Short communication

A central hydrophobic domain of the hepatitis C virus NS4A protein is necessary and sufficient for the activation of the NS3 protease

Liccia Tomei,* Cristina Failla, Rosa Letizia Vitale, Elisabetta Bianchi and Raffaele De Francesco

Istituto di Ricerche di Biologia Molecolare 'P. Angeletti' – 00040 Pomezia (Roma), Italy

The processing at the NS3/4A, NS4A/4B, NS4B/5A and NS5A/5B junctions in the non-structural region of the hepatitis C virus (HCV) polyprotein is performed by a viral serine protease activity contained within the N-terminal 180 amino acids of the NS3 protein. Full protease activity is only achieved upon the interaction of a region at the N terminus of NS3 with the NS4A protein, this region is also involved in the modulation of the protease activity. Using the rabbit reticulocyte expression system, we have defined the minimal domain of NS4A that is necessary to increase the cleavage efficiency of NS3. A synthetic peptide containing the same region, NS4A amino acids 21 to 32, stimulates the proteolytic activity of NS3 at all the trans-cleavage sites.

Hepatitis C virus (HCV) is the major causative agent of transfusion-associated and sporadic non-A, non-B (NANB) hepatitis (Choo et al., 1989). It is an enveloped virus with a (+)-stranded linear RNA genome of about 9.4 kilobases (see Matsuura & Miyamura, 1993 for a review). The viral genome consists of a single open reading frame (ORF) encoding a polyprotein of 3009-3010 amino acids that is processed co- or post-translationally by both cellular and viral proteases into at least nine mature viral polypeptides (Hijikata et al., 1993). The gene order of HCV has been determined to be 5' C-E1-E2-NS2-NS3-NS4A-NS4B-NS5A-NS5B 3'. The putative structural proteins core (C) and the envelope glycoproteins E1 and E2 are released from the N-terminal part of the HCV polyprotein by cellular peptidases (Hijikata et al., 1991), whereas a virus-encoded metalloproteinase is responsible for the cleavage between NS2 and NS3 (Grakoui et al., 1993a; Tomei et al., 1993). The C-terminal remainder of the HCV polyprotein is further processed to give rise to NS3 (68 kDa), NS4A (6 kDa), NS4B (26 kDa), NS5A (56 kDa) and NS5B (65 kDa) proteins by a virus-encoded serine protease, contained in the amino-terminal third of NS3 (Bartenschlager et al., 1993; Eckart et al., 1993; Grakoui et al., 1993c; Hijikata et al., 1993; Manabe et al., 1994; Tomei et al., 1993). In addition to the NS3 serine protease, the NS4A protein is required for cleavage at the NS3/NS4A and NS4B/NS5A sites and it increases the efficiency of cleavage at the NS5A/NS5B and NS4A/NS4B junctions (Bartenschlager et al., 1994; Failla et al., 1994; Lin et al., 1994; Tanji et al., 1995). A stable physical interaction between HCV NS3 and NS4A has been demonstrated, suggesting that the two proteins may associate and that the active form of the protease is a heterodimer (Failla et al., 1995; Lin et al., 1995). A region at the extreme N terminus of NS3 is necessary both for efficient complex formation and for modulation of the proteolytic activity by NS4A, but not for the NS4A-independent serine protease activity of NS3 (Failla et al., 1994; Satoh et al., 1995).

In order to define the regions of NS4A that are implicated in its function as a serine protease cofactor, we decided to study the activity of the HCV NS3/NS4A protease and of a series of N- and C-terminal deletions of the NS4A sequence in a cell-free translation system. Appropriate DNA fragments derived from HCV-BK cDNA (Takamizawa et al., 1991) were inserted under the control of the T7 promoter into the pCite-1 vector (Novagen), using standard recombinant DNA technology. The constructs used in the present study are described in Fig. 1.

Recombinant plasmids were linearized with restriction enzymes and transcribed by using T7 RNA polymerase (Stratagene) as suggested by the manufacturer. The mRNAs were translated using rabbit reticulocyte lysate (RRL; Promega) in the presence of 30 µCi of [35S]methionine (ProMix; Amersham) or 1 µM-methionine. The translation reactions were stopped by the addition of 2 mM-CaCl₂ to digest the mRNA by reactivating the micrococcal nuclease contained in the reticulocyte lysate. After incubation at room temperature for 10 min, equal amounts of lysate containing unlabelled NS3-containing proteins and that containing 35S-labelled
Fig. 1. Schematic representation of the recombinant expression plasmids used. Construct pCiteNS3 (amino acids 1007–1615) has been described by Tomei et al. (1993). Constructs pCiteNS4A (amino acids 1649–1711) and pCiteNS3/4A (amino acids 991–1711) have been described by Failla et al. (1994). pCiteNS3/4AΔcut derives from pCiteNS3/4A but carries a substitution (solid box) of the sequence D1658-L-E-V-V-T1657 by S-M-A-R-I. The mutants pCiteNS3/4AAC32, AC21 and AC14 are all derivatives of the pCiteNS3/4A construct but have C-terminal deletions up to amino acid residues 1690, 1680 and 1672, respectively. In the construct pCiteNS3/4AΔintN21 the sequence between amino acids 1646 and 1678 is substituted by Ala-Thr. pCiteNS3/4AΔintN21AC32 derives from pCiteNS3/4AΔintN21 but a stop codon is present following residue 1690. pCiteNS3AH/4A carries a deletion of the putative RNA-helicase domain between residues 1237 and 1635. The plasmid pCiteNS3AH/4AΔC32 carries the same deletion in the NS3 region but the NS4A sequence ends at residue 1690. The annotation of the corresponding plasmids is shown on the left. A thinner line represents a deletion.

The reaction was stopped with SDS loading buffer and directly analysed by SDS–PAGE. The constructs encoding NS3, NS4A or their derivatives were also routinely translated in the presence of [35S]methionine to check the stability of the various proteins (Fig. 2a, lanes 1–3 and Fig. 2b, lanes 1–7).

We first tested the trans-cleavage activity of the HCV-encoded serine protease by incubating NS3 (Fig. 2a, lane 1) with the NS5A/5BAC51 substrate. As shown in Fig. 2(a) (lane 5), the radiolabelled NS5A/5BAC51 substrate was not processed to detectable levels by NS3 alone. Moreover, processing could not be restored by co-translating NS4A with NS3 prior to the incubation of the protease with the substrate (Fig. 2a, lane 6). Conversely, when the NS3 protease was synthesized as an NS3/NS4A precursor, both cis-cleavage at the NS3/4A site and efficient trans-cleavage of the NS5A/5B substrate were observed (Fig. 2a, lanes 3 and 7).

Since the in vitro activity of NS3 on the substrate containing the NS5A/5B cleavage site appeared to be strictly dependent on the presence of NS4A in the protease precursor, we decided to define the regions of NS4A responsible for its cofactor activity on the NS3 protease by constructing a series of C- and N-terminal deletion mutants of the NS4A sequence in the context of the NS3/4A polyprotein. The mutant constructs (Fig. 1) were expressed in RRL (Fig. 2b, lanes 1–7) and challenged with the labelled NS5A/5BAC51 substrate protein.

Among the NS3/4AAC constructs, in the case of NS3/4AAC32 mutant a protein product co-migrating with mature NS3 was detectable (Fig. 2b, lane 3), suggesting that the NS3/4AAC32 precursor was still capable of cis-cleavage at the NS3-NS4A junction. As to the remaining NS3/4AAC mutants, we could not unequivocally determine the occurrence of cis-cleavage because of the small difference in molecular mass between the precursors and the mature NS3 protein (Fig. 2b, lanes 4 and 5). The use of lower percentage acrylamide gels did not facilitate the interpretation of the results (not shown). In the case of the NS3/4AAC21, cis-cleavage could not occur (Fig. 2b, lanes 6 and 7) because the sequence of the NS3–NS4A junction was mutated. In order to rule out the possibility that any mutation that
affected the cis-cleavage at the NS3/4A site would also alter the processing efficiency at the trans-cleavage sites, we decided to investigate first whether an uncleaved NS3/4A polyprotein still retained the protease trans-cleavage activity. As shown in Fig. 2(b), lane 2, the mutant protein NS3/4AΔcut, in which the sequence in the cis-cleavage site has been altered, migrated with a molecular weight of 76 kDa corresponding to that of the uncleaved precursor protein. However, the NS3/4AΔcut protein still retained trans-protease activity, since mature
NS5A protein was cleaved off the 35S-labelled NS5A/5BAC51 precursor molecule (Fig. 2b, lane 10).

C-terminal deletions of the NS4A sequence up to residue 32 (mutant ΔC32) did not affect the ability of the NS3 protease to process the NS5A/5B cleavage sites (Fig. 2b, lane 11). Conversely, further deletions toward the N terminus of NS4A (mutants ΔC21 and ΔC14) appeared to interfere with the trans-protease activity (Fig. 2b, lanes 12 and 13).

We have previously reported that the C-terminal 33 residues, from amino acid 22 to 55, of the NS4A protein were sufficient to complement the NS3 activity (Failla et al., 1994). We now asked whether the N-terminal 21 residues of NS4A were dispensable even if the remaining portion of the molecule is directly fused to NS3. Therefore, we constructed the mutant NS3/4AintN21 and tested its ability to process the NS5A/5B site. Fig. 2(b) shows that this mutant is stably produced in the RRL (lane 6) and that it is still as active as the wild-type NS3/4A protease, since the same amount of mature NS5A protein was produced from the NS5A/5BAC51 precursor (lane 14).

The present series of C- and N-terminal deletion mutants allowed us to define the region of the NS4A protein which contains the protease cofactor activity as being between amino acids 21 and 32. This sequence is in fact sufficient to drive the NS4A activity, since a double NS3/4A mutant which bears both a C-terminal deletion up to residue 32 and an N-terminal deletion up to residue 21 still retained the ability to process the NS5A/5B cleavage site (mutant ΔIntN21ΔC32; Fig. 2b, lanes 7 and 15).

Previous studies have shown that NS3 is a tripartite protein. The minimal protease domain lies within the N-terminal 181 residues (Failla et al., 1995; Tanji et al., 1994), whereas the remainder has been shown to contain an RNA-dependent ATPase (Suzich et al., 1993). The N-terminal protease domain is in turn composed of two subdomains: a short sequence at the N terminus (amino acids 1027–1049) which is necessary for the interaction with NS4A and a downstream region that contains the NS4A-independent serine protease activity (Failla et al., 1995; Satoh et al., 1995).

We wanted to explore whether it was possible to express a minimal NS3/4A protease by removing the RNA-dependent ATPase domain from the NS3/4A precursor. To this end, we constructed the plasmid pCiteNS3ΔAH/4A in which the N-terminal 209 residues of NS3 were directly fused to the full-length NS4A sequence and the plasmid pCiteNS3ΔAH/4AAC32 that bears a C-terminal deletion of NS4A up to residue 32 (Fig. 1). We decided to leave intact the N terminus of NS4A up to residue 21, even though this is dispensable for the NS4A activity (see above), hoping that it would function as a linker sequence allowing the intramolecular interaction between the active domains of NS4A and NS3. For the same reason we chose to fuse NS4A to an NS3 fragment slightly longer than the minimal protease domain. The two mutant proteins were expressed in RRL and then incubated with the NS5A/5BAC51 [35S]methionine-labelled substrate to assess their trans-cleavage activity. The results, reported in Fig. 2(b) lanes 2 and 3 show that both the proteins were able to process the NS5A/5B site since mature NS5A protein was produced. The activation of the trans-cleavage efficiency of the NS3 protease by NS4A could be obtained either if a productive interaction is established between the NS3 and NS4A moieties of the same molecule (intramolecular interaction), or if the NS4A protein of one molecule is acting in trans on the NS3 protease domain of a second NS3/4A mutant. We excluded the latter possibility, since co-translation of wild-type NS3 with a NS3AH/4A protein carrying a mutation of the catalytic serine did not result in the processing of the NS5A/5B precursor (data not shown). These data suggest that NS4A can activate the serine protease domain of NS3 when fused directly to its C terminus via an intramolecular interaction between NS4A and the N-terminal portion of NS3.

In the cell-free translation system used, efficient cleavage by NS3 was obtained only if the NS4A cofactor was expressed in cis, i.e. as a NS3/4A precursor protein. These results differ from our previous observations while expressing the HCV proteins in a cell transfection system, where NS4A was shown to restore proteolytic cleavage by NS3 when supplied in trans (Bartenschlager et al., 1994; Failla et al., 1994; Lin et al., 1994). The apparent discrepancy may be due to diminished stability, incorrect folding, or an exceedingly low concentration of NS4A in the RRL. Addition of microsomal membranes, detergents or co-translation of NS3 and NS4A mRNAs together with the substrate molecule were all inefficient in allowing effective NS3–NS4A interaction (not shown).

In order to test whether the NS3 protease synthesized in the RRL was intrinsically capable of interacting with NS4A offered in trans, we synthesized a 14 residue peptide spanning amino acids 21 to 34 of NS4A (pep4A; Fig. 3a).

We preincubated the in vitro transcribed unlabelled NS3 protease domain [NS3(1007–1236)] with 30 μM peptide before adding the labelled substrate proteins containing the trans-cleavage sites NS4A/4B, NS4B/5A and NS5A/5B, respectively. None of the substrates were processed by NS3 alone (Fig. 3b, lanes 2–4), but preincubation with pep4A made the NS3 protease able to cut at all the trans-cleavage sites, as judged from the appearance of mature products (Fig. 3b, lanes 5–7).

Secondary structure prediction algorithms predicted the region of the NS4A protein encompassed by the
pep4A synthetic peptide to be folded in the \( \beta \)-strand conformation (Fig. 3a). In order to address this issue, we studied the conformation of pep4A by circular dichroism spectroscopy and observed that, while the peptide was primarily found in the random-coil conformation at concentrations below 0.1 mM, it acquired a substantial degree of \( \beta \)-sheet secondary structure when the concentration was raised above 0.5 mM (Fig. 3a). These findings indicate that the pep4A synthetic peptide is unfolded at a concentration sufficient to activate NS3. However, at higher concentrations it reveals a high propensity to fold in \( \beta \)-strands, most likely by forming intermolecular sheets.

The data reported in this paper demonstrate that a central, hydrophobic domain of NS4A (amino acid 21 to 32) is implicated in its function as a serine protease cofactor. A very recent study, expressing deletion and site-directed NS4A mutants in a cell-based transient expression system, indicated the same region as being critical for the serine protease cofactor activity of NS4A (Lin et al., 1995). Furthermore, we show here that a synthetic peptide spanning residues 21 to 34 of NS4A is sufficient to stimulate the NS3 serine proteolytic cleavage of the NS4A/4AB, NS4B/5A and NS5A/5B junctions in a cell-free trans-cleavage assay.

The mechanism by which HCV NS4A activates the NS3 serine protease is as yet unknown. Many possibilities have been suggested from targeting of the NS3 protease to the membranes (Hijikata et al., 1993; Tanji et al., 1995) to the NS4A action as a chaperonin aiding the correct folding of the protease domain. A third, more likely, possibility is that NS4A represents a true protease subunit, stabilizing NS3 in an active conformation. Given the observation that a synthetic peptide encompassing the cofactor domain of NS4A showed a high propensity to form \( \beta \)-strands, it is tempting to speculate that NS4A might be part of a heterodimeric protease by forming an intermolecular \( \beta \)-sheet with NS3. Efforts are still needed to answer this question.

Lin et al. (1995) have also recently reported that a similar NS4A-derived synthetic peptide, based on the sequence of a different HCV genotype, can promote the proteolytic cleavage of the NS4B/5A precursor by the NS3 protease produced in a cell-free in vitro translation system. The differences in sequence between the NS4A-derived peptide we describe and the one used by Lin et al. (1995) are only conservative ones: serine 22 and arginine 34 are replaced by cysteine and lysine, respectively, in the HCV-H genotype. Furthermore, amino acids at positions 22 and 34 of NS4A were shown, by Lin et al. (1995) not to affect that cofactor activity of NS4A in cell transfection experiments. Given the high sequence conservation of the NS3 and NS4A proteins in all the HCV genotypes thus far characterized (see Lin et al., 1994 for alignment), it is likely that the modulation of the NS3 protease activity by the central, hydrophobic region of NS4A is a feature common to all the HCV variants. The

![Graph](image_url)
interaction between NS3 and NS4A is thus likely to be essential for viral replication and thus constitutes an attractive alternative target for antiviral therapy. The cell-free protease assay reported in the present paper represents the basis for developing an *in vitro* assay useful for the testing and development of compounds as competitive inhibitors of the NS3/4A protease activity, as well as compounds that could lead to the disruption of the interaction between NS3 and NS4A.

References


(Received 25 September 1995; Accepted 12 December 1995)