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

CHRISTIAN PROBST, MONIKA JECHT, AND VERENA GAUSS-MÜLLER*

Institute for Medical Molecular Biology, Medical University of Lübeck, D-23538 Lübeck, Germany

Received 29 July 1996/Accepted 4 December 1996

Two prominent features distinguish hepatitis A virus (HAV) from other members of the picornavirus family. A C-terminally prolonged precursor of the structural protein VP1 is incorporated into assembly intermediates (e.g., the provirion), and a single proteinase is contained within the HAV polyprotein. Using an in vivo expression system, we show that proteolytic liberation of VP1 from its precursors P1-2A and VP1-2A is catalyzed by the virus-encoded proteinase 3C^{pro}. Among the proposed cleavage sites within VP1-2A, the Glu/Ser pair found at VP1 amino acid position 273/274 of most HAV strains is efficiently processed, whereas proteolysis of the Val/Ser site of the attenuated HM175 strain is protracted. Two mutations within VP1-2A (Lys[297]Arg and Ser[330]Asn) had no effect on 3C^{pro}-mediated cleavage at this site. Additional sites in this region of VP1-2A can also be utilized as substrates by the proteinase, yet less efficiently, and might give rise to smaller and larger VP1 polypeptides also detected in HAV-infected cells.

The picornavirus RNA genome contains a single open reading frame encoding one large polyprotein from which the structural proteins (VP0, VP3, and VP1) as well as the functional proteins (2A, 2B, 2C, 3A, 3B, 3C^{pro}, and 3D^{pol}) are produced by proteolytic cleavages. Most of the processing events are catalyzed by the main viral proteinase, 3C^{pro}. Yet the primary cleavage step leading to the liberation of the structural proteins' precursor (P1 or P1-2A) is mediated by the nonstructural protein 2A. Two different proteolytic mechanisms have evolved in the enteroviruses and rhinoviruses and cardio/apthoviruses for this processing step, both of which apparently ensure a regulatory control distinct from the secondary cleavages (6, 20). Hepatitis A virus (HAV), the only member of the hepatovirus genus, differs from all other picornaviruses in that no proteolytic activity is found within the genomic region of 2A and that the sole viral proteinase, 3C^{pro}, seems to mediate the primary as well as all secondary proteolytic cleavages (14–16, 21, 22). However, the proteinase involved in the formation of mature VP1 and the exact location of the cleavage site are still open to debate.

Another unique feature of HAV relates to the viral capsid proteins. The smallest structural polypeptide, VP4, which is presumably liberated by an autocatalytic maturation cleavage, has not yet been detected in HAV capsids (9, 17, 23). Furthermore, the correct molecular mass of the main structural protein VP1 has for long been puzzling (1, 3, 17). In contrast to all other picornaviruses, it was observed that HAV provirions and procapsids as well as the 14-S-pentameric assembly precursor contain VP1-2A, the C-terminally extended form of VP1, whereas mature virions contain VP1 instead. Neither VP1 nor VP1-2A (also termed PX) migrates with the apparent molecular mass predicted from the nucleotide sequence first described (4). The electrophoretic mobilities of 3C^{pro} cleavage products have led to the genomic localization of VP1-2A, which comprises amino acid residues 491 to 836 of the HAV

polyprotein (10, 17, 19). In vitro experiments suggested that the cleavage within VP1-2A is also mediated by 3C^{pro} (21). Considering the homology to other picornaviruses and the preferred substrate recognition sequence of proteinase 3C, several sites within VP1-2A (Q/S, L/N, E/S, and E/S at amino acid positions 254, 264, 273, and 285 of VP1, respectively) have recently been proposed as potential cleavage points (Fig. 1) (5). Among those, the Glu residue at amino acid position 273 was suggested to be the C-terminal residue of mature VP1. In this study, the cleavability of this and/or the additional potential sites by HAV 3C^{pro} or a host proteinase was not addressed. In infected cells, VP1 proteins with different electrophoretic mobilities were observed, and thus, the possibility that cleavage at multiple sites within VP1-2A is the reason for this heterogeneity cannot be ruled out (1, 3, 17). In addition to the elusive C terminus of VP1, the biological role of HAV protein 2A and the various forms of HAV VP1-2A during the infectious cycle are still completely unknown (11).

As an alternative approach to characterize the primary structure of HAV VP1 in vivo and to elucidate its generation from precursor polypeptides, we have studied the proteolytic processing of precursor P1-2A by 3C^{pro} in eukaryotic cells. Comparison of the amino acid sequences at the putative VP1-2A junction from various HAV strains and variants had revealed a mutation at position 273 where a Glu residue in the wild-type variant of strain HM175 and all others (including HAS-15) is changed to a Val in the attenuated HAV/7 variant of HAV strain HM175 (Fig. 1) (4). Here we tested whether this mutation reduces the efficiency of catalysis by 3C^{pro} and whether mutations in 2A have any influences on the cleavability of VP1-2A.

Construction and expression of plasmids. To compare the cleavability of HAV polypeptide VP1-2A and its precursor P1-2A by 3C^{pro} at the proposed sites, P1-2A substrates containing either the amino acid sequence of the attenuated HAV strain HM175 or that of HAS-15 were expressed in vivo in the presence of HAV proteinase 3C. To facilitate subsequent cloning, the recipient plasmid pEXT7-HM-VP4-Δ3D, encoding almost the entire HAV polyprotein in frame with the internal ribosomal entry site (IRES) and the first N-terminal amino

* Corresponding author. Mailing address: Institute for Medical Molecular Biology, Medical University of Lübeck, Ratzeburger Alle 160, 23538 Lübeck, Germany. Phone: 49-451-500 4085. Fax: 49-451-500 3637. E-mail: gaussmue@molbio.mu-luebeck.de.

FIG. 1. Amino acid sequence alignment of the attenuated HAV variant HM175 and strain HAS-15 at the C terminus of VP1 and the N terminus of 2A. The putative 3C^{pro} cleavage sites at VP1 positions 254, 264, 273, and 285 (boxed) were proposed by Dotzauer et al. (5). The VP1-2A cleavage site at position 273 was suggested on the basis of comigration studies with expression products derived from different VP1 constructs. Amino acid substitutions in strain HAS-15 are shown above the sequence of the attenuated variant of strain HM175. Numbers in square brackets correspond to amino acid positions in VP1 of HAV strain HM 175 (attenuated) (4).

acids of the encephalomyocarditis virus (EMCV), was constructed by ligating the *Xba*I (blunt ended with mung bean exonuclease)-*Xho*I fragment of pHAV/7 (4) into pEXT7-HIV-env (24) which was *Nco*I restricted, blunt ended with mung bean exonuclease, and cut with *Xho*I. To generate subclones representing P1-2A and P1 precursors with defined C termini, the following primers were used for PCR amplification of VP1 and VP1-2A of the attenuated strain HM175 and wild-type HAS-15, which carries the same amino acid sequence at the VP1/2A cleavage site as wild-type HM175: primer 1, 5'-GGG GTACCCACAACCATGGTTGGAGATGATTCTGGAGG T-3' (primer 5'-VP1); primer 2, 5'-CCGCTCGAGAAGCTTA TTGTGAAAACAGTCCCTTCAT-3' (primer 3'-2A); primer 3, 5'-CCGCTCGAGTTATTGTTCTGTGACAGACAAATA ACAACTAAAAGACAAG-3' (primer 3'-VP1[254]); and primer 4, 5'-CCGCTCGAGTTACTCAGTGGACAACATAG CTTTGAATTTAATGGA-3' (primer 3'-VP1[273]). The *Xho*I sites are marked in italics, and the antisense stop codon is underlined.

With primers 1 and 2 and pT7-HAV1 as the template (12), the VP1-2A region of attenuated HM175 was amplified by PCR. After restriction with *Nco*I and *Xho*I, the 0.5-kb fragment representing nucleotides 2809 to the 3' end of the cDNA fragment was inserted into pEXT7-HM-VP4-Δ3D cut with the same enzymes, resulting in pEXT7-HM/HM-P1-2A(V/S) (Fig. 2). For the construction of pEXT7-HM/HAS-P1-2A(E/S) containing the E/S cleavage site and 2A of HAS-15, the same primers and cloning strategy, but with pGEM-HAS-15 as the template, were used. The reciprocal exchange of the 1-kb *Eam*1105I restriction fragments of pEXT7-HM/HM-P1-2A(V/S) and pEXT7-HM/HAS-P1-2A(E/S) resulted in the construction of pEXT7-HM/HM-P1-2A(E/S) and pEXT7-HM/HAS-P1-2A(V/S). pEXT7-HM-P1[254] was produced by inserting the PCR-amplified (with primers 1 and 3 and pT7-HAV1 as the template) and *Nco*I-*Xho*I-restricted 0.2-kb fragment into the recipient clone. pEXT7-HM-P1[273] was generated accordingly with primer 4 instead of primer 3 and pGEM-HAS-15 as the template. The cDNA clone pEXT7-LA-3C for expression of active 3C^{pro} was produced by inserting the blunt-ended *Nco*I (mung bean exonuclease-treated)-*Sal*I-restricted 0.6-kb fragment derived from clone pHAV-3CEX (18) into pEXT7-HIV-env which was treated with *Nco*I, mung bean exonuclease, and *Xho*I. Expression of the HAV coding sequences is driven by

the T7 promoter which is followed by the IRES of EMCV to improve translation initiation (7). All constructs and the amino acid sequences near the VP1-2A cleavage site are depicted in Fig. 2. Twenty-four hours after transfecting COS-7 cells with a total of 1 μg of cDNA and subsequent infection with vaccinia virus vTF7-3, the expression products were analyzed by immunoblot using antisera directed against purified HAV VP1 (9) and recombinant His2A.

Products of HAV P1-2A processing by 3C^{pro}. Expression of the various P1 and P1-2A constructs resulted in products of the expected molecular masses which were detected by anti-VP1 (Fig. 3, lanes 1 to 6). Some minor products of different molecular masses might have arisen by incorrect initiation or termination during overexpression. Expression of the same cDNA constructs together with the cDNA of HAV proteinase 3C resulted in highly specific proteolytic cleavage giving rise to polypeptides of the calculated molecular masses listed in Table 1 (Fig. 3, lanes 7 to 20). Processing products P1, VP3-VP1-2A, VP3-VP1, VP1-2A, and VP1 were identified by anti-VP1 (lanes 7 to 13). P1-2A and VP1-2A were also recognized by anti-2A (lanes 14 to 20). VP3-VP1-2A could not be specifically detected by anti-2A due to the serum's cross-reactivity with comigrating bovine serum albumin. VP1-2A cleaved off P1-2A of HM175 by 3C^{pro} (lanes 7, 8, 14, and 15) and HM/HAS (lanes 9, 10, 16, and 17) migrated to the position predicted from the calculated molecular mass. The slightly faster mobilities of the HAS-15 2A-containing processing intermediates compared to those of the respective HM175 (attenuated) proteins can be explained by the Lys-to-Arg and Ser-to-Asn mutations in 2A (Fig. 1). In analogy, the Val residue at position 273 leads to a slightly faster electrophoretic mobility of the VP1-containing proteins compared to mobilities of those containing a Glu instead (Fig. 3, lanes 7 and 8 or 9 and 10). The expected additional processing products derived from P1 and P1-2A, such as VP0-VP3, VP0, and VP3, were identified by anti-VP0 and anti-VP3 (reference 9 and data not shown). 2A as an 8-kDa end product of 3C^{pro}-mediated cleavage was not detected by immunoblot using anti-2A. No further cleavage of VP0 into VP4 and VP2 was observed.

Efficiency of the VP1-2A cleavage. The mobility and amount of VP1 cleaved off the various precursor polypeptides were of particular interest. VP3-VP1[254] and VP1[254] liberated from P1[254] (lane 12) migrated with higher mobilities compared to those of VP3-VP1[273] and VP1[273] derived from P1[273] (lane 11), respectively. The latter processing product (VP1[273]) comigrated with VP1 produced in infected cells (lane 13) and the major protein cleaved off P1-2A (lanes 8 and 10), supporting the notion that the VP1-2A cleavage site is located at position 273 (5). Comparison of lanes 7 and 8 and of lanes 9 and 10 clearly showed that the main cleavage products of precursor P1-2A (V/S) were 2A-containing polypeptides whereas those with E/S at the same position comprise about equal amounts of VP1 and VP1-2A. Since these proteins differ only in their amino acid sequences at the presumed cleavage site, it can be concluded that the E/S junction in HAS-15 (or wild-type HM175) is a much better substrate for 3C^{pro} than the V/S site in the attenuated HM175 strain. In agreement with this notion, more P1 and VP3-VP1 are produced from P1-2A (E/S) (lanes 8 and 10) than from the corresponding V/S-containing substrates (lanes 7 and 9). Mutations in 2A did not influence the accessibility of 3C^{pro} to the cleavage site at position 273/274, as identical processing patterns were observed in lanes 7 and 9 or 8 and 10. These results clearly support our earlier observation that 3C^{pro} mediates cleavage at the VP1/2A junction (21) and confirm the predicted site at amino acid position 273/274 of VP1 (5). Efficient processing occurs only

FIG. 2. Schematic presentation of HAV cDNA constructs used for transient expression in the vaccinia virus T7 system described before (7). pEXT7 has been described before (24) and is similar to pTM 1. Additional amino acid residues at the N terminus of the open reading frames are derived from vector-encoded sequences downstream of the EMCV IRES. The open reading frames are followed by a stop codon and the T7 termination signal. The relevant amino acid sequences near the VP1-2A cleavage site are shown. pEXT7-HM/HM-P1-2A(V/S) contains the V/S position 273 cleavage site, whereas pEXT7-HM/HM-P1-2A(E/S) contains an E/S pair at this position. Instead of the 2A sequence of HM175 (attenuated), pEXT7-HM/HAS-P1-2A(E/S) and pEXT7-HM/HAS-P1-2A(V/S) contain that of HAS-15. Substrates of 3C^{pro} liberating authentic VP1[254] and VP1[273] without additional amino acids in the VP1-coding sequence are derived from pEXT7-HM-P1[254] and pEXT7-HM-P1[273].

when the dipeptide sequence at the cleavage site is E/S, whereas V/S found in attenuated HM175 is cleaved only to a small extent. This finding is not surprising in the light of the preferred cleavage site sequences identified for HAV 3C^{pro} and might explain the different yields of products described before (15, 21).

Both VP1 and VP1-2A derived from recombinant precursor polypeptides comigrated with the proteins of the HAV-infected cell extract (Fig. 3, lane 13). The relatively larger proportion of VP1 in the infected compared to the recombinant extracts is probably due to the different time points of harvest. Among the P1-2A processing products carrying the V/S cleavage site (Fig. 3, lanes 7 and 9), small amounts of three proteins with electrophoretic mobilities similar to that of VP1 and with molecular masses differing by approximately 1 kDa were detected by anti-VP1. The band in the middle of the triplet comigrated with VP1[273] in lanes 8 and 10 and VP1 of infected cells (lane 13). Possibly 3C^{pro} or a 3C^{pro}-induced host proteinase cleaves at additional sites which are N and C ter-

minal to amino acid position 273. Cleavage at the additional sites is also detected in all other P1-2A substrates, because proteins both migrating slightly faster (VP1f) or slower (VP1s) than VP1[273] were found. Suboptimal 3C^{pro} cleavage sites (L/N and E/S) have been considered at positions 264/265 and 285/286 in VP1-2A, which would give rise to proteins that migrate slightly faster and slower than VP1[273]. Cleavages at these sites are more easily detectable when processing at position 273/274 is inefficient, as is the case when P1-2A(V/S) is used as substrate for 3C^{pro}. Interestingly, an alternative proteolytic step, similar to the one reported here, was observed in mengovirus particles, where three amino acid residues are trimmed by an unknown mechanism (2). Processing at additional sites seemed to be induced by HAV 3C^{pro} and not by a virus-independent host proteinase, since eukaryotic expression of VP1[273] in the absence of 3C^{pro} led to only one product (5). At this point, the possibility that a 3C^{pro}-induced host proteinase catalyzes this cleavage or that overexpression of the recombinant proteinase and its substrate results in a star ac-

FIG. 3. Recombinant expression products derived from the cDNA constructs shown in Fig. 2. COS-7 cells (3×10^5) in 35-mm-diameter dishes were transfected with a total of 1 μ g of cDNA as indicated at the top of the figure. Twenty-four hours after the subsequent infection with vaccinia virus T7 (vTF7-3), the cells were lysed in 250 μ l of phosphate-buffered saline-Tween 20 (0.05%) and the extracts (80 μ l per 6-mm slot) were analyzed by immunoblot using rabbit polyclonal anti-VP1 and anti-2A as indicated at the bottom of the figure. An HAV (strain HM175 [attenuated]) infected extract of FRhK4 cells harvested 18 days postinfection was used as the comigration standard. The anti-2A serum shows cross-reactivity with bovine serum albumin (BSA). Immunologically reactive products are marked on the right side; positions of molecular mass standards are shown on the left.

tivity of 3C^{pro} leading to VP1f and VP1s in the vaccinia virus system cannot be ruled out. In contrast to most other users of the vaccinia virus T7 system, we identified the HAV expression products by immunoblot, which is highly sensitive and thus might detect minor 3C^{pro} processing products of P1-2A.

The variations observed in 3C^{pro} cleavage efficiency within VP1-2A may provide one explanation for the low replication rate of some HAV variants in cell culture as well as in experimental animals (8). Less-efficient processing within VP1-2A of the attenuated HAV strain HM175, which is the strain most commonly used, might result in lower yields of mature VP1.

the structural protein predominantly found in natural infectious capsids of HAV. Immunoreactive VP1 polypeptides of variable molecular masses have been observed in extracts of HAV-infected cells, though only rarely in preparations of purified particles, suggesting that alternative cleavage sites are utilized during particle assembly and maturation (3, 17). Whether the precursor and intermediate polypeptides of VP1 play distinct biological roles during the infectious cycle of HAV needs to be determined by other experimental approaches.

We thank H.-G. Kraeusslich for providing pEXT7-HIVenv and B. Moss for vaccinia virus vTF7-3.

C.P. was a recipient of a fellowship granted by the Studienstiftung des Deutschen Volkes, and M.J. is supported by a grant of the state of Schleswig-Holstein. The work was supported by the Deutsche Forschungsgemeinschaft (SFB 367, project B7).

REFERENCES

1. **Anderson, D. A., and B. Ross.** 1990. Morphogenesis of hepatitis A virus: isolation and characterization of subviral particles. *J. Virol.* **64**:5284-5289.
2. **Boege, U., and D. G. Scraba.** 1989. Mengo virus maturation is accompanied by C-terminal modification of capsid protein VP1. *Virology* **168**:409-412.
3. **Cho, M. W., and E. Ehrenfeld.** 1991. Rapid completion of replication cycle of hepatitis A virus subsequent to reversal of guanine inhibition. *Virology* **180**:770-780.
4. **Cohen, J. I., J. R. Ticehurst, R. H. Purcell, A. Buckler-White, and B. M. Baroudy.** 1987. Complete nucleotide sequence of wild-type hepatitis A virus: comparison with different strains of hepatitis A virus and other picornaviruses. *J. Virol.* **61**:50-59.

TABLE 1. Molecular mass of VP1-containing processing products

Protein	Calculated molecular mass (kDa)
VP3-VP1[285].....	59
VP3-VP1[273].....	58
VP3-VP1[264].....	57
VP3-VP1[254].....	56
VP1-2A HM/HM.....	38
VP1-2A HM/HAS.....	38
VP1[285].....	31
VP1[273].....	30
VP1[264].....	29
VP1[254].....	28

5. Dotzauer, A., A. Vallbracht, and G. M. Keil. 1995. The proposed gene for VP1 of HAV encodes for a larger protein than that observed in HAV-infected cells and virions. *Virology* **213**:671–675.
6. Dougherty, W. G., and B. L. Semler. 1993. Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. *Microbiol. Rev.* **57**:781–822.
7. Elroy-Stein, O., T. R. Fuerst, and B. Moss. 1989. Cap-independent translation of mRNA conferred by encephalomyocarditis virus 5' sequence improves the performance of the vaccinia/bacteriophage T7 hybrid expression system. *Proc. Natl. Acad. Sci. USA* **86**:6126–6130.
8. Emerson, S. U., M. Lewis, S. Govindarajan, M. Shapiro, T. Moskal, and R. H. Purcell. 1992. cDNA clone of hepatitis A virus encoding a virulent virus: induction of viral hepatitis by direct nucleic acid transfection of marmoset. *J. Virol.* **66**:6649–6654.
9. Gauss-Müller, V., F. Lottspeich, and F. Deinhardt. 1986. Characterization of hepatitis A virus structural proteins. *Virology* **155**:732–736.
10. Gosert, R., P. Cassinotti, G. Siegl, and M. Weitz. 1996. Identification of hepatitis A virus non-structural protein 2B and its release by the major protease 3C. *J. Gen. Virol.* **77**:247–255.
11. Harmon, S. A., S. U. Emerson, Y. K. Huang, D. F. Summers, and E. Ehrenfeld. 1995. Hepatitis A viruses with deletions in the 2A gene are infectious in cultured cells and marmosets. *J. Virol.* **69**:5576–5581.
12. Harmon, S. A., O. C. Richards, D. F. Summers, and E. Ehrenfeld. 1991. The 5'-terminal nucleotides of hepatitis A virus RNA, but not poliovirus RNA, are required for infectivity. *J. Virol.* **65**:2757–2760.
13. Jewell, D. A., W. Swietnicki, B. Dunn, and B. Malcolm. 1992. Hepatitis A virus 3C proteinase substrate specificity. *Biochemistry* **31**:7862–7869.
14. Jia, X.-Y., E. Ehrenfeld, and D. F. Summers. 1991. Proteolytic activity of hepatitis A virus 3C protein. *J. Virol.* **65**:2595–2600.
15. Jia, X.-Y., D. F. Summers, and E. Ehrenfeld. 1993. Primary cleavage of the HAV capsid protein precursor in the middle of the proposed 2A coding region. *Virology* **193**:515–519.
16. Jürgensen, D., Y. Y. Kusov, M. Fäcke, and V. Gauss-Müller. 1992. Proteinase 3C of hepatitis A virus: identification and characterisation of activity using *in vitro* transcription and translation. *J. Gen. Virol.* **74**:677–683.
17. Kusov, Y. Y., Y. A. Kazachkov, G. K. Dzagurov, G. A. Khozinskaya, M. S. Balayan, and V. Gauss-Müller. 1992. Identification of precursors of structural proteins VP1 and VP2 of hepatitis A virus. *J. Med. Virol.* **37**:220–227.
18. Malcolm, B. A., S. M. Chin, D. A. Jewell, R. J. Stratton-Thomas, K. B. Thudium, R. Ralston, and S. Rosenberg. 1992. Expression and characterization of recombinant hepatitis A virus 3C proteinase. *Biochemistry* **31**:3358–3363.
19. Martin, A., N. Escriou, S.-F. Choa, M. Girard, S. M. Lemon, and C. Wychowski. 1995. Identification and site-directed mutagenesis of the primary (2A/2B) cleavage site of the hepatitis A virus polyprotein: functional impact on the infectivity of HAV RNA transcripts. *Virology* **213**:213–222.
20. Palmenberg, A. C. 1990. Proteolytic processing of picornaviral polyproteins. *Annu. Rev. Microbiol.* **44**:603–623.
21. Schultheiß, T., Y. Y. Kusov, and V. Gauss-Müller. 1994. Proteinase 3C of hepatitis A virus (HAV) cleaves the HAV polyprotein P2-P3 at all sites including VP1/2A and 2A/2B. *Virology* **198**:275–281.
22. Schultheiß, T., W. Sommergruber, Y. Y. Kusov, and V. Gauss-Müller. 1995. Cleavage specificity of purified recombinant hepatitis A virus 3C proteinase on natural substrates. *J. Virol.* **69**:1727–1733.
23. Tesar, M., X.-Y. Jia, D. F. Summers, and E. Ehrenfeld. 1993. Analysis of a potential myristoylation site in hepatitis A virus capsid protein VP4. *Virology* **194**:616–626.
24. Wilk, T., H. Mierswa, H.-G. Kräusslich, J. J. Dunn, and V. Bosch. 1992. Expression of biologically active HIV glycoproteins using a T7 RNA polymerase-based eucaryotic vector system. *Virus Genes* **6**:229–246.