Arch Virol (1996) 141: 1123–1131

Archives of Virology

Structural and functional analysis of the S proteins of two human coronavirus OC43 strains adapted to growth in different cells

Brief Report

F. Künkel* and **G. Herrler**

Institut für Virologie, Philipps-Universität Marburg, Marburg, Federal Republic of Germany

Accepted February 6, 1996

Summary. The receptor-binding activity of strain CU (grown in MDCK I cells) and of strain VA (adapted to Vero cells) of human coronavirus OC43 was analyzed and compared with the binding activity of bovine coronavirus (BCV) and of the OC43 strain provided by the American Type Culture Collection (AT). Results obtained with resialylated erythrocytes indicated that the ability of the viruses to recognize 9-O-acetylated sialic acid in an α2,6-linkage decreased in the following order: AT > CU > BCV > VA. Only minor differences were observed with respect to the α2,3-linkage. The amino acid sequence of the S protein of strain CU and VA was derived from the nucleotide sequence of the cloned gene. Strain VA differed from strain CU in 34 positions, 18 in the S1 and 16 in the S2 subunit.

*

Coronaviruses are members of the family *Coronaviridae* that comprises large, single-stranded RNA viruses with positive genome polarity. The virions are composed of a nucleocapsid protein (N) and two or three glycoproteins, respectively. The glycoproteins include the surface protein (S), the hemagglutinin-esterase (HE) and the membrane protein (M). The presence of an HE protein is a characteristic feature of a serological subgroup of coronaviruses that includes human coronavirus OC43 (HCV-OC43), bovine coronavirus (BCV), mouse hepatitis virus (MHV) and hemagglutinating encephalomyelitis virus (for re-

*Present address: Maco Pharma International GmbH, Langen, Federal Republic of Germany.

views see [11, 12]). Human coronaviruses infect cells of the upper respiratory tract and are known to be a causative agent of common colds. By contrast, the closely related bovine coronavirus causes severe diarrhoea in newborn calves. The S protein is a main target of the immune response to coronavirus infections and plays an important role for the binding of virions to cell surface receptors and for the subsequent penetration by fusion of the viral envelope with the cellular membrane. BCV and HCV-OC43 have been shown to use N-acetyl-9-O-acetylneuraminic acid ($Neu5,9Ac_2$) as a receptor determinant for binding to cells [4, 8, 13]. Though the HE protein is also able to recognize this type of sugar, the S protein has been shown to be the major sialic acid-binding protein of BCV and HCV-OC43 [4, 8]. It is a typical class I membrane protein with an amino-terminal signal peptide that is cleaved during protein processing, and a carboxy-terminal hydrophobic membrane anchor. The spike proteins of BCV and MHV are cleaved by an intracellular protease into two subunits. HCV-OC43 has been reported to contain an uncleaved S protein [3, 4]. The lack of cleavage may be due to an altered cleavage site with a glycine at position-1 instead of an arginine [4, 6].

Here we report the sequence of another isolate of HCV-OC43, as well as of a variant that has been adapted to growth in Vero cells. The different viruses were compared with respect to their ability to use 9-O-acetylated sialic acid as a receptor determinant for binding to cells. One strain of OC43 was provided by the former British Common Cold Unit to Dr. Czerny (Institut für Medizinische Mikrobiologie, Infektions- und Seuchenmedizin, University of Munich), who adapted the virus to growth in Vero cells by eleven consecutive passages [2]. We prepared virus stocks in our laboratory by one additional passage in Vero cells and designated the virus OC43-VA. Another OC43-isolate was provided by the former British Common Cold Unit to Dr. Siddell (Institut für Virologie und Immunobiologie) as suckling mouse brain material and given to us without any further passage. The virus grew readily in MDCK I cells and virus stocks were prepared after 3 passages in these cells. This virus has been designated OC43-CU. Strain HCV-OC43-AT was obtained from the American Type Culture Collection and was grown in HRT cells as described recently [4]. Strains OC43-CU and OC43-AT were harvested two days p.i.; maximum titers of OC43-VA were obtained 72 h p.i.

The different strains were compared with respect to their ability to use 9-O-acetylated sialic acid as a receptor determinant for attachment to cells. For this purpose, endogenous sialic acid was removed from the surface of one day-old chicken erythrocytes by treatment with neuraminidase from *Vibrio cholerae*. The asialo-cells were then incubated with either of two sialyltransferases (Galβ1, 3GalNAc α2,3-sialyltransferase or Galβ1,4GlcNAc α2,6-sialyltransferase, respectively, from Boehringer Mannheim) and CMP-$Neu5,9Ac_2$ as described previously [9]. This resulted in the attachment of 9-O-acetylated sialic acid to the cell surface in two different linkage types (α2,3 or α2,6-linked to galactose, respectively). A further variation was obtained by using different amounts of CMP-activated sialic acid. In this way, batches of erythrocytes were obtained

that differed in the amount of $Neu5,9Ac_2$ present on the cell surface. The different strains of OC43 were analyzed for their ability to agglutinate the resialylated cells. Strain OC43-AT has been reported recently to recognize α2,6-linked $Neu5,9Ac_2$ more efficiently than does BCV. As shown in Table 1, strain OC43-CU was somewhat superior to BCV agglutinating cells that had been resialylated in the presence of 1–2 nmol of CMP-activated 9-O-acetylated sialic acid compared to 4 nmol required for BCV. With strain OC43-AT, however, hemagglutination was observed when the erythrocytes had been resialylated with amounts of CMP-$Neu5,9Ac_2$ as low as 0.25 nmol. The Vero cell adapted strain OC43-VA, on the other hand, appeared to be unable to recognize 9-O-acetylated sialic acid in an α2,6-linkage, because it failed to agglutinate cells that had been resialylated at the highest amount of CMP-$Neu5,9Ac_2$ tested, 16 nmol. The differences between individual strains were less pronounced when the recognition of α2,3-linked sialic acid was analyzed. Strain OC43-VA and BCV were able to agglutinate erythrocytes after resialylation in the presence of 1 nmol of CMP-$Neu5,9Ac_2$, whereas OC43-AT and OC43-CU required 2 nmol.

In order to explain the functional differences between the strains of HCV-OC43 on a molecular level, the nucleotide sequence of the S gene of OC43-CU and OC43-VA was determined and compared with the corresponding sequence of strain OC43-AT reported recently [4]. Virus was purified by discontinuous sucrose gradient centrifugation as described recently [4]. The final virus pellet was resuspended in a small volume of PBS. RNA was prepared from purified virus after proteinase-K digestion and phenol-chloroform extraction. After precipitation with 3M NaAc, pH 5.3, and 99% EtOH, vRNA was dissolved in DEPC-treated water as reported previously [4]. The S genes of the strains OC43-CU and OC43-VA were cloned by transcribing vRNA into two overlapping cDNA fragments using MoMuLV reverse transcriptase (Boehringer Mannheim). The product was amplified by PCR and then cloned into pBluescript (Stratagene).

Full length S genes were generated as described recently [4] and two clones were independently sequenced in both directions. In addition to sequencing the cloned S genes, the viral RNA from OC43-CU was directly sequenced twice with a modification of the dideoxy chain termination method [7] using AMV reverse transcriptase (Boehringer Mannheim). Hydridization of vRNA and primers was performed in a total volume of 8.5 µl using a minimum of 2 µg RNA and 20 ng primer (18–26 nt long) in the presence of a final concentration of 205 mM KCl. After 90 sec of heating to 95°C, the mixture was cooled over a period of 30 min to room temperature. The following steps were done at 4 °C. Two µl ^{35}S-dATP, 2.5 µl 10 × reaction buffer (600 mM Tris/HCl, pH 8.3, 100 mM $MgCl_2$ and 100 mM DTT) and 1 unit AMV reverse transcriptase were added to the annealing mixture. Aliquots of 3 µl were added to four nucleotide mixtures containing 50 µM dATP, 300 µM dCTP, 200 µM dGTP, 200 µM dTTP; and either of the dideoxy-NTPs (2 µM in the case of ddATP, and 10 µM in the case of ddCTP, ddGTP or ddTTP, respectively). After incubation for 20 min at 42 °C, 1 µl,

```
                    (T)
OC43-CU      1  MFLILLLISLPMAFAVIGDLKCT**TVSINDIDTGAPSISTDIVDVTNGLG
OC43-VA                   L          **  A   V   V       T
OC43-AT                  T           SD SY   K   P P     T

OC43-CU     51  TYYVLDRVYLNTTLLLNGYYPTSGSTYRNMALKGTLLLSRLWFKPPFLSD
OC43-VA                                                  T
OC43-AT                     F                    SV

OC43-CU    101  FINGIFAKVKNTKVIKKGVMYSEFPAITIGSTFVNTSYSVVVQPHTTN**
OC43-VA                         H                              **
OC43-AT                        DR                          R J ST

OC43-CU    151  LD**NKLQGLLEISVCQYTMCEYPHTICHPNLGNRRVELWHWDTGVVSCL
OC43-VA        **                    N
OC43-AT       Q GY        V     N     Q          H K    L

OC43-CU    201  YKRNFTYDVNADYLYFHFYQEGGTFYAYFTDTGVVTKFLFNVYLGTVLSH
OC43-VA                             I                          Y
OC43-AT                                                     MA

OC43-CU    251  YYVLPLTCNS****AMTLEYWVTPLTSKQYLLAFNQDGVIFNAVDCKSDF
OC43-VA         M      ****
OC43-AT         M      KVKNGF           R          I       M

OC43-CU    301  MSEIKCKTLSIAPSTGVYELNGYTVQPIADVYRRIPNLPDCNIEAWLNDK
OC43-VA
OC43-AT              Q    P                    K    N

OC43-CU    351  SVPSPLNWERKTFSNCNFNMSSLMSFIQADSFTCNNIDAAKIYGMCFSSI
OC43-VA
OC43-AT

OC43-CU    401  TIDKFAIPNGRKVDLQLGNLGYLQSFNYRIDTTATSCQLYYNLPAANVSV
OC43-VA
OC43-AT

OC43-CU    451  SRFNPSTWNRRFGFTEQSVFKPQPVGVFTHHDVVYAQHCFKAPTNFCPCK
                                  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
OC43-VA             I                   A   D
OC43-AT                K    I D     R A  L N             K

OC43-CU    501  LDGSLCVGNGPGIDAGYKNSGIGTCPAGTNYLTCHNAAQCDCLCTPDPIT
                ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
OC43-VA                                          V  N
OC43-AT          NGS    S   *****  N            ****** N

OC43-CU    551  SKSTGPYKCPQTKYLVGIGEHCSGLAIKSDYCGGNPCTCQPQAFLGWSVD
                ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
OC43-VA
OC43-AT          F A  T       S           V        S   R        A
```

Fig. 1 *(continued)*

100 µM dATP was added and incubation was continued at 42 °C for 10 minutes. The reaction was stopped with 10 µl stop-solution (95% formamide, 20 mM EDTA, 0.05% bromphenol blue and 0.05% xylene cyanol ff). Sequencing of DNA was performed using Sequenase 2.0 (USB) as described recently [4].

```
OC43-CU   601  SCLQGDRCNIFANFILHDVNSGTTCSTDLQKSNTDIILGVCVNYDLYGIT
               ~~~~~~~~~
OC43-VA
OC43-AT             K               L        A                 L

OC43-CU   651  GQGIFVEVNAPYYNSWQNLLYDSNGNLYGFRDYLTNRTFMIRSCYSGRVS
OC43-VA
OC43-AT               NAT                      I

                                    (S)
OC43-CU   701  AAFHANSSEPALLFRNIKCNYVFNNTLSRQLQPINYFDSYLGCVVNADNS
OC43-VA
OC43-AT                       S           S T               A   Y

OC43-CU   751  TSSVVQTCDLTVGSGYCVDYSTKRRSRRAITTGYRFTNFEPFTVNSVNDS
OC43-VA         A A
OC43-AT         AIS                  KN    G

OC43-CU   801  LEPVGGLYEIQIPSEFTIGNMEEFIQTSSPKVTIDCSAFVCGDYAACKSQ
OC43-VA         H                                           C
OC43-AT                                           A

OC43-CU   851  LVEYGSFCDNINAILTEVNELLDTTQLQVANSLMNGVTLSTKLKDGVNFN
OC43-VA
OC43-AT         G

OC43-CU   901  VDDINFSPVLGCLGSACNKVSSRSAIEDLLFSKVKLSDVGFVEAYNNCTG
OC43-VA           V           E                 R
OC43-AT                       E S A           D

OC43-CU   951  GAEIRDLICVQSYNGIKVLPPLLSVNQISGYTLAATSASLFPPWSAAAGV
OC43-VA         G                     D             N
OC43-AT                     K         E                   T

                                              (S)
OC43-CU  1001  PFYLNVQYRINGIGVTMDVLSQNQKLIANAFNNALDAIQEGFDATNSALV
OC43-VA
OC43-AT                     L                       Y

                                         (G)
OC43-CU  1051  KIQAVVNANAEALNNLLQQLSNRFGAISSSLQEILSRLDALEAQAQIDRL
OC43-VA                 D
OC43-AT                                    A              E

OC43-CU  1101  INGRLTALNAYVSQQLSDSTLVKFSAAQAMEKVNECVKSQSSRINFCGNG
OC43-VA                 D
OC43-AT

OC43-CU  1151  NHIISLVQNAPYGLYFIHFSYVPTKYVTAKVSPGLCIAGDRGIAPKSGYF
OC43-VA
OC43-AT
```

Fig. 1 (*continued*)

The comparison of vRNA and cDNA sequence allowed us to determine the extent of mutations during reverse transcription, PCR and the various cloning steps. Compared to the vRNA, the cDNA sequence of the S gene from OC43-CU showed 4 point mutations within the entire length of the gene resulting in

```
OC43-CU  1201  VNVNNTWMFTGSGYYYPEPITGNNVVVMSTCAVNYTKAPDVMLNISTPNL
OC43-VA                    R
OC43-AT                Y            E                 Y    T I

OC43-CU  1251  HDFKEELDQWFKNQTSVAPDLSLDYINVTFLDLQDEMNRLQEAIKVLNQS
                                                    ----------------
OC43-VA        P             L
OC43-AT        P                                  V

OC43-CU  1301  YINLKDIGTYEYYVKWPWYVWLLIGFAGVAMLVLLFFICCCTGCGTSCFK
               ----
OC43-VA
OC43-AT                                  CL

OC43-CU  1351  ICGGCCDDYTGHQELVIKTSHDD
OC43-VA        K                   EG
OC43-AT        K          Y
```

Fig. 1. Comparison of the predicted amino acid sequences of the S proteins from the isolates CU, VA and AT of HCV-OC43. For OC43-CU, the complete sequence is presented, for OC43-VA and OC43-AT, only the amino acid differences are shown. The amino-terminal signal peptide and the carboxy-terminal membrane anchoring domain are underlined and printed in bold letters. Potential sites for N-glycosylation (Asn-Xxx-Ser/Thr) are indicated in the same way. Amino acids which are inserted or missing, are marked by asterisks and the predicted polymorphic region in the S1 subunit of BCV [14] is marked by a wavy line. The proteolytic cleavage site is printed in bold letters and underlined twice. The amino acids in parentheses above the sequence indicate the differences found in the cDNA compared to the vRNA sequence of OC43-CU

4 amino acid exchanges (Fig. 1). The structural properties of the two S proteins were similar to those reported for OC43-AT [4]. Both open reading frames have a coding capacity of 1363 amino acids corresponding to a Mr of nearly 150 kDa in the unprocessed form. They are two amino acids longer than the S protein of OC43-AT and identical in length to the corresponding protein of BCV-Mebus [1]. OC43-VA and OC43-CU have 18 potential sites for N-glycosylation, 9 in each subunit, in contrast to the S protein of OC43-AT, which has 20 of them. The two additional glycosylation sites are located in the S1 subunit (Fig. 1). The three OC43 isolates differ also in the sequence of the potential cleavage site. The AT isolate has the sequence *KNRRSRG* in contrast to *TKRRSRR* in the two other isolates. The latter sequence is identical to that found in BCV-Mebus [1]. In contrast to the S proteins of OC43-CU and OC43-VA, which were cleaved to an extent of nearly 40% in infected cells, cleavage products of the S protein of the AT-isolate were never detected irrespective of the cell line used (data not shown).

The sequence of the S protein of OC43-CU was found to be very similar to the corresponding sequence of BCV-Mebus differing in only 12 amino acids, 7 in the S1 and 5 in the S2 subunit (Table 2). By contrast, the S proteins of strains CU and AT of HCV-OC43 differed from each other in 115 amino acid positions, the majority of them being located in the S1 subunit. The S protein of strain OC43-VA differed from OC43-CU in 34 amino acids, but it was still more closely related to BCV than to OC43-AT (Table 2). As OC43-CU and OC43-VA are

Table 1. Comparison of the efficiency of BCV and the HCV-OC43 strains AT, CU and VA in recognizing 9-O-acetylated sialic acid as a receptor determinant

		HA-activity (HAU/ml)			
Chicken erythrocytes	CMP-Neu5,9Ac$_2$ (nmol)	BCV	AT	CU	VA
asialo resialylated	–	<2	<2	<2	<2
α2,3	0.5	<2	<2	<2	<2
	1	32	<2	<2	64
	2	128	4	32	128
	4	256	16	128	128
α2,6	0.25	<2	32	<2	<2
	0.5	<2	32	<2	<2
	1	<2	32	8	<2
	2	<2	64	128	<2
	4	128	64	1024	<2
	8	512	128	1024	<2
	16	1024	128	1024	<2

Table 2. Comparison of the amino acid homology between the CU, VA and VT isolates of HCV-OC43 and the mebus strain of BCV [1]

	Amino acid homology in %							
	BCV Mebus		OC43-CU		OC43-AT		OC43-VA	
	S1	S2	S1	S2	S1	S2	S1	S2
BCV-Mebus			99.1%	99.2%	88%	95.3%	97.3%	96.6%
OC43-CU	7	5			82.2%	96%	97.7%	97.3%
OC43-AT	92	28	91	24			87.9%	94.5%
OC43-VA	21	20	18	16	93	33		
Number of amino acid differences								

derived from the same source, we assume that most of the differences between both strains are due to mutations that occurred during the adaptation to Vero cells.

The Vero-adapated strain recognized 9-O-acetylated sialic acid in an α2,3-linkage as efficiently as did the other strains tested; however, it was unable to agglutinate erythrocytes resialylated to contain α2,6-linked sialic acid – at least at the range of substrate concentrations tested. Assuming that the receptor-binding site is located in the S1 subunit, one or more of the 18 amino acids differing between OC43-CU and OC43-VA are expected to be part of the sialic

acid binding site. They should be helpful in future attempts to localize the binding site on the S protein. As the sialic acid binding activity of OC43-VA is diminished rather than improved, the change in this activity does not explain the ability of this virus to grow in Vero cells. Therefore, the mutations selected during the adaptation process are expected to affect an additional property in the S protein of the virus. Apart from the primary attachment to the cell surface, the S protein is also responsible for the fusion of the viral envelope with the cell membrane. The events involved in the fusion reaction of coronaviruses are not well characterized. A concept that has been proposed suggests that, after primary attachment to sialic acid-containing receptors, the S protein may interact with a post-adsorption receptor and this interaction may induce the fusion activity [10]. If this concept is true, the sequence changes during the adaptation process may have enabled strain OC43-VA to recognize such a post-adsoption receptor on Vero cells. Future work has to determine the property of HCV-OC43 that is critical for growing in Vero cells.

The S proteins of strains OC43-CU and OC43-VA are more related to BCV than to OC43-AT by several characteristics: (i) the sequence similarity, (ii) the length of the open reading frame, (iii) the number of sites for N-linked glycosylation. This might suggest a contamination by BCV. This is, however, unlikely for several reasons. The seventh brain passage of HCV-OC43 in suckling mice has been reported by McIntosh et al. [5] at a time when BCV has not yet been described. At Marburg, we have obtained brain material derived from the eighth passage from the former British Common Cold Unit. The virus grew readily in MDCK I cells and did not require adaptation. The stock virus used for the present analysis was derived from the third passage in cell culture. Independently, HCV-OC43 obtained from the Common Cold Unit was adapted at Munich to growth in Vero cells [2]. The fact that this strain (OC43-VA) as well as strain OC43-CU are both more related to BCV than to OC43-AT argues against a laboratory contamination. Strain OC43-AT obtained from the American Type Culture Collection has also been described as originating from the suckling mouse brain material reported by McIntosh et al [5]. An explanation of our findings may be that the mouse brain material contained a mixture of viruses and that different viruses were selected for by growth in different cell cultures (MDCK I cells for the CU, Vero cells for the VA, and HRT cells for the AT isolate). Alternatively, the mutations might have been introduced during the different passage history. In this context it should be noted that OC43 obtained from the American Type Culture Collection has been sequenced independently in two laboratories [4, 6]. The amino acid sequence derived from these nucleotide sequences differed in 26 amino acids, 23 in S1 and 3 in S2, including 11 amino acids present in our virus and missing in the sequence reported by Mounir and Talbot [6]. These differences may be due to different passage history in the two laboratories. Analysis of more HCV isolates is required to determine the extent of sequence variation within human coronaviruses and the similarity to BCV. The Vero cell adapted strain OC43-VA should be helpful in characterizing functional domains of the S protein.

Acknowledgements

We thank Dr. C.-P. Czerny for kindly providing HCV-OC43-VA and Prof. Dr. S. Siddell for HCV-OC43-CU. Financial support was provided by a grant from Deutsche Forschungsgemeinschaft (He 1168/2-2 and 2-3). Sequences have been submitted to EMBL data library under accession number Z32768 (OC43-CU) and Z32769 (OC43-VA).

References

1. Abraham S, Kienzle TE, Lapps W, Brian DA (1990) Deduced sequence of the coronavirus spike protein and identification of the internal proteolytic cleavage site. Virology 176: 296–301
2. Czerny C-P (1986) doctoral thesis. Ludwigs-Maximilian-Universität, München
3. Hogue BG, Brian DA (1986) Structural proteins of human coronavirus OC43. Virus Res 5: 131–144
4. Künkel F, Herrler G (1993) Structural and functional analysis of the surface protein of human coronavirus OC43. Virology 195: 195–202
5. McIntosh K, Becker WB, Chanock RM (1967) Growth in suckling-mouse brain of "IBV-like" viruses from patients with upper respiratory tract disease. Proc Natl Acad Sci USA 58: 2268–2273
6. Mounir S, Talbot PJ (1993) Molecular characterization of the S protein gene of human coronavirus OC43. J Gen Virol 74: 1981–1987
7. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467
8. Schultze B, Gross HJ, Brossmer R, Herrler G (1991) The S protein of bovine coronavirus is a hemagglutinin recognizing 9-O-acetylated sialic acid as a receptor determinant. J Virol 65: 6232–6237
9. Schultze B, Gross HJ, Brossmer R, Klenk H-D, Herrler G (1990) Hemagglutinating encephalomyelitis virus attaches to N-acetyl-9-O-acetylneuraminic acid containing receptors on erythrocytes: comparison with bovine coronavirus and influenza C virus. Virus Res 16: 185–194
10. Schultze B, Herrler G (1994) Recognition of cellular receptors by bovine coronavirus. In: Brinton MA, Calisher CH, Rueckert R (eds) Positive-strand RNA viruses. Springer, Wien New York, pp 451–459 (Arch Virol [Suppl] 9)
11. Siddell S, Wege H, ter Meulen V (1983) The biology of coronaviruses. J Gen Virol 64: 761–776
12. Spaan W, Cavanagh D, Horzinek MC (1988) Coronaviruses: structure and genome expression. J Gen Virol 69: 2939–2952
13. Vlasak R, Luytjes W, Spaan W, Palese P (1988) Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses. Proc Natl Acad Sci USA 85: 4526–4529
14. Yoo D, Parker MD, Song J, Cox GJ, Deregt D, Babiuk LA (1991) Structural analysis of the conformational domains involved in neutralization of bovine coronavirus using deletion mutants of the spike glycoprotein S1 subunit expressed by recombinant baculoviruses. Virology 183: 91–98

Authors' address: Dr. Georg Herrler, Institut für Virologie, Philipps-Universität Marburg, Robert-Koch-Strasse 17, D-35037 Marburg, Federal Republic of Germany.

Received October 27, 1995