VIROLOGY 195, 195–202 (1993)

Structural and Functional Analysis of the Surface Protein of Human Coronavirus OC43

FRANK KÜNKEL AND GEORG HERRLER[1]

Institut für Virologie, Philipps-Universität Marburg, 3550 Marburg, Germany

Received January 20, 1993; accepted March 9, 1993

The two surface glycoproteins S and HE of human coronavirus OC43 (HCV-OC43) were isolated from the viral membrane and purified. Only the S protein was able to agglutinate chicken erythrocytes, indicating that this viral protein is the major hemagglutinin of HCV-OC43. The receptor determinant recognized by this virus on the surface of erythrocytes is *N*-acetyl-9-*O*-acetylneuraminic acid (Neu5,9Ac_2) which is also used by bovine coronavirus for attachment to cells. By analyzing erythrocytes containing different amounts of Neu5,9Ac_2 in either of two linkage types, it was found that there are subtle differences in the affinity of both viruses for 9-*O*-acetylated sialic acid. Bovine coronavirus was more efficient in recognizing low amounts of Neu5,9Ac_2 α2,3 linked to galactose, whereas HCV-OC43 was superior with respect to the α2,6 linkage. The gene coding for the S protein of HCV-OC43 was cloned and sequenced. A large open reading frame predicts a polypeptide of 150 kDa in the unglycosylated form. A protein of about 190 kDa is expected if the 20 potential glycosylation sites are used for attachment of N-linked oligosaccharide side chains. These predictions were confirmed by *in vitro* transcription and translation of the gene in the presence or absence of canine pancreatic microsomal membranes. A high degree of sequence homology was found between the S proteins of HCV-OC43 and bovine coronavirus. Structural and functional analyses of more strains should help to identify the location of the sialic acid-binding site.

INTRODUCTION

Coronaviruses are a family of large, enveloped viruses, which contain a single-stranded RNA genome of positive polarity with a size of about 30 kilobases. The virion contains two to three structural glycoproteins: the membrane protein (M); the spike, surface, or peplomer protein (S); and the hemagglutinin/esterase protein (HE). The presence of a HE protein is a characteristic feature of a serological subgroup of coronaviruses comprising human coronavirus (HCV) strain OC43, bovine coronavirus (BCV), hemagglutinating encephalomyelitis virus (HEV), and mouse hepatitis virus (MHV). Coronaviruses, which lack an HE protein and which have no antigenic relationship to the above-mentioned viruses, include human coronavirus HCV-229E, porcine transmissible gastroenteritis virus (TGEV), and avian infectious bronchitis virus (IBV) (for a review see Siddell *et al.*, 1983; Spaan *et al.*, 1988). Human coronaviruses are known as respiratory pathogens; in addition, they have been associated with diarrhea and multiple sclerosis (McIntosh, 1974; Resta *et al.*, 1985; Burks *et al.*, 1980).

The S protein is a typical class I membrane protein, with an amino-terminal signal peptide, which is cleaved during protein processing and a carboxy-terminal hydrophobic membrane anchor. The spike proteins of BCV, MHV, and IBV are cleaved by an intracellular protease into two subunits of comparable size. Furthermore the S protein is expected to have an overall structure similar to the hemagglutinin of influenza A virus (Wilson *et al.*, 1981; De Groot *et al.*, 1987; Rasschaert and Laude, 1987). The S protein is a major target of the cellular immune response to coronaviruses and plays an important role in the initial stage of infection. It mediates the attachment of the virus to the cell surface receptors and induces the fusion of the viral and cellular membranes (reviewed by Spaan *et al.*, 1988).

Different types of virus receptors have been identified for coronaviruses. A member of the carcinoembryonic antigen family of proteins is a receptor for MHV (Williams *et al.*, 1991); TGEV and HCV-229E recognize aminopeptidase N as a receptor (Delmas *et al.*, 1992; Yeager *et al.*, 1992). BCV uses N-acetyl-9-O-acetylneuraminic acid as a receptor determinant to initiate infection (Schultze and Herrler, 1992). Results obtained with erythrocytes indicate, that the attachment of HCV-OC43 and HEV is also dependent on the presence of Neu5,9Ac_2 on the cell surface (Vlasak *et al.*, 1988; Schultze *et al.*, 1990). Although the HE protein has an affinity for 9-*O*-acetylated sialic acid, too, the S protein has been shown to be the major sialic acid-binding protein of BCV (Schultze *et al.*, 1991b).

In the present report we show that the S protein is also the major sialic acid-binding protein of HCV-

[1] To whom reprint requests should be addressed at Institut für Virologie, Philipps-Universität Marburg, Robert-Koch-Str. 17, 3550 Marburg, Germany.

0042-6822/93 $5.00

OC43. The spike protein was further analyzed (i) by determining the nucleotide sequence of the corresponding gene, (ii) by *in vitro* transcription and translation of the cloned gene, and (iii) by analyzing the efficiency in recognizing Neu5,9Ac_2.

MATERIALS AND METHODS

Viruses and cells

HCV-OC43 was obtained from the American Type Culture Collection (ATCC) and passaged three times in human rectal tumor cells (HRT-18). Three days p.i. hemagglutination titers of 256 HAU/ml were determined in the supernatant using 0.5% chicken erythrocytes for HA-titration (Schultze *et al.*, 1990). HRT-18 cells were obtained from ATCC and grown in Dulbecco's minimal essential medium supplemented with 10% fetal calf serum.

Purification of virus

Two to 4 days p.i. supernatants of infected cells were harvested and centrifuged for 30 min at 8,000 rpm and 4°. Virus was pelleted from the clarified medium by centrifugation for 2.5 hr at 25,000 rpm and 4° using an SW28 rotor. The virus pellet was homogenized in phosphate-buffered saline deficient in magnesium and calcium ions (PBS def.) by forcing it through a syringe needle and loaded on a discontinuous sucrose gradient consisting of 60, 50, and 30% sucrose in PBS def. After centrifugation for 1.5 hr at 40,000 rpm and 4° in an SW41 rotor, the virus was collected at the 30–50% interface. The virus was diluted with PBS def. and pelleted for 1 hr at 40,000 rpm and 4° in an SW41 rotor.

Purification of viral glycoproteins

The glycoproteins S and HE were isolated from the viral membrane by treatment with octylglucoside and purified by sucrose gradient centrifugation as described recently for BCV (Schultze *et al.*, 1991a,b). The glycoproteins were analyzed for HA activity and acetylesterase activity as described recently (Schultze *et al.*, 1990, 1991b).

Resialylation of erythrocytes

Neu5,9Ac_2 was attached to cell surface glycoconjugates in two linkage types by incubating erythrocytes from 1-day-old chicken with CMP-activated sialic acid and either of two sialyltransferases (Galβ1,3GalNAc α2,3-sialyltransferase or Galβ1,4GlcNAc α2,6-sialyltransferase, respectively; Boehringer-Mannheim) as described recently (Schultze *et al.*, 1990, 1992).

Preparation of viral RNA

The pellet of purified virus was homogenized in TE buffer, pH 7.6, and subjected to an overnight proteinase K digestion (Sambrook *et al.*, 1990). Viral RNA was extracted twice with phenol/chloroform 1/1 and twice with chloroform/isoamylalcohol 24/1 followed. The upper phase was precipitated by 3 *M* NaAc, pH 5.3, and EtOH at −20° for 48 hr. After pelleting and washing with 70% EtOH, vRNA was dissolved in DEPC-treated water.

DNA sequencing

The cDNA of the cloned HCV-OC43 S gene was sequenced with a modification of the dideoxy chain termination method (Sanger *et al.*, 1977) using Sequenase 2.0 (USB). Sequencing primers were between 18 and 26 nucleotides in length and were synthesized with an oligonucleotide synthesizer (Applied Biosystems). Part of them were designed on the basis of the published S gene sequence for BCV strain Mebus (Abraham *et al.*, 1990). The reaction products were analyzed on 5 and 6% acrylamide gels containing 7 *M* urea.

Cloning of the S gene

The vRNA was transcribed into DNA using MoMuLV reverse transcriptase (Boehringer-Mannheim). Because of the inability of the enzyme to produce a full-length cDNA, the S gene was transcribed in two overlapping parts (Fig. 1). Primer Reb-1 binds to nucleotides 2244–2264 and Rbx-2 to the 3'-end of the S gene. Both of them contain at the 5'-end a *Bam*HI linker, Rbx-2 an additional *Xho*I linker (Reb-1 5' CGCGGATCCTACCTACTGTGAGATCACATG 3'; Rbx-2 5' CCTCGAGGGATCCCACGAACTTAGTCGTCATGTG 3'). First-strand cDNA synthesis was carried out in a volume of 20 μl containing 10 m*M* DTT, 6 m*M* $MgCl_2$, 1 μg vRNA, 100 μg/ml BSA, 50 m*M* Tris–HCl, pH 8.3, 40 m*M* KCl, dNTP at 1 m*M* each, 0.75 μ*M* primer, 10 U MoMuLV reverse transcriptase. Denaturation for 3.5 min was followed by annealing for 30 min at room temperature. Polymerization was carried out at 42° for 1 hr. Afterward the enzyme was denatured by incubation at 99° for 7 min. This step was followed by PCR using the Gene Amp Kit and a thermal cycler (Perkin–Elmer Cetus) according to the manufacturer's instructions.

The sense primers were used at 0.15 μ*M* and contain at their 5'-end a *Bam*HI linker (Pbam-1) or an *Xho*I linker (Pxho-2), respectively. (Pbam-1: 5' CGCGGATCCGCTGCATGATGCTTAGACCA 3'; Pxho-2: 5' CCTCGAGGACCAGCATTGCTATTTCGGAATAT 3'). There were 25 cycles performed each consisting of denaturation for 1 min at 95°, annealing for 1 min at 59° and a 3 min extension step at 72°. In each cycle the polymerization step was automatically extended

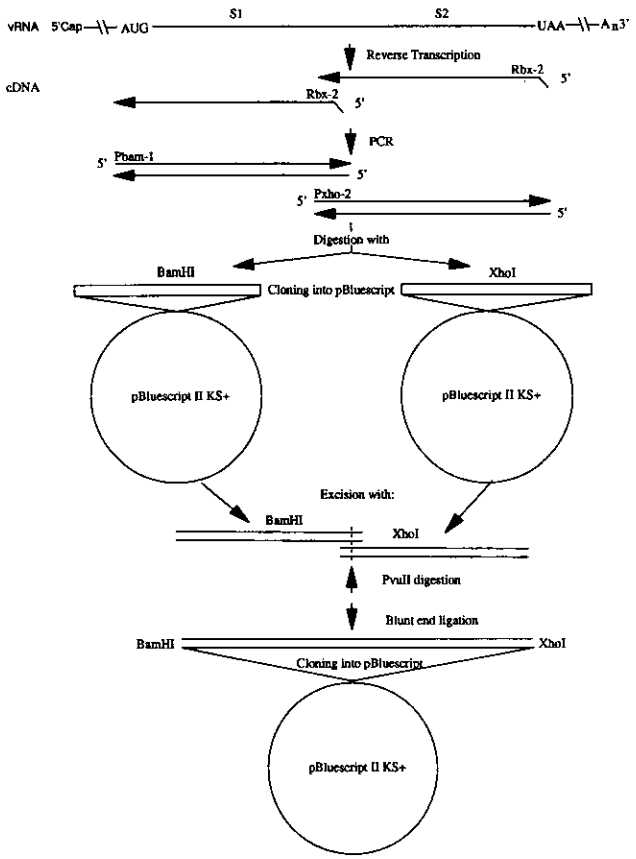

FIG. 1. Flow chart of the procedure chosen for cloning the S gene of HCV-OC43.

for 2 sec. Five independent reverse transcription and PCR reactions were done and mixed afterward. To remove the Taq polymerase which might have been bound to the reaction products (Crowe *et al.*, 1991), the whole mixture was digested with proteinase K (17 μg/ml) for 1 hr at 37°, extracted with phenol/chloroform/isoamylalcohol (50/48/2), and precipitated. After digestion with *Bam*HI or *Xho*I, respectively, the products were separated on a 1% agarose gel and eluted using Quiaex (Diagen). The two S gene fragments were subcloned into pBluescript (Stratagene). Strain XI-1 Blue (Stratagene) of *Escherichia coli* was transformed with this plasmid using standard procedures. The two overlapping inserts were prepared by digesting the recombinant plasmids with either *Bam*HI or *Xho*I, respectively, followed by agarose gel electrophoresis and elution. They were dephosphorylated with 8 units calf intestinal alkaline phosphatase (Boehringer-Mannheim) for 1 hr at 37°, extracted with phenol/chloroform/isoamylalcohol, and precipitated. Then both fragments were blunt-ended with PvuII (Fig. 1) and subjected a second time to electrophoresis and elution. Both fragments were ligated with 12.5 units T4-ligase (Boehringer-Mannheim) for 12 hr at 14°. An aliquot of this ligation mixture was incubated with *Bam*HI- and *Xho*I-digested pBluescript and 5 units T4-ligase at 16° for an additional 12 hr. After transformation, three positive clones were sequenced at both ends of the S gene (300 nucleotides in both directions). No nucleotide difference was found. One of the positive clones was sequenced completely in both directions.

In vitro transcription and translation

The recombinant plasmid pBSKS$^+$/atS8 was linearized with *Xho*I, digested with proteinase K, extracted with phenol/chloroform, precipitated, and dissolved in DEPC-treated TE-buffer, pH 7.5, at 1 μg/μl. *In vitro* transcription was done in a total volume of 25 μl, using 2 μg of linearized plasmid DNA and the mCAP kit (Stratagene), which yielded capped mRNAs. These mRNAs were subsequently treated for 5 min at 37° with 10 units RNase-free DNaseI and then extracted once with phenol/chloroform. After precipitation the mRNA was resuspended in TE buffer. *In vitro* translation was performed with a commercial kit (IN VITRO EXPRESS, Stratagene), which consisted of a pretreated white rabbit reticulocyte lysate. The reaction was performed according to the manufacturer's instructions with 1–3 μg mRNA and ^{35}S-methionine. To obtain proteins which were processed and core-glycosylated, translation reactions were done using canine pancreatic microsomal membranes (Promega). The α-mating factor from *S. cerevisiae* served as a control for signal processing and the β-lactamase from *E. coli* as a control for glycosylation (data not shown).

Immunoprecipitation

After translation, the proteins were immunoprecipitated with a rabbit antisera directed against the S protein of BCV. Incubation for 1 hr on ice was followed by incubation for 30 min at 4° with protein A–Sepharose (Serva) on a shaker and three cycles of washing at 4° with PBS def. Afterward, the immunoprecipitate was resuspended in sample buffer containing 100 m*M* DTT, heated to 95°, subjected to SDS–PAGE on 8% gels (Laemmli, 1970), and prepared for fluorography.

RESULTS

The S protein of HCV-OC43 is a hemagglutinin

In order to assign the hemagglutinating activity of HCV-OC43 to one of the surface proteins, the glycoproteins were isolated in purified form. The membrane proteins were solubilized by detergent treatment and separated by sucrose gradient centrifugation as described recently for BCV (Schultze *et al.*, 1991a,b). A peak of the S protein was detected in fraction 3, whereas HE was recovered from fraction 6. A similar sedimentation behavior has been reported for the glycoproteins of BCV (Schultze *et al.*, 1991b). These frac-

TABLE 1

ANALYSIS OF THE GLYCOPROTEINS S AND HE OF HCV-OC43 FOR ACETYLESTERASE AND HEMAGGLUTINATING ACTIVITY

Glycoprotein[a]	HA activity (HA-units/ml)	Acetylesterase (OD405)
S protein	1024	<0.01
HE protein	<2	0.56

[a] The glycoproteins of HCV-OC43 were purified by sucrose gradient centrifugation. Fractions 3 and 6 representing the peak fractions of S and HE, respectively, were dialyzed and analyzed for HA and esterase activity.

tions were analyzed for acetylesterase and hemagglutinating activity. As shown in Table 1, only fraction 6 was able to cleave *p*-nitrophenyl acetate, confirming that the HE protein is the esterase of HCV-OC43. This protein was, however, unable to agglutinate chicken erythrocytes. On the other hand, the S protein present in fraction 3 was found to be a very potent agglutinating agent. This result indicates that the S protein is the major hemagglutinin of HCV-OC43. This finding is in agreement with the data reported for BCV (Schultze *et al.*, 1991b). A difference between both viruses was observed in the electrophoretic behavior of the S protein. While the surface protein of BCV is cleaved by a cellular protease into two subunits, S1 and S2, the glycoprotein of HCV-OC43 is present on virions in the uncleaved form (not shown). As has been shown previously, cleavage is possible by incubation with trypsin (Hogue and Brian, 1986).

Comparison of the sialic acid-binding activity of HCV-OC43 and BCV

The sensitivity of the erythrocyte receptors for HCV-OC43 to acetylesterase suggested that 9-*O*-acetylated sialic acid is a receptor determinant for this virus (Vlasak *et al.*, 1988). Therefore, it can be concluded from the data presented in Table 1 that the S protein is not only the major hemagglutinin but also the major sialic acid-binding protein. Direct evidence for the importance of Neu5,9Ac_2 was obtained by resialylation of erythrocytes from 1-day-old chicken. These cells lack 9-*O*-acetylated sialic acid and are resistant to agglutination by HCV-OC43. Neu5,9Ac_2 was attached to the cell surface by incubation of neuraminidase-treated erythrocytes with sialyltransferase and CMP-activated sialic acid. In order to determine whether there is a preference of the virus for a certain linkage type, two different transferases were used, which allow to compare the following oligosaccharide structures: Neu5,9$Ac_2\alpha$2,3Galβ1,3GalNAc, and Neu5,9$Ac_2\alpha$2,6Galβ1,4GlcNAc, respectively. A further variation was obtained by applying different concentrations of CMP-sialic acid. In this way batches of erythrocytes were obtained which differed in their content of Neu5,9Ac_2. In the range of concentrations used in the present work, a proportional amount of Neu5,9Ac_2 is transferred to the cell surface (Herrler *et al.*, 1992). As shown in Table 2, HCV-OC43 was able to recognize 9-*O*-acetylated sialic acid in both the α2,3 and the α2,6 linkage. The exact amount of sialic acid transferred was not determined, because CMP-Neu5,9Ac_2 was not available in radioactive form. Using erythrocyte preparations with different amounts of 9-*O*-acetylated sialic acid on the surface, it was possible to compare the efficiency of different viruses in recognizing Neu5,9Ac_2. HCV-OC43 was found to require less sialic acid in the α2,6-linkage than BCV for agglutination of erythrocytes. The α2,3-linkage, on the other hand, was recognized more efficiently by BCV. As far as the preference for one of the two linkages is concerned, HCV-OC43 was found to resemble influenza C virus (strain Johannesburg/1/66) more than BCV (Table 2).

Sequence analysis of the S gene

To determine the sequence of the S gene of HCV-OC43, viral RNA was isolated, transcribed into cDNA, and amplified by PCR. The DNA obtained was cloned

TABLE 2

COMPARISON OF THE EFFICIENCY OF HCV-OC43, BCV, AND INFLUENZA C VIRUS (JOHANNESBURG/1/66) IN RECOGNIZING 9-*O*-ACETYLATED SIALIC ACID AS A RECEPTOR DETERMINANT

Linkage type	CMP–sialic acid[a] (nmol)	HA activity (HA units/ml) BCV	HCV-OC43	Influenza C
α2,3	4	1024	128	512
	2	512	32	128
	1	64	<2	<2
	0.5	16	<2	<2
α2,6	8	512	128	512
	4	128	64	512
	2	<2	64	512
	1	<2	32	512
	0.5	<2	32	512
	0.25	<2	32	512
Erythrocytes from adult chicken	—	1024	128	1024

[a] Neuraminidase-treated erythrocytes from 1-day-old chicken were incubated as a 50% suspension in a volume of 100 μl with either 1,5 mU of Galβ1, 3GalNAc α2,3 or 5 mU of Galβ1, 4GlcNAc α2,6-sialyltransferase and the amount of CMP-activated Neu5, 9Ac_2 indicated. The resialylated cells were used to determine the HA titer of BCV, HCV-OC43, and influenza C virus. For comparison the HA titer of the virus suspensions obtained with untreated erythrocytes from adult chicken is given at the bottom of the table.

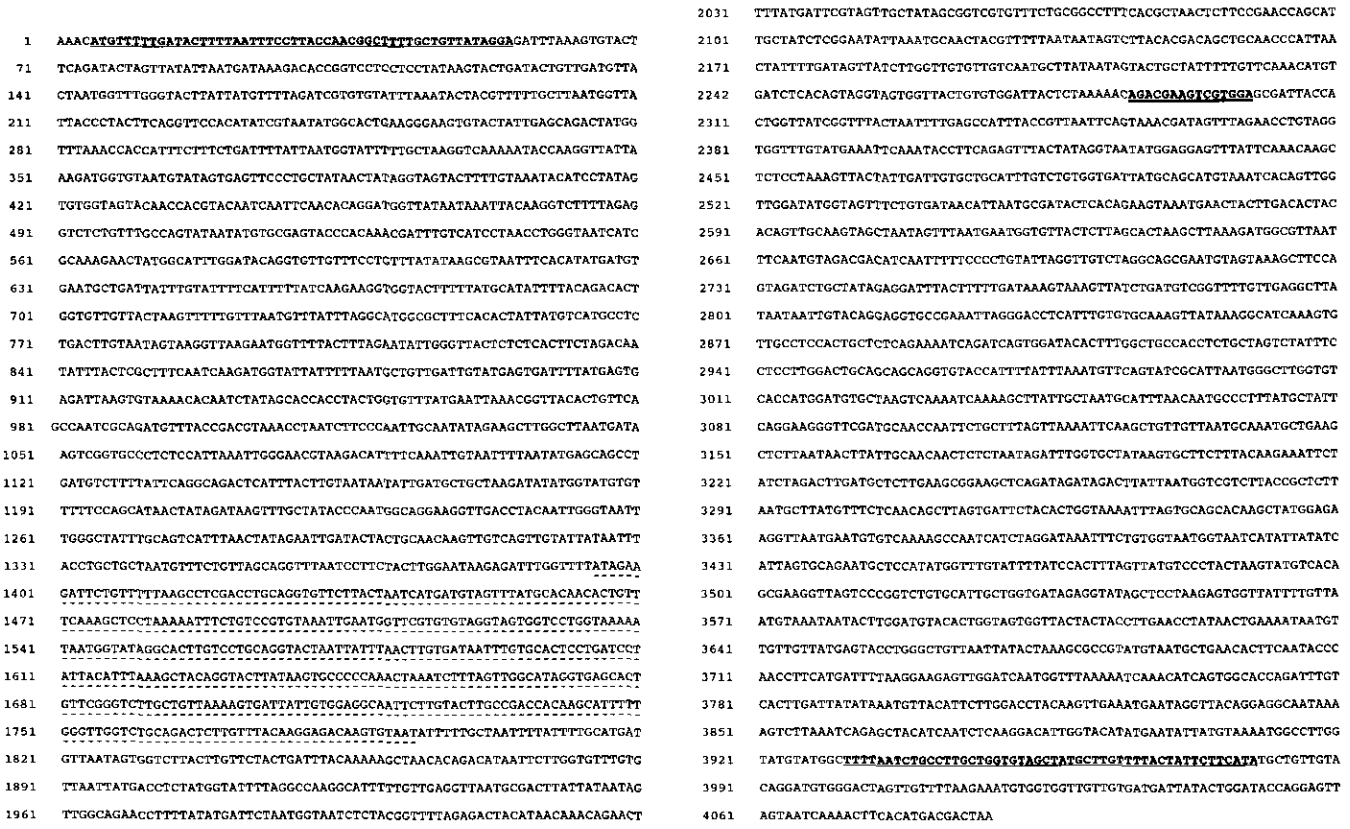

FIG. 2. Nucleotide sequence of the S gene from HCV-OC43. Sequences coding for the N-terminal signal peptide and the C-terminal membrane anchor, respectively, are underlined and printed in bold letters. The sequence coding for the potential cleavage site is bold-printed and underlined twice. The predicted polymorphic region in the S1 subunit of BCV (Yoo *et al.*, 1991) is indicated by a wavy line.

in two steps into pBluescript and sequenced. The nucleotide sequence determined for the S gene of HCV-OC43 contains a single large open reading frame (ORF), which extends from the first ATG codon (nucleotides 5–7) to the TAA stop codon at nucleotides 4086–4088 (Figs. 2 and 3). This ORF has a coding capacity of 1361 amino acids corresponding to a M_r of 150 kDa in the unprocessed form. There are 20 potential sites for N-glycosylation in the sequence of HCV-OC43 in contrast to 19 in the S protein of BCV-Mebus (Abraham *et al.*, 1990). The predicted N-glycosylation site which is missing in BCV is located at amino acid position 502–504. Assuming a mean M_r of 2100 Da per carbohydrate side chain (Hunter *et al.*, 1983), a M_r of about 190 kDa is expected for the glycosylated protein. Like the corresponding proteins of other coronaviruses, the S protein of HCV-OC43 shows the characteristics of a typical class I membrane protein. It has an N-terminal signal peptide with a hydrophobic core of 13 amino acids and a helix-breaking glycine at position 17 (Watson, 1984), suggesting that the signal peptidase cleaves between amino acids 17 and 18 (von Heijne, 1986). Close to the C-terminus, there is a stretch of hydrophobic amino acids (1322 to 1338) which is predicted to form an α-helix and to function as a membrane anchor (Chou-Fassman prediction; HUSAR-Program, EMBL Heidelberg). By comparison with the S protein of related coronaviruses, the sequence 772-Arg-Arg-Ser-Arg-Gly-778 is suggested to be the site for the proteolytic cleavage into the subunits S1 and S2 (Fig. 3). This sequence is not identical to the proteolytic cleavage site determined for different strains of BCV and MHV, which contain either two pairs of arginines or a single arginine separated by two nonbasic amino acids from a pair of arginines (Schmidt *et al.*, 1987; Luytjes *et al.*, 1988; Boireau *et al.*, 1990; Parker *et al.*, 1990; Abraham *et al.*, 1990).

HCV-OC43 has 120 amino acids, 92 in S1 and 28 in S2, which differ from the S protein of BCV-Mebus (Abraham *et al.*, 1990) including insertions and deletions. The resulting amino acid homology for S1 is 88% and for the S2 subunit 95.3%. The insertions found in HCV-OC43 are located within the first 256 amino acids, whereas deletions were found within amino acids 494–531. To see whether the cloned S gene is in a functional form, *in vitro* transcription/translation experiments were performed. The capped mRNAs obtained with the transcription kit were incubated with white rabbit reticulocyte lysates in the presence or absence of canine pancreatic microsomal membranes. [^{35}S]Methionine-labeled proteins were immunoprecipitated with a polyclonal antiserum directed against the

S protein of BCV and analyzed on 8% SDS–polyacrylamide gels (Fig. 4). Translation products were obtained with mRNAs derived from plasmids containing either the intact S gene or a fragment coding for S1. The faster migrating bands probably represent premature reaction products. The sizes of the largest proteins detected (indicated by arrows) are in good agreement with the molecular weights calculated for the glycosylated and unglycosylated form of the S protein (see above) and the S1 protein, respectively. Cleavage of S into the two subunits S1 and S2 was not observed in this system. These results indicate that the cloned S gene of HCV-OC43 is in a functional form.

DISCUSSION

A high degree of homology was found between the S genes of HCV-OC43 and BCV. The overall relatedness between these two viruses had already been indicated by oligonucleotide fingerprinting which predicted a sequence homology of greater than 96% (Lapps and Brian, 1985). Compared to the N protein with 97.5% (Kamahora *et al.*, 1989) and the HE protein with 95% amino acid homology (Zhang *et al.*, 1992), the S protein of HCV-OC43 is less conserved (91.2%), especially the S1 subunit with a homology of 88%. The fact that the S2 portion of the molecule is more conserved than S1 was also found in other coronaviruses, e.g., in IBV, MHV, and FIPV (for a review see Spaan *et al.*, 1988). Interestingly, HCV-OC43 shows insertions and deletions in the S1 part of the molecule, which, however, counterbalance each other and lead only to a loss of two amino acids compared to BCV. The deletions are all located in a part of S1 (amino acids 455–599), which in the case of BCV has been described as a polymorphic region (Yoo *et al.*, 1991). It is the same region which in MHV-JHM is almost completely and in MHV-A59 partly missing (Schmidt *et al.*, 1987; Luytjes *et al.*, 1988). Taken all the sequence changes in the S protein of HCV-OC43 and BCV together, 25.5% of the amino acids are different within the polymorphic region compared to 12.2% in the rest of the S1 subunit and only 7.7% in S2.

The S protein of most strains of BCV and MHV is cleaved into the subunits S1 and S2 by a cellular protease acting at the C-terminus of the sequence N-Arg-Arg-Xxx-Arg-Arg-C (Schmidt *et al.*, 1987; Luytjes *et al.*, 1988; Boireau *et al.*, 1990; Parker *et al.*, 1990; Abraham *et al.*, 1990; Zhang *et al.*, 1990). The substitution of Gly for Arg at position -1 of the cleavage site of HCV-OC43 may be responsible for the reduced cleavability reported for the spike protein of this virus (Hogue and Brian, 1986). We also observed that S was only cleaved to an extent of 10%, when OC43 was propagated in

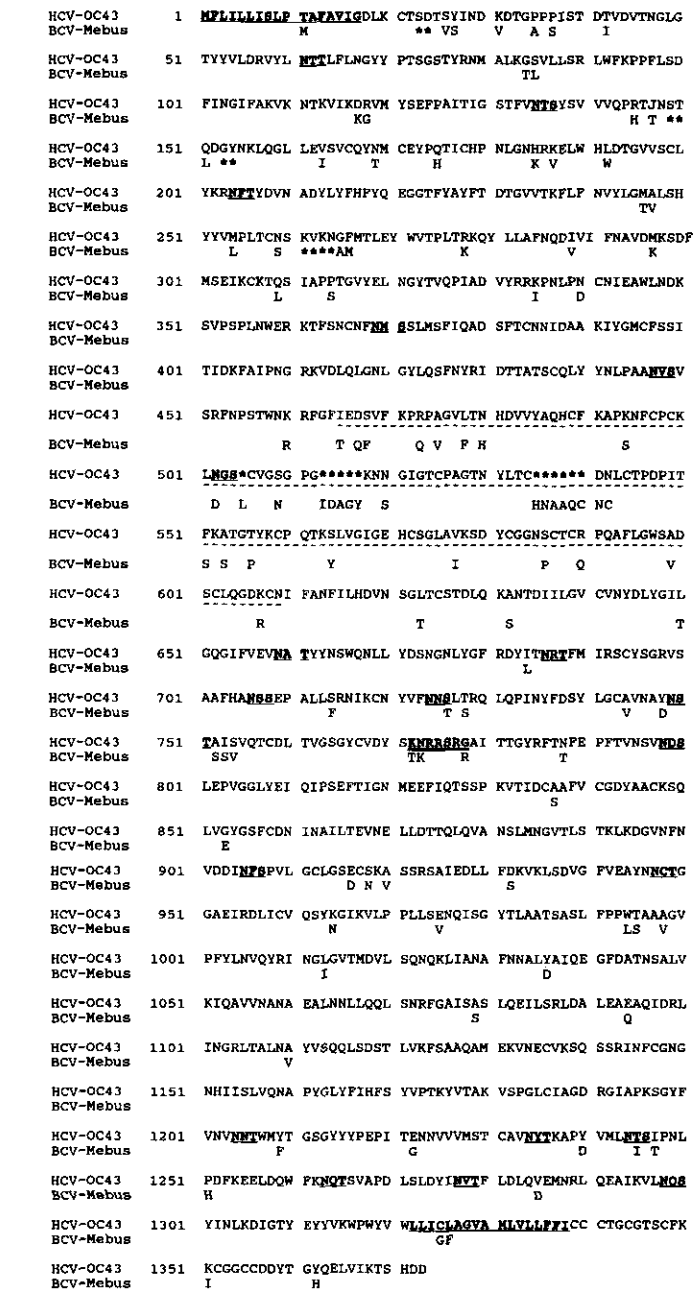

FIG. 3. The predicted amino acid sequence of the S protein from HCV-OC43 compared to strain Mebus of BCV (Abraham *et al.*, 1990). For HCV-OC43, the complete sequence is presented; for BCV, only the amino acid differences are shown. The amino-terminal signal peptide and the carboxy-terminal membrane 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 (Yoo *et al.*, 1991) is marked by a wavy line. The proteolytic cleavage site is printed in bold letters and underlined twice.

HRT-18 cells in contrast to BCV grown in this cell line, which showed nearly 80% cleavage of the S protein (data not shown). In the case of influenza viruses, the sequence Arg-Xxx-Lys/Arg-Arg has been determined as the consensus motif for the cleavage of the hemag-

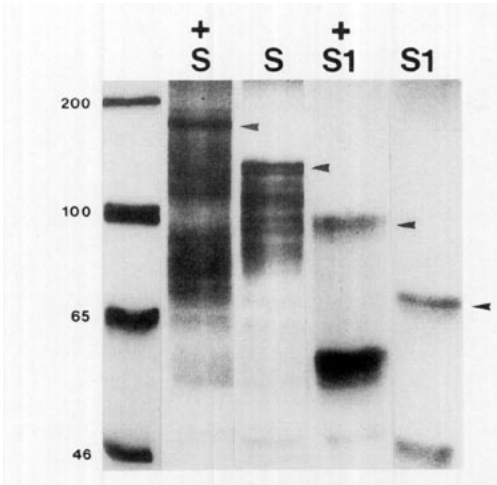

FIG. 4. Analysis of the polypeptides obtained by *in vitro* transcription/translation of the S gene of HCV-OC43 and a fragment coding for the S1 subunit. The ^{35}-methionine-labeled proteins were immunoprecipitated with a polyclonal antiserum directed against the S protein of BCV and subjected to SDS–PAGE under denaturing conditions. The samples incubated in the presence of canine pancreatic microsomal membranes are indicated by (+). The size of molecular weight markers (kDa) is indicated on the left.

glutinin by the cellular protease furin (Vey *et al.*, 1992; Stienecke-Gröber *et al.*, 1992). This consensus sequence can also account for the differential cleavage of the S protein of HCV-OC43, BCV, and MHV, respectively. The proteolytic cleavage of the influenza hemagglutinin is an absolute requirement for the fusion activity, and consequently for the infectivity, of the virus. As far as coronaviruses are concerned, proteases may have an enhancing effect (St. Cyr-Coats *et al.*, 1988). However, fusion activity and infectivity appear not to be strictly dependent on the proteolytic cleavage of S. The S protein of TGEV and related viruses completely lacks a motif for proteolytic cleavage and is only found in the uncleaved form (Rasschaert and Laude, 1987). Moreover, after site-directed mutagenesis of the cloned S gene, the surface protein of MHV was obtained in an uncleaved form, which nevertheless was fusion-active (Stauber *et al.*, 1993).

We presented evidence that the S protein of HCV-OC43 is a hemagglutinin which interacts with Neu5,9Ac$_2$ present on the surface of erythrocytes. The HE protein was found to be unable to agglutinate chicken erythrocytes. Similar findings have been reported recently for BCV (Schultze *et al.*, 1991b), where the HE protein only agglutinates mouse or rat erythrocytes which are very rich in surface-bound 9-*O*-acetylated sialic acid. Thus, the S protein is the major sialic acid-binding protein for both HCV-OC43 and BCV. This may not be surprising because of the sequence similarity. There are, however, subtle differences in the affinity of both viruses for sialic acid. While BCV is more efficient in recognizing Neu5,9Ac$_2$ in an α2,3-linkage, HCV-OC43 is more effective toward an α2,6-linkage. In this respect, HCV-OC43 was similar to influenza C virus. As both viruses infect humans, one might speculate that the difference between these viruses and BCV in the preference for a certain linkage type is related to the host tropism. However, more strains have to be analyzed for conclusive evidence. Assuming that the binding site for 9-*O*-acetylated sialic acid is part of the S1 subunit, there are 92 amino acid changes which could be responsible for the difference between HCV-OC43 and BCV. This number is too high to allow a conclusion about the receptor-binding site. However, by extending this approach to more strains, it should be possible to get a clue about the location of the binding site for 9-*O*-acetylated sialic acid.

ACKNOWLEDGMENTS

The technical assistance of Brigit Döll is gratefully acknowledged. We thank Prof. Dr. Dr. R. Brossmer for kindly providing CMP-Neu5,9Ac$_2$. Part of this work was conducted by F.K. in partial fulfillment of the requirements for the Dr. rer. nat. degree from FB17, Philipps-Universität, Marburg. Financial support was provided by a grant from Deutsche Forschungsgemeinschaft (He 1168/2-2).

REFERENCES

ABRAHAM, S., KIENZLE, T. E., LAPPS, W., and BRIAN, D. A. (1990). Deduced sequence of the bovine coronavirus spike protein and identification of the internal proteolytic cleavage site. *Virology* **176,** 296–301.

BINNS, M. M., BOURSNELL, M. E. G., TOMLEY, F. M., and BROWN, T. D. K. (1986). Comparison of the spike precursor sequences of coronavirus IBV strains M41 and 6/82 with that of IBV Beaudette. *J. Gen. Virol.* **67,** 2825–2831.

BOIREAU, P., CRUCIERE, C., and LAPORTE, J. (1990). Nucleotide sequence of the glycoprotein S gene of bovine enteric coronavirus and comparison with the S proteins of two mouse hepatitis virus strains. *J. Gen. Virol.* **71,** 487–492.

BURKS, J. S., DE BALD, B. D., JANKOVSKY, L. C., and GERDES, J. (1980). Two coronaviruses isolated from central nervous system tissue of two multiple sclerosis patients. *Science* **209,** 933–936.

CROWE, J. S., COOPER, H. J., SMITH, M. A., SIMS, M. J., PARKER, D., and GEWERT, D. (1991). Improved cloning efficiency of polymerase chain reaction (PCR) products after proteinase K digestion. *Nucleic Acids Res.* **19,** 184.

DE GROOT, R. J., LUYTJES, W., HORZINEK, K. V., VAN DER ZEIJST, B. A. M., SPAAN, W. J. M., and LENSTRA, J. A. (1987). Evidence for a coiled-coil structure in the spike proteins of coronaviruses. *J. Mol. Biol.* **196,** 963–966.

DELMAS, B., GELFI, J., L'HARIDON, R., VOGEL, L. K., SJÖSTRÖM, H., NOREN, O., and LAUDE, H. (1992). Aminopeptidase N is a major receptor for the enteropathogenic coronavirus TGEV. *Nature* **357,** 417–420.

HERRLER, G., GROSS, H.-J., IMHOF, A., BROSSMER, R., MILKS, G., and PAULSON, J. C. (1992). A synthetic sialic acid analogue is recognized by influenza C virus as a receptor determinant but is resistant to the receptor-destroying enzyme. *J. Biol. Chem.* **267,** 12501–12505.

HOGUE, B. G., and BRIAN, D. A. (1986). Structural proteins of human coronavirus OC43. *Virus Res.* **5,** 131–144.

HUNTER, E., HILL, E., HARDWICK, M., BROWN, A., SCHWARTZ, D. E., and TIZARD, R. (1983). Complete sequence of the Rous sarcoma virus env gene: Identification of structural and functional regions of its product. *J. Virol.* **46,** 920–936.

KAMAHORA, T., SOE, L. H., and LAI, M. M. C. (1989). Sequence analysis of nucleocapsid gene and leader RNA of human coronavirus OC43. *Virus Res.* **12,** 1–9.

LAEMMLI, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227,** 680–685.

LAPPS, W., and BRIAN, D. A. (1985). Oligonucleotide fingerprints of antigenically related bovine coronavirus and human coronavirus OC43. *Arch. Virol.* **86,** 101–108.

LUYTJES, W., STURMAN, L. S., BREDENBEEK, P. J., CHARITE, J., VAN DER ZEIJST, B. A. A., HORZINEK, M. C., and SPAAN, W. J. M. (1987). Primary structure of the glycoprotein E2 of coronavirus MHV-A59 and identification of the trypsin cleavage site. *Virology* **161,** 479–487.

MCINTOSH, K. (1974). Coronaviruses: A comparative review. *Curr. Top. Microbiol. Immunol.* **63,** 85–129.

PARKER, M. D., YOO, D., COX, G. J., and BABIUK, L. A. (1990). Primary structure of the S peplomer gene of bovine coronavirus and surface expression in insect cells. *J. Gen. Virol.* **71,** 263–270.

RAABE, T., SCHELLE-PRINZ, B., and SIDDELL, S. (1990). Nucleotide sequence of the gene encoding the spike glycoprotein of human coronavirus HCV 229E. *J. Gen. Virol.* **71,** 1065–1073.

RASSCHAERT, D., and LAUDE, H. (1987). The predicted primary structure of the peplomer protein E2 of the porcine coronavirus transmissible gastroenteritis virus. *J. Gen. Virol.* **68,** 1883–1890.

RESTA, S., LUBY, J. P., ROSENFELD, C. R., and SIEGEL, J. D. (1985). Isolation and propagation of a human enteric coronavirus. *Science* **229,** 978–981.

SAMBROOK, J., FRITSCH, E. F., and MANIATIS, T. (1990). "Molecular Cloning: A Laboratory Manual," 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

SANGER, F., NICKLEN, S., and COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.

SCHMIDT, I., SKINNER, M., and SIDDELL, S. (1987). Nucleotide sequence of the gene encoding the surface projection glycoprotein of coronavirus MHV-JHM. *J. Gen. Virol.* **68,** 47–56.

SCHULTZE, B., GROSS, H.-J., BROSSMER, R., KLENK, H.-D., and 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.

SCHULTZE, B., WAHN, K., KLENK, H.-D., and HERRLER, G. (1991a). Isolated HE protein from hemagglutinating encephalomyelitis virus and bovine coronavirus has receptor-destroying and receptor-binding activity. *Virology* **180,** 221–228.

SCHULTZE, B., GROSS, H.-J., BROSSMER, R., and HERRLER, G. (1991b). The S protein of bovine coronavirus is a hemagglutinin recognizing 9-*O*-acetylated sialic acid as a receptor determinant. *J. Virol.* **65,** 6232–6237.

SCHULTZE, B., and HERRLER, G. (1992). Bovine coronavirus uses N-acetyl-9-O-acetylneuraminic acid as a receptor determinant to initiate the infection of cultured cells. *J. Gen. Virol.* **73,** 901–906.

SCHULTZE, B., CAVANAGH, D., and HERRLER, G. (1992). Neuraminidase treatment of avian infectious bronchitis coronavirus reveals a hemagglutinating activity that is dependent on sialic acid-containing receptors on erythrocytes. *Virology* **189,** 792–794.

SIDDELL, S., WEGE, H., and TER MEULEN, V. (1983). The biology of coronaviruses. *J. Gen. Virol.* **64,** 761–776.

SPAAN, W., CAVANAGH, D., and HORZINEK, M. C. (1988). Coronaviruses: structure and genome expression. *J. Gen. Virol.* **69,** 2939–2952.

STAUBER, R., PFLEIDERER, M., and SIDDELL, S. (1993). Proteolytic cleavage of the murine coronavirus surface glycoprotein is not required for fusion activity. *J. Gen. Virol.,* **74,** 183–191.

ST. CYR-COATS, K., STORZ, J., HUSSAIN, K. A., and SCHNORR, K. L. (1988). Structural proteins of bovine coronavirus strain L9: Effects of the host cell and trypsin treatment. *Arch. Virol.* **103,** 35–45.

STIENEKE-GRÖBER, A., VEY, M., ANGLIKER, H., SHAW, E., THOMAS, G., ROBERTS, C., KLENK, H.-D., and GARTEN, W. (1992). Influenza virus hemagglutinin with multibasic cleavage site is activated by furin, a subtilisin-like endoprotease. *EMBO J.* **11,** 2407–2414.

VEY, M., ORLICH, M., ADLER, S., KLENK, H.-D., ROTT, R., and GARTEN, W. (1992). Hemagglutinin activation of pathogenic avian influenza viruses of serotype H7 requires the protease recognition motif R-S-K/R-R. *Virology* **188,** 408–413.

VLASAK, R., LUYTJES, W., SPAAN, W., and PALESE, P. (1988). Human and bovine coronaviruses recognize sialic acid containing receptors similar to those of influenza C virus. *Proc. Natl. Acad. Sci. USA* **85,** 4526–4529.

VON HEINJE, G. (1984). How signal sequences maintain cleavage specificity. *J. Mol. Biol.* **173,** 243–251.

WATSON, M. E. E. (1984). Compilation of published signal sequences. *Nucleic Acids Res.* **12,** 5145–5146.

WILLIAMS, R. K., JIANG, G. S., and HOLMES, K. V. (1991). Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. *Proc. Natl. Acad. Sci. USA* **88,** 5533–5536.

WILSON, I. A., SKEHEL, J. J., and WILEY, D. C. (1981). Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **289,** 366–373.

YEAGER, C. L., ASHMUN, R. A., WILLIAMS, R. K., CARDELLICHIO, C. B., SHAPIRO, L. H., LOOK, A. T., and HOLMES, K. V. (1992). Human aminopeptidase N is a receptor for human coronavirus 229E. *Nature* **357,** 420–422.

YOO, D., PARKER, M. D., SONG, J., COX, G. J., DEREGT, D., and BABIUK, L. A. (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.

ZHANG, X., KOUSOULAS, K. G., and STORZ, J. (1992). The hemagglutinin/esterase gene of the human coronavirus strain OC43: Phylogenetic relationships to bovine and murine coronaviruses and influenza C virus. *Virology* **186,** 318–323.