19

CHARACTERIZATION OF A PAPAIN-LIKE CYSTEINE-PROTEINASE ENCODED BY GENE 1 OF THE HUMAN CORONAVIRUS HCV 229E

Jens Herold, Volker Thiel, and Stuart G. Siddell

Institute of Virology and Immunology
University of Würzburg
Versbacher Strasse 7
97078 Würzburg, Germany

1. ABSTRACT

Expression of the coronaviral gene 1 polyproteins, pp1a and pp1ab, involves a series of proteolytic events that are mediated by virus-encoded proteinases similar to cellular papain-like cysteine-proteinases and the 3C-like proteinases of picornaviruses. In this study, we have characterized, *in vitro,* the human coronavirus HCV 229E papain-like cysteine-proteinase PCP 1. We show that PCP 1 is able to mediate cleavage of an aminoterminal polypeptide, p9, from *in vitro* translation products representing the aminoproximal region of pp1a/pp1ab. Mutagenesis studies support the prediction of Cys1054 and His1278 as the catalytic amino acids of the HCV 229E PCP 1, since mutation of these residues abolishes the proteolytic activity of the enzyme.

2. INTRODUCTION

The HCV 229E genome is a single-stranded, positive sense RNA comprised of approximately 27000 nucleotides. Gene 1 encompasses the 5'-two thirds of the genome and encodes for proteins thought to be involved in viral replication and transcription (Herold et al., 1993). Gene 1 contains two large overlapping open reading frames, ORF 1a and ORF 1b, with the potential to encode for polypeptides of approximately 450 kDa and 300 kDa, respectively (Herold et al., 1993). In vitro studies suggest that the ORF 1b gene product is expressed as a fusion protein by a (-1) ribosomal frameshifting event mediated by an element located at the junction of ORF 1a and ORF 1b (Herold and Siddell, 1993).

Coronaviruses and Arteriviruses, edited by Enjuanes *et al.*
Plenum Press, New York, 1998

Computer-assisted analyses of the HCV 229E gene 1 sequence has revealed a number of putative functional domains within pp1a and pp1ab. These include an RNA-dependent RNA-polymerase domain and a metal-binding/helicase domain, both encoded in ORF 1b, as well as domains indicative of papain-like cysteine-proteinases and a chymotrypsin-like cysteine-proteinases (3C-like proteinases) encoded by ORF 1a (Herold et al., 1993).

Recently, we have shown that an autoproteolytic activity is encoded within the aminoterminal 1315 amino acids of the HCV 229E pp1a/pp1ab, a region that includes the predicted PCP 1 domain (Herold, 1995). The proteolytic activity could be blocked *in vitro* by $ZnCl_2$ but not by leupeptin.

In this study, we have produced a polyclonal rabbit antiserum specific for the amino acids 41 to 250 of HCV 229E pp1a/pp1ab. Using this antiserum, we have been able to show that the HCV 229E PCP 1 proteinase is responsible for the release of the aminoterminal polypeptide, p9, from pp1a/pp1ab and we have determined, by mutagenesis analysis, the putative catalytic amino acids of this proteinase.

3. MATERIALS AND METHODS

3.1. Preparation of Antigen and Antiserum

A 632 bp SphI/KpnI cDNA fragment, representing the nucleotides 412–1043 of the genomic RNA of HCV 229E, was excised from the plasmid J12E6 (Herold et al., 1993) and ligated with Sph/KpnI digested DNA of the bacterial expression vector plasmid QE30 (Diagen, Germany). The resulting plasmid, I1a.1, was transformed into competent JM109 bacteria and characterized by restriction enzyme analysis and sequencing of the cloning sites.

For bacterial expression of the recombinant polypeptide, plasmid I1a.1 was transformed into M15/pRep4 bacteria. The recombinant protein expressed from this plasmid comprises 230 amino acids: 12 amino acids at the aminoterminus that are encoded by the expression vector (including 6 consecutive histidines), 210 amino acids encoded by the HCV 229E RNA polymerase gene, corresponding to amino acids 41–250 of pp1a/pp1ab and two vector-derived amino acids at the carboxy terminus. Expression and purification of this fusion protein, pHis(41–250), and immunization of rabbits was done as described previously (Ziebuhr et al., 1995). The resulting pHis(41–250) protein-specific antiserum is referred to as IS1720.

3.2. Plasmids

The construction of plasmid Pap has been described (Herold et al., 1996). Briefly, a cDNA representing the nucleotides 224–4793 of the genomic RNA of HCV 229E was amplified by RT-PCR from poly(A) selected RNA from HCV 229E infected C16 cells and cloned into a BluescriptIISK+ vector. The nucleotide sequence of Pap was determined and several PCR-derived nucleotide misincorporations were identified. In the context of this study, the change of a lysine codon AAA (Lys-1316) to a termination codon TAA is relevant.

3.3. PCR

Pap DNA (1ng) was used as a template for four different PCR reactions using the Elongase amplification system (Life Technologies, USA) as recommended by the manu-

facturer. Oligonucleotide T7 (5'-AATACGACTCACTATAG-3') was the upstream primer, and the oligonucleotides I (5'-CACAAGTCACAGTGGTTGG-3'), II (5'-GTGCTGATTGAATAGTCTTAC-3'), III (5'-GTTAGTCTGGTAATGACCAC-3') and IV (5'-GCAAGGTTCTCATTAGCA-3') were used as downstream primers. The cycle conditions were: initial denaturation, 94 °C for 30 s; 30 cycles at 94 °C for 30 s, 50 °C for 30 s, 68 °C for 75 s/1 kb to be amplified; final elongation step 72 °C for 10 min. The PCR DNAs are referred to as PCR I to PCR IV and the encoded polypeptides are called, according to their length, pI-613, pII-956, pIII-1209 and pIV-1309.

3.4. *In Vitro* Transcription and Translation

RNAs were synthesized in vitro using the MEGAscript T7 Kit (Ambion, USA) and translated (100 ng per µl translation reaction) in a reticulocyte lysate in the presence of [^{35}S]-methionine as described earlier (Herold et al., 1996). Aliquots (2 µl) of the translation reaction were used for immunoprecipitation with IS1720 as described previously (Ziebuhr et al., 1995).

3.5. Site-Directed Mutagenesis

Site specific mutations were introduced into pPap by *in vivo*-recombination PCR as described (Herold and Siddell., 1993).

4. RESULTS

We have expressed an antigen, pHis(41–250), representing amino acids 41–250 of the HCV 229E pp1a/pp1ab, which allows for a fast, single-step purification of the recombinant protein by affinity chromatography on Ni-NTA-agarose columns (Figure 1). The polypeptide has been used to immunize rabbits, resulting in the antiserum IS1720.

A recombinant plasmid, Pap, containing a T7 RNA polymerase promotor followed by the coding sequence of the aminoterminal 1315 amino acids of pp1a/pp1ab was constructed. A series of PCR-DNAs (I-IV) were derived from pPap and subsequently transcribed and translated *in vitro*.. The translation products were then immunoprecipitated with IS1720. As expected, and as is shown in Figure 2, RNA derived from PCR I, PCR II and PCR III encoded for proteins with apparent molecular masses of 66 kDa (pI-613), 104 kDa (pII-956) and 133 kDa (pIII-1209), respectively. PCR IV encodes for the aminoterminal 1.309 amino acids of the HCV 229E pp1a/pp1ab and contains the predicted catalytic amino acids of the first papain-like cysteine-proteinase domain PCP 1. Upon translation of PCR IV-RNA a 9 kDa polypeptide was detected, together with a number of higher molecular mass polypeptides. These polypeptides include products with apparent molecular masses of 143 kDa, the expected size of the PCR IV-RNA primary translation product, and 134 kDa, the expected size of the processed product.

Our interpretation of the data shown in Fig. 2 is that PCP 1 recognizes the full length translation product of PCR IV-RNA, releasing p9 and a processed polypeptide of 134 KDa. Due to the position of pHis(41–250), relative to pp1a/pp1ab, we concluded that p9 represents the aminoterminal polypeptide of pp1a/pp1ab. Using this *in vitro* approach, we are also able to position the carboxyterminal border of the active PCP 1 domain between the amino acids 1.209 and 1.309 of pp1a/pp1ab.

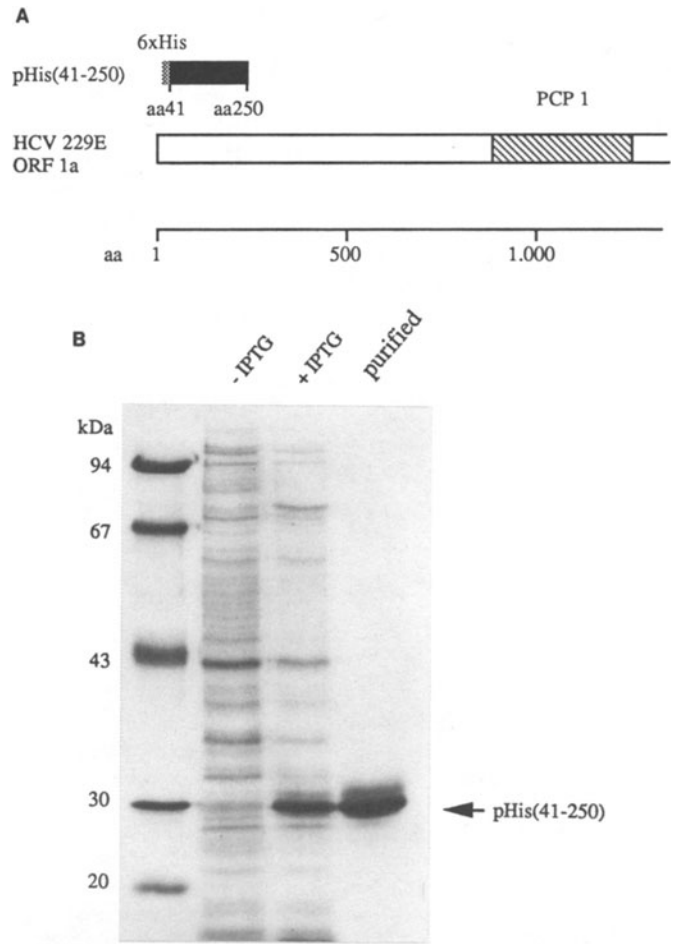

Figure 1. Expression and purification of pHis(41–250). (A) Schematic representation of pHis(41–250) in relation to the aminoproximal region of the HCV 229E pp1a/pp1ab. (B) Cell lysates from non-induced and IPTG-induced bacteria and purified pHis(41–250) were separated by SDS-PAGE in a 15% polyacrylamide gel.

In a further series of experiments, the codons for the cysteine residues Cys962 and Cys1054 and the histidine residues His1205 and His1278 have been mutated and the effect of these changes on proteolytic activity have been monitored (Figure 3). Upon transcription and translation of pPap-DNA, p9 and a processed form of the full length translation product (primary translation product 144 kDa, processed form 135 kDa) are detected . When Cys962 is changed to Gly (Cys962Gly) or His1278 is changed to Gly (His1278Gly) or Val (His1278Val), the proteolytic processing remained unaffected. In contrast, when Cys1054 was changed to Arg (Cys1054Arg), Gly (Cys1054Gly), or Ser (Cys1054Ser) and His1205 was changed to Ala (His1205Ala) or Gly (His1205Gly), the generation of p9 was

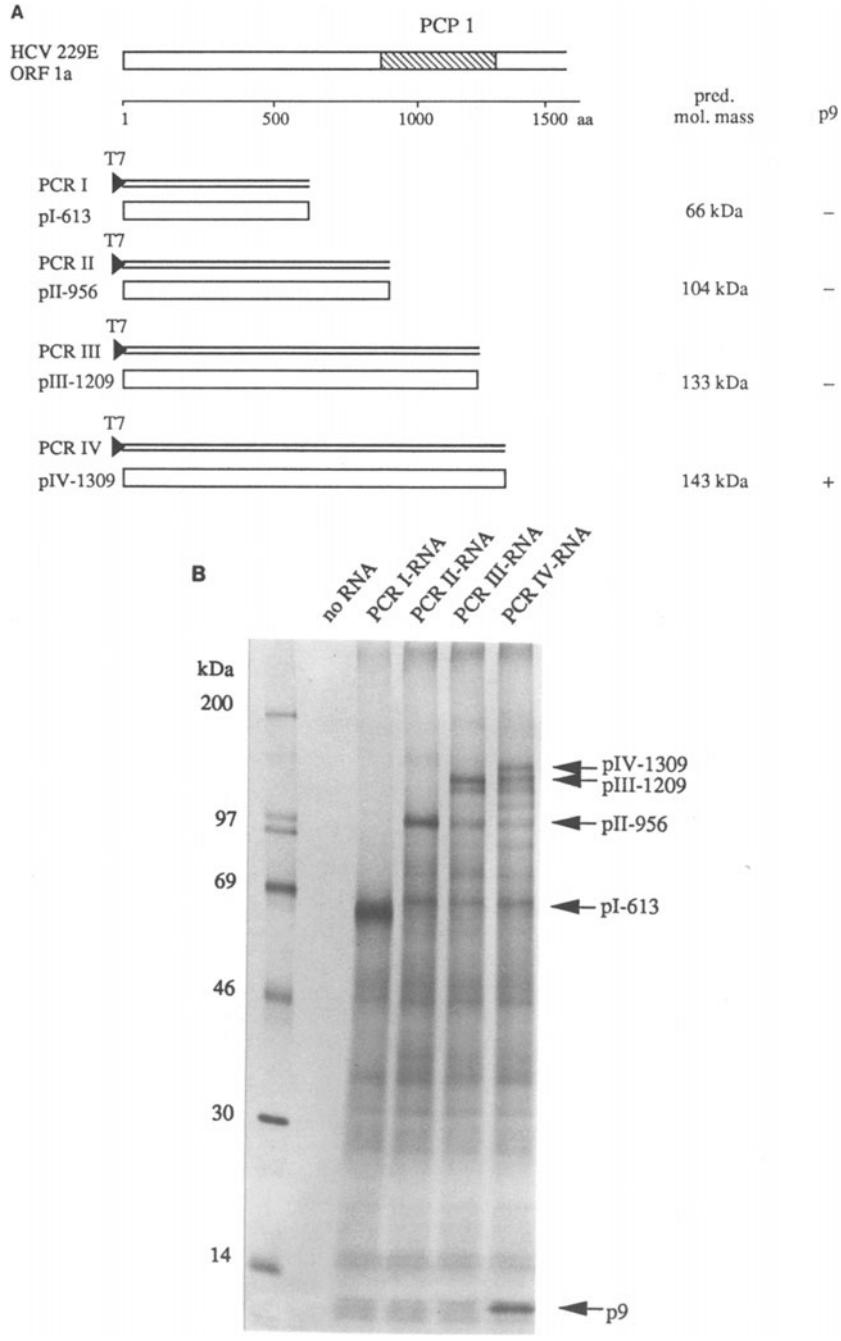

Figure 2. Mapping of the domain necessary for p9 processing. (A) Schematic representation of the HCV 229E ORF 1a, the PCR-DNAs and their predicted primary translation products. (B) In vitro transcription/translation of PCR-DNAs and immunoprecipitation of the translation products. The immunoprecipitated poylpeptides were separated by SDS-PAGE in 10–17.5%-polyacrylamide gels. Protein molecular weight markers are indicated.

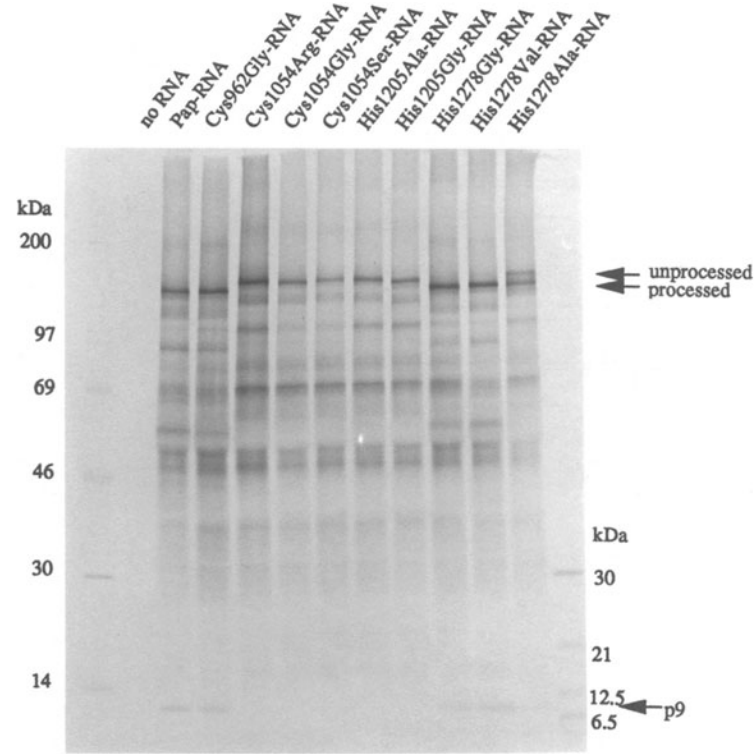

Figure 3. Mutagenic analysis of the HCV 229E-PCP 1. Pap RNA and RNAs derived from mutated DNAs were translated in reticulocyte lysates and the translation products were immunoprecipitated with IS1720. The proteins were analyzed by electrophoresis in a 10–17.5% SDS-polyacrylamide gradient gel. Protein molecular weight markers are indicated.

completely abolished. The change His1278Ala decreased the extent of proteolytic processing. These data support the previous prediction of Cys1054 and His1205 as the catalytic residues of HCV 229E PCP 1.

5. DISCUSSION

Proteolytic processing of non-structural gene products is thought to play a major role in the genesis of funtional replication complexes for many positive stranded RNA viruses (Dougherty and Semmler, 1993; Gorbalenya and Snijder, 1996). Although an adequate processing map of the coronaviral replicase polyproteins has not yet been established, considerable progress has been made in identifying the proteinases involved and their cleavage sites within the polyproteins. In this study, we have investigated proteolytic processing events within the aminoterminal part of the human coronavirus 229E pp1a/pp1ab.

Broadly, our results confirm a number of earlier studies that investigated the MHV PCP1 activity (Baker et al., 1989). It appears that the coronavirus PCP1 activity is responsible at least for the cleavage of a small polypeptide from the aminoterminus of pp1a/pp1ab. In the case of MHV this polypeptide is p28. In the case of HCV 229E it is p9. At the present time, however, there is no indication of the possible function of these proteins. Immunofluorescence assays of HCV 229E-infected cells using the IS1720 antiserum has revealed a punctated pattern of staining in the perinuclear region. This may suggest that at least one of these proteins recognized by this serum remains associated with the virus replication complex (Heusipp et al., 1997).

REFERENCES

Baker, S. C., Shieh, C. K., Soe, L. H., Chang, M. F., Vannier, D. N., and Lai, M. M. C., 1989, Identification of a domain required for autoproteolytic cleavage of murine coronavirus gene A polyprotein, *J. Virol.* **63:**3693–3699.

Dougherty, W. G., and Semler, B. L., 1993, Expression of virus-encoded proteinases: functional and strutural similarities with cellular enzymes, *Micribiol. Rev.* **57:**781–822.

Gorbalenya, A. E., and Snijder, E. J., 1996, Viral cysteine proteinases, *Perspectives in Drug Discovery and Design* **6:** 64–86.

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, 1995, Organization and Expression of RNA Polymerase Locus of the Human Coronavirus 229E, Ph.D. Thesis., University of Würzburg, Germany.

Herold, J., Siddell, S.G., and Ziebuhr, J., 1996, Characterization of coronavirus RNA polymerase gene products, *Methods Enzymol.* **275:**68–89.

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

Heusipp, G., Grötzinger, C., Herold, J., Siddell, S. G., and Ziebuhr, J., 1997, Identification and subcellular localization of a 41-kDa, polyprotein 1ab processing product in human coronavirus 229E-infected cells, *J. Gen. Virol.* in press.

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.