Journal of General Virology (2002), **83**, 359–368. Printed in Great Britain

Hepatitis A virus polyprotein processing by *Escherichia coli* proteases

Rosa M. Pintó,[1] Susana Guix,[1] Juan F. González-Dankaart,[1]† Santiago Caballero,[1] Gloria Sánchez,[1] Ke-Jian Guo,[1]‡ Enric Ribes[2] and Albert Bosch[1]

Department of Microbiology[1] and Department of Animal and Plant Cell Biology[2], University of Barcelona, Av. Diagonal 645, 08028 Barcelona, Spain

Hepatitis A virus (HAV) encodes a single polyprotein, which is post-translationally processed. This processing represents an essential step in capsid formation. The virus possesses only one protease, 3C, responsible for all cleavages, except for that at the VP1/2A junction region, which is processed by cellular proteases. In this study, data demonstrates that HAV polyprotein processing by *Escherichia coli* protease(s) leads to the formation of particulate structures. P3 polyprotein processing in *E. coli* is not dependent on an active 3C protease: the same processing pattern is observed with wild-type 3C or with several 3C mutants. However, this processing pattern is temperature-dependant, since it differs at 37 or 42 °C. The bacterial protease(s) cleave scissile bonds other than those of HAV; this contributes to the low efficiency of particle formation.

Introduction

Hepatitis A virus (HAV) is classified as the type species of the genus *Hepatovirus* within the family *Picornaviridae* (Murphy *et al.*, 1995) and is a health-significant hepatotropic virus. The virion capsid is composed of structural proteins VP1, VP2, VP3 and, possibly, VP4 (Lemon & Robertson, 1993). The major neutralization epitopes of HAV appear to be discontinuous (Lemon & Robertson, 1993). The immunodominant neutralization antigenic site is composed of closely related epitopes: some of them are detected on 14S pentameric subunits, while others are formed by structural changes during the assembly of the 14S structures into 70S structures and intact particles (Stapleton *et al.*, 1993). The assembly of capsid proteins into subvirus or virion structures may then be necessary for the generation of efficient HAV-neutralizing epitopes. Autoproteolytic processing of the viral polyprotein by means of the viral protease (Borovec & Anderson, 1993) seems to be necessary for the formation of pentamers (14S) and procapsids (70S). The 3C protease of HAV has been characterized extensively in terms of its biochemical and structural properties (Gauss-Müller *et al.*, 1991; Jia *et al.*, 1991a; Harmon *et al.*, 1992; Malcolm *et al.*, 1992; Bergmann *et al.*, 1997). Up to the present time, this protease was thought to be responsible for all cleavages required for polyprotein processing. However, recent evidence (Graff *et al.*, 1999; Martin *et al.*, 1999) indicates that VP1 protein maturation may be dependent solely on a host protease.

In the present study, the production of recombinant subvirus and/or virion structures of HAV was attempted in *Escherichia coli* by expressing either the complete open reading frame (ORF) of HAV (pTHAVF construct) or the region encoding the capsid proteins only (pTHAVP1 construct).

Author for correspondence: Albert Bosch.
Fax +34 934034629. e-mail albert@porthos.bio.ub.es

† **Present address:** Biokit, SA, Lliçà d'Amunt, 08186 Barcelona, Spain.

‡ **Present address:** Department of Hepatitis, Institute of Virology, Chinese Academy of Preventive Medicine, 100052 Beijing, China.

Methods

■ **Cells and virus.** FRhK-4 cell cultures were used to propagate and assay the cytopathogenic HM-175 strain of HAV (courtesy of T. Cromeans, Centers for Disease Control, Atlanta, GA, USA) (Cromeans *et al.*, 1987). Virus titrations were performed, as described previously (Pintó *et al.*, 1994), by calculating the most probable number of cytopathogenic units per ml (MPNCU/ml). This was carried out by infecting cell monolayers grown in 96-well microtitre plates. A total of 16 wells was infected for each dilution and 20 µl of inoculum was added to each well. Data were then processed using an MPN computer program (Hurley & Roscoe, 1983).

0001-7995

Table 1. Primers used for the generation of 3C mutants by a PCR site-directed mutagenesis

Nucleotides indicated in bold correspond to those changed for the purpose of mutagenesis. The numeration of the sequences corresponds to that of the pHAV7 cDNA clone of the HM-175 strain of HAV, except for P3-4387, which corresponds to the numeration of pBTac2.

Primer	Nucleotide Sequence (5′ → 3′)
P3-5531	^{5531}GGGATTCCAGGATGTTGTTC^{5550}
P3-5817-μ172	^{5817}AGGCCCCACCA**GC**CATTCCAGGAA^{5794}
P3-5794-μ172	^{5794}TTCCTGGAATG**GC**TGGTGGGGCCT^{5817}
P3-5870-μ191	^{5870}AGCAACA**CC**GATGCCCAAGA^{5851}
P3-5856-μ191	^{5856}GGCATC**GG**TGTTGCTGGAGG^{5875}
P3-4387	^{4387}GGAGACCCCACACTACCATC^{4406}

■ **Plasmid constructs.** Three DNA plasmid constructs of the HAV genome were generated using the expression vector pBTac2 (Boehringer Mannheim) and standard cloning procedures. These constructs contained the complete HAV ORF (pTHAVF), the coding region for the polyprotein precursor of the viral structural proteins (pTHAVP1) or the coding region for the 3C protease (pTHAV3C). The pBTac high-expression vector contains the *Tac* promoter and the *lac*Z ribosome-binding site followed by the ATG initiation codon and strong transcription terminators. pTHAVF was constructed by cleaving pHAV7 (Cohen *et al.*, 1987) from the *Xba*I site in the HAV genome to the *Hae*II site in the pGEM backbone vector. The resulting 6·7 Kb fragment was blunt-end ligated into pBTac2. pTHAVP1 was created by cloning the 2·7 Kb fragment from the *Xba*I site (745 bp) to the *Ava*II site (3493 bp) of pHAV7 into the *Bam*HI site of pBTac2, after blunt-end generation of both DNAs. This construct contained the P1 and 2A regions plus 33% of the 2B region. pTHAV3C was generated after cloning the 0·92 Kb *Pst*I fragment (5135–6062 bp) of pHAV7 into the *Pst*I site of pBTac2. This construct contained 38% of the 3A region (28 aa), the whole 3B region (23 aa), the 3C region (219 aa) and 8% of the 3D region (38 aa). *Escherichia coli* strain JM109 was transformed and positive clones were selected by hybridization to digoxigenin-labelled probes (corresponding to the different HAV fragments cloned in each construct) and ulterior restriction analysis using sites regenerated after cloning.

Three different mutant constructs of the 3C protease were generated by site-directed mutagenesis. The general approach, similar to that described by Jia *et al.* (1991b), was based on the replacement of the 383 bp *Apa*I–*Hin*dIII fragment of pTHAV3C by a 383 bp PCR fragment containing a single codon substitution encoding an amino acid substitution. pTHAVμ3C(172) incorporated an alanine residue in place of the cysteine residue. The replacement of cysteine by alanine, glycine or serine is thought to induce loss of proteolytic activity (Malcolm *et al.*, 1992; Gosert *et al.*, 1997). In pTHAVμ3C(191), the glycine residue was replaced by a histidine residue, since this substitution is associated with loss of protease substrate-binding capacity (Jia *et al.*, 1991a, b). Finally, a double mutant construct incorporating both mutations, pTHAVμμ3C (172,191), was generated by creating the cysteine to alanine replacement in pTHAVμ3C(191). To generate the substituting fragments, two complementary oligonucleotides (Table 1) containing the two replaced base pairs were used as PCR primers to obtain two overlapping fragments. These overlapping fragments were then hybridized and used as templates in a third PCR, where primers from the ends were employed to generate a single DNA fragment by overlap extension. All mutations introduced were confirmed by nucleotide sequencing.

An additional construct, pTHAVVP1, corresponding to the VP1 gene fragment was made in pBTac2 by PCR. The primers used were VP1A, 5′ TCCACTGCAGAGTTGGAGATGAT, and VP1B, 5′ AGGCAAGCTT-CTCAAATCTTTT, which contain *Pst*I and *Hin*dIII sites, respectively.

■ **Induction of protein synthesis.** All HAV-derived constructs were expressed in JM109 cells grown in M9 medium supplemented with 0·4% glucose. When the OD_{600} was approximately 0·6, protein synthesis of the genomes under the control of the *Tac* promoter was induced by the addition of 1 mM IPTG. After 4–16 h of induction, bacterial cells pelleted from 50 ml of culture were resuspended in 500 μl TNE buffer (50 mM Tris–HCl, 150 mM NaCl and 1 mM EDTA, pH 7·4) and treated with 1 mg/ml lysozyme for 1 h. After three freeze–thaw cycles at −70 °C, $MgCl_2$ was added to a final concentration of 10 mM and cell extracts were incubated with 10 μg/ml DNase I for 2 h at 4 °C. After centrifugation of the bacterial lysates at 11000 ***g*** for 10 min, two different fractions were recovered: an insoluble protein fraction in the form of inclusion bodies, corresponding to the pellet, and a soluble protein fraction, corresponding to the supernatant.

Proteins were resolved by SDS–PAGE and stained with coomassie blue. The relative concentrations and molecular masses of the proteins were determined using ImageMaster 1D, version 2.0 (Pharmacia).

In some experiments, the protease inhibitor *N*-ethylmaleimide (NEM) was added to a final concentration of 10 μM.

■ **Antibodies.** The following monoclonal antibodies (mAbs) against HAV were used: K3-4C8, K2-4F2 (Commonwealth Serum Laboratories) and 33Z/37/39 (generously provided by Z.-M. Yun, Institute of Virology, Beijing, China). A convalescent serum, HCS-2 (generously provided by R. Lluna, Hospital Militar, Barcelona, Spain), which recognizes HAV at a 1/1000000 dilution, was also used. A polyclonal ascites antibody (anti-HAV_s) was obtained after the immunization of mice with intact HAV particles. Another polyclonal ascites antibody (anti-3C) was obtained after the immunization of mice with a synthetic peptide (SEGPLKMEEKATYV; a sequence derived from HAV 3C) coupled to KLH. The use of this sequence to generate anti-3C antibodies had been described previously (Gauss-Müller *et al.*, 1991). The anti-3C antibodies were used at a 1/10 dilution for Western blotting.

■ **Immunoprecipitation.** Protein was expressed from pTHAVF or pTHAVP1 for 16 h, after which 500 μl samples from the soluble fractions were immunoprecipitated overnight at 4 °C with mAbs K2-4F2 and K3-4C8 (diluted 1/250 and 1/500, respectively) in order to recover viral structures. Immune complexes were harvested by the addition of protein A–agarose and incubation at 4 °C for 2 h, followed by centrifugation at 10000 ***g*** for 1 min. Pellets were washed twice with TNMg buffer (20 mM Tris–HCl, 10 mM NaCl and 50 mM $MgCl_2$, pH 6·7) and resuspended in 20 μl of the same buffer. Samples were disrupted by adding 5 μl Laemli buffer and boiling for 10 min. Proteins were then resolved by 12–24% gradient SDS–PAGE and examined by Western blot analysis.

■ **Sucrose gradient analysis.** After three 30 s sonication cycles at 70 W in the presence of 0·5% sodium lauryl sarcosine, 500 μl of the soluble fraction extracted after the 16 h expression of pTHAVF or pTHAVP1 were layered onto a 5–45% sucrose gradient in TNMg buffer and spun at 205000 ***g*** for 165 min. Fractions of 500 μl were collected and the presence of HAV antigenic material and refraction indexes were determined for each fraction. Sedimentation markers comprised human IgM (19S) and IgG (7S) antibodies, as well as the different HAV structures generated after virus infection of cells.

For Western blot analysis, fractions corresponding to the 70S or 14S peaks from six different gradients were pooled and concentrated by methanol precipitation to a final volume of 500 µl.

■ **Western immunoblotting.** Samples (20 µl) from the immunoprecipitated supernatants and concentrated sucrose fractions from pTHAVF and pTHAVP1 were resolved by 12–24% gradient SDS–PAGE. Purified inclusion bodies (10 µl) from the 3C constructs were pelleted and resuspended in 10 µl 8 M urea and 5 µl 20% SDS, boiled for 5 min in Laemmli buffer and resolved by 10% SDS–PAGE. Proteins were then electroblotted onto nitrocellulose membranes. Membranes were blocked overnight at room temperature in 5% non-fat milk powder in TBS (10 mM Tris–HCl and 150 mM NaCl, pH 7·5) (blocking solution) and then incubated for 2 h at room temperature with either the anti-HAV_s or the anti-3C antibodies. After extensive washes, a second incubation of 2 h with a sheep anti-mouse IgG (The Binding Site) was performed. Bound antibodies were then detected using a donkey anti-sheep IgG conjugated to alkaline phosphatase (The Binding Site). X-phosphate and NBT (Roche) were used as substrates. Samples (15 µl) of the soluble fraction from cultures harbouring pTHAVVP1 and 15 µl samples of an HAV cell-infected extract were assayed as positive controls.

■ **ELISA.** Two different ELISAs, a direct ELISA and a sandwich ELISA, were performed for the detection of HAV antigenic material in crude supernatants after the expression of pTHAVF and pTHAVP1. In the direct ELISA, antigenic material was coated directly onto the microtitre wells and the HAV-related material was detected using anti-HAV_s ascites fluid. In the sandwich ELISA, HAV structures were captured by HCS-2 convalescent serum and detected using anti-HAV_s antibodies. HAV-infected and mock-infected FRhK-4 cell extracts were used as positive and negative controls, respectively. HAV-related antigens were also assayed in sucrose gradient fractions by sandwich ELISA consisting of HAV capture by HCS-2 convalescent serum, followed by detection with mAb K2-4F2. Sucrose gradient fractions of HAV-infected cell extracts were used as positive controls.

■ **N-terminal sequencing of proteins.** Proteins from inclusion bodies of the different constructs containing 3C sequences were resolved by SDS–PAGE, transferred to Immobilon membranes (Millipore) and stained with coomassie blue. The required protein bands were removed from the gel and subjected to an automated Edman degradation in a Beckman LF3000 sequencer (Beckman).

■ **Electron microscopy.** Samples from supernatants of pBTac2, pTHAVF or pTHAVP1 bacterial lysates were observed by transmission electron microscopy (TEM) with a Hitachi MT-800 electron microscope, after negative staining with 3% phosphotungstic acid, pH 6·5 (KPTA). Supernatant samples were also observed by immunoelectron microscopy (IEM). Briefly, 20 µl of the supernatant samples were incubated for 4 h at 37 °C with mAbs K2-4F2 and K3-4C8, diluted 1/1000 and 1/5000, respectively, and placed onto 2% agarose in the wells of a microtitre plate. Formvar-coated grids were inverted and placed over each drop of sample. After a 2 h incubation at 37 °C, the grids were removed from the agarose and stained with 3% KPTA. A third procedure based on immunoprecipitation was also performed. Samples (500 µl) of supernatant were incubated with mAbs K2-4F2 and K3-4C8, diluted 1/1000 and 1/5000, respectively, for 4 h at 37 °C. Immune complexes were then incubated with protein A conjugated to 10 nm gold particles for 4 h at 4 °C and collected by centrifugation at 10000 ***g*** for 10 min. The pellet obtained was resuspended in 30 µl KPTA and applied to the grids.

As a control, HAV suspensions were submitted to all of the procedures described above.

Results and Discussion

Characterization of the anti-HAV_s ascites fluid

The antigen used for the characterization of the anti-HAV_s ascites fluid was an HAV-infected cell extract that contained intact virus particles, subvirus particles and individual proteins, since it was collected as early as 7 days post-infection. The anti-HAV_s antibodies were characterized for HAV recognition by sandwich ELISA (detecting mainly HAV virus and subvirus particles), direct ELISA (detecting all kinds of HAV material present in the antigenic preparation) and Western blotting (detecting denatured proteins). The maximum recognition titres observed for these assays were 1/1000000, 1/10000 and 1/50, respectively. These results suggested that the anti-HAV_s ascites fluid contained antibodies against all kinds of HAV antigens, although in higher proportions against structured (particles and subparticles) material. In order to know the relative proportion of antibodies to the denatured proteins, this ascites fluid was tested (at a 1/50 dilution) using both ELISA techniques against an HAV-infected cell extract, both native and denatured by boiling (Table 2). In all instances, antigenicity was lost after boiling, indicating that the proportion of antibodies against the denatured HAV proteins was very low and insufficient for ELISA. Western blots of HAV-infected cell extracts revealed that the ascites fluid detected mainly VP1 (Fig. 1). VP1 recognition was confirmed by co-running an HAV-infected cell extract with the recombinant VP1 protein obtained after pTHAVVP1 expression (Fig. 1).

Formation of virus-like particles (VLPs) by expression of either pTHAVF or pTHAVP1

Bacterial cells harbouring pBTac2, pTHAVP1 and pTHAVF were named pBTac, P1 and F, respectively. Bacteria were

Table 2. Anti-HAV response in cell-free extracts from recombinant *E. coli* after 16 h of induction

ELISA using antibodies diluted 1/50 for FRhK-4 cells and HAV and diluted 1/15 for the recombinant proteins. Positive results are indicated in bold. ND, Not determined.

Antigen	Sandwich ELISA (OD_{492})	Direct ELISA (OD_{492})
FRhK-4 cells	0·177 ± 0·01	0·387 ± 0·01
HAV	**1·008 ± 0·04**	**0·813 ± 0·08**
Boiled FRhK-4 cells	ND	ND
Boiled HAV	0·166 ± 0·01	0·398 ± 0·02
pBTac	0·160 ± 0·03	0·301 ± 0·05
pTHAVP1	**0·312 ± 0·02**	**0·606 ± 0·07**
pTHAVF	**0·388 ± 0·04**	**0·653 ± 0·08**
Boiled pBTac	0·197 ± 0·04	0·423 ± 0·06
Boiled pTHAVP1	0·201 ± 0·04	0·431 ± 0·06
Boiled pTHAVF	0·221 ± 0·06	0·474 ± 0·04

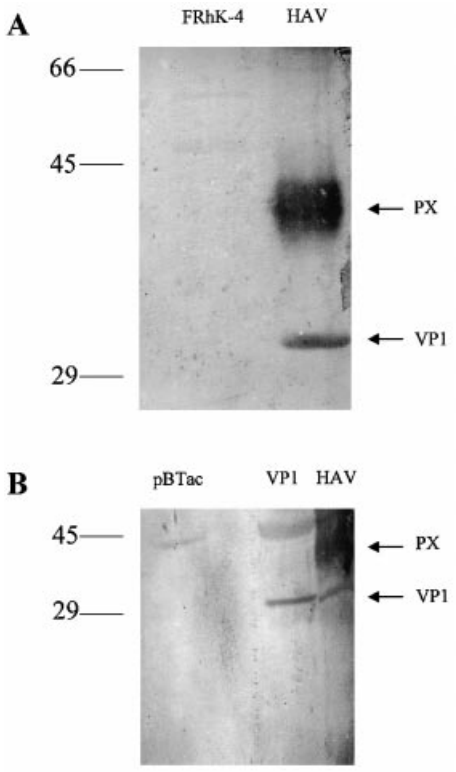

Fig. 1. Western blot analysis of mock- and HAV-infected cell extracts (A) and recombinant HAV proteins (B) using anti-HAV_s antibodies. FRhK-4, mock-infected FRhK-4 cells; HAV, HAV-infected FRhK-4 cells; pBTac, soluble fractions from *E. coli* harbouring pBTac2; VP1, soluble fractions from *E. coli* harbouring pTHAVVP1. Molecular size standards are shown on the left.

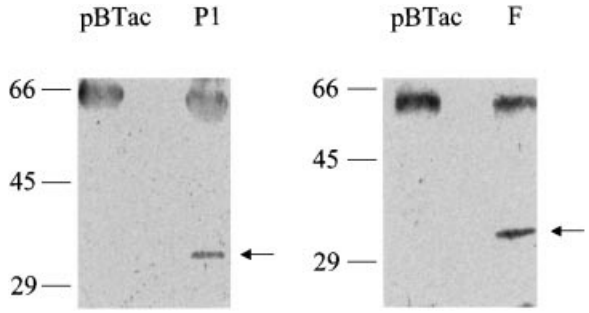

Fig. 2. *E. coli* harbouring either pBTac 2, pTHAVP1 or pTHAVF were induced with IPTG for 16 h and the soluble proteins extracted by lysozyme treatment, immunoprecipitated with mAbs K2-4F2 and K3-4C8, resolved by 12–24% gradient SDS–PAGE and visualized by Western blot using anti-HAV_s antibodies. pBTac, immunoprecipitated soluble fractions from *E. coli* harbouring pBTac2; P1, immunoprecipitated soluble fractions from *E. coli* harbouring pTHAVP1; F, immunoprecipitated soluble fractions from *E. coli* harbouring pTHAVF. Arrows indicate immunoreactive bands. The band of approximately 60–65 kDa, appearing in pBTac, P1 and F corresponds to the heavy chain of the mouse mAbs employed for immunoprecipitation. Molecular size standards are shown on the left.

treated with IPTG in order to induce protein synthesis (under the control of the *Tac* promoter). After 4 h of induction at 37 °C, the most prominent expression effect was the synthesis of inclusion bodies in both F and P1 strains, as has been described previously (Bosch *et al.*, 1997); these inclusion bodies contained HAV-related proteins of approximately 95–97 kDa and 100 KDa, respectively. The molecular mass of the protein generated with pTHAVP1 corresponded to the molecular mass of a polyprotein, including the entire P1 and 2A regions and 33% of the 2B region (also cloned in this construct). However, the molecular mass of the protein produced by pTHAVF corresponded to the structural proteins of HAV (P1 region) plus the 2A region included in the PX precursor (Probst *et al.*, 1997; Martin *et al.*, 1999). This result revealed that a proteolytic cleavage occurred in *E. coli*, since the molecular mass of the entire polyprotein coded by this insert is considerably higher.

To establish whether the expression of the complete HAV ORF in *E. coli* resulted in the synthesis of VLPs, the soluble proteins produced after 16 h of induction at 37 °C were analysed by direct and sandwich ELISAs. HAV antigenic material was detected using anti-HAV_s antibodies (diluted 1/15) in supernatants free from inclusion bodies from F and, surprisingly, P1 cells (Table 2) in both kinds of ELISA techniques. Higher dilutions of this antibody failed, indicating a very low concentration of HAV-related material in the soluble fractions. Since the antibody used for this detection was generated against intact virus particles and its reactivity against denatured proteins was demonstrated to be much lower than against whole viruses (Table 2), it was suspected that some structured material could be present in F and also P1 cells. To recover and concentrate the structures present in the samples, a particle-specific immunoprecipitation with mAbs K2-4F2 and K3-4C8 was performed. mAb K2-4F2 specifically recognizes 14S epitopes present in both pentamers and procapsids, while mAb K3-4C8 recognizes 70S epitopes present only in procapsids (Stapleton *et al.*, 1993). After separating the immunoprecipitated proteins by SDS–PAGE and revealing their presence by Western blot using anti-HAV_s antibodies, a clear band of approximately 33 kDa, probably VP1, could be resolved in both P1 and F cells (Fig. 2), indicating that the same kind of processing had occurred in both constructs. To confirm the presence of structured material in P1 and F cell extracts, sucrose density gradient centrifugations were performed. Two major peaks of antigenicity were detected in both P1 and F extracts, corresponding to sedimentation coefficients of 13–14S (P1) and 70S (F) (Fig. 3B). These results suggested that, after expression of either pTHAVP1 or pTHAVF in *E. coli*, both pentamers and procapsids are formed. In order to investigate whether protein processing occurred in both types of structures from either type of construct, Western blot analysis was performed (Fig. 3C). In all cases, a band of approximately 33 kDa was detected, indicating that both pentamers and procapsids were proteolytically processed. Probst *et al.* (1999) have indicated recently that part of the 2A region is essential for the assembly of pentameric structures and that VP4 is required for an efficient

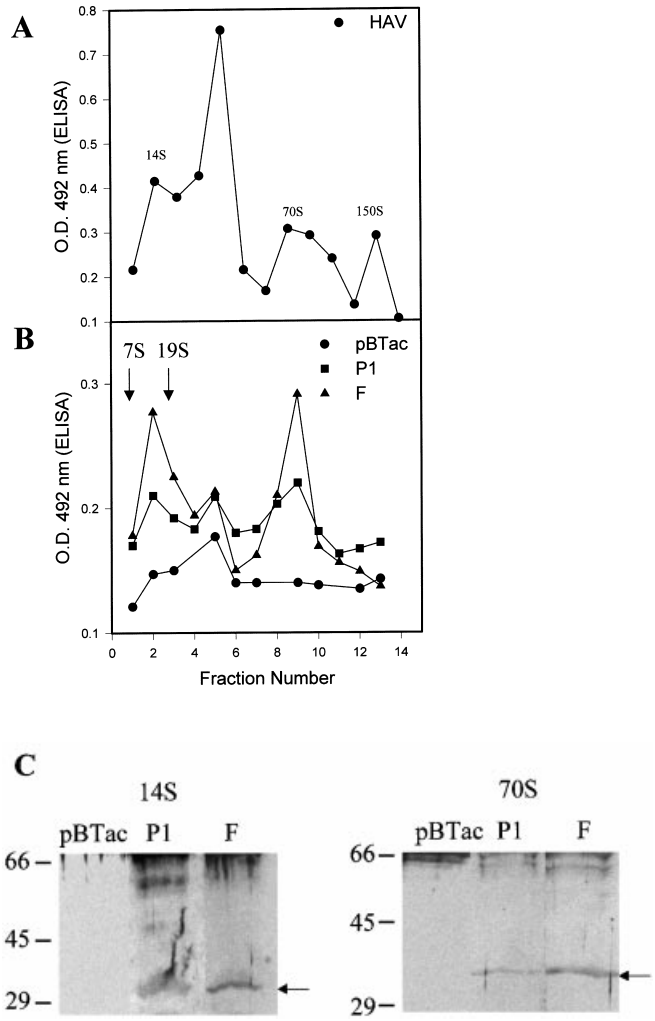

Fig. 3. *E. coli* harbouring pBTac 2, pTHAVP1 or pTHAVF were induced with IPTG for 16 h. The soluble proteins were then extracted by lysozyme treatment and fractionated on a 5–45% sucrose gradient. Gradient fractions were collected from the bottom of the tube and analysed by sandwich ELISA using the convalescent serum for the HAV-related antigen capture, followed by detection with mAb K2-4F2. (A) Results obtained when HAV-infected cell extracts were fractionated under the same conditions. (B) Results obtained with the fractionation of cell extracts from the recombinant bacterial strains after induction. IgG (7S) and IgM (19S) markers fractionated simultaneously sedimented in the indicated fractions. (C) Western blot analysis of samples corresponding to the 14S and 70S peaks. Fractions from six different gradients were pooled, concentrated by methanol precipitation, separated by 12–24% gradient SDS-PAGE and visualized by Western blot using anti-HAV_s antibodies. pBTac, purified fractions from *E. coli* harbouring pBTac2; P1, purified fractions from *E. coli* harbouring pTHAVP1; F, purified fractions from *E. coli* harbouring pTHAVF. Arrows indicate immunoreactive bands. Molecular size standards are shown on the left.

formation of procapsids; both requirements are accomplished even in pTHAVP1. Concentration of both kinds of subvirus structures was lower in P1 than in F cultures, while the growth of the P1 cells was of a higher magnitude than that of F cells (data not shown); this indicated that processing and maturation of the HAV-related structures is more efficient when the P3 region is present. In this same direction, Kusov & Gauss-Müller (1999) have suggested recently that accumulation of uncleaved 3AB and/or 3ABC is pivotal for both virus replication and efficient particle formation. HAV particle production was confirmed by TEM. Isolated single particles of around 30 nm were observed, although always in low numbers, in extracts from pTHAVP1 and pTHAVF, but not pBTac (Fig. 4A, B, H and I). To confirm its virus origin, an IEM was performed and, again, the same kind of VLPs were detected, although on some occasions, these were surrounded by antibodies or in the form of pairs (Fig. 4, C and D). However, aggregates were not observed, again suggesting a very small concentration of particles. In order to concentrate the virus particles and to facilitate their detection, immunoprecipitation together with immunolabelling was performed. Gold-decorated small aggregates or single particles of icosahedral VLPs, with sizes ranging from 30 to 35 nm, both in P1 or F extracts (Fig. 4, E, F, J–L, N and O), could be detected by this procedure. The fact that only electron-dense particles were visualized is surprising, since we could expect that KPTA penetrates into the empty capsid shell. However, although it is always true that whole particles are not permeable to the stain, the contrary does not always apply: on some occasions, empty particles may not be permeable to KPTA, as may be observed in micrographs of HAV VLPs produced through the vaccinia virus expression system (Winokur *et al.*, 1991). Additionally, actual HAV 70S empty structures sometimes, under EM, appear as electron-dense particles (Fig. 4R). The exclusive presence of electron-dense particles in our bacterial extracts suggests that the folding of the capsid-like structures may not be the same as those of 70S HAV particles.

Proteolytic processing of the HAV P3 region

Since VLPs could be synthesized from pTHAVP1 in *E. coli*, which do not express the viral protease, it was suspected that a bacterial protease(s) could be responsible for processing the HAV polyprotein. The viral 3C protease is carried in the P3 polyprotein surrounded by VPg (3B) and the viral RNA polymerase (3D), and its catalytic activity in *cis* or in *trans* is responsible for its release. For this reason, the P3 polyprotein may be employed as a model for the study of the proteolytical processing of the viral 3C protease. Four plasmid constructs were generated for this purpose: pTHAV3C, which encodes the wild-type protease; pTHAVμ3C(172), which encodes an active site mutant; pTHAVμ3C(191), which encodes a substrate-binding site mutant; and pTHAVμμ3C(172,191), which encodes the double mutant. All nucleotide sequences were confirmed after cloning. The predicted molecular masses of the different uncleaved and cleaved products were calculated by applying the GeneRunner software package (Hastings Software). These were as follows: 34·2 for the uncleaved precursor Δ3A3B3CΔ3D; 31·1 for 3B3CΔ3D; 29·7 for Δ3A3B3C; 28·6 for 3CΔ3D; 26·6 for 3B3C; 24·1 for 3C; 5·6 for Δ3A3B; 4·5 for

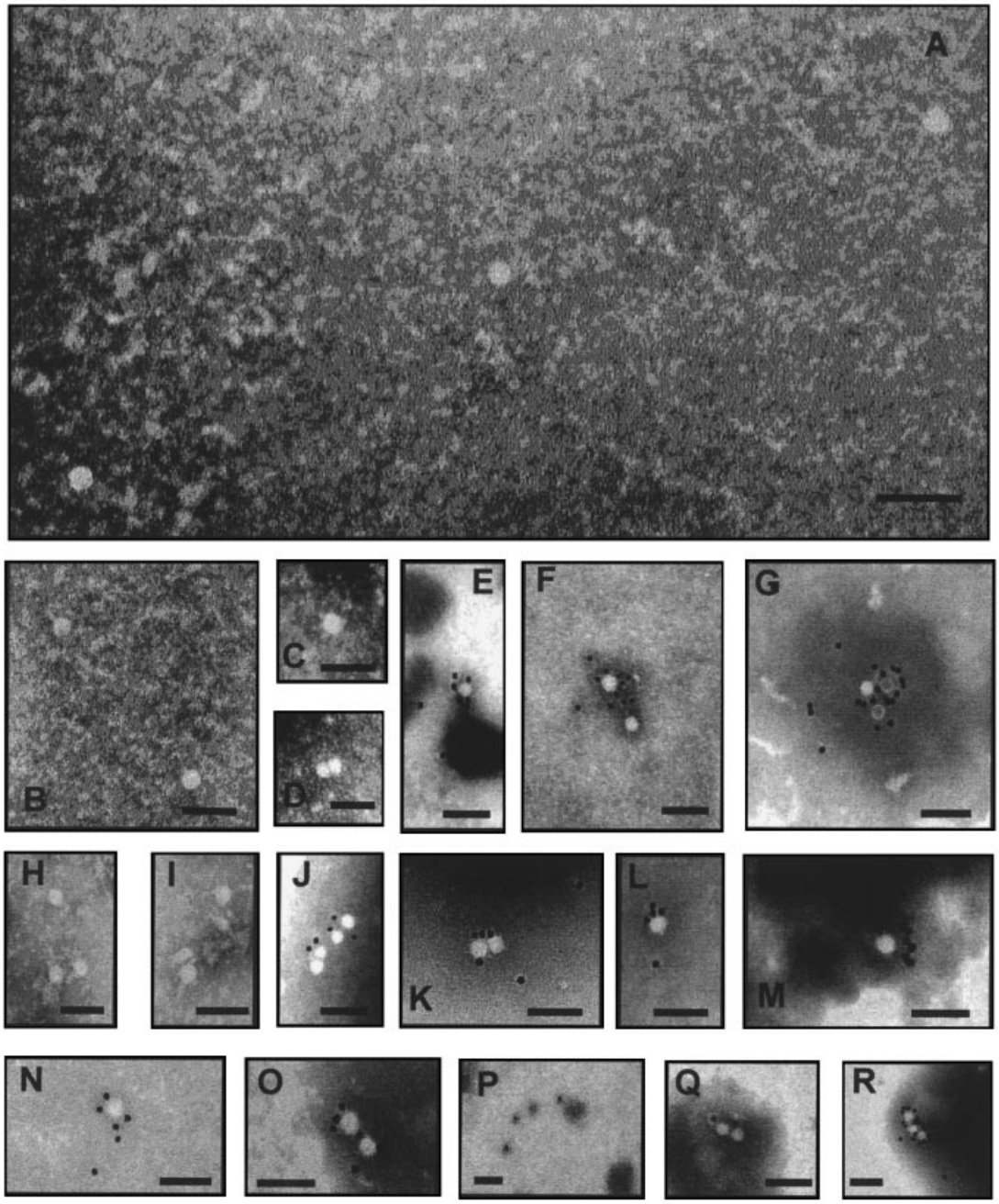

Fig. 4. Electron microscopy of HAV VLPs recovered after 16 h of expression. (A, B) VLPs extracted from *E. coli* harbouring pTHAVF. (C, D) VLPs from pTHAVF after IEM using mAbs K2-4F2 and K3-4C8. (E, F) and (J–L, N and O) VLPs from pTHAVF and pTHAVP1, respectively, after protein A–gold immunoprecipitation using mAbs K2-4F2 and K3-4C8. (H, I) VLPs recovered from *E. coli* harbouring pTHAVP1. (G, M and Q) Immunoprecipitation applied to an HAV suspension, an HAV 70S fraction (R) and a pBTac extract (P). Bars, 100 nm.

Δ3D; 3·1 for Δ3A; and 2·5 for 3B. After 4 h of expression at 37 °C, the formation of inclusion bodies could be observed in all constructs. Analysis of the protein composition of these inclusion bodies revealed the presence of two proteins with calculated molecular masses of around 30–32 and 24–26 kDa (Fig. 5A), whose HAV nature was confirmed by Western blot analysis using anti-3C antibodies (Fig. 5B). Since the entire cloned protein, approximately 35 kDa, could not be detected and the 24–26 kDa protein was quantitatively much more important than the 30–32 kDa protein, it was suggested that proteolytical processing had occurred. The same pattern of banding was observed in both the wild-type and the mutant constructs. If these proteins are the result of proteolytical processing, a bacterial protease should be responsible for it, as

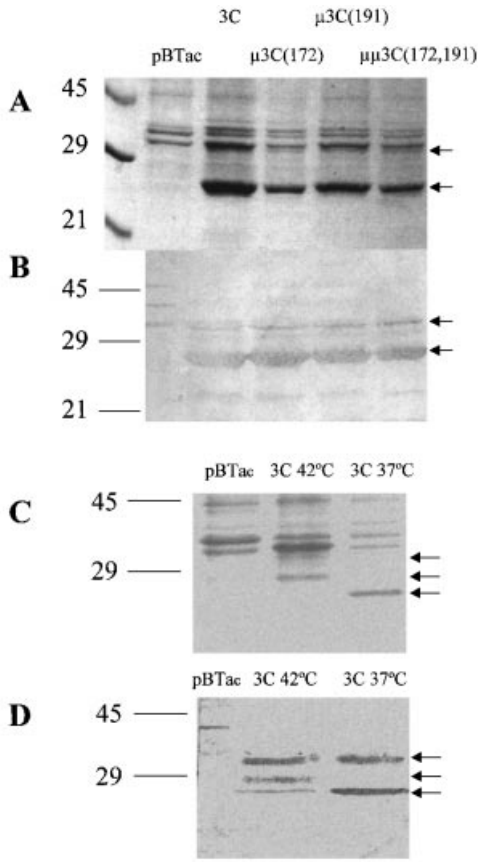

Fig. 5. SDS–PAGE and immunoblot analysis of HAV proteins synthesized in *E. coli* from plasmids containing P3 sequences. After 4 h of induction, the inclusion bodies formed were collected by lysozyme extraction and centrifugation at 11000 ***g***. Pellet fractions were solubilized in 8 M urea and 7% SDS. The proteins extracted were separated by 10% SDS–PAGE. (A) Coomassie blue staining of proteins obtained at 37 °C. (B) Western blot using anti-3C antibodies of proteins obtained at 37 °C. (C) Coomassie blue staining of proteins obtained either at 37 °C or at 42 °C. (D) Western blot using anti-3C antibodies of proteins obtained either at 37 °C or at 42 °C. pBTac, insoluble fraction from *E. coli* harbouring pBTac2; 3C, inclusion bodies from *E. coli* harbouring pTHAV3C; μ3C(172), inclusion bodies from *E. coli* harbouring pTHAVμ3C(172); μ3C(191), inclusion bodies from *E. coli* harbouring pTHAVμ3C(191); μμ3C(172,191), inclusion bodies from *E. coli* harbouring pTHAVμμ3C(172,191). Arrows indicate protein (A, C) and immunoreactive (B, D) bands corresponding to P3 expression. Molecular size standards are shown on the left.

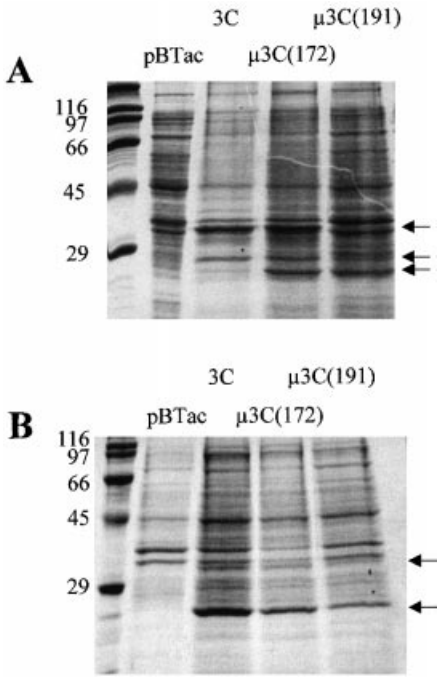

Fig. 6. SDS–PAGE of HAV proteins synthesized in *E. coli* from plasmids containing P3 sequences. After 4 h of induction at 42 °C in the absence (A) or presence (B) of NEM, proteins contained in the inclusion bodies were extracted and separated as described in the legend to Fig. 5. pBTac, insoluble fraction from *E. coli* harbouring pBTac2; 3C, inclusion bodies from *E. coli* harbouring pTHAV3C; μ3C(172), inclusion bodies from *E. coli* harbouring pTHAVμ3C(172); and μ3C(191), inclusion bodies from *E. coli* harbouring pTHAVμ3C(191). Arrows indicate P3 protein bands. Molecular size standards are shown on the left.

the mutant constructs are inactive (Jia *et al.*, 1991a, b; Gosert *et al.*, 1997; Malcolm *et al.*, 1992). To determine the sites at which proteolysis had occurred, both proteins were subjected to N-terminal sequencing. It was impossible to elucidate the sequence of the larger protein; the N-terminal sequence of the smaller protein was MMEFY in all of the constructs. This sequence is not compatible with any of the cleavages described for the P3 polyprotein; instead, this sequence, in the middle of the 3C region, has been described as an internal initiation product (Harmon *et al.*, 1992). However, in this latter work, the primary precursor was detected mainly in the mutant constructs and several proteolytic events were described to be associated exclusively with the wild-type protease. To investigate further whether any difference could exist between the processing patterns of the wild-type and mutant proteases, expression at different temperatures (20, 30 and 42 °C) was performed. At 20 and 30 °C, no difference in the processing pattern could be observed in comparison with the 37 °C pattern, except for the lower amount of protein generated (data not shown). However, at 42 °C, the protein pattern changed with respect to that observed at 37 °C (Fig. 5, C and D; Fig. 6A), although not significantly among the different constructs. The accumulation of the 24–26 kDa protein decreased significantly (from 15% at 37 °C to 1% at 42 °C of the total protein in the case of the 3C construct). However, concomitantly, a protein of approximately 29 kDa always appeared (12% of the total protein in the case of the 3C construct), suggesting that this new protein could be its actual precursor. On the other hand, a protein of around 35 kDa was sometimes detected (Fig. 6A): this could correspond to the entire cloned protein. Nevertheless, the existence of a temperature-dependent internal initiation of translation cannot be ruled out, as the structure of the RNA may, itself, be temperature-dependent. However, in this case, at 42 °C, the

band of 24–26 kDa should disappear without the appearance of the 29 and 35 kDa bands. Conversely, two clear bands of similar protein concentration were obtained, regardless of the temperature (37 or 42 °C), when expressing pTHAVVP0 (VP0 cloned into pBTac2 vector), which contains an actual Shine–Dalgarno sequence (GAGGAAG, −12 to −8) instead of that of the mRNA of the 24–26 kDa protein (GAGAA, −15 to −13) (data not shown). None of the 3C proteins obtained at 42 °C could be sequenced. To try to produce the 29 kDa protein in conditions other than growth at 42 °C, experiments in the presence of the protease inhibitor NEM at 37 °C were performed in order to reduce the activity of cysteine proteases. Under these conditions, the processing pattern was more or less the same than that without NEM for all of the constructs, although the 29 kDa protein could be recovered (3% of the total protein). The N-terminal sequence of the 29 kDa protein, which could be determined only for pTHAVμ3C(172), was STLE, confirming the cleavage between 3B and 3C (VESQSTLE), even in the absence of an active viral protease. Taken together, these results suggest that the 24–26 kDa protein probably corresponds to a truncated Δ3CΔ3D protein, while the 29 kDa corresponds to a truncated 3CΔ3D protein. Surprisingly, when expression was tested at 42 °C in the presence of NEM, the typical pattern of bands with the predominant 24–26 kDa protein (Fig. 6B) was observed for all of the constructs. The first conclusion that could be drawn from these results was the exclusion of the internal initiation of translation hypothesis, since NEM should interact with proteases rather than with the mRNA. Since the same pattern of processing was observed among the different constructs under any condition tested, two possibilities could exist: either the different mutants were still active *in vivo* in *E. coli* or a bacterial protease was responsible for processing. This second possibility is likely to be the correct one, as, in the case of pTHAVP1, we could also observe processing in the absence of the viral protease. On the other hand, it could be concluded from the results obtained with the expression in the presence of NEM that this protease inhibitor did not affect, to a great extent, the bacterial protease responsible for this processing. Consequently, it should in some way interact with the HAV P3 molecule(s), since the pattern of processing at 42 °C differs depending on the presence or absence of this inhibitor. Assuming that a bacterial protease is responsible for the described processing, one question arises: why this activity has not been described earlier in other works on the expression of HAV protease in *E. coli* (Gauss-Müller *et al.*, 1991; Harmon *et al.*, 1992). The most prominent difference between the constructs made in the aforementioned studies and our constructs is that, in our work, the HAV sequences are not fused to bacterial genes, whereas in the aforementioned work, they were fused to the *β*-galactosidase gene or the TrpE-coding sequences. The synthesis of fusion proteins is a well-known method to avoid bacterial proteolysis of recombinant proteins. On the other hand, the 3C-derived protein with the amino acid sequence MMEFY has been interpreted to be a result of an internal initiation process (Harmon *et al.*, 1992) rather than proteolysis. This latter conclusion can only be drawn after expression at 42 °C, a temperature that was not tested in the aforementioned work. To assess whether some of the well-known *E. coli* proteases were responsible for this processing, expression was performed in the BLB21(DE3) strain of *E. coli*, deficient for the omp T and lon proteases: the results obtained were identical to those with the JM109 strain (data not shown), thus ruling out the participation of these proteases in P3 processing. The fact that the final product of 24–26 kDa decreased drastically at 42 °C in favour of its 29 kDa precursor suggests that the conformation of this cleavage sequence may be temperature-dependent. On the other hand, the change of pattern of processing at 42 °C when including NEM in the expression conditions suggested that this inhibitor interacted with the 3C-containing molecules, thereby constricting the conformational change induced by the high temperature. Some characteristics of the crystal structure of the HAV 3C protease indicates that the residues surrounding the cleavage site (P5–P4–P3–P2–P1–P1′) that produces the Δ3CΔ3D protein are located on the surface of the molecule in the form of a reverse-turn helix 3_{10} (Bergmann *et al.*, 1997). In the case of the picornaviral 3C proteases substrate recognition, determinants other than amino acid pairs at the scissile bond should exist, since not all cognate pairs of amino acids encoded in their polyproteins are cleaved by 3C. Additional substrate determinants may include the secondary or tertiary structure through their recognition by or their accessibility to the protease (Dougherty & Semler, 1993). Extrapolations from structural data suggest that authentic 3C sites are presented typically in flexible, turn-coil surface configurations (Palmenberg, 1990). If we assume that a bacterial protease is capable of processing the HAV polyprotein, its activity and requirements should be similar to HAV 3C and then the structural determinants described for picornaviruses would be applicable to this unidentified bacterial protease. In this sense, the structure of the cleavage site described above, in the context of the 3C protein, appears to be conformationally adequate, although we should bear in mind that our uncleaved precursor protein possibly corresponds to a 3CΔ3D protein rather than 3C alone and that the additional amino acids could induce conformational changes. When comparing the primary sequence of this unusual cleavage site with the actual sites of the HAV polyprotein, from residues P5 to P5′, several conclusions may be drawn: the glutamic acid of position P5 is shared with the sites 2BX2C and 2CX3A; methionine at position P1′ is shared with site VP2XVP3; methionine at position P2′ is shared with sites VP2XVP3 and VP1X2A; glutamic acid at position P3′ is shared with site VP4XVP2; and, more importantly, glutamic acid at position P1 is shared with sites VP1X2A and 3AX3B. The glutamic acid at position P1 has been described for several picornavirus 3C proteases (Palmenberg, 1990) and 3C-like proteases of caliciviruses (Boniotti *et*

al., 1994; Wirblich *et al.*, 1995; Sosnovtsev *et al.*, 1998; Liu *et al.*, 1999). Interestingly, the positions P3–P2–P1 are conserved with the X-TCP site of the calicivirus rabbit haemorrhagic virus and the P2–P1 residues are conserved with the TCP–Pol site of the same calicivirus member (Wirblich *et al.*, 1995). Since this protease may process sites other than those of the actual HAV 3C protease, it is likely that the efficiency of the formation of VLPs will be low. On the other hand, most of the protein is accumulated as inclusion bodies, which also contributes to the low level production of VLPs. The synthesis of HAV subvirus structures and VLPs in two eukaryotic expression systems, baculovirus (Stapleton *et al.*, 1991; Rosen *et al.*, 1993) and vaccinia virus (Winokur *et al.*, 1991), has been reported. In both systems, generation of 70S empty particles has been accomplished. However, similar approaches for the generation of empty capsids of picornaviruses in prokaryotic systems have been described only for foot-and-mouth disease virus (Lewis *et al.*, 1991) and, in this case, the synthesis of particles is accomplished only with the collaboration of the viral protease. Recently, the existence of proteases other than 3C, i.e. host cell proteases, which play a role in the processing and maturation of the HAV capsid polyprotein, has been proposed (Martin *et al.*, 1999; Graff *et al.*, 1999). Since the picornaviral 3C proteinases (including the HAV 3C) are defined as cysteine proteinases with a serine proteinase-like folding pattern, and the bacterial protease appears to be resistant to NEM, their similarity may rely upon similar serine-like folding. In any case, the biological significance of an enteric bacteria protease capable of processing the HAV polyprotein is unknown.

We are grateful to Dr R. Lluna, Hospital Militar and Dr Z.-M. Yun, Institute of Virology, Beijing for the gift of antibodies. We acknowledge the technical expertise of the Serveis Científic-Tècnics of the University of Barcelona. This work was supported in part by grants BIO95-0061 and BIO99-0455 from the CICYT, Spain, and grants 1997SGR 00224 and 1999SGR 00022 from the Generalitat de Catalunya. J. F. González-Dankaart and K. J. Guo were recipients of FPI and DGICYT fellowships from the Spanish Ministry for Education and Science, respectively.

References

Bergmann, E. M., Mosimann, S. C., Chernaia, M. M., Malcolm, B. A. & James, M. N. G. (1997). The refined crystal structure of the 3C gene product from hepatitis A virus: specific proteinase activity and RNA recognition. *Journal of Virology* **71**, 2436–2448.

Boniotti, B., Wirblich, C., Sibilia, M., Meyers, G., Thiel, H.-J. & Rossi, C. (1994). Identification and characterization of a 3C-like protease from rabbit hemorrhagic disease virus, a calicivirus. *Journal of Virology* **68**, 6487–6495.

Borovec, S. & Anderson, D. A. (1993). Synthesis and assembly of hepatitis A virus-specific proteins in BS-C-1 cells. *Journal of Virology* **67**, 3095–3102.

Bosch, A., Guo, K.-J., González-Dankaart, J. F., Arnijas, X., Guix, S., Sánchez, G., Ribes, E. & Pintó, R. M. (1997). Recombinant hepatitis A virus polyprotein expressed in *E. coli* assembles in subviral structures. In *Viral Hepatitis and Liver Disease*, pp. 27–31. Edited by M. Rizzetto, R. H. Purcell, J. L. Gerin & G. Verme. Turin: Edizioni Minerva Medica.

Cohen, J. I., Ticehurst, J. R., Purcell, R. H., Buckler-White, A. & Baroudy, B. M. (1987). Complete nucleotide sequence of wild-type hepatitis A virus: comparison with different strains of hepatitis A virus and other picornaviruses. *Journal of Virology* **61**, 50–59.

Cromeans, T., Sobsey, M. D. & Fields, H. A. (1987). Development of a plaque assay for a cytopathogenic, rapidly replicating isolate of a hepatitis A virus. *Journal of Medical Virology* **22**, 45–56.

Dougherty, W. G. & Semler, B. L. (1993). Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. *Microbiological Reviews* **57**, 781–822.

Gauss-Müller, V., Jürgensen, D. & Deutzmann, R. (1991). Autoproteolytic cleavage of recombinant 3C proteinase of hepatitis A virus. *Virology* **182**, 861–864.

Gosert, R., Dollenmaier, G. & Weitz, M. (1997). Identification of active-site residues in protease 3C of hepatitis A virus by site-directed mutagenesis. *Journal of Virology* **71**, 3062–3068.

Graff, J., Richards, O. C., Swiderek, K. M., Davis, M. T., Rusnak, F., Harmon, S. A., Jia, X.-Y., Summers, D. F. & Ehrenfeld, E. (1999). Hepatitis A virus capsid protein VP1 has a heterogeneous C terminus. *Journal of Virology* **73**, 6015–6023.

Harmon, S. A., Updike, W., Jia, X.-Y., Summers, D. & Ehrenfeld, E. (1992). Polyprotein processing in *cis* and in *trans* by hepatitis A virus 3C protease cloned and expressed in *Escherichia coli*. *Journal of Virology* **66**, 5242–5247.

Hurley, M. A. & Roscoe, M. E. (1983). Automated statistical analysis of microbial enumeration by dilution series. *Journal of Applied Bacteriology* **55**, 159–164.

Jia, X.-Y., Ehrenfeld, E. & Summers, D. (1991a). Proteolytic activity of hepatitis A virus 3C protein. *Journal of Virology* **65**, 2595–2600.

Jia, X.-Y., Scheper, G., Brown, D., Updike, W., Harmon, S., Richards, O., Summers, D. & Ehrenfeld, E. (1991b). Translation of hepatitis A virus RNA *in vitro*: aberrant internal initiations influenced by 5′ noncoding region. *Virology* **182**, 712–722.

Kusov, Y. & Gauss-Müller, V. (1999). Improving proteolytic cleavage at the 3A/3B site of the hepatitis A virus polyprotein impairs processing and particle formation, and the impairment can be complemented in *trans* by 3AB and 3ABC. *Journal of Virology* **73**, 9867–9878.

Lemon, S. M. & Robertson, B. H. (1993). Current perspectives in the virology and molecular biology of hepatitis A virus. *Seminars in Virology* **4**, 285–295.

Lewis, S. A., Morgan, D. O. & Grubman, M. J. (1991). Expression, processing, and assembly of foot-and-mouth disease virus capsid structures in heterologous systems: induction of a neutralizing antibody response in guinea pigs. *Journal of Virology* **65**, 6572–6580.

Liu, B. L., Viljoen, G. J., Clarke, I. N. & Lambden, P. R. (1999). Identification of further proteolytic cleavage sites in the Southampton calicivirus polyprotein by expression of the viral protease in *E. coli*. *Journal of General Virology* **80**, 291–296.

Malcolm, B. A., Chin, S. M., Jewell, D. A., Stratton-Thomas, J. R., Thudium, K. B., Ralston, R. & Rosenberg, S. (1992). Expression and characterization of recombinant hepatitis A virus 3C proteinase. *Biochemistry* **31**, 3358–3363.

Martin, A., Benichou, D., Chao, S. F., Cohen, L. M. & Lemon, S. M. (1999). Maturation of the hepatitis A virus capsid protein VP1 is not dependent on processing by the $3C^{pro}$ proteinase. *Journal of Virology* **73**, 6220–6227.

Murphy, F. A., Fauquet, C. M., Bishop, D. H. L., Ghabrial, S. A., Jarvis, A. W., Martelli, G. P., Mayo, M. A. & Summers, M. D. (editors) (1995). *Virus Taxonomy. Sixth Report of the International Committee on Taxonomy of Viruses*. Vienna & New York: Springer-Verlag.

Palmenberg, A. C. (1990). Proteolytic processing of picornaviral polyprotein. *Annual Review of Microbiology* **44**, 603–623.

Pintó, R. M., Diez, J. M. & Bosch, A. (1994). Use of the colonic carcinoma cell line CaCo-2 for *in vivo* amplification and detection of enteric viruses. *Journal of Medical Virology* **44**, 310–315.

Probst, C., Jecht, M. & Gauss-Müller, V. (1997). Proteinase 3C-mediated processing of VP1–2A of two hepatitis A virus strains: *in vivo* evidence for cleavage at amino acid position 273/274 of VP1. *Journal of Virology* **71**, 3288–3292.

Probst, C., Jecht, M. & Gauss-Müller, V. (1999). Processing of proteinase precursors and their effect on hepatitis A virus particle formation. *Journal of Virology* **72**, 8013–8020.

Rosen, E., Stapleton, J. T. & McLinden, J. (1993). Synthesis of immunogenic hepatitis A virus particles by recombinant baculoviruses. *Vaccine* **11**, 706–712.

Sosnovtsev, S. V., Sosnovtseva, S. A. & Green, K. Y. (1998). Cleavage of the feline calicivirus capsid precursor is mediated by a virus-encoded proteinase. *Journal of Virology* **72**, 3051–3059.

Stapleton, J. T., Rosen, E. & McLinden, J. (1991). Detection of hepatitis A virus capsid proteins in insect cells infected with recombinant baculoviruses encoding the entire hepatitis A virus open reading frame. In *Viral Hepatitis and Liver Disease*, pp. 50–54. Edited by F. B. Hollinger, S. M. Lemon & H. S. Margolis. Baltimore: Williams & Wilkins.

Stapleton, J. T., Raina, V., Winokur, P. L., Walters, K., Klinzman, D., Rosen, E. & McLinden, J. H. (1993). Antigenic and immunogenic properties of recombinant hepatitis A virus 14S and 70S subviral particles. *Journal of Virology* **67**, 1080–1085.

Winokur, P. L., McLinden, J. H. & Stapleton, J. T. (1991). The hepatitis A virus polyprotein expressed by a recombinant vaccinia virus undergoes proteolytic processing and assembly into viruslike particles. *Journal of Virology* **65**, 5029–5036.

Wirblich, C., Sibilia, M., Boniotti, M. B., Rossi, C., Thiel, H.-J. & Meyers, G. (1995). 3C-like protease of rabbit hemorrhagic disease virus: identification of cleavage sites in the ORF1 polyprotein and analysis of cleavage specificity. *Journal of Virology* **69**, 7159–7168.

Received 26 July 2001; Accepted 11 October 2001