

## Antigenic Heterogeneity of the Hepatitis C Virus NS5A Protein

Xiao-Guang Dou,<sup>1,2</sup> Ganesh Talekar,<sup>1</sup> Joy Chang,<sup>1</sup> Xing Dai,<sup>1</sup> Lixia Li,<sup>1</sup> Maria-Teresa Bonafonte,<sup>3</sup> Brian Holloway,<sup>1</sup> Howard A. Fields,<sup>1</sup> and Yury E. Khudyakov<sup>1\*</sup>

Hepatitis Branch, Division of Viral and Rickettsial Diseases,<sup>1</sup> and Biotechnology Core Facility, Scientific Resources,<sup>2</sup> National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, and VA Medical Center, Decatur,<sup>3</sup> Georgia, and Department of Infectious Disease, China Medical University, Shenyang, China<sup>2</sup>

Received 14 November 2000/Returned for modification 22 March 2001/Accepted 9 October 2001

The effect of sequence variability between different types of hepatitis C virus (HCV) on the antigenic properties of the NS5 protein was studied by using recombinant proteins. A strong antigenic region was identified within the HCV NS5A protein at amino acids 2212 to 2313. Forty-five unique sequences encompassing this region were selected from GenBank and were compared to each other. The results of this analysis showed that the primary structure of this strong antigenic region is highly variable. Percent homology between different genotype sequences varied from 40.4 to 72.5%. Thirteen representative sequences from all six HCV genotypes were selected to design synthetic genes coding for this antigenic region. These genes were assembled by PCR from synthetic oligonucleotides and expressed in *Escherichia coli* as hybrid proteins with glutathione S-transferase. All 13 fusion proteins were purified from bacterial lysates and used to test a panel of anti-HCV positive sera ( $n = 91$ ) obtained from patients infected with HCV genotypes 1 through 6. All but two proteins immunoreacted with 62 to 93% of HCV anti-NS5-positive serum samples. Although a variable degree of genotype-specific antigenic reactivity was detected, only one protein demonstrated a noticeable preference to immunoreact with antibodies against the homologous HCV genotype. On the other hand, closely related proteins derived from the same subtype or genotype immunoreacted with significantly different efficiency with HCV antibodies. Thus, sequence variability has a profound effect on the antigenic properties of the NS5A immunodominant regions. This observation should be taken into consideration in the development of diagnostic tests for the efficient detection of anti-HCV activity in serum specimens.

The hepatitis C virus (HCV) is associated with the vast majority of cases of posttransfusion hepatitis and sporadic non-A, non-B hepatitis worldwide (6). The HCV genome is a single-stranded positive RNA coding for a polyprotein of ~3,000 amino acids (aa) (7) which is processed into 10 viral proteins, namely, core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B (14, 23). The HCV genome is very heterogeneous and has been classified into six genotypes (18).

HCV serologic diagnosis is achieved by detecting specific antibody by enzyme immunoassays (EIAs) that use antigens derived from core, NS3, NS4, and NS5 proteins. Although screening assays have been effective in reducing the risk of posttransfusion hepatitis C, these assays do not detect anti-HCV activity in all patients with HCV infection. For example, the second-generation EIA (EIA-2) detects anti-HCV in ~90% of HCV-infected patients (2). In accordance with this observation, 3 to 10% of anti-HCV-negative donors transmitted HCV to their recipients (1, 3, 28). In addition, these assays are prone to false-positive results, especially noticeable in low-risk populations (9). For example, approximately 40 to 50% of specimens from healthy blood donors which were positive by EIA-2 could not be confirmed by a supplemental assay (12). A recently released third generation of EIA (EIA-3) includes the NS5 protein of 942 aa. EIA-3 is slightly more sensitive than EIA-2, but most of the increased sensitivity appears to be

attributable to improvements in detecting anti-C33 (4, 11, 27). The detection rate of anti-NS5 (59.3%) is much lower than the detection rates of anti-C22 (92.1%) and C33 (88.1%) (11). Commercial HCV tests based upon proteins or peptides derived solely from HCV genotype 1 may be less effective for the detection of antibody against HCV of other genotypes (8, 17, 24). Dow et al. (10) found that sera from donors infected with different genotypes showed different patterns of reactivity by the third-generation confirmatory assay. As more data are accumulated, it has become increasingly evident that further improvement of HCV diagnostic tests will rely on the careful selection of sequence variants of antigens capable of detecting antibodies against different HCV genotypes with equal efficiencies.

The antigenic compositions of many HCV proteins have been studied by using recombinant proteins (19) and synthetic peptides (13, 24, 25, 26). Studies of the major HCV antigenic regions showed that sequence heterogeneity has an impact on the antigenic properties of these regions (17, 19). However, the antigenic properties of the NS5 protein have not been carefully studied, and, therefore, no attempts have been made to improve the diagnostic value of this antigen.

The present paper describes the immunoreactivity of 13 proteins containing the major NS5A antigenic region (aa 2212 to 2313) derived from all six HCV genotypes. The data showed that the antigenic reactivity of these proteins is significantly affected by sequence heterogeneity.

### MATERIALS AND METHODS

**Serum specimens.** A genotyped panel of 91 anti-HCV-positive serum specimens, 61 of which are anti-NS5 positive, was obtained from the United States

\* Corresponding author. Mailing address: Hepatitis Branch, MS A-33, Division of Viral and Rickettsial Diseases, Centers for Disease Control and Prevention, 1600 Clifton Rd., Atlanta, GA 30333. Phone: (404) 639-2610. Fax: (404) 639-1563. E-mail: yek0@cdc.gov.

(genotypes 1, 2, and 3), Egypt (genotype 4), Venezuela (genotypes 1 and 3), Australia (genotypes 1, 3, and 4), South Africa (genotype 5), Thailand (genotype 6), and Vietnam (genotype 1, 3, and 6). Eighteen specimens from this panel representing HCV genotypes 1 to 5 were also obtained from Boehringer Mannheim GmbH (Penzberg, Germany). This panel includes specimens from patients infected with HCV genotype 1 ( $n = 25$ ), 2 ( $n = 24$ ), 3 ( $n = 19$ ), 4 ( $n = 8$ ), 5 ( $n = 5$ ), and 6 ( $n = 10$ ). The HCV genotype was confirmed by direct sequencing of PCR fragments derived from the 5'-end-noncoding region or by restriction fragment length polymorphism analysis (5). Anti-HCV-negative serum specimens ( $n = 40$ ) were obtained from healthy blood donors. The anti-HCV status was determined by second-generation anti-HCV EIA (Abbott Laboratories, Chicago, Ill.) and confirmed by third-generation recombinant immunoblot assay (RIBA-3; Chiron Corporation, Emeryville, Calif.).

**Synthetic gene assembly.** Each gene was assembled from four oligonucleotides: A1, A2, B1, and B2. The length of each oligonucleotide varied from 83 to 104 nucleotides (nt). As the first step, pairs A1-A2 and B1-B2 were annealed by their 3'-end complementary sequences and converted into a complete double-stranded-form A1A2 or B1B2 by the large fragment of DNA polymerase I from *Escherichia coli* (Promega, Madison, Wis.). The full-size gene was assembled by PCR using A1A2 and B1B2 fragments and two PCR primers. These PCR primers were designed from 5'-terminal sequences of oligonucleotides A1 and B2 by adding an *Xho*I site or a stop codon TGA and *Bam*HI site. The PCR conditions were as follows: 94°C for 45 s, 50°C for 40 s, and 72°C for 1 min for 30 cycles. Each fragment was digested with *Xho*I and *Bam*HI and cloned with pGEX-4T-2 (Promega). The primary structure of each insert was confirmed by direct sequencing.

**Gene expression and protein purification.** The transformed *E. coli* JM109 cells were grown until an optical density (OD) of 0.6 to 1.0 at 600 nm was obtained. Gene expression was activated by the addition of isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) (Sigma Chemical Co.) at a final concentration of 1 mM. The bacteria were harvested 3 h after IPTG activation. Cell lysates were subsequently prepared for protein purification. The glutathione *S*-transferase (GST) hybrid proteins were purified by ligand-affinity chromatography (Pharmacia Biotech Inc., Piscataway, N.J.).

**EIA.** To optimize plate coating, purified proteins were titrated to determine an optimal concentration (ranging from 20 to 70 ng per well) at which a pool of five positive serum specimens from patients infected with HCV genotype 1 and negative controls were maximally discriminated. Proteins were then adsorbed onto microtiter wells (Immuno II; Dynatech Industries, Inc., Chantilly, Va.) in 0.1 M phosphate-buffered saline (PBS), pH 7.5, overnight at room temperature. The wells were washed with PBS-Tween buffer (0.1 M PBS containing 0.05% Tween 20) five times. Serum specimens were diluted 1:200 in blocking solution and incubated in the microtiter wells with the preadsorbed recombinant protein for 1 h at 37°C. After the microtiter wells were washed, goat anti-human immunoglobulin conjugated to horseradish peroxidase diluted 1:4,000 in blocking solution was added to the wells and incubated for 1 h at 37°C. The microtiter wells were washed again five times and then subjected to color development by adding *o*-phenylenediamine according to the provided protocol (Abbott Laboratories). The reaction was stopped with 50  $\mu$ l of 2 N sulfuric acid, and the OD of the reaction was measured at 492 nm. The proteins were evaluated by calculating a signal/cutoff (S/C) value, where S is the OD value of positive sera and C is the cutoff value statistically determined as the mean OD value of 40 negative control serum samples plus 3.5 standard deviations of the mean.

**Computer-assisted sequence analysis.** Amino acid sequences were aligned by using the MegAlign program from Lasergene software package (DNASar Inc., Madison, Wis.). All sequences were collected in November 1998. GenBank accession numbers for the amino acid sequences of the NS5A region (aa 2182 to 2343) are as follows: genotype 1: 1, D90208; 2, D63857; 3, D50480; 4, D85516; 5, D10934; 6, D50482; 7, D89872; 8, U45476; 9, U16362; 10, X61596; 11, D14484; 12, D50481; 13, M58335; 14, D10750; 15, U01214; 16, D50485; 17, D45172; 18, E07579; 19, D50483; 20, D50484; 21, L02863; 22, S66587; 23, U89019; 24, AF034151; 25, M84754; 26, AF11751; 27, AF11752; 28, M32084; 29, M67463; 30, D10749; 31, D14853; genotype 2: 32, D50409; 33, D00944; 34, M95105; 35, D10988; 36, M95104; genotype 3: 37, X76918; 38, D28917; 39, D17763; 40, D49374; 41, D63821; genotype 4: 42, Y11604; genotype 5: 43, Y13184; and genotype 6: 44, D63822 (11); 45, Y12083.

## RESULTS

**Epitope mapping within the NS5A region.** The region of the HCV genome at nt 6246 to 7934 encoding the entire NS5A protein (aa 1973 to 2419) and the N-terminal region of the

NS5B protein (aa 2420 to 2535) was spanned with 17 overlapping PCR fragments of  $\sim$ 300 bp each (Table 1). PCR fragments were amplified from two large DNA fragments (nt 6076 to 7596; nt 7223 to 8220) cloned in *E. coli* as a template. Both fragments were obtained from a patient infected with HCV genotype 1b (sequence accession number AY044867). PCR primers (Table 1) were designed based on the HCV genotype 1b sequence with the GenBank accession number D90208. All 17 PCR fragments were cloned with vector pGEX4t-2 and expressed in *E. coli* as hybrid proteins with GST (data not shown). Each protein was tested by immunoblotting against a pool of five HCV-positive serum specimens. Two proteins comprising regions at aa 2183 to 2290 and aa 2212 to 2313, regions 34 and 35, respectively (Table 1), strongly immunoreacted with this pool (data not shown). These two proteins were purified from bacterial extracts and tested for HCV-specific antigenic reactivity by EIA against a panel of 44 anti-HCV NS5A-positive and 20 anti-HCV-negative serum specimens. Both proteins demonstrated strict HCV-specific immunoreactivity, although protein 34 was found to be slightly less immunoreactive than protein 35. Protein 34 immunoreacted with 37 anti-HCV-positive serum specimens with an average S/C ratio equal to 10.1, whereas protein 35 immunoreacted with 39 anti-HCV-positive serum specimens with an average S/C ratio of 11.5. Based on these data, the sequence of protein 35 was selected for further antigenic analysis.

**Design and construction of synthetic genes.** Thirteen sequences out of 45 obtained from GenBank (Fig. 1) were selected for further antigenic analysis based on two criteria: (i) representation of all six HCV genotypes and (ii) maximal evolutionary distances between sequences. Five sequences were selected from genotype 1, two sequences were selected from genotype 2, two sequences were selected from genotype 3, one sequence was selected from genotype 4, one sequence was selected from genotype 5, and two sequences were selected from genotype 6 (Fig. 1). Percent similarity between these amino acid sequences varied from 40.4 to 86.3% (Table 2). It is interesting that sequences 11 and 6a belong to the same HCV genotype, genotype 6. However, these two sequences share only 46.1% similarity, which is significantly lower than the percent similarity between some other sequences from different genotypes (Table 2). Each of the 13 sequences was reverse translated into nucleotide sequence using optimal *E. coli* codons. These nucleotide sequences were used to design synthetic oligonucleotides. Synthetic genes were assembled from these oligonucleotides by PCR (see Materials and Methods), cloned with vector pGEX4T-2, and expressed in *E. coli* as hybrid proteins with GST. All recombinant proteins were purified from *E. coli* extracts by ligand affinity chromatography and used to test the antigenic reactivity by EIA as described previously (5, 13).

**Immunoreactivity of recombinant proteins.** All recombinant proteins were tested against a panel of anti-HCV NS5-positive ( $n = 61$ ) and anti-HCV-negative ( $n = 40$ ) serum specimens. The anti-HCV-positive serum specimens were obtained from patients infected with HCV genotypes 1 to 6 (see Materials and Methods). The results of these experiments are presented in Table 3. The analysis of these results demonstrated that, in general, there is significant variation in immunoreactivity between different proteins. For example, proteins NS5A-2b and

TABLE 1. Locations of PCR and protein fragments within the HCV genome and sequences of primers

| Fragment no. | Location of:      |                       | Primers  |
|--------------|-------------------|-----------------------|--|
|              | PCR fragment (nt) | Protein fragment (aa) |  |
| 27           | 6246–6530         | 1973–2067             | 5'-TCCGGCTCGTGGCTAAAGGAT,<br>5'-CGTGGTGTATGCGTTGATGGCCTT |
| 28           | 6345–6647         | 2006–2106             | 5'-CTCCCTTTCCTGTCATGCCAA,<br>5'-GGTCATGCCCCGTCACGTAGTGG  |
| 29           | 6423–6764         | 2032–2145             | 5'-GGAGCACAGATCACCGGACAT,<br>5'-GACCTCCTCTCGTAGGAGAGG    |
| 30           | 6507–6812         | 2060–2161             | 5'-TTCCCCATCAACGCATACACC,<br>5'-TGGGAGCTGTGACCCGACCA     |
| 31           | 6582–6899         | 2085–2190             | 5'-GTGGCTGCTGAGGAGTACGTGGA,<br>5'-CAGCCTACGCTTGGCCGTCTC  |
| 32           | 6705–7016         | 2126–2229             | 5'-GATGGAGTACGGTTGCACAGG,<br>5'-GGCCTCGATGAGGTCAGCGTCC   |
| 33           | 6798–7109         | 2157–2260             | 5'-GGGTACAGCTCCCATGTGAG,<br>5'-CCGAATCGGATCGAAAGAGTCC    |
| 34           | 6876–7199         | 2183–2290             | 5'-GCAGAGACGGCCAAGCGTAGGCT,<br>5'-GCGTGCCCATATGGGCAACGCT |
| 35           | 6963–7268         | 2212–2313             | 5'-AAGGCGACATGTACTACCCATC,<br>5'-CCCGTGTACCACCGGGGGACGT  |
| 36           | 7047–7349         | 2240–2340             | 5'-AACATCACCCGTGTGGAGTCAGA,<br>5'-CACGGTGGACTCTGTCAGGAC  |
| 37           | 7140–7445         | 2271–2372             | 5'-CCGGCGGAGATCCTGCGAAAAC,<br>5'-GGCCTGATCGGGAGGGCCAGT   |
| 38           | 7197–7490         | 2290–2387             | 5'-CGCCCGGATTACAACCCTCC,<br>5'-GGAGGAGTACGACTCAACGTC     |
| 39           | 7287–7586         | 2320–2419             | 5'-AAGGCCCCCAATACCACCTC,<br>5'-GCAGCAGACGACGTCCTCACC     |
| 40           | 7377–7685         | 2350–2452             | 5'-AAGACCTTTGGCAGTCCGGGT,<br>5'-GTGACGCAGCAAAGAGTTGCT    |
| 41           | 7470–7769         | 2381–2480             | 5'-GACGTTGAGTCGTACTCCTCCA,<br>5'-GTCCAGGACTTGCAGTCTGTC   |
| 42           | 7566–7898         | 2413–2523             | 5'-GGTGAGGACGTCGTCGCTGC,<br>5'-CCCGTAGCCAAATTTGGATTT     |
| 43           | 7632–7934         | 2435–2535             | 5'-GCGGAGGAGAGCAAGTTGCCCA,<br>5'-GACGGCCCTGCTGGATAGGCT   |

NS5A-6a immunoreacted with only 16 to 18% of anti-HCV-positive sera tested, and protein NS5A-1a showed an intermediate immunoreactivity of 62%, whereas all the other proteins were immunoreactive with 77 to 93% of sera (Table 3). The other interesting observation is that not one protein demonstrated strict genotype-specific immunoreactivity. The only protein that demonstrated a significantly greater efficiency of binding antibodies from serum specimens from patients infected with the homologous HCV genotype was protein NS5A-2b (Table 3). Variations in immunoreactivity between proteins of different subtypes derived from same HCV genotype may be explained, to some extent, by an unequal representation of subtypes in serum specimens used in this study. However, the data presented in Table 3 demonstrate that proteins having lower immunoreactivity with serum specimens containing the homologous genotype are also less immunoreactive with specimens containing different HCV genotypes, and, therefore, are less immunoreactive in general.

Three proteins (NS5A-1b-2, NS5A-3a, and NS5A-4a) were found to be the most immunoreactive among all proteins tested in this study. These proteins immunoreacted with 92 to 93% of all anti-HCV-positive serum specimens. Protein NS5A-1b-2 was slightly more immunoreactive than the other two proteins. This protein not only immunoreacted with 93% of all sera (Table 3), but also immunoreacted with the highest S/C

value (data not shown). However, this protein did not detect HCV antibody in three serum specimens belonging to genotype 2 and one specimen belonging to genotype 6. Two of these genotype 2 specimens were immunoreactive only with the protein NS5a-2b. The third genotype 2 specimen was immunoreactive only with protein NS5A-1c or NS5A-4a. The genotype 6 specimens, which were not immunoreactive with protein NS5A-1b-2, immunoreacted with several proteins including proteins NS5A-3a and NS5A-4a. Therefore, the combination of three proteins—namely, NS5A-1b-2, NS5A-2b, and NS5a-3A or NS5A-4a—detected anti-HCV activity in all anti-NS5-positive serum specimens used in this study. This observation suggests a different strategy for choosing antigenic targets for test development. Instead of using different antigenic epitopes from one sequence variant, antigenic epitopes may be used from several different HCV sequence variants. However, it should be noted that by increasing the size of the panel, more protein variants may be required to detect all anti-NS5-positive serum specimens.

It is interesting that three proteins derived from the HCV subtype 1b demonstrated significant variations in immunoreactivity with serum specimens of genotype 5 (Table 3). Whereas protein NS5A-1b-2 detected HCV antibody in all genotype 5 specimens, protein NS5A-1b-3 did not immunoreact with any of these specimens, despite sharing almost 90% of

Genotype

| Consensus | KATCTTHHPH | --DSPDADLI | EANLLWROEM | GGNITRVESE | NKVWILDSFD | PLRAEEEDER | EVSVPAEILR | KSPRKFPFAL | PIWARPDYNP | PLLESWKDPD | YVPPVHHG |
|-----------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|----------|
| M58335    | ..V        | ..V        | ..V        | ..V        | ..V        | ..V        | ..K        | ..A        | ..M        | ..A        | ..M      |
| D1484     | ..I        | ..I        | ..E        | ..I        | ..I        | ..V        | ..R        | ..A        | ..M        | ..A        | ..M      |
| D90208    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D50480    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D85516    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D63857    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D89872    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| L02863    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| S66587    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D10934    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D50482    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| U45476    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D50485    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D50481    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| X61596    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| U16362    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D10750    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| AF034151  | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D50484    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| U01214    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D45172    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| AF011751  | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| AF001151  | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| M32084    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| AF011752  | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| M67463    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D10749    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D14893    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| U89019    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D50483    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D07579    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| M95104    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D10988    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| M95105    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D00944    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D50409    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| X76918    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D28917    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D17763    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D49374    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D63821    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| X11604    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| X1E184    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| D63822    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |
| X12083    | ..I        | ..I        | ..I        | ..I        | ..I        | ..V        | ..I        | ..A        | ..M        | ..A        | ..M      |

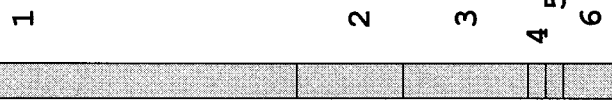


FIG. 1. Alignment of 45 sequences of the HCV NSSA major antigenic region at aa 2212 to 2313. Each sequence is identified by a GenBank accession number. The sequences used for constructing synthetic genes are shaded and provided with short identifiers.

TABLE 2. Similarities between NS5A regions (aa 2212 to 2313) used for constructing synthetic genes

| Protein | % Similarity with protein: |      |      |      |      |      |      |      |      |      |      |      |
|---------|----------------------------|------|------|------|------|------|------|------|------|------|------|------|
|         | 1b-1                       | 1b-2 | 1b-3 | 1c   | 2a   | 2b   | 3a   | 3b   | 4a   | 5a   | 11   | 6a   |
| 1a      | 75.5                       | 77.2 | 86.3 | 83.3 | 55.1 | 49   | 66.7 | 70.6 | 58.6 | 67.6 | 72.5 | 53.9 |
| 1b-1    |                            | 74.3 | 80.4 | 76.5 | 48   | 51   | 69.6 | 68.6 | 55.6 | 59.8 | 68   | 50.5 |
| 1b-2    |                            |      | 86.1 | 77.2 | 48   | 49   | 66.3 | 68.3 | 53.3 | 63.4 | 70.3 | 50.5 |
| 1b-3    |                            |      |      | 86.3 | 55.1 | 55.1 | 71.6 | 73.5 | 61.6 | 68.6 | 72.5 | 51   |
| 1c      |                            |      |      |      | 57.1 | 57.1 | 70.6 | 73.5 | 60.6 | 69.6 | 73.5 | 51   |
| 2a      |                            |      |      |      |      | 70.4 | 52   | 50   | 48   | 48   | 49   | 42.9 |
| 2b      |                            |      |      |      |      |      | 53.1 | 50   | 45.9 | 48   | 41.8 | 41.8 |
| 3a      |                            |      |      |      |      |      |      | 87.3 | 53.5 | 65.7 | 75.5 | 52.9 |
| 3b      |                            |      |      |      |      |      |      |      | 50.5 | 66.7 | 87.3 | 52   |
| 4a      |                            |      |      |      |      |      |      |      |      | 48.5 | 50.5 | 40.4 |
| 5a      |                            |      |      |      |      |      |      |      |      |      | 48.5 | 46.1 |
| 11      |                            |      |      |      |      |      |      |      |      |      |      | 46.1 |

its amino acid sequence. Thus, despite this significant similarity between these two proteins, they drastically differ in their reactivities with antibodies from genotype 5.

Another important observation is that some proteins could detect anti-HCV activity in all serum specimens of genotypes 1, 3, 4, 5, or 6, whereas none of the 13 proteins could detect HCV antibody in all serum specimens containing HCV genotype 2 (Table 3). The best immunoreactivity with genotype 2 sera was observed with protein NS5A-1b-2, which detected HCV antibody in 82% of genotype 2 serum specimens (Table 3). Also, some proteins do not always demonstrate broad and strong immunoreactivity with serum specimens from patients infected with the corresponding HCV genotype. For example, HCV genotype 2 serum specimens immunoreacted the most efficiently with a protein of HCV genotype 1b. Similarly, genotype 5 serum specimens were most immunoreactive with proteins of genotype 1, 4, and 6, and genotype 6 serum specimens were most immunoreactive with proteins of genotypes 1, 3, and 4 (Table 3).

**Immunoreactivity with RIBA anti-NS5-negative serum specimens.** A small panel of 30 RIBA-confirmed anti-HCV-positive serum specimens, which did not demonstrate any anti-NS5 activity by RIBA, was used to test the immunoreactivity of the 13 NS5 proteins. Ten of these serum specimens contained antibody immunoreactive with at least one of these recombinant proteins. Protein NS5A-1b-2 detected anti-NS5 activity in 7 out of the 30 serum specimens. Another three specimens, which were not immunoreactive with this protein, immunore-

acted with protein NS5A-2b. Two of these three specimens were immunoreactive with proteins NS5A-1b-1, NS5A-3b-1, and NS5A-4a. It is interesting that these three specimens, which were not detected by NS5A-1b-2, were obtained from patients infected with HCV genotypes 1 and 5. However, proteins immunoreactive with these three specimens were derived from different HCV genotypes. Therefore, the poor immunoreactivities of these serum specimens with several recombinant proteins cannot be convincingly explained by genotype specificity of antibodies.

DISCUSSION

Third-generation immunoassays for the detection of anti-HCV activity include almost the complete NS5 protein as one of four antigenic targets (16, 20). Different studies performed with these assays have shown that the NS5 antigen is not as immunoreactive as the other antigens (4, 15) and is potentially less specific than any of the other HCV antigens included in the third generation of commercially available diagnostic assays (11).

Regions of nonspecific antigenic reactivity can be found within different HCV proteins, including the NS5 protein (11). One study showed that more than 30% of the 543 synthetic peptides derived from different regions of the HCV polyproteins of different genotypes were recognized by antibodies from healthy blood donors (24). This observation suggests that potential nonspecific immunoreactivity of HCV antigenic

TABLE 3. Percent immunoreactivity with NS5A recombinant antigens derived from different HCV sequence variants of serum specimens obtained from patients infected with different genotypes

| Serum genotype | % Immunoreactivity with NS5A proteins of indicated genotype |      |      |      |     |    |    |     |     |     |     |     |    |
|----------------|---|------|------|------|-----|----|----|-----|-----|-----|-----|-----|----|
|                | 1   |      |      |      | 2   |    | 3  |     | 4   | 5   | 6   |     |    |
|                | 1a  | 1b-1 | 1b-2 | 1b-3 | 1c  | 2a | 2b | 3a  | 3b  | 4a  | 5a  | 11  | 6a |
| 1 (n = 17)     | 88  | 94   | 100  | 94   | 94  | 88 | 6  | 100 | 94  | 100 | 88  | 88  | 11 |
| 2 (n = 17)     | 10  | 65   | 82   | 71   | 59  | 71 | 52 | 76  | 71  | 65  | 59  | 53  | 18 |
| 3 (n = 16)     | 87  | 93   | 100  | 93   | 88  | 88 | 0  | 100 | 93  | 93  | 100 | 100 | 25 |
| 4 (n = 5)      | 80  | 100  | 100  | 100  | 100 | 80 | 0  | 100 | 100 | 100 | 100 | 100 | 20 |
| 5 (n = 3)      | 67  | 67   | 100  | 0    | 67  | 67 | 0  | 67  | 67  | 100 | 67  | 100 | 33 |
| 6 (n = 3)      | 33  | 67   | 67   | 67   | 100 | 67 | 0  | 100 | 100 | 100 | 67  | 67  | 0  |
| Total (n = 61) | 62  | 84   | 93   | 84   | 82  | 77 | 16 | 92  | 87  | 92  | 82  | 84  | 18 |

epitopes may exist. Therefore, the use of shortened antigens that contain a reduction in the number of antigenic epitopes, some of which may demonstrate cross-immunoreactivity with antibodies from individuals without HCV infection, should result in improving the specificity of diagnostic assays. Similarly, an improvement in the specificity of the C22 and C100 antigens included in third-generation HCV diagnostic assays was attributed to the replacement of recombinant proteins with shorter synthetic peptides. Previous studies (13, 19, 24, 25), as well as the data presented in this paper, demonstrate that the NS5 protein, especially the NS5A protein, contains several strong diagnostically relevant antigenic epitopes that can be efficiently modeled with short synthetic peptides or recombinant proteins. Therefore, it is conceivable that the specificity of anti-NS5 detection would be improved if short antigenically active regions of proteins were used as antigenic targets.

In the present study the NS5A immunodominant antigenic region was localized at aa 2212 to 2313 by using a set of recombinant overlapping proteins. This finding confirms a previous observation that the region at aa 2238 to 2320 contains a strongly immunoreactive antigenic epitope that can be functionally modeled with recombinant proteins (19). The primary structure of this NS5A antigenic region is very variable. It is conceivable that such significant sequence variability may significantly affect the antigenic properties of this region. In support of this concept other scientists have shown that the frequency of immunoreactivity of a similar NS5A region with type-homologous serum specimens was greater than with type-heterologous specimens (19). However, our study failed to strongly support this previous observation. The NS5A-2b protein was the only NS5A protein used in the present study that demonstrated better immunoreactivity with genotype 2 serum specimens than with other genotypes. Moreover, our data suggest the opposite effect (i.e., a genotype 6 protein weakly immunoreacted with serum specimens of genotype 1 to 5 and did not immunoreact with genotype 6 specimens). NS5A-1a is another example, although less striking, of some genotype dependence. In general, our data demonstrate that the NS5A major antigenic region does not have strict genotype-specific antigenic reactivity.

Despite the absence of strict genotype-specific antigenic reactivity with any of the NS5A proteins used in this study, sequence heterogeneity was clearly shown to have a prominent effect on the antigenic properties of these proteins. The most remarkable example of how small an amino acid change may affect the antigenic properties of proteins is the immunoreactivity of two proteins, NS5A-1b-2 and NS5A-1b-3. Despite a sequence homology of 86%, one protein immunoreacted with all genotype 5 specimens, whereas the other protein did not immunoreact with a single specimen of this genotype. These two proteins differ by only 12 aa positions.

In conclusion, the data obtained in the present study strongly suggest that sequence heterogeneity profoundly affects the antigenic properties of the NS5A immunodominant region. However, since different sequence variants derived from the same genotype, or even subtype, may display noticeably different rates of immunoreactivity, this effect is not genotype specific, but rather sequence specific. This observation is important when selecting antigenic targets for the develop-

ment of diagnostic tests, suggesting that antigenic targets should be carefully selected from different sequence variants based on breadth and strength of immunoreactivity. The data show that when carefully selected, the NS5A antigen may be used to detect anti-HCV activities with almost equal efficiencies from serum specimens obtained from patients infected with different HCV genotypes obtained from different parts of the world.

#### REFERENCES

1. Aach, A. D., C. E. Stevens, F. B. Hollinger, J. W. Mosley, D. A. Peterson, P. E. Taylor, R. G. Johnson, L. H. Barbosa, and G. J. Nemo. 1991. Hepatitis C virus infection in post-transfusion hepatitis. An analysis with first- and second-generation assays. *N. Engl. J. Med.* **325**:1325-1329.
2. Alter, M. J. 1993. The detection, transmission, and outcome of hepatitis C virus infection. *Infect. Agents Dis.* **2**:155-166.
3. Barrera, J. M., M. Bruguera, M. G. Ercilla, J. M. Sanchez-Tapia, M. P. Gil, J. Costal, A. Gelabert, J. Rodes, and R. Gasto. 1991. Incidence of non-A, non-B hepatitis after screening blood donors for antibodies to hepatitis C virus and surrogate markers. *Ann. Intern. Med.* **115**:596-600.
4. Barrera, J. M., B. Francis, M. G. Ercilla, M. Nelles, D. Achord, J. Darner, and S. R. Lee. 1995. Improved detection of anti-HCV in post-transfusion hepatitis by a third generation ELISA. *Vox Sang.* **68**:15-18.
5. Chang, J. C., C. Seidel, B. Ofenloch, D. L. Jue, H. A. Fields, and Y. E. Khudyakov. 1999. Antigenic heterogeneity of the hepatitis C virus NS4 protein as modeled with synthetic peptides. *Virology* **257**:177-190.
6. Choo, Q. L., G. Kuo, A. J. Weiner, L. R. Overby, D. W. Bradley, and M. Houghton. 1989. Isolation of cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* **244**:359-362.
7. Choo, Q. L., K. H. Richman, J. H. Han, K. Berger, C. Lee, C. Dong, C. Gallegos, D. Coit, A. Medina-Selby, P. J. Barr, A. J. Weiner, D. W. Bradley, G. Kuo, and M. Houghton. 1991. Genetic organization and diversity of hepatitis C virus. *Proc. Natl. Acad. Sci. USA* **88**:2451-2455.
8. Dhaliwal, S. K., L. E. Prescott, B. C. Dow, F. Davidson, H. Brown, P. L. Yap, E. A. C. Follett, and P. Simmonds. 1996. Influence of viraemia and genotype upon serological reactivity in screening assay for antibody to hepatitis C virus. *J. Med. Virol.* **48**:184-190.
9. Dow, B. C., E. C. A. Follett, H. Munro, I. Buchanan, K. Roy, F. McOmish, P. L. Yap, and P. Simmonds. 1994. Failure of 2nd- and 3rd-generation HCV ELISA and RIBA to detect HCV polymerase chain reaction-positive donations. *Vox Sang.* **67**:236-237.
10. Dow, B. C., H. Munro, I. Buchanan, E. C. A. Follett, F. Davidson, P. L. Yap, and P. Simmonds. 1996. Third-generation recombinant immunoblot assay: comparison of reactivity according to hepatitis C virus genotype. *Transfusion* **36**:547-551.
11. Dow, B. C., I. Buchanan, H. Munro, E. A. C. Follett, F. Davidson, L. E. Prescott, P. L. Yap, and P. Simmonds. 1996. Relevance of RIBA-3 supplementary test to HCV PCR positive and genotypes for HCV confirmation of blood donors. *J. Med. Virol.* **49**:132-136.
12. Gretch, D. R. 1997. Diagnostic test for hepatitis C. *Hepatology* **26**(Suppl. 1):43s-47s.
13. Khudyakov, Y. E., N. S. Khudyakova, D. L. Jue, S. B. Lambert, S. Fang, and H. A. Fields. 1995. Linear B-cell epitopes of the NS3-NS4-NS5 proteins of the hepatitis C virus as modeled with synthetic peptides. *Virology* **206**:666-672.
14. Kolykhalov, A. A., E. V. Agapov, and C. M. Rice. 1994. Specificity of the hepatitis C virus NS3 serine protease: effects of substitutions at the 3/4A, 4A/4B, 4B/5A, and 5A/5B cleavage sites on polyprotein processing. *J. Virol.* **68**:7525-7533.
15. Lee, S. R., C. L. Wood, M. J. Lane, B. Francis, C. Gust, C. M. Higgs, M. J. Nelles, A. Polito, R. DiNello, and D. Achord. 1995. Increased detection of hepatitis C virus infection in commercial plasma donors by a third-generation screening assay. *Transfusion* **35**:845-849.
16. Lok, A. S. F., and N. T. Gunaratnam. 1997. Diagnosis of hepatitis C. *Hepatology* **26**(Suppl.):48s-56s.
17. McOmish, F., S. W. Chan, B. C. Dow, J. Gillon, W. D. Frame, R. J. Crawford, P. L. Yap, E. A. C. Follett, and P. Simmonds. 1993. Detection of three types of hepatitis C virus in blood donors: investigation of type-specific differences in serologic reactivity and rate of alanine aminotransferase abnormalities. *Transfusion* **33**:7-13.
18. Mellor, J., E. C. Holmes, L. M. Jarvis, P. L. Yap, P. Simmonds, and the International HCV Collaborative Study Group. 1995. Investigation of the pattern of hepatitis C virus sequence diversity in different geographical regions: implications for virus classification. *J. Gen. Virol.* **76**:2493-2507.
19. Neville, J. A., L. E. Prescott, V. Bhattacharjee, N. Adams, I. Pike, B. Rodgers, A. El-Zayadi, S. Hamid, G. M. Dushiko, A. A. Saeed, G. H. Haydon, and P. Simmonds. 1997. Antigenic variation of core, NS3, and NS5 proteins among genotypes of hepatitis C virus. *J. Clin. Microbiol.* **35**:3062-3070.
20. Pawlotsky, J. M., A. Bastie, C. Pellet, J. Remire, F. Darthuy, L. Wolfe, C.

- Sayada, J. Duval, and D. Dhumeaux. 1996. Significance of indeterminate third-generation hepatitis C virus recombinant immunoblot assay. *J. Clin. Microbiol.* **34**:80–83.
21. Pereboeva, L. A., A. V. Pereboev, and G. E. Morris. 1998. Identification of antigenic sites on three hepatitis C virus proteins using phage-displayed peptide libraries. *J. Med. Virol.* **56**:105–111.
22. Pereboeva, L. A., A. V. Pereboev, L. E. Wang, and G. E. Morris. 2000. Hepatitis C epitopes from phage-displayed cDNA libraries and improved diagnosis with a chimeric antigen. *J. Med. Virol.* **60**:144–151.
23. Purcell, R. 1997. The hepatitis C virus: overview. *Hepatology.* **26**(Suppl.): 11s–14s.
24. Rodríguez-López, M., J. I. Riezu-Boj, M. Ruiz, C. Berasain, M. P. Civeira, J. Prieto, and F. Borrás-Cuesta. 1999. Immunogenicity of variable regions of hepatitis C virus proteins: selection and modification of peptide epitopes to assess hepatitis C virus genotypes by ELISA. *J. Gen. Virol.* **80**:727–738.
25. Rosa, C., S. Osborne, F. Garetto, S. Griva, A. Rivella, G. Calabresi, R. Guaschino, and F. Bonelli. 1995. Epitope mapping of the NS4 and NS5 gene products of hepatitis C virus and the use of a chimeric NS4-NS5 synthetic peptide for serodiagnosis. *J. Virol. Methods* **55**:219–232.
26. Simmonds, P., K. A. Rose, S. Graham, S. W. Chan, F. McOmish, B. C. Dow, E. A. C. Follett, P. L. Yap, and H. Marsden. 1993. Mapping of serotype-specific, immunodominant epitopes in the NS-4 region of hepatitis C virus (HCV): use of type-specific peptides to serologically differentiate infections with HCV types 1, 2, and 3. *J. Clin. Microbiol.* **31**:1493–1503.
27. Vernelen, K., H. Claeys, H. Verhaert, A. Volckaerts, and C. Vermylen. 1994. Significance of NS3 and NS5 antigens in screening for HCV antibody. *Lancet* **343**:853–854.
28. Yuki, N., N. Hayashi, A. Kasahara, H. Hagiwara, K. Ueda, H. Fusamoto, and T. Kamada. 1992. Antibodies to a putative hepatitis C virus polyprotein in Japanese patients with chronic liver disease. *J. Med. Virol.* **38**:86–91.
29. Zhang, Z. X., M. Chen, A. Sonnerborg, and M. Sallberg. 1994. Antigenic structure of the complete nonstructural (NS2) and 5 proteins of hepatitis C virus (HCV) anti-HCV NS2 and NS5 antibody in relation to HCV serotype, presence of HCV RNA, and acute HCV infection. *Clin. Diagn. Lab. Immunol.* **1**:290–294.