Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 4942–4946, June 1992
Biochemistry

Sequence analysis of the 5′ noncoding region of hepatitis C virus

(non-A, non-B hepatitis/genetic heterogeneity/virus evolution)

JENS BUKH, ROBERT H. PURCELL*, AND ROGER H. MILLER

Hepatitis Viruses Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892

Contributed by Robert H. Purcell, February 14, 1992

ABSTRACT **We have determined the nucleotide sequence of the 5′ noncoding (NC) region of the hepatitis C virus (HCV) genome in 44 isolates from around the world. We have identified several HCV isolates with significantly greater sequence heterogeneity than reported previously within the 5′ NC region. The most distantly related isolates were only 90.1% identical. Nucleotide insertions were seen in three isolates. Analysis of the nucleotide sequence from 44 HCV isolates in this study combined with that of 37 isolates reported in the literature reveals that the 5′ NC region of HCV consists of highly conserved domains interspersed with variable domains. The consensus sequence was identical to the prototype HCV sequence. Nucleotide variations were found in 45 (16%) of the 282 nucleotide positions analyzed and were primarily located in three domains of significant heterogeneity (positions −239 to −222, −167 to −118, and −100 to −72). Conversely, there were three highly conserved domains consisting of 18, 22, and 63 completely invariant nucleotides (positions −263 to −246, −199 to −178, and −65 to −3, respectively). Two nucleotide domains within the 5′ NC region, conserved among all HCV isolates studied to date, shared statistically significant similarity with pestivirus 5′ NC sequences, providing further evidence for a close evolutionary relationship between these two groups of viruses. Additional analysis revealed the presence of short open reading frames in all HCV isolates. Our sequence analysis of the 5′ NC region of the HCV genome provides additional information about conserved elements within this region and suggests a possible functional role for the region in viral replication or gene expression. These data also have implications for selection of optimal primer sequences for the detection of HCV RNA by the PCR assay.**

The etiological agent of most posttransfusion non-A, non-B hepatitis cases, hepatitis C virus (HCV), is a positive-stranded RNA virus with a linear genome ≈9.5 kilobases (kb) in length (1, 2). The structure and organization of the HCV genome are similar to that of pesti- and flaviviruses (3–5). Published sequence data indicate that the 5′ noncoding region (NC) of 324–341 nucleotides is generally highly conserved among different HCV isolates (6–8) and, furthermore, is the most highly conserved region of the HCV genome (3–5, 9, 10). Also the 5′ NC region of the HCV genome shares sequence similarity with the 5′ NC region of pestiviruses (5, 7). The high degree of sequence conservation has made this region the choice for primer selection in reverse transcription and amplification of HCV RNA by PCR (cDNA PCR) (for review, see ref. 11). Four short open reading frames (ORFs) have been described in the 5′ NC region of HCV (5, 7), but the significance of these ORFs is unknown. Overall, these findings suggest an important functional role of the 5′ NC region of the HCV genome in virus replication or gene expression.

In a recent study we tested sera from 114 individuals positive for antibodies to HCV (anti-HCV) from around the world for the presence of HCV RNA with four different primer sets in a cDNA PCR assay (12). In this study we have determined the nucleotide sequence of the 5′ NC region of the HCV genome of 44 of these HCV isolates.† The isolates selected for analysis were chosen because they are representative of the different geographical locations and of the different patterns of reactivity to the primers used in the previous study (12). We find that, contrary to previously published data, the 5′ NC region of the HCV genome possesses significant heterogeneity among different HCV isolates.

MATERIALS AND METHODS

Serum samples used in this study were from 44 anti-HCV-positive individuals from 12 countries [Denmark (DK), Dominican Republic (DR), Germany (D), Hong Kong (HK), India (IND), Italy (S), Peru (P), South Africa (SA), Sweden (SW), Taiwan (T), United States (US), and Zaire (Z)]. These samples were used in a recent study in which sera from 114 anti-HCV-positive individuals were tested for HCV RNA in a cDNA PCR assay with four primer sets (12). Primer set *a* was from within the 5′ NC region of the HCV genome; primer set *b* spanned the 5′ end of the 5′ NC region to the 5′ end of the putative core gene sequence; primer set *c* was from the 3′ end of the 5′ NC region to the 3′ end of the core gene region; and primer set *d* was from the nonstructural protein 3-like gene of HCV. We selected for sequence analysis HCV isolates that represented each of the 12 countries and that reflected probable heterogeneity as measured by the different patterns of reactivity with primer sets *a–d*. Specifically, isolates DR4, DK7, HK5, S9, SW2, T3, and US11 were positive with primer sets *a*, *b*, *c*, and *d*; isolates D3, D6, DK11, DK13, IND8, P10, SA1, SA7, SA10, SW3, T10, US6, and Z4 were positive with primer sets *a*, *b*, and *c*; isolate DK9 was positive with primer sets *a*, *b*, and *d*; isolates DK12, HK2, HK10, IND3, IND5, P8, S45, S52, S54, S83, SA11, T8, T9, US3, Z1, Z5, Z6, and Z8 were positive with primer sets *a* and *b*; and isolates SA3, T4, US1, US10, and Z7 were positive only with primer set *a*. Viral RNA was extracted from serum, reverse-transcribed, and the resulting cDNA was amplified in a nested PCR assay as described (12). For 39 HCV isolates that were detected with primer set *b* we sequenced that PCR product, a 321-nucleotide DNA fragment that spanned 282 nucleotides of the 5′ NC region and 39 nucleotides of the core gene region of HCV (i.e., positions −282 to 39). In five HCV isolates that could be amplified only with primer set *a*, that PCR product, a 196-nucleotide DNA fragment from the 5′ NC region (i.e., positions −246 to −51)

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: HCV, hepatitis C virus; NC, noncoding; ORF, open reading frame; BVDV, bovine viral diarrhea virus.

*To whom reprint requests should be addressed.

†The sequences reported in this paper have been deposited in the GenBank data base (accession nos. M84822–M84865).

was sequenced. A number of standard precautions were taken to reduce the risk of contamination with exogenous RNA, and a negative control was included for every serum sample tested in the RNA extraction, reverse transcription, and PCR amplification to rule out contamination as a source of false positive results (12). Amplified DNA for sequencing was purified by gel electrophoresis as described (12), and ≈100 ng of DNA was used for direct sequencing by the dideoxynucleotide chain-termination method (13, 14) with phage T7 DNA polymerase (Sequenase, United States Biochemical). Serum containing the prototype HCV isolate (HCV-1, refs. 5 and 7), provided by D. W. Bradley (Centers for Disease Control, Atlanta), was used as a positive control in PCR and sequencing reactions. Computer analysis of the sequences of HCV, flavivirus, and pestivirus genomes was done as described (15, 16).

RESULTS AND DISCUSSION

Nucleotide Sequence of the 5′ Noncoding Region from 44 HCV Isolates. A primary goal of this investigation was to analyze the nucleotide sequence of the 5′ NC region of the HCV genome from a large number of isolates obtained throughout the world. Therefore, we reverse-transcribed HCV RNA, PCR-amplified the resultant cDNA, and directly sequenced the product to obtain the "consensus" sequence in each serum sample. An alignment of the nucleotide sequence of the 5′ NC region from the 44 HCV isolates we studied is presented in Fig. 1. Previously published multiple sequence alignments of others (6–8) demonstrated that the 5′ NC region was highly conserved and 98% identical with prototype isolate HCV-1. In contrast, we have identified several HCV isolates with significantly more sequence variability. (*i*) We found that three isolates had nucleotide insertions. Isolate HK2 had two separate nucleotide insertions of one and two nucleotides, whereas isolates Z5 and Z8 each had a single nucleotide insertion. We did not detect nucleotide deletions. (*ii*) Nucleotide variation among all of the additional HCV isolates was as high as 9.9% (HK10, S52, S54 versus DK11, T8, SW3), and the nucleotide variation from HCV-1 was as high as 6.4% (DK12, HK10, S52, and S54) within the region of 282 nucleotides sequenced. Thus, contrary to the findings of others, we have demonstrated significant sequence variation within the 5′ NC region of the HCV genome.

Houghton and coworkers (11) analyzed the degree of sequence heterogeneity of HCV isolates and, based on this,

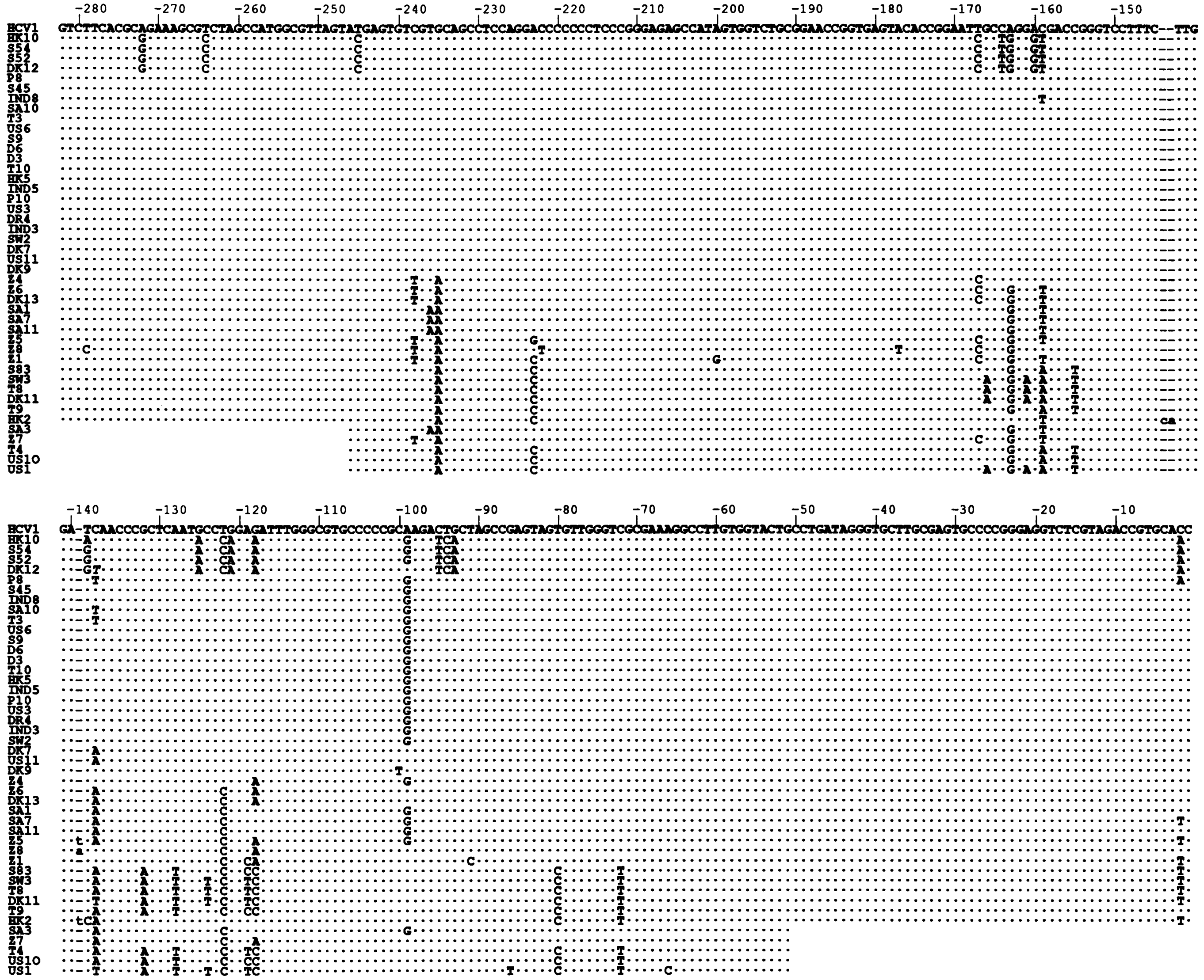

FIG. 1. Alignment of nucleotide sequences of the 5′ NC region of 44 HCV isolates from around the world. The sequences are compared to the prototype HCV sequence (HCV-1, refs. 5 and 7, and resequenced in this study) shown on the top line. Nucleotide substitutions are indicated as uppercase letters, and identical nucleotides are shown as dots. Nucleotide insertions seen in three isolates (HK2, Z5, and Z8) are shown as lowercase letters. A single site of microheterogeneity is shown in italics, at position −138 in isolate DK12 (i.e., C and T).

segregated them into three groups (HCV I, II, and III). Other published sequences can be placed in these groups (9, 10). Group I includes isolates HCV-1 (5, 7), HC-J1 (6), and HCV-H (10); group II includes isolates HC-J4 (6), HCV-J (4), HCV-BK (3), and HCV-K1 (17); and group III includes isolates HCV-K2a (17), HCV-K2b (17), and HC-J6 (9). Our analysis, based entirely on sequence of the 5′ NC region, shows that isolates from the present study represent a broad spectrum of sequence patterns that cannot all be placed within these groups. The predominant sequence pattern, seen in isolates D3, D6, DK7, DK9, DR4, HK5, IND3, IND5, IND8, P8, P10, S9, S45, SA10, SW2, T3, T10, US3, US6, and US11, was most similar to the sequence of the prototype isolate HCV-1 and closely related sequences (HCV groups I and II). In addition, there were several isolates (DK11, S83, SW3, T4, T8, T9, US1, and US10) with sequences similar to those of HCV group III. However, in the remaining isolates (DK12, DK13, HK2, HK10, S52, S54, SA1, SA3, SA7, SA11, Z1, Z4, Z5, Z6, Z7, and Z8) the 5′ NC sequence was significantly different from those of reported sequences (Fig. 1). It is noteworthy that the DK12, HK10, S52, and S54 sequences were 5–10% different from any other isolate. Thus, we have observed patterns of nucleotide sequence in the HCV 5′ NC region significantly different from the patterns assigned to groups I, II, and III (11). We were unable to segregate the sequence patterns that we observed in the 5′ NC region to defined geographical regions (Fig. 1). Further sequence analysis will demonstrate how the heterogeneity observed in the 5′ NC region among different HCV isolates relates to sequence differences elsewhere in the HCV genome.

It is well known that the primary sequence around the AUG initiation codon of a gene is important for initiation of translation (for review, see ref. 18). In this context we find it interesting that (*i*) the polyprotein start codon occurs at the same location relative to the prototype sequence in all 39 HCV isolates studied and (*ii*) the nucleotides surrounding the AUG codon are particularly well-conserved. Except for the nucleotide variation at position −2, the nucleotide sequence at positions +8 to −65 is invariant among all studied isolates. It is noteworthy that the nucleotide at position −3 contains adenine because a purine at this position is regarded as key for initiation of translation (18). Our data imply that the position of the AUG initiation codon and the surrounding sequence is crucial to the translation of the HCV polyprotein.

Different Open Reading Frames of the 5′ NC Region of the HCV Genome. Han and coworkers (7) recently described short ORFs in the 5′ NC region of the HCV and pestivirus genomes. We find that all HCV isolates examined in this study also have short ORFs. The different patterns of short ORFs observed within the 5′ NC region of the various isolates of HCV are shown in Fig. 2. We cannot comment on the presence or absence of ORF 1, described previously (5, 7), in our analysis because the initiation codon of ORF 1 in HCV-1 is 5′ to the region that we sequenced. However, most HCV isolates included in this study have three short ORFs (ORFs 2–4) identical to those described (5, 7) in the prototype HCV sequence. Interestingly, different patterns of these ORFs were seen in several HCV isolates. An initiation codon beginning at nucleotide position −160 was found in 11 isolates (DK13, HK2, IND8, SA1, SA3, SA7, SA11, Z1, Z5, Z6, and Z7). This ORF (ORF 5) was 7 amino acids in length and was fused with ORF 3 in several isolates (Fig. 2). Four isolates (DK12, HK10, S52, and S54) had only a single ORF, a version of ORF 4 longer at the 3′ end by 21 amino acids for a total of 25 amino acids. Isolate Z1 possessed the most unusual arrangement of ORFs: ORF 2 was elongated at the 3′ end by 14 amino acids for a total of 29 amino acids, and ORFs 3 and 5 were fused, creating an ORF of 49 amino acids. Thus, this isolate had ORFs spanning most of the 5′ NC region and may reflect the genomic organization of a putative ancestral virus that encoded a polyprotein extending into what is now an untranslated region. The functional status of the ORFs in the 5′ NC region of the HCV genome is unknown. We find it interesting that all 44 HCV isolates included in this study, as well as 35 HCV isolates reported by others (3–10, 16, 17, 19), have at least one ORF within the 5′ NC region. These data are consistent with the hypothesis that these ORFs are maintained because of a role in control of translation (7).

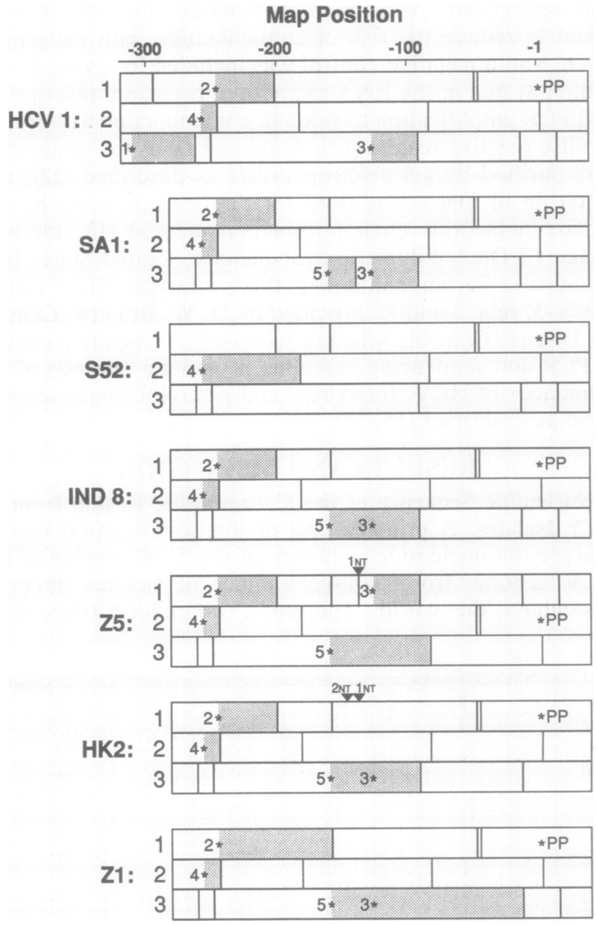

FIG. 2. Short ORFs within the 5′ NC region of the HCV genome. Shown at the top are the short ORFs described previously in the HCV prototype sequence (5, 7). All three translation frames of representative isolates are shown. The shadowed areas represent the short ORFs, defined as an initiation codon followed by codons specifying amino acids and not termination codons. Initiation codons are depicted as stars and numbered 1–5. Termination codons are depicted as vertical lines. The polyprotein start codon is marked PP. Inverted triangles show the position of nucleotide insertions. An ORF pattern similar to that of isolate SA1 was seen in isolates DK13, SA7, SA11, and Z6, and an ORF pattern similar to that of S52 was seen in isolates DK12, HK10, and S54.

Consensus Sequence of the 5′ NC Region of the HCV Genome of 81 HCV Isolates. To determine the extent of sequence variability within the 5′ NC region of the HCV genome, we combined our data on 44 HCV isolates (Fig. 1) with that of 37 reported HCV isolates (3–10, 16, 17, 19, 20) and performed a multiple sequence alignment on all HCV 5′ NC sequences currently available. The resulting consensus sequence, identical to that of the prototype sequence (5, 7), is shown in Fig. 3 as a histogram, illustrating the percent of sequences different from the consensus sequence at each nucleotide position in 282 nucleotides of the 5′ NC region of the HCV genome. Our data confirm that the 5′ NC region of the HCV genome is well conserved among HCV isolates

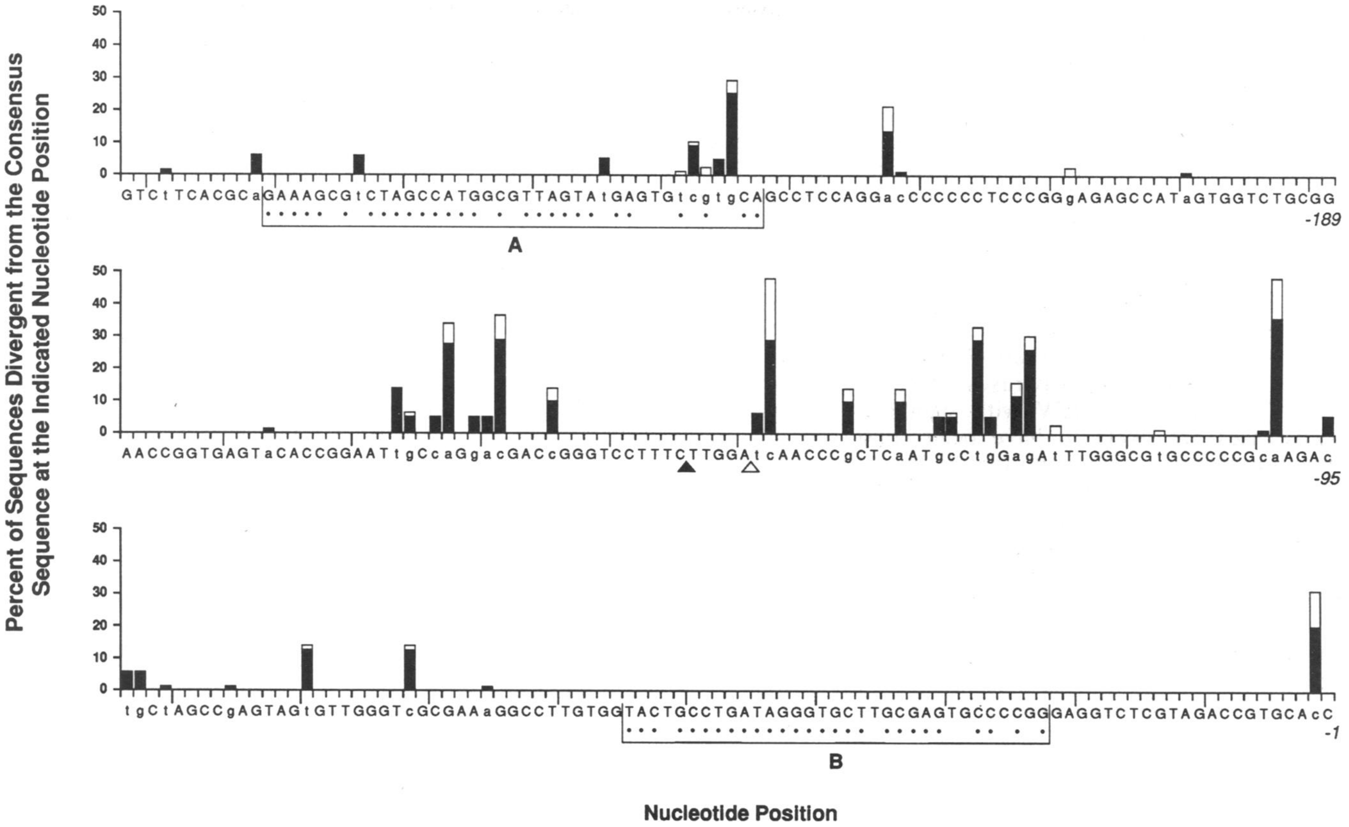

FIG. 3. Histogram of the percent of sequences different from the consensus sequence at each nucleotide position in 282 nucleotides of the 5′ noncoding region among 81 HCV isolates (37 published sequences and 44 sequences from this study) from around the world. The sequence of all 282 nucleotide positions was available in 53 isolates, of which 39 were from this study; the remaining sequences were partial sequences. The consensus sequence and the percent of sequences divergent from it are given at each nucleotide position for all sequences available at that site. Open bars represent contribution of published sequences; closed bars represent data from this study. Uppercase and lowercase letters indicate nucleotide positions that are invariant and variable, respectively. The closed triangle indicates the position of an insertion of two nucleotides seen in isolate HK2, whereas the open triangle indicates the position of insertions of a single nucleotide seen in isolates HK2, Z5, and Z8. The two domains of HCV-1 with significant similarity to bovine viral diarrhea virus (BVDV) are boxed (spaces between adjacent nucleotides required for optimal alignment of HCV-1 and BVDV are omitted in this figure). Dots under a nucleotide within boxes A and B indicate identical nucleotide matches between HCV-1 and BVDV.

because the overall nucleotide variation, defined as the total number of nucleotides different from the consensus sequence, is only 2.0%. However, we find that the 5′ NC region of HCV consists of highly conserved domains interspersed between variable domains. Nucleotide variations from the consensus sequence are found in 45 (16%) of the 282 nucleotide positions. The most variable domain spans 50 nucleotides (positions −167 to −118) and has an overall nucleotide variation of 6.0% with variation from the consensus sequence at 18 (36%) of the nucleotide positions. In addition, the nucleotide insertions observed in three HCV isolates in this study are located within this domain (Fig. 1). It is noteworthy that the nucleotide identity in this variable domain between two different HCV isolates (HK10, S52, S54 versus DK11, SW3, T8, US1, and DK12 versus SW3, T8) is as low as 68%, and the identity to HCV-1 is as low as 76% (DK11, SW3, T8, US1) (Fig. 1). Furthermore, this variable domain has two subdomains with even greater variability flanking a region that is invariant, except for the nucleotide insertions observed in three isolates. The overall nucleotide variability of these subdomains (positions −167 to −155 and −139 to −118) is 9.3% and 8.1%, respectively, and nucleotide variations from the consensus sequence were seen at 8 (61.5%) and 10 (45.5%) of the nucleotide positions, respectively. We have thus defined a variable domain of 50 nucleotides within the 5′ NC region of HCV with significant heterogeneity among different HCV isolates. Two other variable domains could be identified in the HCV 5′ NC region (positions −239 to −222 and −100 to −72) that displayed an overall nucleotide variation of 4.0% and 3.4%, respectively. Nucleotide variations from the consensus sequence were seen at 7 (38.9%) and 9 (31%) of these nucleotide positions, respectively. The high degree of nucleotide changes seen within the variable domains suggests that functional constraints are low in these regions.

The remaining 185 of the 282 nucleotides of the HCV 5′ NC region analyzed are highly conserved with an overall nucleotide variation of only 0.3%. Nucleotide variations from the consensus sequence were seen at only 11 (5.9%) of the 185 nucleotide positions. Interestingly, there are three long stretches with completely invariant nucleotide sequences of 18, 22, and 63 bases (positions −263 to −246, −199 to −178, and −65 to −3, respectively) among all studied HCV isolates. Most impressive is the stretch of 63 invariant nucleotides immediately upstream of the polyprotein start codon. This region contains a domain with significant similarity to pestiviruses (see below). The 5′ NC region of a viral genome typically contains regulatory elements, so it is likely that this region in HCV is conserved because it contains cis-acting elements involved in replication of the viral genome (e.g., RNA packaging signal, etc.) or expression of viral genes (e.g., translation initiation signal, etc.).

Previously reported sequence analysis suggests a distant evolutionary relationship among certain proteins of HCV, flaviviruses, and pestiviruses (15, 21). We used computer-assisted nucleotide sequence analysis to look for similarity within the 5′ NC regions of these viruses. Using the program SEQ, we found no such similarity between HCV and flaviviruses. However, two nucleotide domains within the 5′ NC region of the HCV-1 genome showed statistically significant similarity to 5′ NC sequences of pestiviruses. (*i*) A domain in the 5′ end of the 5′ NC region (box A in Fig. 3) of HCV-1 was found to have statistically significant similarity ($P = 0.003$) with BVDV (22) but not with hog cholera virus (23, 24). (*ii*) Next, a domain in the 3′ end of the 5′ NC region (box B in Fig. 3) of HCV-1 was found to have significant similarity ($P = 0.001$) with BVDV (22) and two strains of hog cholera virus ($P = 0.002$, ref. 23 and $P = 0.005$, ref. 24). These two nucleotide domains correspond to conserved regions I and IV described by Han and coworkers (7) in an alignment of the 5′ NC region of HCV and pestiviruses. The two domains within the 5′ NC region of HCV with significant similarity to pestiviruses are likely to have been conserved in virus evolution because of an important biological role. This hypothesis is supported by our finding that these domains are conserved among many different HCV isolates (Fig. 3).

The findings in this study have important implications for the selection of primers in cDNA PCR assays to detect HCV RNA because genetic heterogeneity among different HCV strains results in false negative results because of primer and template mismatch. The 5′ NC region, previously shown to be the most conserved region of the HCV genome, was a natural first choice for designing primers for cDNA PCR assays (for detailed review, see ref. 11). However, in this comprehensive analysis, we have identified additional variable sequences within the 5′ NC region of the HCV genome that should be avoided in the design of primers for cDNA PCR assays. Conversely, our data highlight regions within the 5′ NC region of the HCV genome that are especially well conserved among a large number of different HCV isolates from around the world. The long stretches of invariant nucleotides are preferred for primer design. We have recently shown in a study comparing primers that a cDNA PCR assay with primers designed from the two domains of the HCV 5′ NC region that share statistically significant similarity with BVDV sequences (i.e., boxes A and B in Fig. 3) is both sensitive and specific for detecting HCV RNA (12). We have now further demonstrated that these domains are highly conserved in a large number of HCV isolates from around the world.

In summary, we have demonstrated significant nucleotide sequence variation within the 5′ NC region of HCV in several HCV isolates. Furthermore, we have defined highly conserved domains within the 5′ NC region of HCV, which suggest that these domains have crucial functional roles.

We thank Ms. T. Tsareva for synthesis of oligonucleotides and Mr. T. J. Heishman for assistance with graphics and manuscript preparation. We gratefully acknowledge the generosity of the following investigators for providing sera for this study: Drs. H. Alter, V. Arankalle, D. W. Bradley, D.-S. Chen, P. Farci, M. C. Kew, K. Krogsgaard, A. Lok, T. Quinn, F. Renger, M. Sjogren, and A. Widell. Computer-assisted nucleic acid analysis was through the GenBank Online Service.

1. Choo, Q.-L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W. & Houghton, M. (1989) *Science* **244**, 359–362.
2. Kuo, G., Choo, Q.-L., Alter, H. J., Gitnick, 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. & Houghton, M. (1989) *Science* **244**, 362–364.
3. Takamizawa, A., Mori, C., Fuke, I., Manabe, S., Murakami, S., Fujita, J., Onishi, E., Andoh, T., Yoshida, I. & Okayama, H. (1991) *J. Virol.* **65**, 1105–1113.
4. Kato, N., Hijikata, M., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S., Sugimura, T. & Shimotohno, K. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9524–9528.
5. 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) *Proc. Natl. Acad. Sci. USA* **88**, 2451–2455.
6. Okamoto, H., Okada, S., Sugiyama, Y., Yotsumoto, S., Tanaka, T., Yoshizawa, H., Tsuda, F., Miyakawa, Y. & Mayumi, M. (1990) *Jpn. J. Exp. Med.* **60**, 167–177.
7. Han, J. H., Shyamala, V., Richman, K. H., Brauer, M. J., Irvine, B., Urdea, M. S., Tekamp-Olson, P., Kuo, G., Choo, Q.-L. & Houghton, M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 1711–1715.
8. Cha, T.-A., Kolberg, J., Irvine, B., Stempien, M., Beall, E., Yano, M., Choo, Q.-L., Houghton, M., Kuo, G., Han, J. H. & Urdea, M. S. (1991) *J. Clin. Microbiol.* **29**, 2528–2534.
9. Okamoto, H., Okada, S., Sugiyama, Y., Kurai, K., Iizuka, H., Machida, A., Miyakawa, Y. & Mayumi, M. (1991) *J. Gen. Virol.* **72**, 2697–2704.
10. Inchauspe, G., Zebedee, S., Lee, D.-H., Sugitani, M., Nasoff, M. & Prince, A. M. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 10292–10296.
11. Houghton, M., Weiner, A., Han, J., Kuo, G. & Choo, Q.-L. (1991) *Hepatology* **14**, 381–388.
12. Bukh, J., Purcell, R. H. & Miller, R. H. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 187–191.
13. Bachmann, B., Luke, W. & Hunsmann, G. (1990) *Nucleic Acids Res.* **18**, 1309.
14. Winship, P. R. (1989) *Nucleic Acids Res.* **17**, 1266.
15. Miller, R. H. & Purcell, R. H. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2057–2061.
16. Ogata, N., Alter, H. J., Miller, R. H. & Purcell, R. H. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 3392–3396.
17. Nakao, T., Enomoto, N., Takada, N., Takada, A. & Date, T. (1991) *J. Gen. Virol.* **72**, 2105–2112.
18. Kozak, M. (1991) *J. Cell Biol.* **115**, 887–903.
19. Fuchs, K., Motz, M., Schreier, E., Zachoval, R., Deinhardt, F. & Roggendorf, M. (1991) *Gene* **103**, 163–169.
20. Takeuchi, K., Kubo, Y., Boonmar, S., Watanabe, Y., Katayama, T., Choo, Q.-L., Kuo, G., Houghton, M., Saito, I. & Miyamura, T. (1990) *J. Gen. Virol.* **71**, 3027–3033.
21. Choo, Q.-L., Han, J., Weiner, A. J., Overby, L. R., Bradley, D. W., Kuo, G. & Houghton, M. (1991) in *Viral Hepatitis C, D, and E*, eds. Shikata, T., Purcell, R. H. & Uchida, T. (Elsevier, Amsterdam), pp. 47–52.
22. Collett, M. S., Larson, R., Gold, C., Strick, D., Anderson, D. K. & Purchio, A. F. (1988) *Virology* **165**, 191–199.
23. Meyers, G., Rumenapf, T. & Thiel, H.-J. (1989) *Virology* **171**, 555–567.
24. Moormann, R. J. M., Warmerdam, P. A. M., van der Meer, B., Schaaper, W. M. M., Wensvoort, G. & Hulst, M. M. (1990) *Virology* **177**, 184–198.