RAPID COMMUNICATION Jpn. J. Cancer Res. 81, 1092–1094, November 1990

A Structural Protein Encoded by the 5′ Region of the Hepatitis C Virus Genome Efficiently Detects Viral Infection

Nobuyuki Kato, Makoto Hijikata, Yuko Ootsuyama, Masanori Nakagawa, Showgo Ohkoshi and Kunitada Shimotohno

Virology Division, National Cancer Center Research Institute, 5-1-1, Tsukiji, Chuo-ku, Tokyo 104

A structural region of the hepatitis C virus genome was molecularly cloned. A protein expressed *in vitro* by transcription followed by translation was precipitated immunologically by sera from patients with chronic hepatitis or hepatocellular carcinoma that were positive for antibody against the non-structural protein, C100, of hepatitis C virus, but not by sera from healthy persons. Thus, this structural protein should be useful for detection of infection with this virus.

Key words: Hepatitis C virus (HCV) — HCV expression — Anti-HCV

Since the development of a kit for diagnosis of hepatitis C virus (HCV) infection by Chiron Corporation,[1] a major cause of post-transfusion-associated hepatitis has been shown to be HCV infection. In fact more than 60% of patients with chronic hepatitis give a positive reaction for the antibody against a protein, C100, encoded by a non-structural region (NS3 to NS4 region) of the HCV genome.[1, 2] This antibody has also been found in the sera of patients with other hepatic diseases, suggesting the significance of this virus in not only chronic hepatitis but also liver cirrhosis and hepatocellular carcinoma.[1, 3–5] The high incidence (86%) of the antibody in patients with hepatocellular carcinoma that is not associated with hepatitis B virus (HBV) suggests a close association of infection with this virus and hepatic cancer. This possibility is supported by the frequent detection of HCV RNA in patients with hepatic cancer.[6, 7]

Using two techniques, detection of antibody against HCV and detection of HCV RNA we found a good correlation between the presence of the antibody and HCV RNA in patients with chronic hepatitis, although we obtained inconsistent results in some cases.[6] The inconsistencies could be explained by failure to detect anti-C100 because of a low immune response, a different amino acid sequence in the epitope of the C100 antigen in some Japanese isolates of HCV (HCV-J), failure to amplify the HCV gene by the polymerase chain reaction (PCR) due to a sequence mismatch between primer and viral template because of variation of the HCV sequence or the presence of too little viral RNA in the specimens for its detection by the PCR.

For examination of these possibilities, a more sensitive system is required for detecting viral antibodies than the conventional anti-C100 assay system. For this purpose, we molecularly cloned a gene that is supposed to encode the structural protein of HCV and demonstrated expression of the gene product *in vitro*. We found that antibody against this protein could be detected in many sera that were positive for antibody against C100, but not in sera from healthy persons.

A cDNA library from the sera of patients with chronic hepatitis was constructed using as a primer a nucleotide sequence that coded for a non-structural region of HCV-J. The library was screened for the virus sequence that contained a structural region (Kato *et al.*, manuscript in preparation). One of the resultant clones, pPS4S3, in which an insert of 564 nucleotides (nt) of the virus sequence was subcloned into the *Eco*RI site of the original vector, pTZ19R, was used in this study. The nucleotide sequence of the insert was determined by the dideoxy chain termination method and the results, together with the deduced amino acid sequence of the longest open reading frame, are shown in Fig. 1. An open reading frame starting from position −42 continues through the end of this insert. A possible initiator methionine is present at position 1. The hydropathicity profile of this region matches that of the core protein of other flaviviruses.[8] Therefore, we conclude that the insert of pPS4S3 encodes the core protein of HCV. It is noteworthy that the first 120 amino acids encoded by this open reading frame are rich in basic amino acids (arginine and lysine, 23%) and proline (13%), like those in the core proteins of other flaviviruses.[8] The sequence of the 5′ terminus of HCV has been reported by another group.[9] We carried out an *in vitro* transcription and translation experiment to determine whether this open reading frame expresses the virus protein. Since the longest open reading frame in the insert continues through the flanking sequence into the vector and therefore, consists of 214 amino acid residues, a protein of about 23 kd was expected to be expressed. RNA was synthesized *in vitro* by T7 RNA polymerase in the presence of components of

```
(GAATTCC) GTGTTGGGTCGCGAAAGGCCTTGTGGTACTGCCTGATAGGGTGCTTGCGAGTGCCCCG

  1                              Met Ser Thr Asn Pro Lys Pro Gln Arg Lys
-23  GGAGGTCTCGTAGACCGTGCATC ATG AGC ACA AAT CCT AAA CCT CAA AGA AAA

 11  Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp Val Lys Phe Pro Gly
 31  ACC AAA CGT AAC ACC AAC CGC CGC CCA CAG GAC GTT AAG TTC CCG GGC

 27  Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro
 79  GGT GGT CAG ATC GTT GGT GGA GTT TAC CTG TTG CCG CGC AGG GGC CCC

 43  Arg Leu Gly Val Arg Ala Thr Arg Lys Thr Ser Glu Arg Ser Gln Pro
127  AGG TTG GGT GTG CGC GCG ACT AGG AAG ACT TCC GAG CGG TCG CAA CCT

 59  Arg Gly Arg Arg Gln Pro Ile Pro Lys Ala Arg Arg Pro Glu Gly Arg
175  CGT GGA AGG CGA CAA CCT ATC CCC AAG GCT CGC CGG CCC GAG GGT AGG

 75  Thr Trp Ala Gln Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly
223  ACC TGG GCT CAG CCC GGG TAC CCT TGG CCC CTC TAT GGC AAC GAG GGT

 91  Met Gly Trp Ala Gly Trp Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser
271  ATG GGG TGG GCA GGA TGG CTC CTG TCA CCC CGT GGC TCT CGG CCT AGT

107  Trp Gly Pro Thr Asp Pro Arg Arg Arg Ser Arg Asn Leu Gly Lys Val
319  TGG GGC CCC ACA GAC CCC CGG CGT AGG TCG CGT AAT TTG GGT AAG GTC

123  Ile Asp Thr Leu Thr Cys Gly Phe Ala Asp Leu Met Gly Tyr Ile Pro
367  ATC GAT ACC CTT ACA TGC GGC TTC GCC GAC CTC ATG GGG TAC ATT CCG

139  Leu Val Gly Ala Pro Leu Gly Gly Ala Ala Arg Ala Leu Ala His Gly
415  CTT GTC GGC GCC CCC CTA GGG GGC GCT GCC AGG GCC CTG GCA CAT GGT

155  Val Arg Val Leu Glu Asp Gly Val Asn
463  GTC CGG GTT CTG GAG GAC GGC GTG AAC T (GGAATTC)
```

Fig. 1. Nucleotide sequence of the 5′region of the HCV-J genome. The sequence of the insert of pPS4S3 is shown. The putative amino acid sequence from the longest open reading frame is also shown.

RNA and magnesium salt. The reaction mixture, containing 40 m*M* Tris (pH 8.0), 15 m*M* $MgCl_2$, 5 m*M* dithiothreitol, 0.1 m*M* GTP, 0.5 m*M* each of ATP, CTP and UTP, 2 m*M* m^7GpppG, 0.5 m*M* bovine serum albumin, 1 μg of plasmid and 10 units of T7 RNA polymerase in a volume of 50 μl, was incubated at 37°C for one hour. The RNA was separated by phenol-chloroform extraction followed by ethanol precipitation. Then it was translated with a rabbit reticulocyte lysate in the presence of ^{35}S-methionine. The reaction mixture, containing 40 μl of rabbit reticulocyte lysate, 3 μl of RNA (1.5 μg/μl) and 7 μl of ^{35}S-methionine (10 μCi/μl, 37 TBq/mmol), was incubated for one hour at 30°C, then incubated with 5 μl samples of sera from patients with post-transfusion non-A, non-B chronic hepatitis, hepatocellular carcinoma or healthy individuals for 12 h, and further incubated with protein A Sepharose for 5 h. The immunocomplex bound to protein A Sepharose was extracted with 2% sodium dodecyl sulfate (SDS) and 0.1 *M* beta-mercaptoethanol, and was analyzed by SDS-15% polyacrylamide gel electrophoresis.[10] The gel was dried and autoradiographed. As expected from the molecular weight of the putative amino acid sequence, a band of about 23 kd was detected when the product was immunoprecipitated with sera from patients with chronic hepatitis, but was not detected after immunoprecipitation with sera from healthy individuals (Fig. 2). Many sera from patients with chronic hepatitis used in this experiment gave a positive reaction for antibody against C100 by the enzyme-linked immunosorbent assay (ELISA) established by Ortho Diagnostic Systems. These results indicate that the protein expressed *in vitro* was related to

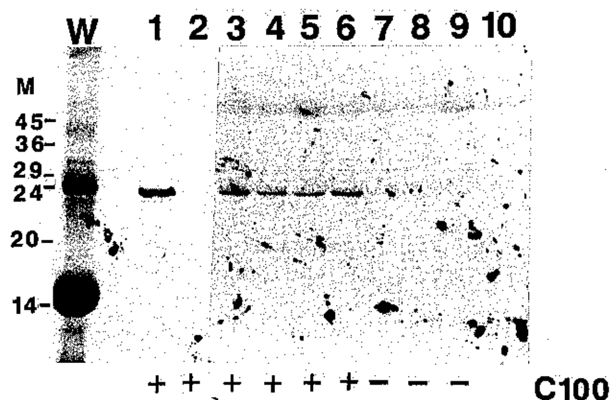

Fig. 2. Detection of antibody against the core protein synthesized in the *in vitro* transcription-translation system. Lane W, whole reaction mixture; lanes 1 and 2, immunoprecipitation with sera from patients with hepatocellular carcinoma; lanes 3 to 6, immunoprecipitation with sera from patients with chronic hepatitis; lanes 7 to 9, immunoprecipitation with sera from healthy persons; lane 10, immunoprecipitation of reaction mixture, carried out without the viral RNA, with serum from a patient with chronic hepatitis. M, size markers (×1000). The results of the ELISA of C100 are shown at the bottom of this figure.

HCV infection and probably reflected the expression of at least part of a structural viral protein.

Other methods to detect HCV infection include the ELISA system from Ortho Diagnostic Systems and amplification of part of the HCV genome by reverse transcription followed by PCR. Both of these methods detect viral infection, but they are not effective in some cases.[6] Positive sera found by the method described in this paper were also positive for HCV RNA by RT-PCR (unpublished result). Therefore, detection of antibody against HCV using structural protein should be useful for the diagnosis of viral infection. Recently, we succeeded in expressing this virus protein efficiently in bacteria and have shown that it efficiently detects the antibody against HCV, indicating its value for diagnosis of HCV infection.

This work was supported by a Grant-in-Aid from the Ministry of Education, Science and Culture, and a Grant-in-Aid for the Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health and Welfare of Japan. M.H. is the recipient of a Research Resident Fellowship from the Foundation for Promotion of Cancer Research.

(Received July 23, 1990/Accepted August 15, 1990)

REFERENCES

1) Kuo, G., Choo, Q-L., Alter, H. J., Githick, G. L., Redeker, A. G., Purcell, R. H., Miyamura, T., Dienstag, J. L., Alter, M. J., Stevens, C. E., Tegtmeier, G. E., Bonino, F., Colombo, M., Lee, W-S., Kuo, C., Berger, K., Shuster, J. R., Overby, L. R., Bradley, D. W. and Houghton, M. An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science*, **244**, 362–364 (1989).

2) Choo, Q-L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W. and Houghton, M. Isolation of a cDNA clone derived from a blood borne non-A, non-B viral hepatitis genome. *Science*, **244**, 359–362 (1989).

3) Brouix, J., Barrera, J. M., Calvet, X., Ercilla, G., Costa, J., Sanchez-Tapias, J. M., Ventura, M., Vall, M., Bruguera, M., Bru, C., Castillo, R. and Rodes, J. Prevalence of antibodies to hepatitis C virus in Spanish patients with hepatocellular carcinoma and hepatic cirrhosis. *Lancet*, **ii**, 1004–1006 (1989).

4) Colombo, M., Kuo, G., Choo, Q-L., Danato, M. F., Minno, E. D., Tommasini, M. A., Dioguardi, N. and Houghton, M. Prevalence of antibodies to hepatitis C virus in Italian patients with hepatocellular carcinoma. *Lancet*, **ii**, 1006–1008 (1989).

5) Ohkoshi, S., Kojima, H., Tawaraya, H., Miyajima, T., Kamimura, T., Asakura, H., Satoh, A., Hirose, S., Hijikata, M., Kato, N. and Shimotohno, K. Prevalence of antibody against non-A, non-B hepatitis virus HCV in Japanese patients with hepatocellular carcinoma. *Jpn. J. Cancer Res.*, **81**, 550–553 (1990).

6) Ohkoshi, S., Kato, N., Kinoshita, T., Hijikata, M., Ootsuyama, Y., Okazaki, N., Ohkura, H., Hirohashi, S., Honma, A., Ozaki, T., Yoshikawa, A., Kojima, H., Asakura, H. and Shimotohno, K. Detection of hepatitis C virus DNA in sera and liver tissues of non-A, non-B hepatitis patients using the polymerase chain reaction. *Jpn. J. Cancer Res.*, **81**, 862–865 (1990).

7) Kaneko, S., Unoura, M., Kobayashi, K., Kuno, K., Murakami, S. and Hattori, N. Detection of serum hepatitis C virus RNA. *Lancet*, **i**, 976 (1990).

8) Rice, C. M., Strauss, E. G. and Strauss, J. H. Structure of the falvivirus genome. *In* "The Togaviridae and Flaviviridae," ed. S. Schlesinger and M. J. Schlesinger, pp. 279–326 (1986). Plenum Publishing Corp., New York.

9) Okamoto, H., Okada, S., Sugiyama, Y., Yotsumoto, T., Tanaka, T., Yoshizawa, H., Tsuda, F., Miyakawa, Y. and Mayumi, M. The 5′-terminal sequence of the hepatitis C virus genome. *Jpn. J. Exp. Med.*, **60**, 167–177 (1990).

10) Shimotohno, K., Miwa, M., Slamon, D. J., Chen, I. S. Y., Hoshino, H., Takano, M., Fujimoto, M. and Sugimura, T. Identification of new gene products coded from X region of human T-cell leukemia viruses. *Proc. Natl. Acad. Sci. USA*, **82**, 302–306 (1985).