Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

VIROLOGY 222, 227–235 (1996)
ARTICLE NO. 0413

Characterization of a 105-kDa Polypeptide Encoded in Gene 1 of the Human Coronavirus HCV 229E

CARSTEN GRÖTZINGER,*,† GERHARD HEUSIPP,† JOHN ZIEBUHR,† UTE HARMS,*,† JOCHEN SÜSS,* and STUART G. SIDDELL†,[1]

*Department of Viral Zoonoses, Federal Institute for Health Protection of Consumers and Veterinary Medicine, 12277 Berlin, Germany; and †Institute of Virology, University of Würzburg, 97078 Würzburg, Germany

Received February 21, 1996; accepted May 22, 1996

Gene 1 of the human coronavirus HCV 229E encompasses approximately 20.7 kb and contains two overlapping open reading frames, ORF 1a and ORF 1b. The downstream ORF 1b is expressed by a mechanism involving (−1) ribosomal frameshifting. Translation of mRNA 1, which is thought to be equivalent to the viral genomic RNA, results in the synthesis of two large polyproteins, pp1a and pp1ab. These polyproteins contain motifs characteristic of papain-like and 3C-like proteinases, RNA-dependent RNA polymerases, helicases, and metal-binding proteins. In this study, we have produced pp1ab-specific monoclonal antibodies and have used them to detect an intracellular, 105-kDa viral polypeptide that contains the putative RNA polymerase domain. Furthermore, using *trans* cleavage assays with bacterially expressed HCV 229E 3C-like proteinase, we have demonstrated that the 105-kDa polypeptide is released from pp1ab by cleavage at the dipeptide bonds Gln-4068/Ser-4069 and Gln-4995/Ala-4996. These data contribute to the characterization of coronavirus 3C-like proteinase-mediated processing of pp1ab and provide the first identification of an HCV 229E ORF 1ab-encoded polypeptide in virus-infected cells.

INTRODUCTION

Human coronaviruses (HCV) are a major cause of upper respiratory tract illness in man, and their involvement in lower respiratory tract illness and gastroenteritis is probable (Zhang *et al.,* 1994; Myint, 1995; Johnston and Holgate, 1996). To date, two antigenic groups of HCV have been identified and they are represented by the prototype strains HCV 229E and HCV OC43. Since HCV 229E is easier to isolate and propagate in tissue culture, it has been more extensively studied.

The HCV 229E genome is a positive-strand RNA of approximately 27,000 nucleotides. Gene 1, which is located at the 5′ end of the genome, is composed of two large, overlapping ORFs, ORF 1a and ORF 1b (Herold *et al.,* 1993). The upstream ORF 1a encodes a polyprotein pp1a, with a calculated molecular mass of 454 kDa. The downstream ORF 1b is expressed as a fusion protein with pp1a by a mechanism involving (−1) ribosomal frameshifting (Herold *et al.,* 1993; Herold and Siddell, 1993). The ORF 1ab gene product has a calculated molecular mass of 754 kDa and is referred to as polyprotein 1ab or pp1ab.

The analysis of murine hepatitis virus temperature-sensitive mutants provides strong evidence that the synthesis of the coronavirus RNA polymerase involves the activity of proteinases that cleave pp1a and pp1ab into smaller, functional polypeptides (Leibowitz *et al.,* 1982; Schaad *et al.,* 1990; Baric *et al.,* 1990; Fu and Baric, 1994). Sequence motifs characteristic of both papain-like and picornavirus 3C proteinases have been identified in the predicted sequences of pp1a and pp1ab and motifs that are associated with replicative functions (such as RNA polymerase, helicase, and metal-binding domains) have been conserved in the pp1ab sequence (Eleouet *et al.,* 1995; Gorbalenya *et al.,* 1989; Herold *et al.,* 1993; Lee *et al.,* 1991). An obvious goal of research in this area is, therefore, to define the role of coronavirus proteinases in the processing of the polymerase gene products.

Recently, considerable progress has been made in characterizing the activities of coronaviral 3C-like proteinases. Lu *et al.* (1995) have shown that the MHV 3C-like proteinase synthesized *in vitro* is able to cleave MHV pp1a and pp1ab at the dipeptide bond Gln-3334/Ser-3335[2] and they have identified Cys-145 and His-41 as proteinase residues that are essential for this activity. Using a vaccinia virus-T7 expression system, Liu *et al.* (1994) have shown that the IBV 3C-like proteinase is involved in the processing of IBV pp1ab to produce a 100-kDa polypeptide that includes the RNA polymerase motif. Mutational analyses suggest that the 100-kDa poly-

[1] To whom correspondence and reprint requests should be addressed at Institute of Virology, University of Würzburg, Versbacher Strasse 7, 97078 Würzburg, Germany. Fax: +49-931-201-3934. E-mail: siddell@rzbox.uni-wuerzburg.de.

[2] The numbering of RNA polymerase amino acids used here follows the conventions adopted by individual authors.

0042-6822/96 $18.00

peptide is released by cleavage at the dipeptide bonds Gln-3928 (1a)/Ser-3929 (1a) and Gln-891 (1b)/Ser-892 (1b). The IBV 3C-like proteinase residues Cys-2922 and His-2820 are essential for this activity (Liu and Brown, 1995). In our own studies, both the autoproteolytic activity and *trans* cleavage activity of a bacterially expressed HCV 229E 3C-like proteinase have been demonstrated. Sequence analysis has shown that the amino-terminal cleavage of the HCV 229E proteinase domain occurs at the HCV pp1a and pp1ab dipeptide bond Gln-2965/Ala-2966. Finally, we have also been able to detect a 34-kDa, 3C-like proteinase polypeptide in HCV 229E-infected cells (Ziebuhr *et al.,* 1995).

In the current study, we report the detection of a 105-kDa polypeptide in HCV 229E-infected cells using monoclonal antibodies specific for epitopes near the polymerase domain of pp1ab. To further characterize the termini of this polypeptide, we have used bacterially synthesized, HCV 229E 3C-like proteinase to cleave both *in vitro* synthesized and bacterially synthesized substrates and we have done N-terminal sequencing of the reaction products. These studies lead us to conclude that the RNA-dependent RNA polymerase domain of the human coronavirus 229E is expressed within a 105-kDa polypeptide encompassing amino acids 4069–4995 of the polyprotein 1ab.

MATERIALS AND METHODS

Cells and viruses

MRC-5 cells (ECACC 84101801) and HeLa cells (ATCC CCL2) overexpressing human aminopeptidase N (HeLa-CD13; Siddell, unpublished) were grown in monolayers in minimal essential medium with Earle's salts containing 10% heat-inactivated fetal bovine serum, 25 m*M* HEPES, GLUTAMAX I, antibiotics, and nonessential amino acids (MEM10). The HCV 229E isolate used in these studies has been described previously (Raabe *et al.,* 1990). The virus was propagated in monolayers of MRC-5 cells and concentrated virus stocks were prepared by NaCl–polyethylene glycol precipitation as described (Wege *et al.,* 1979).

Construction of bacterial expression plasmids

To generate the plasmid pUR-K3, DNA of the vector pUR290 (Ruther and Muller-Hill, 1983) was digested with *Bsp*106I, and the 3256-base-pair (bp) fragment was isolated and ligated to a 1518-bp *Aci*I–*Bsp*106I fragment of DNA from the plasmid pBS-T16D8 (Herold *et al.,* 1993). The plasmid pUR-K3 encodes a fusion protein containing the amino-terminal 280 amino acids of *Escherichia coli* β-galactosidase, 506 amino acids corresponding to amino acids 4084 to 4589 of the HCV 229E pp1ab, and 6 artifactual amino acids at the carboxyl terminus.

To generate the plasmid pT7-F10, the 1518-bp, *Aci*I–*Bsp*106I DNA restriction fragment from plasmid pBS-T16D8 was ligated to *Bsp*106I-linearized DNA of plasmid pT7/7 (Tabor and Richardson, 1985). This construct encodes a fusion protein of 537 amino acids; 19 artifactual amino acids at the amino terminus, 506 amino acids corresponding to amino acids 4084 to 4589 of the HCV 229E polyprotein 1ab, and 12 additional vector-derived amino acids at the carboxyl terminus.

Expression and purification of bacterial fusion proteins

Plasmids pUR290 and pUR-K3 were used to transform *E. coli* BMH 71-18 cells, and plasmid pT7-F10 was used to transform *E. coli* BL21 (DE3) cells. Induction of expression, preparation of cell lysates, separation on preparative SDS–polyacrylamide gels, electroelution of fusion proteins, and dialysis were essentially performed as described (Ziebuhr *et al.,* 1995).

Production of monoclonal antibodies

Three-month-old female BALB/c mice were inoculated six times intraperitoneally with the partially purified K3 fusion protein. Fusion, screening, and cloning of hybridoma lines were performed by using standard techniques (Harlow and Lane, 1988). Hybridomas were screened for HCV 229E polymerase protein-specific antibodies by enzyme-linked immunosorbent assay, using as capture antigens K3 fusion protein, F10 fusion protein, or a crude lysate of proteins from bacteria overexpressing β-galactosidase. Hybridomas secreting polymerase protein-specific antibodies were cloned three times by limiting dilution. For the production of murine monoclonal antibody at concentrations of approximately 1 mg/ml, the TECNOMOUSE hollow fiber bioreactor system (INTEGRA BIOSCIENCES, Fernwald, Germany) was used. Monoclonal antibody isotypes were determined using the Sigma Immunotype kit (092H-4800; Sigma, St. Louis, MO).

Epitope mapping

An antibody–peptide binding assay was done using a library of peptides synthesized on a cellulose membrane support (JERINI BIO-TOOLS, Berlin, Germany). The membranes were incubated with concentrated tissue culture supernatant, washed, and immunostained with an alkaline phosphatase-conjugated, secondary antibody using standard techniques (Harlow and Lane, 1988).

Virus infection, cell lysis, and Western blotting

HeLa–CD13 cells were infected with HCV 229E at an m.o.i. of 10 PFU per cell. At different times after infection, the cells were washed with ice-cold PBS and resuspended in PBS. After centrifugation (800 *g,* 4°, 10 min), the cells were resuspended in lysis buffer (20 m*M* Tris–Cl, pH 7.5, 150 m*M* NaCl, 0.5% sodium deoxycholate,

1% Nonidet-P40, 0.1% SDS) and left for 15 min on ice. Subsequently, the lysates were centrifuged (14,000 *g,* 4°, 10 min) and the supernatants were mixed with 1 volume of 2× SDS–PAGE sample buffer. Proteins were separated on SDS–12.5% polyacrylamide gels and Western blotting and immunostaining were performed using standard procedures (Gallagher *et al.,* 1991).

Preparation of proteinase substrates

In vitro synthesized substrates. Two different strategies were used to express, *in vitro,* HCV 229E pp1ab polypeptides that would serve as substrates for *trans* cleavage assays. First, polyadenylated RNA from HCV 229E-infected cells was isolated and reverse transcription using the oligonucleotide, 5′ TCACATTGCATAAGCATCACTAACAT 3′, was done as described (Herold *et al.,* 1996). Subsequently, a cDNA fragment containing nucleotides 14,599–16,069 of the HCV 229E genome was amplified by a PCR reaction using the RT primer and the oligonucleotide 5′ TAATACGACTCACTATAGGGCCTTGGTTTTGAGCGTTAACTCGTCA 3′. This second primer contains the bacteriophage T7 promoter, which allows for the *in vitro* transcription of the purified PCR product using T7 RNA polymerase without subcloning. The synthetic mRNA produced encodes a polypeptide that encompasses amino acids 4818 to 5259 of pp1ab.

A second pp1ab polypeptide, encompassing the amino acids 3934 to 4237, was synthesized from mRNA derived from the plasmid pBSS3. To introduce mutations into the slippery sequence and the frameshifting site located in the RNA polymerase locus of HCV 229E, recombination PCR technology was used as previously described (Ziebuhr *et al.,* 1995). To generate the required PCR fragments, four oligonucleotides, 5′ TCTTCAGCATCTTTTACTTTC 3′, 5′ GAAAGTAAAAGATGCTGAAGA 3′, 5′ GTTATCTGAATAGGGTCCGGGGCTCTAGTGCCGCT 3′, 5′ CGGACCCTATTCAGATAACTGTTATCAAAACTTTG 3′, and pBS-T16D8 template DNA were used. The mutagenesis procedure exchanged nucleotide 12515 (T for C), nucleotide 12517 (A for G), and nucleotides 12520 to 12522 (CGA for TAG) in relation to the HCV 229E genomic sequence. These mutations and the introduction of an additional nucleotide (G) at position 12523 link the coding sequences of ORFs 1a and 1b to create an ORF that encodes a polypeptide corresponding to a region of the predicted polyprotein 1ab. The mutations introduced to the resultant plasmid, pBS-T16D8ΔFS, were verified by DNA sequence analysis.

Next, the T7 expression plasmid pBS-T was generated by digestion of pBluescript II (KS^+) DNA (Stratagene GmbH, Heidelberg, Germany) with *Sac*I and *Bam*HI and ligation to the annealed oligonucleotides 5′ GATCCATGGTGGACTGCGAGCT 3′ and 5′ CGCAGTCCACCATG 3′. This plasmid provides a translation start codon in an optimal context (Kozak, 1984). Finally, a 908-bp DNA fragment was generated from pBS-T16D8ΔFS DNA by a PCR reaction with the oligonucleotides 5′ ACGGGATCCAATGGCTGGCAAACAGACTGAGTTT 3′ and 5′ ACACACGGTGTATGTCCTCATT 3′. The DNA fragment was treated with T4 DNA polymerase, phosphorylated with polynucleotide kinase, digested with *Bam*HI, and ligated with *Bam*HI/*Eco*RV-digested pBS-T. Synthetic mRNA derived from the resultant plasmid, pBSS3, encodes a polypeptide encompassing amino acids 3934 to 4237 of pp1ab.

Bacterially synthesized substrates. To construct pMal-Npol, pBSS3 DNA was digested with *Bam*HI, treated with T4 DNA polymerase, and digested with *Hin*dIII. The 923-bp DNA fragment was isolated and ligated with *Xmn*I/*Hin*dIII-digested pMal-c2 DNA (800-64S; New England Biolabs; Schwalbach, Germany). pMal-Cpol encodes amino acids 3934 to 4237 of the HCV 229E polyprotein 1ab fused to the *E. coli* maltose-binding protein (MBP).

To generate plasmid pMal-Cpol, a cDNA fragment containing nucleotides 14,599 to 16,069 of the HCV 229E genomic sequence was amplified by PCR using oligonucleotides 5′ CGGGATCCCGTTTTGAGCGTTAACTCG 3′ and 5′ TTGCATAAGCATCACTAACATACACT 3′ and pBS-T13A5 DNA template (Herold *et al.,* 1993). The PCR product was treated with T4 DNA polymerase and polynucleotide kinase, digested with *Hin*cII, and ligated with *Xmn*I-digested pMal-c2 DNA. pMal-Cpol encodes amino acids 4774 to 5259 of the HCV 229E polyprotein 1ab fused to the *E. coli* maltose-binding protein.

Trans cleavage assay using in vitro synthesized substrates. Capped mRNA was synthesized by *in vitro* transcription of PCR-amplified DNA or *Hin*dIII-linearized pBSS3 DNA using standard protocols. Translation reactions were done in a rabbit reticulocyte lysate (Promega/Serva, Heidelberg, Germany) in the presence of [^{35}S]-methionine (Herold and Siddell, 1993). *Trans* cleavage assays were done as previously described (Ziebuhr *et al.,* 1995) except that 10 μg of HCV 229E 3C-like proteinase (1 mg/ml in 20 m*M* Bis–Tris–Cl, pH 7.0), purified by anion exchange and hydrophobic interaction chromatography, was used.

Trans cleavage assay using bacterially synthesized substrates and amino-terminal protein sequence analysis. The plasmids pMal-Npol and pMal-Cpol were used to transform competent *E. coli* TB1 cells. The recombinant proteins were expressed and affinity-purified as described previously (Ziebuhr *et al.,* 1995). The fusion proteins were incubated for 24 hr at 33° with 10 μg of 3C-like proteinase. Following incubation, the cleavage products were separated on SDS–15% polyacrylamide gels and transferred electrophoretically to a polyvinylidene difluoride (PVDF) membrane (Bio-Rad Laboratories, München, Germany). The membrane was stained in 0.025% (w/v) Coomassie brilliant blue R-250 in 40% (v/v) methanol and destained in 50% (v/v) methanol. The appropriate areas of the membrane were isolated and the N-terminal se-

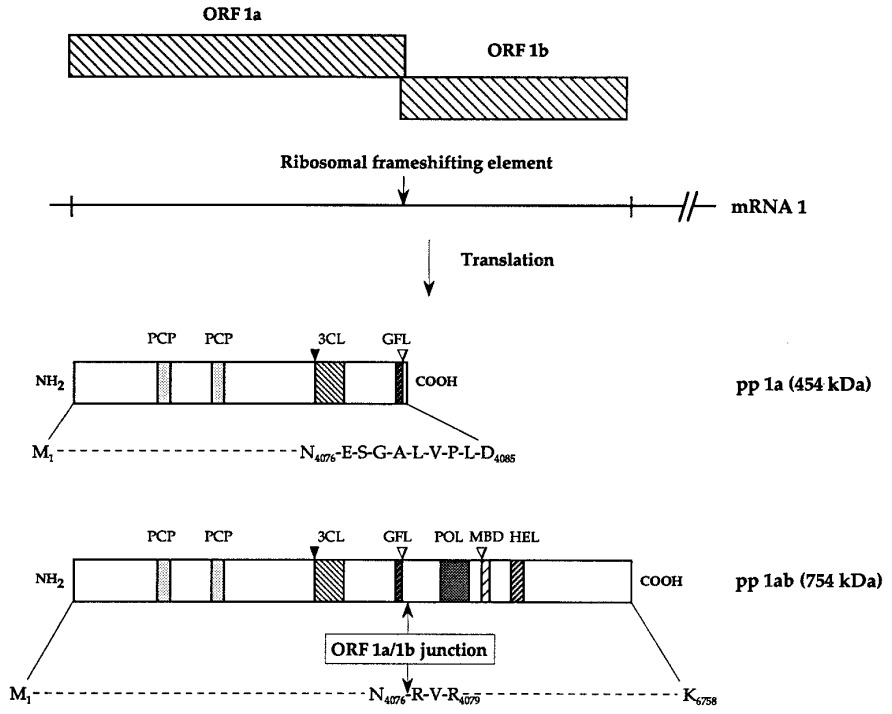

FIG. 1. Schematic presentation of HCV 229E gene 1 expression. Putative functional domains within polyproteins pp1a and pp1ab (Gorbalenya *et al.,* 1989; Herold *et al.,* 1993) are shown. The amino-terminal cleavage site of the 3C-like proteinase domain (Ziebuhr *et al.,* 1995) is designated by solid triangles, and the predicted 3C-like proteinase-mediated cleavages flanking the putative RNA polymerase domain (Gorbalenya *et al.,* 1989; Lee *et al.,* 1991) are represented by open triangles. The amino acid numbering of pp1a and pp1ab including the ORF 1a/1b junction, as predicted to occur (Brierley *et al.,* 1992; Herold and Siddell, 1993), is given in the single-letter code. PCP, papain-like cysteine proteinase; 3CL, 3C-like proteinase; GFL, growth factor/receptor-like domain; POL, RNA polymerase motif; MBD, metal binding domain; HEL, helicase domain.

quences of the bound proteins were determined by standard procedures with a protein sequencer (Model 476A; Applied Biosystems, Weiterstadt, Germany).

RESULTS

Production of polyprotein 1ab-specific monoclonal antibodies

In order to produce monoclonal antibodies (MAbs) specific for the HCV 229E polyprotein 1ab, we constructed the bacterial expression plasmids pUR-K3 and pT7-F10. The plasmid pUR-K3 contains a 1518-nucleotide region of the HCV 229E pol1b ORF, starting 21 nucleotides downstream of the ribosomal frameshifting site. In-frame fusion of this coding region to the first 840 nucleotides of the *E. coli lacZ* gene allowed for the high-level expression of a 90-kDa fusion protein, K3 (Fig. 1). The construct pT7-F10, encoding a minimal number of additional amino acids at both termini of the same HCV sequence, yielded lower levels of fusion protein expression but was nevertheless useful for hybridoma screening.

Using the partially purified K3 protein as antigen, five murine hybridoma cell lines secreting HCV 229E polyprotein 1ab-specific monoclonal antibodies were obtained. All of the MAbs reacted specifically in ELISA and Western blotting with the K3 and F10 fusion proteins but not with a β-galactosidase-containing bacterial lysate (data not shown). The isotypes of the MAbs were determined and are shown in Table 1. To define the epitopes recognized

TABLE 1

Isotype and Specificity of the Polyprotein 1ab-Specific MAbs

MAb	Isotype	Peptides recognized[a]	Position in pp1ab (amino acids)
6.13D7	IgG 2b	KNVDKDDAFYIVK	4123–4135
		DKDDAFYIVKRCI	4126–4138
		DAFYIVKRCIKSV	4129–4141
7.12E3	IgG 1	KNVDKDDAFYIVK	4123–4135
		DKDDAFYIVKRCI	4126–4138
		DAFYIVKRCIKSV	4129–4141
7.12G4	IgG 1	KNVDKDDAFYIVK	4123–4135
		DKDDAFYIVKRCI	4126–4138
		DAFYIVKRCIKSV	4129–4141
8.3E4	IgG 1	SRQDLTKYTMMDL	4177–4189
		DLTKYTMMDLCFA	4180–4192
8.6G5	IgG 1	TDYFEMKNWFDPI	4216–4228
		FEMKNWFDPIENE	4219–4231
		KNWFDPIENEDIH	4222–4234

[a] Consensus recognition sequences common to all of the peptides reacting with a given MAb are shown in boldface.

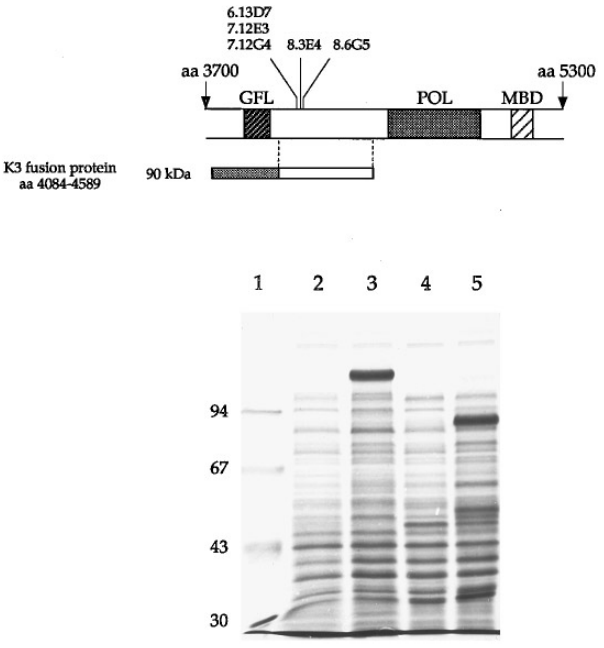

FIG. 2. Expression of the bacterial fusion protein K3 containing a 4084–4589 of the HCV 229E pp1ab. Cell lysates from noninduced (lanes 2 and 4) or IPTG-induced (lanes 3 and 5) *E. coli* BMH 71-18 cells transformed with plasmids pUR290 (lanes 2 and 3) or pUR-K3 (lanes 4 and 5) were separated by SDS–PAGE in a 10% polyacrylamide gel and stained with Coomassie brilliant blue R-250. Protein molecular markers (17-0446-01; Pharmacia Biotech, Freiburg, Germany) are shown in lane 1. The HCV 229E-encoded regions of the fusion protein K3 and the localization of the MAb epitopes within pp1ab is indicated.

by these MAbs, a library of peptides was synthesized consisting of 13-mer peptides, each overlapping its neighbor by 10 amino acids. The library encompasses the 506-amino-acid, polyprotein 1ab region contained in the K3 fusion protein.

The MAbs 6.13D7 and 7.12E3 both reacted with three consecutive peptides covering the amino acid sequence KNVDKDDAFYIVKRCIKSV. This sequence is located at amino acid positions 4123 to 4141 in pp1ab. MAb 7.12G4 reacted with two of the same peptides. MAb 8.3E4 showed binding activity with two consecutive peptides encompassing the sequence SRQDLTKYTMMDLCFA, i.e., amino acid positions 4177 to 4192 in pp1ab. The binding site for MAb 8.6G5 was located within the sequence TDYFEMKNWFDPIENEDIH, which is represented by three neighboring peptides in the library and corresponds to pp1ab amino acids 4216 to 4234. The amino acid sequences of the peptides recognized by individual antibodies, as well as their position in the HCV 229E polyprotein 1ab, are summarized in Table 1 and their locations are indicated in Fig. 2.

Identification of a 105-kDa, gene 1-encoded protein in HCV 229E-infected cells

To study the expression of the HCV 229E gene 1 in virus-infected cells, we mock-infected or infected confluent HeLa–CD13 cells at an m.o.i of 10 PFU per cell, harvested the cells at different times after infection, produced cell lysates, and analyzed them by SDS–PAGE and Western blotting using the panel of MAbs described above. A protein with an apparent molecular mass of 105 kDa was specifically detected by MAb 8.6G5 in lysates from infected, but not from mock-infected, cells (Fig. 2). Using this approach, the earliest time point at which the 105-kDa polypeptide could be detected was 7 hr after infection. The 105-kDa polypeptide was recognized by all of the available pp1ab-specific MAbs but not by a β-galactosidase-specific control MAb 6D2H4 (data not shown; Ziebuhr *et al.,* 1995). Since three of the antibodies used in the immunoblot experiments differ in their specificity, the experiments show that the 105-kDa protein must contain, at least, amino acids 4129 to 4228 of the 1ab polyprotein.

Determination of the termini of the 105-kDa polypeptide

To define the termini of the 105-kDa protein we decided to utilize the *trans* activity of the HCV 229E 3C-like proteinase reported previously (Ziebuhr *et al.,* 1995). First, two *in vitro* synthesized polypeptides, representing amino acids 3934 to 4237 and 4818 to 5259 of the HCV 229E pp1ab, were used as substrates in a cleavage assay. Each polypeptide contained one predicted cleavage site, the dipeptide bonds Gln-4068/Ser-4069 and Gln-4995/Ala-4996, respectively (Gorbalenya *et al.,* 1989; Herold *et al.,* 1993). After incubation with bacterially expressed, purified 3C-like proteinase, each of the primary translation products was cleaved specifically into two polypeptides (Fig. 3). The translation product containing the predicted N-terminal cleavage site of the polymerase domain, with an apparent molecular mass of 36 kDa, was

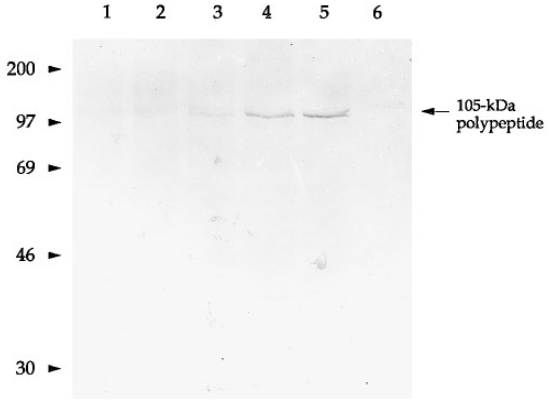

FIG. 3. Detection of a 105-kDa polypeptide in HCV 229E-infected cells. Cell lysates from mock-infected (lane 6) and HCV 229E-infected (lanes 1 to 5) HeLa–CD13 cells (5×10^5) were analyzed by SDS–PAGE and Western blot using the pp1ab-specific MAb 8.6G5. Lysates were obtained at 3 (lane 1), 5 (lane 2), 7 (lane 3), 9 (lane 4), and 11 hr p.i. (lanes 5 and 6). Molecular weight markers (CFA 626; Amersham, Braunschweig, Germany) are shown.

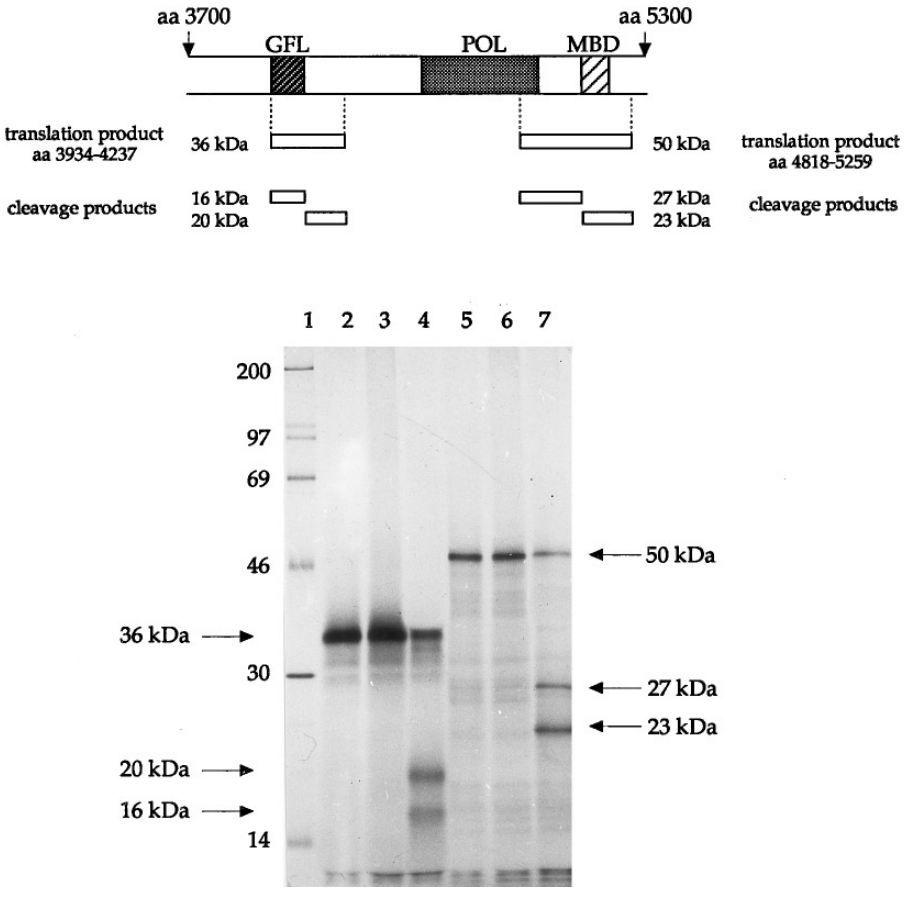

FIG. 4. *Trans* cleavage assay with *in vitro* synthesized substrates containing the predicted amino- and carboxyl-terminal regions of the 105-kDa, ORF 1ab-encoded polypeptide. Protein molecular weight marker (CFA 645; Amersham, Braunschweig, Germany; lane 1); *in vitro* translation product of PCR product-derived mRNA (aa 3934–4237 of pp1ab) without incubation (lane 2), after incubation with buffer (lane 3) or buffer containing 10 μg bacterially expressed 3C-like proteinase (lane 4); translation product of pBSS3-derived mRNA (aa 4818–5259 of pp1ab) without incubation (lane 5), after incubation with buffer (lane 6) or buffer containing 10 μg bacterially expressed 3C-like proteinase (lane 7). The primary translation products and cleavage products are indicated, and their apparent molecular masses are given. Autoradiography was done for 48 hr.

cleaved into two polypeptides with apparent molecular masses of 20 and 16 kDa. This result is consistent with cleavage at the predicted dipeptide bond Gln-4068/Ser-4069. The translation product containing the predicted C-terminal cleavage site of the polymerase domain, with an apparent molecular mass of 50 kDa, was cleaved into two polypeptides with apparent molecular masses of 27 and 23 kDa. This result is consistent with cleavage at the predicted dipeptide bond Gln-4995/Ala-4996. In control reactions with identical buffer conditions, no specific cleavage products were produced. Moreover, a translation product encompassing nearly the complete putative RNA polymerase domain (amino acids 4085 to 4895 of pp1ab) was shown to be resistent to cleavage with recombinant 3C-like proteinase (data not shown).

Second, based upon the data obtained with *in vitro* synthesized substrates, we expressed, in bacteria, HCV 229E pp1ab polypeptides as part of *E. coli* MBP fusion proteins. Incubation of the purified fusion proteins with 3C-like proteinase was then used to obtain specific proteolytic reaction products. Thus, as is shown in Fig. 4, the pMal-Npol fusion protein, containing the predicted N-terminal cleavage site of the polymerase domain with an apparent molecular mass of 76 kDa, was cleaved into two polypeptides with apparent molecular masses of 55 and 21 kDa. The pMal-Cpol fusion protein, containing the predicted C-terminal cleavage site of the polymerase domain with an apparent molecular mass of 94 kDa, was cleaved into two polypeptides with apparent molecular masses of 66 and 28 kDa. These results are consistent with cleavage at the predicted sites and the data shown in Fig. 3. Again, no specific cleavage products were produced in control reactions.

Finally, the 21- and 28-kDa cleavage products of the pMal-Npol and pMal-Cpol fusion proteins were separated on SDS–polyacrylamide gels, transferred to PVDF membranes, isolated, and subjected to N-terminal sequence analysis. The N-terminal protein sequence of the 21-kDa (pMal-Npol) cleavage product was determined to be Ser-Phe-Asp-Asn-Ser-Tyr-Leu, representing amino acids 4069 to 4075 of HCV 229E pp1ab, and the N-terminal protein sequence of the 28-kDa (pMal-Cpol) cleavage

TABLE 2
Cleavage Sites of Coronavirus 3C-like Proteinases

Cleavage sites									Virus	Polyprotein	Amino acids
P5	P4	P3	P2	P1	P1′	P2′	P3′	P4′	P5′			
Thr	Ser	Phe	Leu	Gln	Ser	Gly	Ile	Val	Lys	MHV	pp1a/pp1ab	3330–3339
Gly	Ser	Thr	Leu	Gln	Ala	Gly	Leu	Arg	Lys	HCV	pp1a/pp1ab	2961–2970
Arg	Thr	Ala	Ile	Gln	Ser	Phe	Asp	Asn	Ser	HCV	pp1a/pp1ab	4064–4073
Ser	Thr	Val	Leu	Gln	Ala	Ala	Gly	Leu	Cys	HCV	pp1ab	4991–5000

product was determined to be Ala-Ala-Gly-Leu-Cys-Val-Val, representing amino acids 4996 to 5002 of the HCV 229E 1ab polyprotein (Herold *et al.,* 1993). These data show that two 3C-like proteinase-mediated cleavages, at dipeptide bonds Gln-4068/Ser-4069 and Gln-4995/Ala-4996, release the putative RNA polymerase domain from the HCV 229E 1ab polyprotein in the form of a 926-amino-acid polypeptide with a calculated molecular mass of 105.5 kDa.

DISCUSSION

In this study, we have identified, in virus-infected cells, a 105-kDa viral polypeptide that is encoded in the HCV 229E ORFs 1a and 1b. The identity of the polypeptide has been established by reactivity with monoclonal antibodies of defined specificity. We have also shown that the 105-kDa polypeptide is released from the gene 1-encoded polyprotein precursor, pp1ab, by cleavage at the dipeptide bonds Gln-4068/Ser-4069 and Gln-4995/Ala-4996, a reaction that is mediated by the HCV 229E 3C-like proteinase.

The detection of coronavirus gene 1-encoded polypeptides in the virus-infected cell is proving to be a laborious task. The results shown here, and our previous work on the expression of the HCV 229E 3C-like proteinase (Ziebuhr *et al.,* 1995), indicate that, at least in the early stages of infection, these proteins are present in the cell in exceedingly low amounts. Indeed, the data suggest that it is only possible to detect polymerase translation products after the infection has progressed into the phase of genomic and subgenomic mRNA amplification. Also, as may have been predicted from *in vitro* studies on the level of HCV 229E mRNA 1 ribosomal frameshifting (Herold and Siddell, 1993), we have found it more difficult to identify pp1ab-derived products compared to those derived from pp1a. For example, our attempts to immunoprecipitate the 105-kDa polypeptide using polyclonal antiserum, an approach which has proven successful for the detection of the HCV 229E 3C-like proteinase, have failed to produce convincing results. Obviously, if it is only possible to detect stable translation products as they accumulate in the virus-infected cell, it is going to be difficult to obtain detailed information on the expression and processing of the HCV 229E 1a and 1ab polyproteins. On the basis of the data obtained so far, a quantitative comparison of the amounts of ORF 1a-encoded (Ziebuhr *et al.,* 1995) and ORF 1b-encoded polypeptides is, in our opinion, not possible. For example, the characteristics of the primary and secondary antibodies and the immunological assays that have been used in these studies are very different. Also, the genesis of gene 1-encoded polypeptides involves proteolytic processing, and we do not have any data on the kinetics of this process. Finally, there is no information concerning the stability of different gene 1-encoded polypeptides,

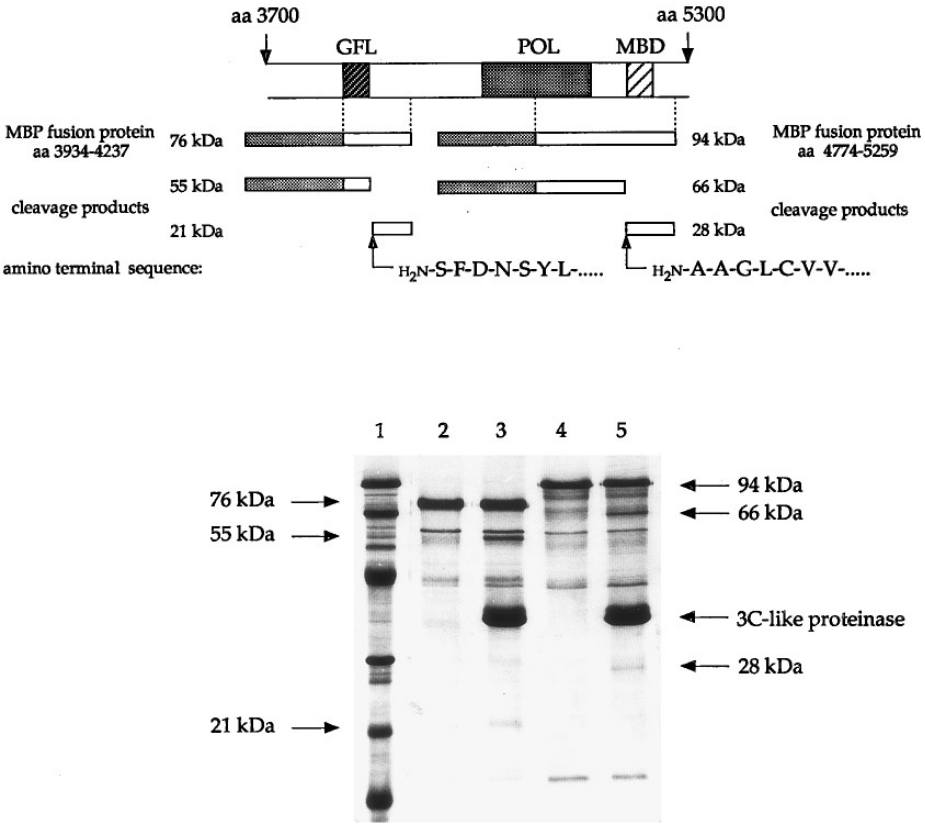

FIG. 5. *Trans* cleavage of bacterially synthesized MBP fusion proteins containing the predicted amino- and carboxyl-terminal regions of the 105-kDa, ORF 1ab-encoded polypeptide. Reaction products were separated on an SDS–15% polyacrylamide gel and stained with Coomassie brilliant blue. Protein molecular weight markers (17-0446-01; Pharmacia Biotech, Freiburg, Germany; lane 1); MBP fusion protein encoded by pMal-NpoI after incubation with buffer (lane 2) or after incubation with 10 μg 3C-like proteinase (lane 3); MBP fusion protein encoded by pMal-CpoI after incubation with buffer (lane 4) or after incubation with 10 μg 3C-like proteinase (lane 5). The fusion proteins and cleavage products are indicated and their apparent molecular masses are given. Additionally, the amino-terminal sequences of the 21- and 28-kDa cleavage products are shown.

nor have any protein–protein interactions that might interfere with antibody recognition been investigated.

The processing of the 105-kDa protein is mediated by the HCV 229E 3C-like proteinase. Information on the substrate specificity of coronaviral 3C-like proteinases is limited, but comparison of the four cleavage sites that have been determined by sequence analysis to date (Table 2; Lu *et al.,* 1995; Ziebuhr *et al.,* 1995) suggests that, besides conservation at the P1 and P1′ sites (Gln/Ala, Ser), the P2 site appears to be represented by hydrophobic residues, mainly Leu. The presence of Ile at the P2 site of the HCV 229E amino-terminal polymerase cleavage site indicates, however, that a certain variability is allowed at this position. It is likely, therefore, that different hydrophobic residues, present at the analogous position of other coronavirus 1a and 1ab polyproteins (Val, Phe, Met), may be part of functional substrate target sites of the 3C-like proteinase. To what extent this variability contributes to the cleavage efficiency at different sites remains to be investigated. It should be noted that the *in vitro* or bacterially synthesized substrates shown in Figs. 4 and 5 contain several Gln/Ala, Gly, Ser dipeptide bonds which are not preceeded by hydrophobic residues and are not recognized as cleavage sites in the *trans* cleavage assays.

The 105-kDa polypeptide contains an RNA polymerase motif that has been identified by computer-aided sequence comparisons with a variety of related cellular and viral enzymes (amino acids 4606 to 4908 of HCV 229E polyprotein 1ab; Gorbalenya *et al.,* 1989; Herold *et al.,* 1993). This suggests that this polypeptide will have a central role in the replication and transcription of viral RNA. Experiments designed to demonstrate an RNA polymerase activity associated with, for example, bacterially expressed 105-kDa polypeptide are in progress. However, even if, for example, a poly(A)-dependent oligo(U)-primed poly(U) polymerase activity could be demonstrated for this polypeptide, this would still represent only one of the many gene functions required for virus replication. Clearly, an integrated approach involving both biochemistry and genetics will be needed to elucidate the structure–function relationships of the coronavirus RNA polymerase complex.

ACKNOWLEDGMENTS

We thank Barbara Schelle and Christine Scheunert for technical help and V. Hoppe for protein sequence data. This work was supported by

grants from the Deutsche Forschungsgemeinschaft (SFB 165/B1 and B8; GK Infektiologie) and the BMFT (01 KI 8838/0).

REFERENCES

Baric, R. S., Fu, K., Schaad, M. C., and Stohlman, S. A. (1990). Establishing a genetic recombination map for murine coronavirus strain A59 complementation groups. *Virology* 177, 646–656.

Brierley, I., Jenner, A. J., and Inglis, S. C. (1992). Mutational analysis of the 'slippery-sequence' component of a coronavirus ribosomal frameshifting signal. *J. Mol. Biol.* 227, 463–479.

Eleouet, J.-F., Rasschaert, D., Lambert, P., Levy, L., Vende, P., and Laude, H. (1995). Complete sequence (20 kilobases) of the polyprotein-encoding gene 1 of transmissible gastroenteritis virus. *Virology* 206, 817–822.

Fu, K., and Baric, R. S. (1994). Map locations of mouse hepatitis virus temperature-sensitive mutants: Confirmation of variable rates of recombination. *J. Virol.* 68, 7458–7466.

Gallagher, S., Winston, S. E., Fuller, S. A., and Hurrell, J. G. R. (1991). Immunoblotting and immunodetection. *In* "Current Protocols in Immunology" (J. E. Coligan *et al.,* Eds.), pp. 8.10.1–8.10.17. Wiley, New York.

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1989). Coronavirus genome: Prediction of putative functional domains in the non-structural polyprotein by comparative amino acid sequence analysis. *Nucleic Acids Res.* 17, 4847–4861.

Harlow, E., and Lane, D. (1988). "Antibodies: A Laboratory Manual." Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Herold, J., Raabe, T., Schelle-Prinz, B., and Siddell, S. G. (1993). Nucleotide sequence of the human coronavirus 229E RNA polymerase locus. *Virology* 195, 680–691.

Herold, J., and Siddell, S. G. (1993). An elaborated pseudoknot is required for high frequency frameshifting during translation of HCV 229E polymerase mRNA. *Nucleic Acids Res.* 21, 5838–5842.

Herold, J., Siddell, S. G., and Ziebuhr, J. (1996). Characterization of coronavirus RNA polymerase gene products. *Methods Enzymol.* 275, in press.

Johnston, S., and Holgate, S. (1996). Epidemiology of viral respiratory tract infections. *In* "Viral and Other Infections of the Human Respiratory Tract" (S. Myint and D. Taylor-Robinson, Eds.), pp. 1–38. Chapman & Hall, London.

Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eucaryotic mRNAs. *Nucleic Acids Res.* 12, 857–872.

Lee, H.-J., Shieh, C.-K., Gorbalenya, A. E., Koonin, E. V., la Monica, N., Tuler, J., Bagdzhadzhyan, A., and Lai, M. M. C. (1991). The complete sequence (22 kilobases) of murine coronavirus gene 1 encoding the putative proteases and RNA polymerase. *Virology* 180, 567–582.

Leibowitz, J. L., DeVries, J. R., and Haspel, M. V. (1982). Genetic analysis of murine hepatitis virus strain JHM. *J. Virol.* 42, 1080–1087.

Liu, D. X., Brierley, I., Tibbles, K. W., and Brown, T. D. K. (1994). A 100-kilodalton polypeptide encoded by open reading frame (ORF) 1b of the coronavirus infectious bronchitis virus is processed by ORF 1a products. *J. Virol.* 68, 5772–5780.

Liu, D. X., and Brown, T. D. K. (1995). Characterization and mutational analysis of an ORF 1a-encoding proteinase domain responsible for proteolytic processing of the infectious bronchitis virus 1a/1b polyprotein. *Virology* 209, 420–427.

Lu, Y., Lu, X., and Denison, M. R. (1995). Identification and characterization of a serine-like proteinase of the murine coronavirus MHV-A59. *J. Virol.* 69, 3554–3559.

Myint, S. H. (1995). Human coronavirus infections. *In* "The Coronaviridae" (S. G. Siddell, Ed.), pp. 389–401. Plenum, New York.

Raabe, T., Schelle-Prinz, B., and Siddell, S. G. (1990). Nucleotide sequence of the gene encoding the spike glykoprotein of human coronavirus HCV 229E. *J. Gen. Virol.* 71, 1065–1073.

Rüther, U., and Müller-Hill, B. (1983). Easy identification of cDNA clones. *EMBO J.* 2, 1791–1794.

Schaad, M. C., Stohlman, S. A., Egbert, J., Lum, K., Fu, K., Wei, T., and Baric, R. S. (1990). Genetics of mouse hepatitis virus transcription: identification of cistrons which may function in positive and negative strand RNA synthesis. *Virology* 177, 634–645.

Tabor, S., and Richardson, C. C. (1985). A bacteriophage T7 polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* 82, 1074–1078.

Wege, H., Wege, H., Nagashima, K., and ter Meulen, V. (1979). Structural polypeptides of the murine coronavirus JHM. *J. Gen. Virol.* 42, 37–47.

Zhang, X. M., Herbst, W., Kousoulas, K. G., and Storz, J. (1994). Biological and genetic characterization of hemagglutinating coronavirus isolated from a diarrhoeic child. *J. Med. Virol.* 44, 152–161.

Ziebuhr, J., Herold, J., and Siddell, S. G. (1995). Characterization of a human coronavirus (strain 229E) 3C-like proteinase activity. *J. Virol.* 69, 4331–4338.