

## Purification and characterization of the NS3 serine protease domain of hepatitis C virus expressed in *Saccharomyces cerevisiae*

William Markland, Richard A. Petrillo, Matthew Fitzgibbon, Ted Fox, Robert McCarrick, Tom McQuaid, John R. Fulghum, Wenyong Chen, Mark A. Fleming, John A. Thomson and Stephen P. Chambers

Vertex Pharmaceuticals Inc., 130 Waverly Street, Cambridge, MA 02139-4242, USA

**cDNA encoding the putative core of the hepatitis C virus NS3 serine protease domain (residues 1–181 of NS3; NS3<sub>181</sub>) was expressed as an N-terminally (His)<sub>6</sub>-tagged fusion protein in *Saccharomyces cerevisiae*. NS3<sub>181</sub> protease activity was found in soluble cell lysates, and the N-terminal metal-chelating domain facilitated the efficient purification of active enzyme, using immobilized metal affinity chromatography. The purified NS3<sub>181</sub> protease activity was characterized by assaying the *trans*-cleavage of *in vitro* transcription–translation generated substrates, and subsequently a previously unobserved cleavage site within the NS5A region was identified. The inhibitory effect of known protease inhibitors was also examined. It is hoped that availability of this method for the expression and purification of the NS3<sub>181</sub> protease will facilitate the development of anti-hepatitis C therapies.**

Hepatitis C virus (HCV), a member of the family *Flaviviridae*, is a major aetiological agent of post-transfusion non-A, non-B hepatitis (Choo *et al.*, 1989). Its single-stranded, positive-sense RNA genome encodes a polyprotein of about 3000 residues, NH<sub>2</sub>-C-E1-E2-NS2-NS3-NS4A-NS4B-NS5A-NS5B-COOH, where C, E and NS are the core, envelope and nonstructural regions, respectively. The nonstructural proteins constitute the critical enzymes that support virus replication. These proteins include RNA replicase components encoding proteases (NS2, NS3), a helicase (NS3) and an RNA-dependent RNA polymerase (NS5B). The structural proteins appear to be cleaved by the action of a host cell signal peptidase, located in the endoplasmic reticulum lumen (Hijikata *et al.*, 1991). Processing of the nonstructural polyprotein requires two virus-encoded proteases: NS2-3, thought to be a metallo-serine protease which appears to cleave between the NS2/3 site

(Grakoui *et al.*, 1993a), and NS3, a serine protease which cleaves the remaining sites downstream of itself. Analysis of polyprotein processing in transient expression systems, using radioactive pulse-chase methods, suggests that the *cis*-cleavage of the NS3/4A junction occurs first, followed by the *trans*-cleavage of the NS4A/4B, NS4B/5A and NS5A/5B sites downstream (Hijikata *et al.*, 1991; Grakoui *et al.*, 1993a, b; Bartenschlager *et al.*, 1994).

Early studies showed a high degree of sequence variation between different HCV isolates (Choo *et al.*, 1991). However, there was sufficient homology among HCV strains, and with the serine proteases of the related flaviviruses and pestiviruses, to enable the identification of a region within NS3 as the putative protease domain (Miller & Purcell, 1990). The NS3 protease domain has since been delineated to approximately 181 residues at the N-terminal third of NS3 (around residues 1027–1207 of the HCV polyprotein), within which are found the conserved residues His-1083, Asp-1107 and Ser-1165, characteristic of the catalytic triad found in trypsin-like serine proteases (Takamizawa *et al.*, 1991). The remaining C-terminal portion of NS3 contains motifs characteristic of a helicase/nucleoside-triphosphate-binding domain (Suzich *et al.*, 1993).

In the absence of an efficient *in vitro* cultivation system for HCV, the proteolytic processing of the NS3 serine proteases has been studied using *in vitro* transcription-translation (IVTT) assay systems (Lin & Rice, 1995), transient protein expression in cultured mammalian cells (Grakoui *et al.*, 1993b) and by following the intracellular processing of substrates in recombinant *Escherichia coli* (Komoda *et al.*, 1994) and insect cell systems (Hirowatari *et al.*, 1995). Some of the most recent studies indicate that the NS3 protease domain does not act independently in cleaving the polyprotein. In fact, the NS3 serine protease activity appears to be regulated by NS4A. The mode of regulation by NS4A is still poorly defined, but evidence suggests that the interaction of NS3 with NS4A may be important for substrate recognition *in vivo* (Bartenschlager *et al.*, 1995). Moreover, small synthetic peptides, spanning the central region of NS4A, have been used to partially mimic this regulatory effect *in vitro* (Lin *et al.*, 1995).

**Author for correspondence:** Stephen P. Chambers.  
Fax +1 617 577 6680. e-mail Chambers@vpharm.com

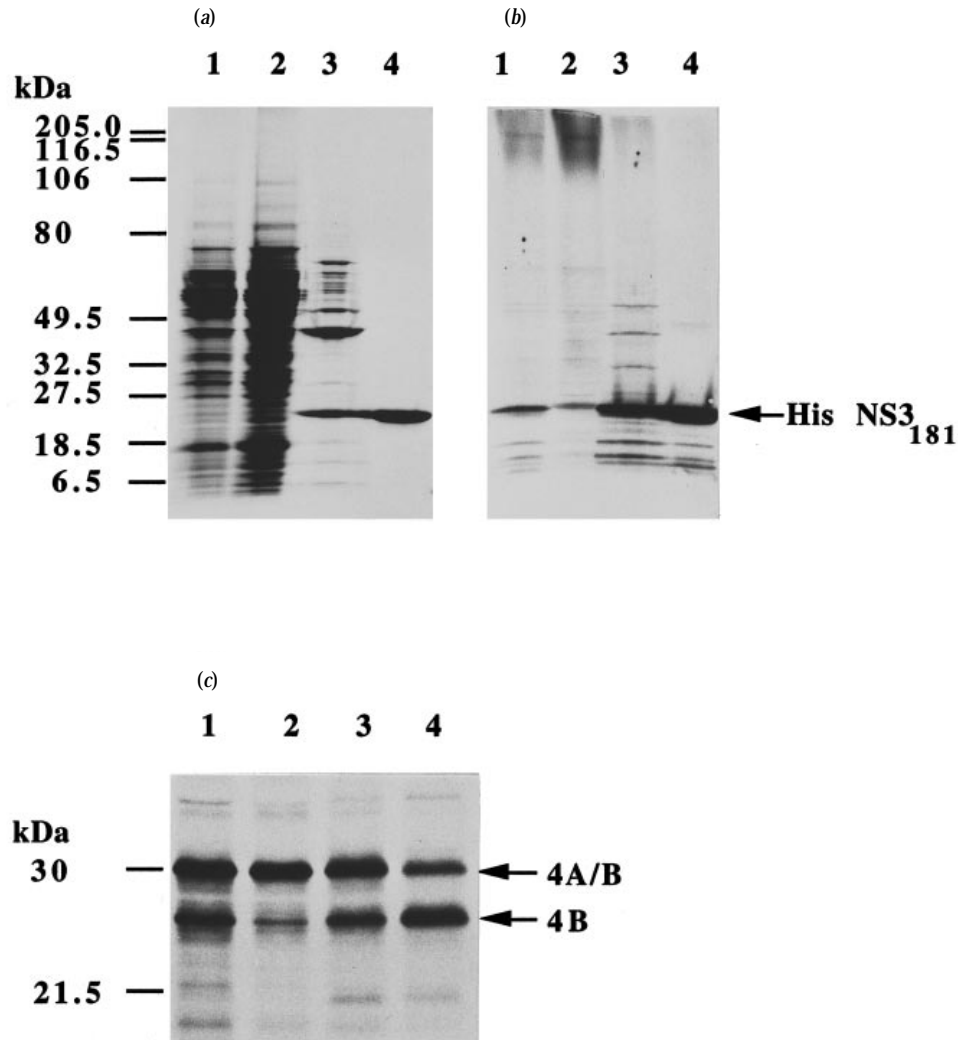


Fig. 1. Purification of His-NS3<sub>181</sub> and analysis of purified fractions. Various purified fractions were analysed by (a) SDS-PAGE and stained with Coomassie blue, and (b) Western blotting with antisera specific for the NS3 protease domain. The activity of the various purified fractions was also examined for *trans*-cleavage activity against the NS4A/B(1658–1972) substrate (c). Lane 1, soluble fraction after cell lysis; lane 2, flow-through from nickel affinity column; lane 3, eluate with 100 mM imidazole; lane 4, eluate with 300 mM NaCl from SP Sepharose column.

All existing evidence suggests that the NS3 serine protease has a critical role in the processing of mature HCV polypeptides and in the overall process of virus replication. This makes it a promising target for anti-viral intervention. The production and purification of large quantities of this enzyme are prerequisites for structural and mechanical studies of the enzyme, and for the development of assays that would lead to potent and selective inhibitors. Expression and purification of recombinant versions of the full-length NS3 have been reported in *E. coli* (D'Souza *et al.*, 1995; Kakiuchi *et al.*, 1995). More recently, a 20 kDa protein encompassing the HCV domain has been purified from recombinant-baculovirus-infected insect cells (Steinkuhler *et al.*, 1996). Here we report a method for the expression and purification of the active recombinant hexahistidine-linked NS3<sub>181</sub> protease domain (His-NS3<sub>181</sub>) in *S. cerevisiae*. We also examined the effect of a

synthetic NS4A peptide on the substrate specificity of the purified protein, and the effects of various commercial inhibitors on its proteolytic activity.

Using standard recombinant DNA techniques (Sambrook *et al.*, 1989), a 619 bp *Xba*I–*Bam*HI cDNA fragment containing the coding sequence for NS3<sub>181</sub> from the HCV subtype strain H (Feinstone *et al.*, 1981), together with an N-terminal His-tag, was isolated from pET8c/HCV/His-NS3<sub>181</sub> provided by C. M. Rice (Washington University School of Medicine, St Louis, Mo., USA). To facilitate its cloning into pYes2 (Invitrogen), additional restriction sites flanking the coding sequence were introduced via subcloning into the *Xba*I–*Bgl*II sites within the polylinker region of pMTL20 (Chambers *et al.*, 1989). This allowed its subsequent cloning as a *Bam*HI–*Xho*I fragment into the *Bam*HI–*Xho*I sites of pYes2. The resulting plasmid, designated pYes/HCV/His-NS3<sub>181</sub>, was transformed into

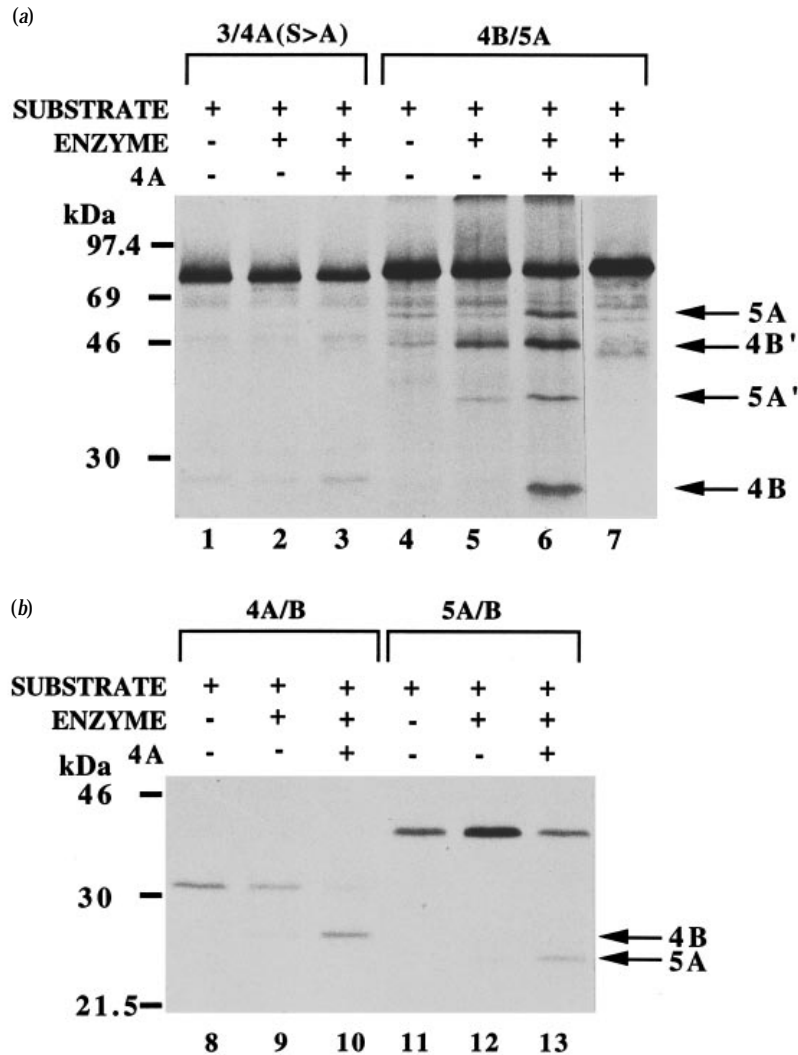


Fig. 2. Processing of cleavage sites by His-NS3<sub>181</sub> and the effect of a NS4A peptide on that activity. [<sup>35</sup>S]Methionine-labelled HCV substrates were prepared using a T7 polymerase IVTT rabbit reticulocyte lysate system, according to the manufacturer's specifications (Promega). The plasmids used to generate the substrates used in this study have previously been described by Lin *et al.* (1994). The labelled substrates NS3(S1165A)/4A(1027–1711) (lanes 1–3), NS4B/5A(1712–2420) (lanes 4–7), NS4A/4B(1658–1972) (lanes 8–10) and NS5A/5B(2285–2508) (lanes 11–13). An inactive His-NS3<sub>181</sub>(S1165A) control was used to confirm the alternative cleavage site within the NS4B/5A substrate was NS3 mediated (lane 7). The positions of the cleaved products are indicated on the right by arrows.

*S. cerevisiae* strain INVSc2 (*MAT $\alpha$* , *his3 $\Delta$ 1*, *leu2*, *trp1-289*, *ura3-52*).

*S. cerevisiae* strain INVSc2, when transformed with the expression vector pYes/HCV/His-NS3<sub>181</sub>, expressed the His-NS3<sub>181</sub> fusion protein upon induction with galactose. Procedures for the cultivation of *S. cerevisiae* and galactose-induced expression were as described by Schneider & Guarente (1991). Fig. 1 shows (a) the purification profile of the individual fractions when analysed on SDS-PAGE, (b) the immunoblot using antisera specific to the HCV NS3 protease and (c) the activity identified by cleavage of [<sup>35</sup>S]methionine-labelled NS4A/B (Lin & Rice, 1995).

Following expression, the cells were harvested by centrifugation at 5000 *g* for 15 min and resuspended in breaking buffer (50 mM Na<sub>2</sub>HPO<sub>4</sub> pH 8.0, 300 mM NaCl, 5% glycerol, 0.25% Tween 20 and 1 mM PMSF). Following addition of an equal vol. of acid-washed 0.45 mm glass beads, the cells were lysed with 10 × 30 s pulses of a vortex mixer, allowing a 30 s cooling period (on ice) in between each pulse. Cell lysate was

removed from the beads and a cleared lysate was obtained by centrifuging at 100 000 *g* for 60 min. The soluble material (Fig. 1, lane 1) was then loaded onto a nickel affinity column which had previously been equilibrated with breaking buffer. The flow-through was collected (Fig. 1, lane 2) and the column washed with 2 column vols of breaking buffer, followed by 2 column vols of wash buffer (50 mM Na<sub>2</sub>HPO<sub>4</sub>, 300 mM NaCl, 5% glycerol pH 6.0). Elution with a step-gradient of 100 mM–300 mM imidazole in breaking buffer eluted the purified protease within the 100 mM fraction (Fig. 1, lane 3). Eluted fractions containing the protease were pooled, diluted (1:1) in wash buffer without NaCl and loaded onto an SP Sepharose cation exchange column equilibrated with breaking buffer containing 150 mM NaCl. Highly purified protease was eluted in breaking buffer with a salt gradient at 300 mM NaCl (Fig. 1, lane 4). The yield was approximately 1–0.5 mg of purified protein from 1 litre of culture. The purity of the final product was estimated to be greater than 90%.

To study the substrate specificity of the purified NS3 serine

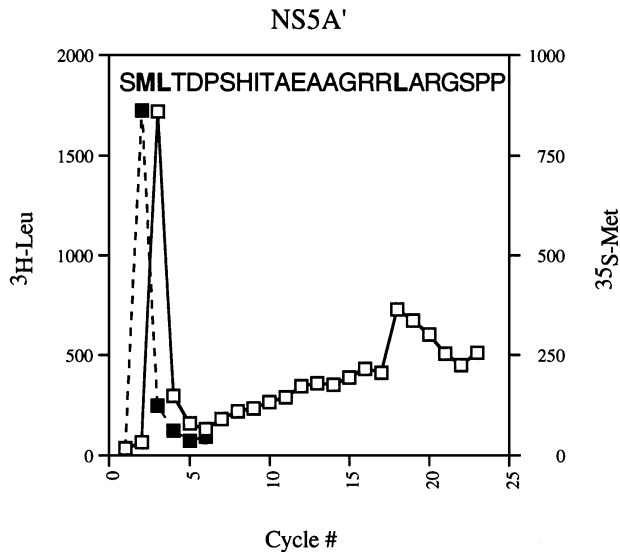


Fig. 3. N-terminal sequencing analysis of HCV NS5A' protein. The graph shows uncorrected c.p.m. released per sequencing cycle. This data identified the sequence for NS5A' beginning at HCV polyprotein residue 2173, as shown at the top of the graph. In this sequence, the amino acid residues shown in boldface type indicate the residues determined, by N-terminal sequence analysis, to be [<sup>3</sup>H]Leu- (□) or [<sup>35</sup>S]Met- (■) radiolabelled.

protease domain, the following substrates – NS3(S1165A)/4A(1027–1711), NS4B/5A(1712–2420), NS4A/B(1658–1972) and NS5A/B(2285–2508) – were examined with and without the addition of a synthetic NS4A peptide, H-Gly-Cys-Val-Val-Ile-Val-Gly-Arg-Ile-Val-Leu-Ser-Gly-Lys-OH (residues 1678–1691) (Fig. 2). The labelled substrates were mixed in a 1:1 (v/v) ratio with purified protease (50 pM) in the presence or absence of the NS4A peptide (15 μM) and incubated for 2 h at 30 °C. The cleavage reaction products were diluted 10-fold with SDS sample buffer, heated to 95 °C for 5 min, and analysed by SDS-PAGE. After electrophoresis, the gels were fixed in methanol-acetic acid, treated with EN<sup>3</sup>HANCE (Du Pont), dried and exposed to photographic film.

The NS3/4A substrate, itself rendered incapable of *cis*-cleavage by virtue of a Ser to Ala mutation at position 1165 (S1165A), produced no cleavage products when incubated with His-NS3<sub>181</sub> (Fig. 2, lanes 1–3). This is consistent with other studies that have proposed that the serine protease domain of the substrate interferes with the accessibility of the NS3/4A cleavage site, thereby inhibiting *trans*-cleavage by the added active protease (Lin *et al.*, 1994).

Incubation of the NS4B/5A substrate with His-NS3<sub>181</sub> alone produced two cleavage products (designated 4B' and 5A' in Fig. 2, lane 5), while NS4A-mediated cleavage resulted in the appearance of two additional cleavage products (designated 4B and 5A in Fig. 2, lane 6). N-terminal sequence analysis identified the cleavage sites observed. The NS4A-mediated cleavage was identified as the authentic NS4B/5A

junction (Grakoui *et al.*, 1993c), while the additional cleavage, which took place both in the presence and absence of the NS4A peptide, was found to be within NS5A, after Thr-2172 (Fig. 3). This latter site is highly conserved among other HCV strains and contains a consensus cleavage site recognition motif (Asp-X-X-X-X-Thr/Ser) similar to that found at the NS3/4A cleavage site. Interestingly, this previously unobserved cleavage site occurs upstream of a hyperphosphorylated region within NS5A, thought to be associated with the sub-cellular localization of NS5A products (Tanji *et al.*, 1995). No cleavage of the NS4B/5A substrate was observed when incubated with an inactive protease domain containing the S1165A mutation (Fig. 2, lane 7).

Cleavage of NS4A/4B and NS5A/5B substrates by His-NS3<sub>181</sub> with and without the addition of NS4A peptide, occurred only at the predicted sites (Fig. 2b). The NS5A/B substrate, a truncated version of the native sequence, lacks the additional non-NS4A mediated cleavage site mentioned above.

Inhibitors of all four classes of proteases (cysteine, serine, metallo and aspartic acid) were tested against His-NS3<sub>181</sub> protease *trans*-cleavage activity. [<sup>35</sup>S]Methionine-labelled NS4A/B was incubated at 30 °C for 2 h with purified protease (50 pM) in the presence of the NS4A peptide (15 μM) and a specific inhibitor. The metalloprotease inhibitors bestatin and EDTA exhibited no inhibitory activity. However, 1,10-phenanthroline was inhibitory at a concentration of 1 mM. This inhibition could possibly be due to Cu<sup>2+</sup> contamination, which had previously been observed by Kobashi & Horecker (1967) and was partially confirmed since Cu<sup>2+</sup> was found to be a potent inhibitor of His-NS3<sub>181</sub> protease activity at 1 mM. Zn<sup>2+</sup> also exhibited weak inhibitory activity, while Mn<sup>2+</sup>, Mg<sup>2+</sup>, Ca<sup>2+</sup> and K<sup>+</sup> showed no inhibition of protease activity at 1 mM.

Pepstatin, an aspartic acid protease inhibitor, had no effect on protease activity. Leupeptin and antipain are microbial aldehyde inhibitors of both serine and cysteine proteases and neither appeared to have any noticeable effect (up to 1 mM) on His-NS3<sub>181</sub> protease activity. Similarly, 1 mM E-64, a potent irreversible inhibitor of cysteine proteases in the papain superfamily, did not inactivate the protease. Investigation of nonspecific serine protease inhibitors showed that no loss of His-NS3<sub>181</sub> protease activity occurred in the presence of 1 mM benzamidin, 4-(2-aminoethyl)-benzenesulfonyl fluoride (AEBSF), PMSF or aprotinin, while chymostatin, a chymotrypsin-like protease inhibitor, inhibited His-NS3<sub>181</sub> protease at high concentrations (5 mM). 3,4-Dichloroisocoumarin (DCIC), a nonspecific irreversible serine protease inhibitor, was a more effective inhibitor of protease activity (0.1 mM). Similarly, elastinal, an elastase-like protease serine protease inhibitor, showed significant inhibition at 0.2 mM.

Highly reactive chloromethyl ketones such as *N*α-*p*-tosyl-L-lysine chloromethyl ketone (TLCK) and *N*-tosyl-L-phenylalanine chloromethyl ketone (TPCK) are potent inhibitors of trypsin and chymotrypsin, respectively. Both were found to

inhibit His-NS3<sub>181</sub> protease activity, with TPCK showing approximately 10-fold better inhibition under the reaction conditions in this study. The inhibitory profile is consistent with the characterization of His-NS3<sub>181</sub> protease as a serine protease, possibly with a chymotrypsin or elastase-like fold. However, the ineffectiveness of many serine protease inhibitors suggests that the geometry of the HCV active site may be significantly different from previously characterized serine proteases.

We would like to thank Charles Rice for providing NS3 antisera, T7 transcription plasmids used to generate the HCV substrates and pET8c/HCV/His-NS3<sub>181</sub>; Scott Harberson for synthesis of the NS4A peptide; and Brett O'Hare for oligonucleotide synthesis and DNA sequencing. We are also grateful to Drs Sato, Livingston, Landro, Morgenstern and Livi for their critical reading and discussion of the manuscript.

## References

- Bartenschlager, R., Ahlborn-Laake, L., Mous, J. & Jacobsen, H. (1994).** Kinetic and structural analyses of hepatitis C virus polyprotein processing. *Journal of Virology* **68**, 5045–5055.
- Bartenschlager, R., Lohmann, V., Wilkinson, T. & Koch, J. O. (1995).** Complex formation between the NS3 serine-type proteinase of the hepatitis C virus and NS4A and its importance for polyprotein maturation. *Journal of Virology* **69**, 7519–7528.
- Chambers, S. P., Prior, S. E., Barstow, D. A. & Minton, N. P. (1989).** The pMTL nic<sup>-</sup> cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. *Gene* **68**, 139–149.
- Choo, Q.-L., Kuo, G., Weiner, A. J., Overby, L. R., Bradely, D. W. & Houghton, M. (1989).** Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* **244**, 359–362.
- Choo, Q.-L., Richman, K. H., Han, J. H., Berger, K., Lee, C., Dong, C., Gallegos, C., Coit, D., Medina-Selby, A., Barr, P. J., Weiner, A. J., Bradley, D. W., Kuo, G. & Houghton, M. (1991).** Genetic organization and diversity of the hepatitis C virus. *Proceedings of the National Academy of Sciences, USA* **88**, 2451–2455.
- D'Souza, E. D. A., Grace, K., Sangar, D. V., Rowlands, D. J. & Clarke, B. E. (1995).** *In vitro* cleavage of hepatitis C virus polyprotein substrates by purified recombinant NS3 protease. *Journal of General Virology* **76**, 1729–1736.
- Feinstone, S. M., Alter, H. J., Dienes, H. P., Shimizu, Y., Popper, H., Blackmore, D., Sly, D., London, W. T. & Purcell, R. H. (1981).** Non-A, non-B hepatitis in chimpanzees and marmosets. *Journal of Infectious Diseases* **144**, 588–598.
- Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M. & Rice C. M. (1993a).** A second hepatitis C virus-encoded proteinase. *Proceedings of the National Academy of Sciences, USA* **90**, 10583–10587.
- Grakoui, A., Wychowski, C., Lin, C., Feinstone, S. M. & Rice C. M. (1993b).** Expression and identification of hepatitis C virus polyprotein cleavage products. *Journal of Virology* **67**, 1385–1395.
- Grakoui, A., McCourt, D. W., Wychowski, C., Feinstone, S. M. & Rice C. M. (1993c).** Characterization of the hepatitis C virus-encoded serine proteinase: determination of the proteinase-dependent polyprotein cleavage sites. *Journal of Virology* **67**, 2832–2843.
- Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M. & Shimotohno, K. (1991).** Gene mapping of the putative structural region of the hepatitis C virus genome by *in vitro* processing analysis. *Proceedings of the National Academy of Sciences, USA* **88**, 5547–5551.
- Hirowatari, Y., Hijikata, M., Tanji, Y. & Shimotohno, K. (1995).** Expression and processing of putative nonstructural proteins of hepatitis C virus in insect cells using baculovirus vector. *Virus Research* **35**, 43–61.
- Kakiuchi, N., Hijikata, M., Komoda, Y., Tanji, Y., Hirowatari, Y. & Shimotohno, K. (1995).** Bacterial expression and analysis of cleavage activity of HCV serine proteinase using recombinant and synthetic substrates. *Biochemical and Biophysical Research Communications* **210**, 1059–1065.
- Kobashi, K. & Horecker, B. L. (1967).** Reversible inactivation of muscle aldolase by o-phenanthroline. *Archives of Biochemistry and Biophysics* **121**, 178–186.
- Komoda, Y., Hijikata, M., Sato, S., Asabe, S.-I., Kimura, K. & Shimotohno, K. (1994).** Substrate requirements of hepatitis C virus serine proteinase for intermolecular polypeptide cleavage in *E. coli*. *Journal of Virology* **68**, 7351–7357.
- Lin, C. & Rice, C. M. (1995).** The hepatitis C virus NS3 serine protease and NS4A cofactor: establishment of a cell-free *trans*-processing assay. *Proceedings of the National Academy of Sciences, USA* **92**, 7622–7626.
- Lin, C., Pragai, B. M., Grakoui, A., Xu, J. & Rice, C. M. (1994).** Hepatitis C virus NS3 serine proteinase: *trans*-cleavage requirements and processing kinetics. *Journal of Virology* **68**, 8147–8157.
- Lin, C., Thomson, J. A. & Rice, C. M. (1995).** A central region in the hepatitis C virus NS4A protein allows formation of an active NS3–NS4A serine proteinase complex *in vivo* and *in vitro*. *Journal of Virology* **69**, 4373–4380.
- 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 groups. *Proceedings of the National Academy of Sciences, USA* **87**, 2057–2061.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989).** *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Schneider, J. C. & Guarente, L. (1991).** Vectors for expression of cloned genes in yeast; regulators, overproduction and underproduction. In *Methods in Enzymology*, pp 373–388. Edited by C. Guthrie & G. R. Fink. Academic Press.
- Steinkuhler, C., Tomei, L. & De Francesco, R. (1996).** *In vitro* activity of hepatitis C virus protease NS3 purified from recombinant baculovirus-infected Sf9 cells. *Journal of Biological Chemistry* **271**, 6367–6373.
- Suzich, J.-A., A., Tamura, J. K., Palmer-Hill, F., Warrenner, P., Grakoui, A., Rice, C. M., Feinstone, S. M. & Collett, M. S. (1993).** Hepatitis C virus NS3 protein polynucleotide-stimulated nucleoside triphosphate and comparison with the related pestivirus and flavivirus enzymes. *Journal of Virology* **67**, 6152–6158.
- Takamizawa, A., Mori, C., Fuke, I., Manabe, S., Murakami, S., Fujita, J., Onishi, E., Andoh, T., Yoshida, I. & Okayama, H. (1991).** Structure and organization of the hepatitis C virus genome isolated from human carriers. *Journal of Virology* **65**, 1105–1113.
- Tanji, Y., Kaneko, T., Satoh, S. & Shimotohno, K. (1995).** Phosphorylation of hepatitis C virus-encoded nonstructural protein NS5A. *Journal of Virology* **69**, 3980–3986.

Received 23 May 1996; Accepted 10 September 1996