Identification of the protease domain in NS3 of hepatitis C virus

Dae Sung Han,1 Bumsuk Hahm,1 Hyune-Mo Rho2 and Sung Key Jang1,*

1 Department of Life Science, Pohang University of Science and Technology, San31 Hyojadong, Pohang, Kyungbuk 790-784, and 2 Department of Molecular Biology and Research Center for Cell Differentiation, Seoul National University, Seoul 151-742, Korea

NS3 of hepatitis C virus (HCV) is a serine protease that carries out the proteolytic processing of the non-structural proteins of the HCV polyprotein. Deletion analysis of the N terminus of NS2,3,4 fusion protein revealed that the N-terminal boundary of the active protease resides between amino acids 1050 and 1083. The processing patterns of internal deletion mutants of NS2,3,4 indicated that the C terminus of the enzymically active protease resides between amino acids 1115 and 1218. The N- and C-terminal boundaries of the protease were also confirmed by determining the trans-cleavage activity of internally deleted NS3,4. NS3 protease activity was inhibited by Cu²⁺ but was slightly enhanced by Zn²⁺. This report provides a possible approach for development of antiviral agents based on protease inhibitors.

Hepatitis C virus (HCV) is known to be implicated in liver cirrhosis and hepatocellular carcinoma (HCC) as well as hepatitis (Alter et al., 1989; Bruix et al., 1989; Kuo et al., 1989; Saito et al., 1990). Due to the similarities in molecular and biological characteristics of flaviviruses and pestiviruses, HCV has been placed in the family Flaviviridae (Francki et al., 1991; Miller & Purcell, 1990). Members of the Flaviviridae contain a positivesense RNA molecule in the virion, and there is no DNA stage in the viral life cycle (Hollinger, 1990). Nevertheless, many epidemiological data suggest that HCV infection can result in HCC after a long persistent infection (Colombo et al., 1989; Saito et al., 1990).

The genomic RNA of HCV contains a long open reading frame encoding about 3000 amino acids. Upon infection into a susceptible cell, the genomic RNA most likely functions as an mRNA to generate structural and nonstructural (NS) proteins as in the case of other positive-sense RNA viruses. Translation of HCV mRNA seems to be carried out via an internal ribosomal entry site (IRES) in the 5′-nontranslated region (5′NTR) of the HCV RNA (Tsukiyama-Kohara et al., 1992; Wang et al., 1993), even though translation by a subgenomic mRNA has not been completely ruled out (Yoo et al., 1992). Synthesis of the HCV polypeptide is initiated at the beginning of the long open reading frame (ORF) and ends at the termination codon. The resulting polypeptide is in turn processed into functional proteins by cellular proteasomes and virus-encoded proteases.

Bazan & Fletterick (1989) and Gorbalenya et al. (1989) have proposed a model for the prediction of serine-type proteases. This protease motif is found in the N-terminal region of flavivirus NS3 and pestivirus p80 (Chambers et al., 1991; Wengler et al., 1991; Wiskerchen & Collett, 1991). As was expected, changes in the amino acids that make up the catalytic triad of bovine viral diarrhea virus p80 and flavivirus NS3 led to inactivation of proteolytic activity (Chambers et al., 1991; Wengler et al., 1991; Wiskerchen & Collett, 1991). Since there is a similar serine protease sequence motif in the N-terminal part of HCV NS3, it has been predicted that NS3 may encode a serine protease.

Study of the proteolytic processing, however, has been hampered by the lack of an efficient in vitro cultivation system for HCV (Shimizu et al., 1992) and the low production of the virions in experimental animals (Alter et al., 1978; Tabor et al., 1978). Recently, the HCV proteolytic processing mechanism has begun to be uncovered by the use of cell culture expression and in vitro translation systems (Bartenschlager et al., 1993; Chiba et al., 1991; Grakoui et al., 1993a; Chiba et al., 1991; Harada et al., 1991; Hjikata et al., 1991, 1993; Kohara et al., 1992; Kumar et al., 1992; Matsuura et al., 1992; Ralston et al., 1993; Selby et al., 1993; Spaete et al., 1992; Tomei et al., 1993). The C/E1, E1/E2 and E2/NS2 junctions, preceded by hydrophobic amino acid stretches, have been shown to be cleaved by a membrane-dependent processing mechanism (Hjikata et al., 1991; Mizushima et al., 1994). This processing, therefore, was suggested to be carried out by a signal peptidase on the host cell membrane. Several independent studies revealed that the NS3/4A, 4A/4B, 4B/5A and 5A/5B junctions are
Short communication

Fig. 1. Effect of NS2 and NS3 on NS3/4 junction cleavage. (a) Schematic diagram of the 5' deletion mutants of pTHE760-1846. The positions of the EMCV 5'NTR, NS2, NS3, NS4A and NS4B, the T7 promoter, the histidine-coding sequence and the polycloning site are indicated on top of the boxes. The solid bar represents the mRNA transcribed from each construct, and the dashed line represents the deleted region of each construct from THE760-1846. The first amino acid at the translational initiation codon of the HCV polyprotein was arbitrarily denoted as number 1. The number of the first HCV-encoded amino acid is indicated at the left side of each mRNA. (b, c) Autoradiographs obtained after SDS–15% PAGE of the translation products directed by the mutant mRNAs. The incubation times of translation reactions were 20 min (b) and 180 min (c). The positions of molecular mass markers are indicated in kDa.

processed by the NS3 serine protease (Bartenschlager et al., 1993; Grakoui et al., 1993a, b; Tomei et al., 1993). It was suggested that the cleavage of the NS2/3 junction is carried out by a novel zinc-dependent metalloprotease (Hijikata et al., 1993).

Plasmid pTHE760-1846 containing the T7 promoter, the 5'NTR of encephalomyocarditis virus (EMCV) and a part of the HCV coding sequence (amino acids 760–1846) was constructed for efficient expression of HCV polyprotein. The HCV-coding sequence was originated from pUC119-NS2, pBM-NS3 and pBM-NS4 containing the HCV-coding sequence 2604–3424, 3360–5123 and 4672–5802, respectively (the numbers indicate nucleotide numbers as described by Hijikata et al., 1991). The HCV cDNA clones, which were obtained by the PCR method from the plasma of a patient infected with group I (type 1b) HCV, were kindly supplied by A. Nomoto and M. Kohara (The Tokyo Metropolitan Institute of Medical Science, Japan). The backbone of the plasmid was derived from pTM1 containing the T7 promoter and the EMCV 5'NTR. The T7 promoter and the EMCV 5'NTR were used for transcription of mRNAs and for efficient translation of the mRNAs in rabbit reticulocyte lysates (RRL), respectively. The
EMCV 5'NTR contains a particular RNA structure directing an efficient cap-independent translation via an IRES (Jang et al., 1988; Jang & Wimmer, 1990). Subcloning of plasmids was carried out by using standard methods (Sambrook et al., 1989). A serial deletion mutagenesis was carried out with pTHE760-1846 in the 5'-to-3' direction. All of the deletion mutants contained the 5'NTR and the translational initiation codon of EMCV and six histidine-coding sequences. The six histidine-coding sequence influenced neither NS3 protease activity nor the effect of divergent ions on NS3 protease activity (B. Hahm & S. Jang, unpublished). This sequence, on the other hand, increased expression of the following genes by an unknown mechanism. Therefore, this sequence was kept in all of the constructs used in this report.

In order to characterize the NS3 protease of HCV, various parts of HCV polypeptides spanning different regions of E2, NS2,3,4 were expressed in in vitro transcription and translation systems. In vitro transcription and translation were carried out by T7 RNA polymerase and in RRL, respectively. The newly synthesized polypeptides were analysed by SDS-PAGE with or without additional incubation of translation reaction mixture for further processing of polypeptides. The fusion polypeptide NS2,3,4 was predicted to provide both protease and substrate by analyses of amino acid sequences of the region (Choo et al., 1991).

Five N-terminal deletion mutants were constructed, and the proteolytic processing of translation products of the mutant RNAs was analysed (see Fig. 1). Upon incubation of the mRNAs in RRL for 20 min or 180 min, precursor polypeptides of the expected sizes were detected (Fig. 1, panels b and c). The deletion mutants containing the N-terminal part of NS3 showed not only precursors but also processed products of about 70, 75, 80 and 100 kDa from polypeptides 1016-1846, 962-1846, 906-1846 and 760-1846, respectively (Fig. 1, lanes 2-5 in panel c). A common band of a 21 kDa protein was detected from these four constructs. This suggests that the NS3/4 junction is cleaved by a protease residing downstream of amino acid 1016 of the HCV polypeptide. Deletion up to amino acid 1083 of HCV polypeptide abolished the NS3/4 junction cleavage (Fig. 1, lanes 1). This indicates that the N-terminal boundary of the HCV protease resides between amino acids 1016 and 1083. It should be noted that the amino acid 1083 is the histidine of the catalytic triad. Longer incubation of the translation mixture at 30 °C increased the amounts of the processed products (compare lanes 2, 3, 4 and 5 of panel a with those of panel c in Fig. 1). The efficiency of NS3/4 junction cleavage was influenced by the presence of NS2. As more of the N-terminus of NS2 was deleted, the efficiency of NS3/4 junction cleavage increased gradually (see lanes 2, 3, 4 and 5 in Fig. 1). This may suggest that cleavage of the NS2/3 junction precedes cleavage of the NS3/4 junction in HCV polypeptide processing. In contrast to reports by C. M. Rice's group (Grakoui et al., 1993a, b), we could not detect NS4A/4B cleavage, which would produce 5.8 and 15.2 kDa proteins instead of a 21 kDa protein. The discrepancy might be due to either the absence of the rest of the HCV proteins, other than E2', NS2,3,4' used in this work, or to differences in the systems used. Similar to our results, Bartenschlager et al. (1993) did not detect NS4A/4B cleavage in an RRL system.

The effects of metal ions on processing of HCV polypeptides were investigated by adding different metal ions at various concentrations to the mRNA translation reaction mixture. Most of the metal ions tested (Ag⁺⁺, Fe⁺⁺, Gd⁺⁺, Cd⁺⁺ and Ca⁺⁺; up to 100 μM) showed little or no effect on processing of HCV polypeptides (data not shown). The most striking effects were with Zn⁺⁺ and Cu⁺⁺ ions. Cu⁺⁺ inhibited NS3/4 junction cleavage almost completely at 25 μM or higher concentrations, but Zn⁺⁺ stimulated NS3/4 junction cleavage slightly at 25 μM or higher concentration (Fig. 2, lanes 4, 5, 7 and 9). The patterns of minor bands with lower molecular masses were not affected by Cu⁺⁺ or Zn⁺⁺. These results strongly suggest that the minor proteins were produced by a mechanism other than NS3-dependent cleavage of the NS3/4 junction. The effects of Zn⁺⁺ and Cu⁺⁺ on the proteolytic processing were not detected in polypeptides lacking the serine protease catalytic triad (Fig. 2, lanes 1 and 2). These results suggest that Zn⁺⁺ and Cu⁺⁺ specifically affect NS3 protease activity.

The N-terminal boundary of the NS3 protease was more precisely determined by making a deletion at the N-terminus of NS3 to give polypeptide 1050-1846. This polypeptide was processed less efficiently than polypeptide 1016-1846, but the cleavage patterns were similar to each other (compare lane 6 with lane 8 in Fig. 2). Zn⁺⁺ enhanced NS3/4 junction cleavage but Cu⁺⁺ blocked cleavage of this polypeptide (compare lanes 7 and 9 with lane 8 in Fig. 2). This indicates that the N-terminal boundary of the NS3 protease resides between amino acids 1050 and 1083 of NS3. Interestingly, the NS4 product was not detected among the processed products of polypeptide 1050-1846 in this particular experiment. When the incubation time for the translation reaction was shortened to 20 min to detect unstable intermediates generated from polypeptide 1050-1846 processing, the NS4 product of 21 kDa was barely detected after a 20 min incubation of the polypeptide (data not shown). This band disappeared after incubation of the translation mixture for 90 min (data not shown). The NS4 product generated from polypeptide 1016-1846, on the other
(a) Schematic diagram of 5' deletion mutants of pTHE760-1846. Symbols and letters are as in the legend to Fig. 1.

(b) In vitro translation was carried out at 30 °C for 90 min in the absence (−) or presence (+) of ZnCl₂ or CuCl₂. Zn²⁺ or Cu²⁺ were added to final concentrations of 100 μM. Three different mRNAs, THE1083-1846, 1016-1846 and 1050-1846 were used in translation reactions. The samples were analysed by SDS-15% PAGE. The positions of protein molecular mass markers are indicated in kDa.

Fig. 2. Effects of Cu²⁺ and Zn²⁺ on the proteolytic processing of polypeptides with different N-terminal ends. (a) In vitro translation was carried out at 30 °C for 90 min in the absence (−) or presence (+) of ZnCl₂ or CuCl₂. Zn²⁺ or Cu²⁺ were added to final concentrations of 100 μM. Three different mRNAs, THE1083-1846, 1016-1846 and 1050-1846 were used in translation reactions. The samples were analysed by SDS-15% PAGE. The positions of protein molecular mass markers are indicated in kDa.

In order to investigate the role of the C-terminal part of NS3 and NS4 in proteolytic processing of HCV polyprotein, we constructed a series of C-terminal deletion mutants of NS3,4 (Fig. 3). The polypeptides containing NS3 and NS4 were cleaved into two products (Fig. 3, lanes 6 and 8). Based on the molecular masses of the processed products, the NS3/4 junction appeared to be cleaved in these polypeptides. Only one processed product with an apparent molecular mass corresponding to NS3 was detected from polypeptide 1016-1776 (Fig. 3, lane 4). This seems to be due to rapid degradation of the NS4 product by cellular protease(s). Cu²⁺ inhibited the NS3-dependent cleavage of the HCV polypeptides tested (see lanes 3, 5 and 7 in Fig. 3). Experiments with the C-
Fig. 3. Effect of C-terminal deletions of NS2,3,4 on self-cleavage. (a) Schematic diagram of the 3' deletion mutants of pTHE1016–1846. Symbols and letters are as in the legend to Fig. 1. (b) Self-cleavage activities of the C-terminal deletion mutants of polypeptide 1016–1846 were tested by in vitro translation of mRNAs THE1016–1647, 1016–1776, 1016–1823 and 1016–1846. The number of the C-terminal amino acid of the polypeptide produced in each translation reaction is given above each lane. In vitro translations were carried out at 30 °C for 90 min in the absence (−) or presence (+) of CuCl2. Cu²⁺ was added to a final concentration of 100 μM. Samples were analysed by SDS–15% PAGE. The positions of protein molecular mass markers are indicated in kDa.

Terminal deletion mutants conclusively showed that the major cleavage site of NS3,4 polypeptides was the NS3/4 junction.

The minimal HCV protease domain was determined by testing proteolytic processing of polypeptides with internal deletions in NS3,4. Various regions of the C-terminal part of NS3 were removed in these mutants, but the N-terminal part of NS3, the NS3/4 junction and NS4 was maintained in all of the constructs. The internal deletion mutant A1219–1634 showed a Cu²⁺-sensitive band of 32 kDa (Fig. 4, lane 6). A polypeptide of this size could be generated by NS4A/4B junction cleavage. It is not known how the NS4A/4B junction is cleaved by NS3 in this particular polypeptide. A further deletion in NS3, between amino acids 1116–1218, abolished proteolytic activity (lane 3, Fig. 4). These results are consistent with the predictions made from comparisons of homologous proteases, since box 3 and box 4 of the trypsin-like protease consensus sequence were included in the deleted region of polypeptide Δ1116–1634. In conclusion, the essential elements of the NS3 protease reside between amino acids 1050–1218.

The N- and C-terminal boundaries of the NS3 protease and the inhibitory effect of Cu²⁺ on the protease activity
Fig. 4. Determination of the C-terminal boundary of NS3 protease. The C-terminal boundary of NS3 protease was determined approximately by analysing the proteolytic activity of the internal deletion mutants of polypeptide 1016–1846. (a) Schematic diagram of the internal deletion mutants of pTHE1016–1846. Symbols and letters are as in the legend to Fig. 1. (b) Self-cleavage activities of the internal deletion mutants of polypeptide 1016–1846 were tested by *in vitro* translation of mRNAs pTHEΔ1116–1846 and Δ1219–1634. The numbers of the deleted amino acids of the polypeptides are given above each lane. *In vitro* translation reactions were carried out at 30 °C for 90 min in the absence (−) or presence (+) of CuCl₂ and ZnCl₂. Cu²⁺ or Zn²⁺ were added to final concentrations of 100 μM. Samples were analysed by SDS–15% PAGE. The positions of protein molecular mass markers are indicated.

were further confirmed by testing the trans-cleavage activity of NS3 at the NS4A/4B junction. [³⁵S]Methionine-labelled substrates synthesized *in vitro* were mixed with polypeptide Δ1219–1634 separately synthesized *in vitro* to test the trans-cleavage activity of the polypeptide Δ1219–1634. The reaction mixtures were incubated at 30 °C for 120 min for proteolytic processing in the presence of cycloheximide to prevent further translation. Incubation of the enzyme and the substrates together resulted in cleavage of substrates into two fragments which could be generated by cleavage at the NS4A/4B junction (Fig. 5, lanes 3, 6 and 9). Interestingly, only the NS4A/4B junction, not the NS3/4A junction, was cleavable in trans-cleavage reactions. This is consistent with the latest report of Bartenschlager et al. (1994) suggesting that the NS3/4A junction is cleaved only in cis. The same patterns of substrate cleavage were also detected in reactions with polypeptide 1016–1846 as an enzyme source (data not shown). This strongly demonstrates that the N-terminal part (about 20 kDa) of HCV NS3 is sufficient for protease activity. Addition of Cu²⁺ also completely blocked proteolytic processing of the substrates in trans-cleavage reactions (Fig. 5, lanes 4, 7 and 10). This clearly demonstrates that protease inhibition by Cu²⁺ is directly related to the proteolytic activity of NS3. A polypeptide spanning amino acids 1016–1218 (without the NS4 region) showed very little trans-cleavage activity, if any, to substrates used in Fig. 5 (data not shown). This suggests that NS4 assists NS3 in trans-cleavage reactions (B. Hahm & S. Jang, unpublished). The helper function of NS4 on NS3 activity was also shown by Failla et al. (1994) and Bartenschlager et al. (1994).

Proteolytic processing of NS proteins of viruses in the family Flaviviridae, which is mostly performed by virus-encoded proteases, is essential for proliferation of the
viruses (Hollinger, 1990). Therefore, development of specific agents which block the proteolytic processing is one of the logical therapeutic approaches to disease caused by the viruses. Since NS3 of HCV is the protease responsible for cleavage of the NS3/4A, 4A/4B, 4B/5A and 5A/5B junctions of HCV polyprotein (Bartenschlager et al., 1993; Grakoui et al., 1993a, b; Tomei et al., 1993), it is worth characterizing the protease in detail and investigating its substrate selection mechanism.

The role of Zn\(^{2+}\) in processing of the NS3/4 junction cannot be elucidated by the present experiments. Zn\(^{2+}\) seems to play an indirect role in the processing, possibly by assisting in producing the correct conformation of the enzyme or the substrate, since the stimulatory effect of Zn\(^{2+}\) was either subtle or not detected in many internal

---

**Fig. 5.** Trans-cleavage of the NS4A/4B junction by THEA1219–1634 in vitro. (a) Schematic diagram of polypeptides used in trans-cleavage reactions. Symbols and letters are as in the legend to Fig. 1. (b) mRNA THEA1219–1634 was translated in RRL to produce unlabelled protease. \(^{35}\)S-Methionine-labelled substrates for protease assays were synthesized in RRL by using mRNAs THE1467–1846, 1404–1846 and 1218–1846. Unlabelled protease was mixed with different substrates in the presence of cycloheximide and then incubated at 30 °C for a further 120 min. Protein samples were separated by SDS–12% PAGE. The positions of protein molecular mass markers are indicated.
deletion mutants. The proteolytic activity of all the self-cleaving polypeptides, on the other hand, is inhibited by Cu²⁺. When Zn²⁺ and Cu²⁺ were tested together, only the Cu²⁺ effect was apparent (data not shown). Therefore, Cu²⁺ seems to act near or on the active site of the HCV protease. This suggests that the effective sites of the two ions are different. We do not yet understand how Cu²⁺ inhibits NS3 protease. Simple oxidation of the active site by Cu²⁺ is not a plausible inhibition mechanism, since addition of β-mercaptoethanol to translation reactions did not overcome the inhibitory effect of Cu²⁺ (data not shown). Inhibition of a protease by Cu²⁺ was also shown for the protease of the human immunodeficiency virus (HIV) (Karlström & Levine, 1991). These authors showed that the presence of cysteine residue(s) in the HIV protease is necessary for inhibition by Cu²⁺. The effect of thiol groups in the HCV protease on Cu²⁺ action remains to be elucidated. Investigations into the mechanism of protease inhibition by Cu²⁺ may lead to the development of anti-HCV drugs.

We are indebted to Akio Nomoto and Michinori Kohara for the supply of HCV cDNA clones and for valuable suggestions. We thank Chi-Bom Chae, Pann G. Sub and Kyoko Tsukiyama-Kohara for critical reading of the manuscript, and Seung K. Yoon, Chul-Joong Kim and Ok K. Song for very useful suggestions and discussions. This research was supported in part by a grant from genetic engineering research program of the Education Ministry and by a grant from KOSEF through SRC for cell differentiation.

References


Miller, R. H. & Purcell, R. H. (1990). Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. Proceedings of the National Academy of Sciences, USA 87, 2057-2061.


(Received 23 May 1994; Accepted 22 November 1994)