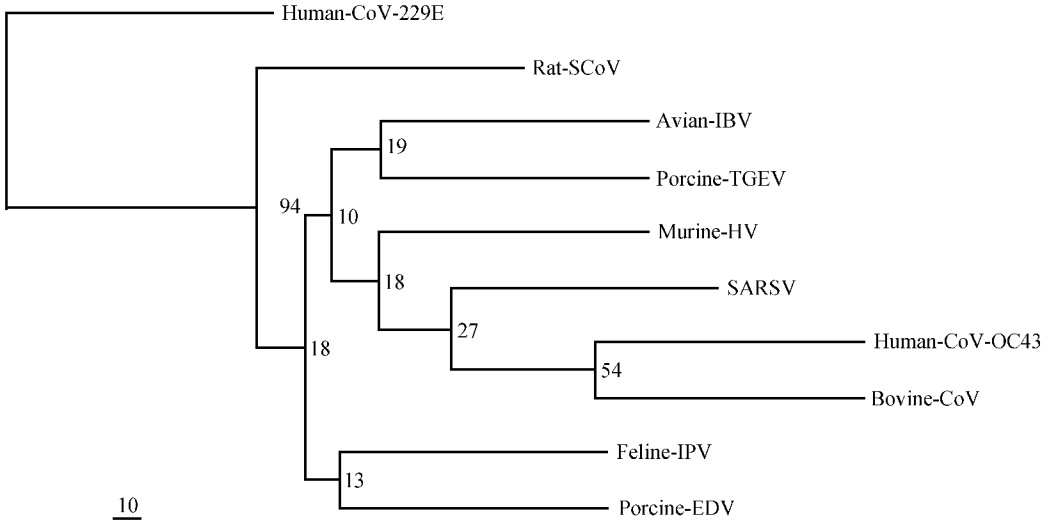


Fig. 2. Phylogenetic analysis of the N-terminus 840aa of the spike protein of different coronavirus species. The tree is constructed by neighbour-joining method using Philip version 3.57c. Horizontal-line distance represents number of sites at which the two sequences compared are different. Bootstrap values deduced from 100 replicates. In addition to human coronavirus isolate 229E and OC43, rat sialodacryoadenitis coronavirus SDAV-681, avian infectious bronchitis virus KB8523, porcine transmissible gastroenteritis virus Miller, murine hepatitis virus JHM, SARS-associated virus CUHK-W1, bovine coronavirus Mebus, feline infectious peritonitis virus 79-1146, and porcine epidemic diarrhea virus Chinju 99 were included in the analysis.

Table 1 The cleavage sites of spike proteins of four different coronavirus

| Coronavirus Strains | Murine HV (JHM) | Bovine CoV (Mebus) | Human CoV (OC43) | Avian IBV (KB8523) |
|---|---|---|---|---|
| Cleavage site (aa) | 769—770 | 768—769 | 768—769 | 537—538 |

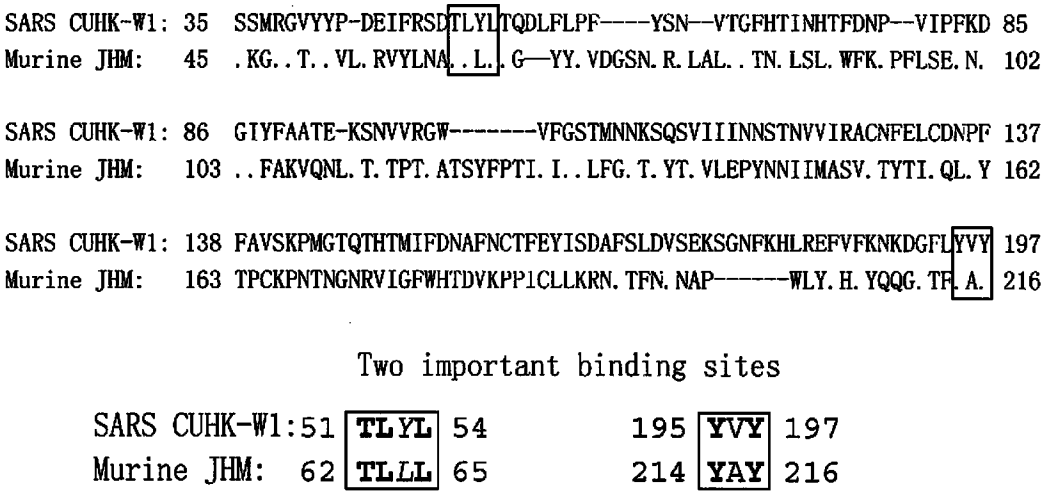


Fig. 3. Partial spike protein comparison of SARS-associated virus and murine hepatitis virus. The N-terminus of two spike proteins have weak homology. However, the important sites (aa62—65, aa214—216) for the receptor-binding of murine hepatitis virus JHM, which were framed, were highly homologous with the spike protein of SARS-associated virus CUHK-W1, namely, 5 of 7 amino acid residues are identical and one amino acid residue site (A and V) is conservative change.

(aa62—65 and aa214—216) on the spike protein of MHV with receptor-binding capacity were highly homologous in comparison with the newly separated SARS-associated virus (the corresponding sites are aa51—54 and aa195—197) (Fig. 3). 5 of 7 amino acid residues are identical and one amino acid residue site (A and V) is conservative change. And the secondary structure predictions of SN840 (CUHK-W1) and SN840 with two amino acids substitution(Y53L and V196A) in the putitive receptor-binding site show little change (data not shown), which indicates that these two amino acids mutation may not cause significant change of the protein structure. Also all 5 SARS-associated viruses isolated by different research groups have identical sequences in these two sites, including the newly separated strain BJ01[12]. This high conservation suggests that the SARS-associated virus might also use the same regions as its receptor-binding domain and might have similar receptor with MHVR, which belong to the carcinoembryonic antigen (CEA) family[8]. Further studies should be carried out to verify our hypothesis. These results from bioinformatics analysis provide important information, which might help us to study the receptor-binding sites of SARS-associated virus and the mechanism of the virus entry into the target cell, and even help us to design effective antiviral drugs and potent vaccines.

**Acknowledgements** This work was supported by the Anti-SARS Fund of Ministry of Education and the National Natural Science Foundation for Outstanding Young Scientist of China (Grant No. 30025038).

## References

1. Siddell, S., Wege H., ter Meulen V., The structure and replication of coronaviruses, Curr. Top. Microbiol. Immunol., 1982, 99: 131—163.
2. Holmes, K. V., Tresnan, D. B., Zelus, B. D., Virus-receptor interactions in the enteric tract, Virus-receptor interactions, Adv. Exp. Med. Biol., 1997, 412: 125—33.
3. Bonavia, A., Zelus, B. D., Wentworth, D. E. et al., Identification of a receptor-binding domain of the spike glycoprotein of human coronavirus HCoV-229E, J. Virol., 2003, 77(4): 2530—2538.
4. Abraham, S., Kienzle, T. E., Lapps, W. et al., Deduced sequence of the bovine coronavirus spike protein and identification of the internal proteolytic cleavage site, Virology, 1990, 176(1): 296—301.
5. Marra, M. A., Jones, S. J. M., Astell, C. R. et al., The genome sequence of the SARS-associated coronavirus, Science, 1085953, 2003.
6. Kubo, H., Yamada, Y. K., Taguchi, F., Location of neutralizing epitopes and the receptor-binding site within the amino-terminal 330 amino acids of the murine coronavirus spike protein, J. Virol., 1994, 68(9): 5403—5410.
7. Taguchi, F., Kubo, H., Suzuki, H. et al., Location of neutralizing epitopes and receptor-binding site in murine coronavirus spike protein, Adv. Exp. Med. Biol., 1995, 380: 359—365.
8. Wessner, D. R., Shick, P. C., Lu, J. H. et al., Mutational analysis of the virus and monoclonal antibody binding sites in MHVR, the cellular receptor of the murine coronavirus mouse hepatitis virus strain A59, J. Virol., 1998, 72(3): 1941—1948.
9. Suzuki, H., Taguchi, F., Analysis of the receptor binding site of murine coronavirus spike glycoprotein, J. Virol., 1996, 70: 2632—2636.
10. Saeki, K., Ohtsuka, N., Taguchi, F., Identification of spike protein residues of murine coronavirus responsible for receptor-binding activity by use of soluble receptor-resistant mutants, J. Virol., 1997, 71(12): 9024—31.
11. Saeki, K., Ohtsuka, N., Taguchi, F., Isolation and characterization of murine coronavirus mutants resistant to neutralization by soluble receptor, Adv. Exp. Med. Biol., 1998, 440: 11—16.
12. Qin, E'de, Zhu, Q. Y., Yu, M. et al., A complete sequence and comparative analysis of a SARS-associated virus (Isolate BJ01), Chinese Sci. Bull., 2003, 48(10): 941—948.

(Received May 7, 2003; accepted May 15, 2003)