

# Novel Nuclear Autoantigen with Splicing Factor Motifs Identified with Antibody from Hepatocellular Carcinoma

Haruhiko Imai,\* Edward K. L. Chan,\* Kendo Kiyosawa,† Xiang-Dong Fu,‡ and Eng M. Tan\*

\*W. M. Keck Autoimmune Disease Center and DNA Core Facility, Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, California 92037; †Second Department of Internal Medicine, Shinshu University School of Medicine, Matsumoto, 390, Japan; and ‡Division of Cellular and Molecular Medicine, University of California, San Diego, California 92193-0651

## Abstract

A patient with liver cirrhosis who progressed to hepatocellular carcinoma was found to develop novel antinuclear antibodies. The serum was used to isolate full-length cDNA clones encoding related proteins of 530 amino acids (representative clone HCC1.4) and 524 amino acids (representative clone HCC1.3). Affinity-purified antibodies eluted from recombinant proteins recognized a 64-kD nuclear protein in Western blotting and decorated the nucleoplasm in a speckled-network fashion in immunofluorescence, colocalizing with antibodies to pre-mRNA splicing factor SC35 and uridine-rich small nuclear RNAs. The deduced amino acid sequence contained an arginine/serine-rich (RS) domain and three-ribonucleoprotein consensus sequence domains, two classes of motifs present in several splicing factors. A repeating octapeptide of Arg-Ser-Arg-Ser-Arg(Lys)-Glu(Asp)-Arg-Lys(Arg) was present in RS region of HCC1. This octapeptide sequence called RS-ERK motif was also found in splicing factors U2AF 35- and 65-kD proteins and 70-kD U1 small nuclear ribonucleoprotein. The molecular features and immunolocalization data suggest that the HCC1 autoantigen may be associated with splicing activities and are consistent with observations that autoantibody responses frequently target molecules involved in important cellular biosynthetic functions. (*J. Clin. Invest.* 1993. 92:2419–2426.) **Key words:** autoantibodies • molecular cloning • RNA-binding proteins • hepatitis • neoplasms

## Introduction

Autoantibodies that are highly specific in their reaction with intracellular antigens are present in the sera of patients with certain diseases such as systemic lupus erythematosus and scleroderma (for review, see reference 1). These naturally occurring antibodies have been helpful in isolating many intracellular molecules engaged in biosynthetic or proliferative cell functions such as transcription, splicing of pre-mRNA, DNA replication, and cell division. A special property of autoantibodies in contrast with experimentally induced antibodies is that the former often recognize antigenic determinants that corre-

spond to functional sites, and have, therefore, been useful in elucidating functional mechanisms, such as the role of proliferating cell nuclear antigen in DNA replication (2) and DNA repair (3), and the role of Ku antigen in a DNA-dependent protein kinase reaction (4).

Autoantibodies to intracellular components are not restricted only to systemic autoimmune diseases but are also found in certain other diseases, usually in lower frequency. Recently, we reported that antinuclear antibodies (ANAs)<sup>1</sup> were detected in 31% of patients with hepatocellular carcinoma (HCC) (5). Three antibody specificities were identified in HCC and were antibodies directed against fibrillarin, a 34-kD protein involved in processing of nascent RNA polymerase I transcripts (6), NOR-90/hUBF, a protein doublet of 89 and 93 kD that are RNA polymerase I transcription factors (7, 8), and nucleophosmin/protein B23, a nucleolar protein associated with mitogen-induced cell proliferation (9). Many other autoantibody specificities in HCC sera remained unidentified.

HCC is a liver malignancy that develops from preexisting chronic liver disease, which in some regions of the world is predominantly caused by viral hepatitis (10). Patients with such forms of chronic hepatitis may develop HCC after many years and are followed at frequent intervals with the hope that malignant transformation can be detected early. We have observed several patients in transition from chronic liver disease to HCC who either developed ANAs de novo or showed increasing titers of ANAs and/or altered specificities in ANA reactivity (11). These changes in immune response during malignant transformation are consistent with the notions that autoantibody responses might be antigen-driven and that the responsible antigens might be involved in important cellular functions (12).

In this study, serum from a patient who developed novel ANAs during progression from chronic liver disease to HCC was used to isolate cDNA clones encoding a nuclear protein called HCC1. This novel nuclear protein contains RNA-binding domains and a long region rich in arginine and serine (RS motif). These sequence features of HCC1 are conserved in a class of pre-mRNA splicing factors. Consistent with the possibility that HCC1 might be a splicing factor, immunofluorescence microscopy using affinity-purified antibodies showed that HCC1 protein was distributed in a speckled network pattern in the nucleus and colocalized with uridine-rich small nuclear ribonucleoprotein (U-snRNP) and non-snRNP spliceosome components.

1. *Abbreviations used in this paper:* ANA, antinuclear antibody; HCC, hepatocellular carcinoma; m<sub>3</sub>G, 2,2,7-trimethylguanosine; NK, patient with liver cirrhosis who developed HCC; RNP-CS, ribonucleoprotein consensus sequence; RS, arginine/serine-rich; U-snRNP, uridine-rich small nuclear ribonucleoprotein; U2AF<sup>65</sup> and U2AF<sup>35</sup>, U2 snRNP auxiliary factor 65 and 35 kD protein, respectively.

Address correspondence to Eng M. Tan, M.D., W. M. Keck Autoimmune Disease Center, The Scripps Research Institute, La Jolla, CA 92037.

Received for publication 23 March 1993 and in revised form 9 June 1993.

J. Clin. Invest.

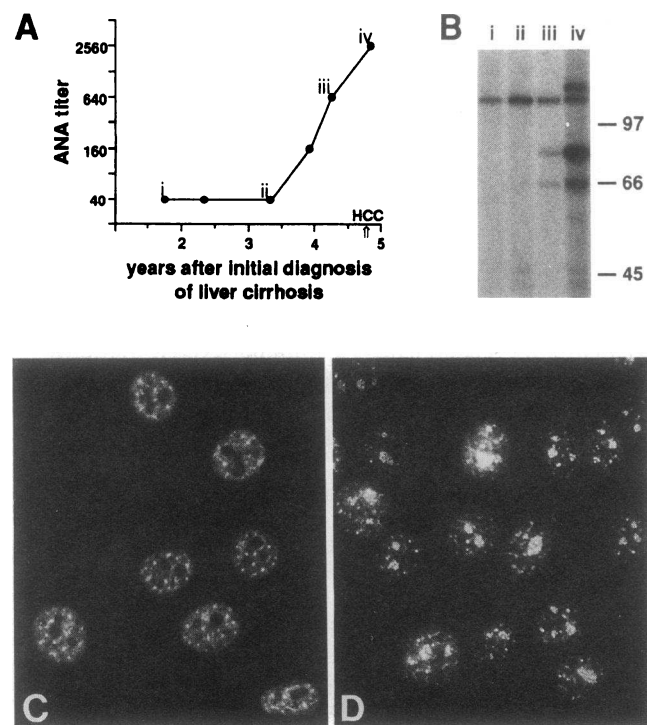
© The American Society for Clinical Investigation, Inc.  
0021-9738/93/11/2419/08 \$2.00

Volume 92, November 1993, 2419–2426

## Methods

**Antibody.** Patient NK had hepatitis B-related liver cirrhosis and developed HCC several years after the diagnosis of liver cirrhosis. Studies of sequential sera showed new specificities of ANA associated with increasing titers several months before diagnosis of HCC (Fig. 1). Prototype autoimmune sera from other diseases were obtained from the W. M. Keck Autoimmune Disease Center laboratory serum bank. Normal human sera were collected from laboratory personnel. Mouse mAb to 2,2,7-trimethylguanosine (m<sub>3</sub>G) (clone K121) (13) was obtained from Oncogene Science (Uniondale, NY). Mouse mAb to non-snRNP splicing factor SC35 was described previously (14). Affinity-purified antibodies to recombinant λZap phage plaques were prepared from NK serum obtained upon diagnosis of HCC by incubating with recombinant phage protein bound to nitrocellulose filters and subsequently eluted at pH 2.5 as described (15). Antibodies to the 64-kD antigen in Western blots were also affinity purified from NK serum as described previously (5).

**Immunofluorescence microscopy.** HEp-2 and HeLa cells were grown on coverslips, fixed for 5 min at -20°C in 100% methanol, and permeabilized for 3 min at -20°C in 100% acetone. Human sera were used at 1:50 dilution or higher. FITC-conjugated goat anti-human IgG (Caltag Laboratories, San Francisco, CA) was used as secondary detecting reagent. For double immunofluorescence, mAb to m<sub>3</sub>G was used at 1:100 dilution, and FITC-conjugated goat anti-human IgG (mouse immunoglobulin adsorbed) and rhodamine-conjugated goat anti-mouse IgG (human immunoglobulin adsorbed) (Caltag Labora-



**Figure 1.** Increasing titers of ANAs associated with new antibody specificities observed in patient NK who developed HCC. (A) ANA patterns changed from weak generalized nucleoplasmic staining (points *i* and *ii*, data not shown) to strong nuclear speckles excluding nucleoli (*iii*, C) and finally to nuclear and nucleolar staining (*iv*, D). Arrow indicates the time point when the diagnosis of HCC was made. (B) Western blotting studies of nuclear extracts from HEp-2 cells were performed using sera collected at time points indicated by *i*, *ii*, *iii*, and *iv*. These studies confirmed the appearance of new antibody specificities in lanes *iii* and *iv* that were not present before the rise in ANA titer. Immunofluorescence on HeLa cells represented sera collected at time points *iii* (C) and *iv* (D).  $\times 500$ .

tories) were used as secondary detecting reagents. Culture supernatant of hybridomas producing mAb to SC35 was used to detect SC35. Nucleic acid digestion of cell substrates was performed in either PBS containing 100  $\mu$ g/ml RNase A (Sigma Immunochemicals, St. Louis, MO) or PBS containing 5 mM MgCl<sub>2</sub> and 100  $\mu$ g/ml DNase I (Sigma Immunochemicals) at 25°C for 2 h. Human autoantibodies to DNA or U1-snRNP were used to monitor the efficiency of DNase I and RNase A digestions, respectively.

**Western blotting.** Nuclear extracts from HEp-2 cells were prepared as described (5), separated by SDS-PAGE, and analyzed by Western blotting using sera diluted 1:100 and <sup>125</sup>I-labeled protein A (ICN Pharmaceuticals, Inc., Irvine, CA) as the detecting reagent. Whole cell extracts from HeLa, HepG2, MOLT-4, rat Novikoff hepatoma cell, and mouse 3T3 were also analyzed for the presence of HCC1 antigen.

**cDNA cloning and sequence analysis.** NK serum obtained at diagnosis of HCC was used for immunoscreening of  $7 \times 10^5$  recombinants from a HepG2 λZap cDNA library (8, 15). Before screening the cDNA library, the serum was extensively adsorbed against bacteria infected with wild type λZap phage. Six clones, λHCC1.1–λHCC1.6, were isolated and subcloned *in vivo* into pBluescript plasmid using R408 helper phage, as recommended in the manufacturer's instructions (Stratagene Inc., La Jolla, CA). DNA sequences were determined in both strands using dye terminator sequencing and an automated sequencer from Applied Biosystems Inc. (Foster City, CA). DNA sequences were aligned and compiled using the SeqEd 675 sequence editor (Applied Biosystems Inc.). DNA and protein sequences were analyzed by the Genetics Computer Group (GCG, University of Madison, Madison, WI) Sequence Analysis Software Package version 7.1-UNIX. U2 snRNP auxiliary factor 65-kD protein (U2AF<sup>65</sup>) cDNA was kindly provided by Dr. M. R. Green (University of Massachusetts Medical Center, Worcester, MA) (16). Isolation of SC35 cDNA was described previously (17).

**In vitro transcription and translation.** Plasmid clones pHCC1.3 and pHCC1.4 were used as DNA templates. RNAs were transcribed *in vitro* from linearized plasmids using T3 RNA polymerase and were translated *in vitro* in rabbit reticulocyte lysate or wheat germ extract in the presence of [<sup>35</sup>S]methionine (Tran<sup>35</sup>S-label; ICN), as described in the manufacturer's instructions (Promega Biotec, Madison, WI).

**Purification of recombinant protein.** Recombinant proteins were prepared from plasmid pHCC1.3 or pHCC1.4 transformed in *Escherichia coli* strain XL1-Blue (Stratagene Inc.). Cultures were grown to OD<sub>600</sub> = 0.6 and induced with isopropyl-β-thiogalactopyranoside (Sigma Immunochemicals) added to a final concentration of 10 mM. The culture was further grown overnight at 37°C before harvest. The inclusion bodies were purified according to the method described previously (18), and aliquots were stored at -70°C. For rabbit immunization, antigens were further purified by preparative SDS-PAGE.

**Rabbit immunization.** Two female New Zealand White rabbits were immunized by subcutaneous injection of 1 mg of pHCC1.3 or pHCC1.4 recombinant protein in complete Freund's adjuvant. Rabbits were boosted twice with a further 1 mg recombinant protein in incomplete Freund's adjuvant at 30-d intervals.

**Antigen depletion.** Antigen extract was prepared by incubating HEp-2 cells in lysis buffer consisting of 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 1% NP-40, 0.1% SDS, 0.5% sodium deoxycholate, 1 mM PMSF, 1  $\mu$ g/ml leupeptin, and 1  $\mu$ g/ml aprotinin. 20  $\mu$ l of extract (~40  $\mu$ g of protein) was mixed with 10  $\mu$ l of NK serum complexed to protein A-Sepharose (Pharmacia, Uppsala, Sweden), and the mixture was allowed to react for 1 h at 4°C in NET2 buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 0.1% SDS, 0.5% sodium deoxycholate). The supernatant was transferred to new tubes and proteins were precipitated by adding trichloroacetic acid to a final concentration of 10%. Protein pellets were rinsed with acetone and solubilized in SDS sample buffer. Protein samples were separated by SDS-PAGE and transferred to nitrocellulose. The membrane was probed with rabbit antiserum and bound antibodies were detected using <sup>125</sup>I-labeled protein A (ICN). Conversely, the same amount of extracts were depleted with rabbit antiserum or preimmune rabbit serum and the resultant depleted extracts were used in Western



pBluescript plasmid. These plasmids were designated as pHCC1.1 to pHCC1.6 with insert sizes of 1.6, 2.2, 2.6, 2.5, 1.9, and 1.5 kb, respectively (Fig. 2A). In Western blotting analysis, recombinant proteins from the six pHCC1 clones were reactive with the original NK serum, but were not reactive with either normal human sera or unrelated autoimmune sera (data not shown).

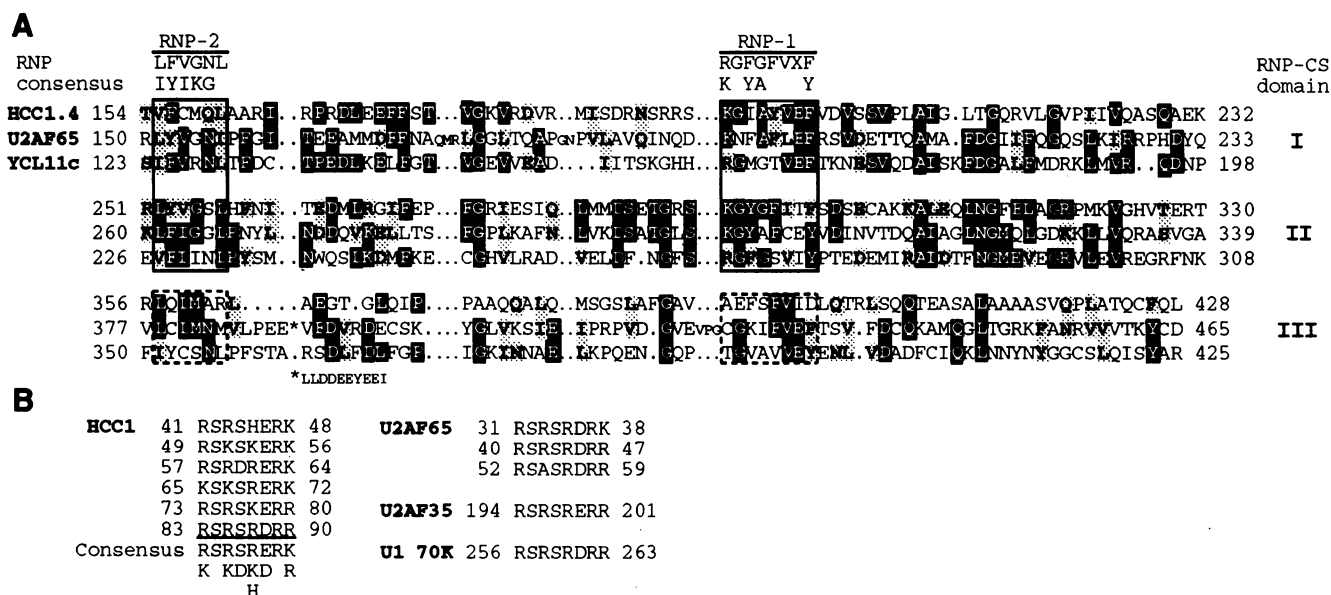
**Sequence analysis.** All six cDNA clones were sequenced completely in both strands using custom-designed primers. These cDNAs were overlapping clones, three of which had an 18-nucleotide deletion (Fig. 2A). There was only one uninterrupted open reading frame and the complete cDNA and deduced amino acid sequences of HCC1.3 (524 amino acids) and HCC1.4 (530 amino acids) are shown in Fig. 2B. Both cDNA sequences were identical, except that HCC1.3 had an 18-nucleotide (six-amino acid) deletion (boxed in Fig. 2B) and a longer 5'-untranslated sequence. The first ATG codons with flanking sequences in good agreement with the Kozak consensus (21) were designated as the methionine initiation codons. There were three polyadenylation signals (AATAAA) in the 3'-untranslated sequences, two of which were shown to be used in the isolated clones (Fig. 2).

HCC1 sequences were not found in GenBank, EMBL, and SwissProt data banks. Two motifs were found in the deduced amino acid sequences, one of which was the arginine/serine-rich (RS) motif located in the NH<sub>2</sub>-terminal region. There were 15 arginine/serine or serine/arginine dipeptides present between residues 31 and 127. The other located in the central region of the protein were those ribonucleoprotein consensus sequence (RNP-CS) motifs (22). The first two displayed good agreement with the RNP-CS each containing a highly conserved octamer RNP-1 and a less well-conserved hexamer

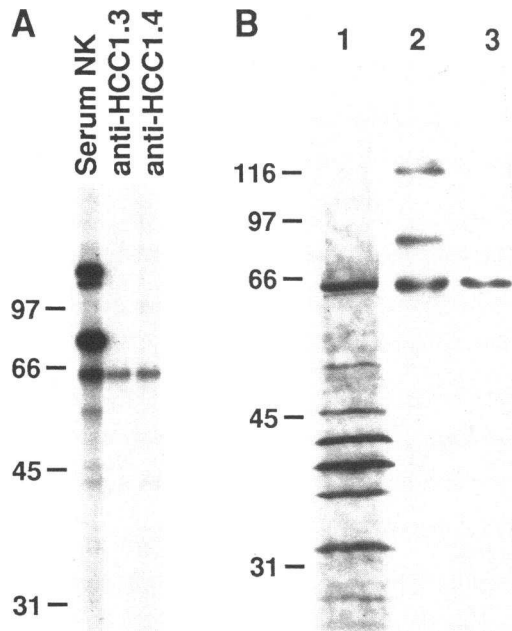
RNP-2, while the third was weakly homologous to the consensus (Fig. 3A). Since the overall domain structure resembled that reported for the U2AF<sup>65</sup> (16), the HCC1 sequence was compared to that of U2AF<sup>65</sup> and the YCL11c gene product that was recently reported as a putative homologue of U2AF<sup>65</sup> in *Saccharomyces cerevisiae* (23). The alignment of the three sequences revealed that conserved stretches were present in the RS region (data not shown) and in the first two RNP-CS domains, but the sequence homology in the third putative RNP-CS domain was less strong (Fig. 3A). Percent identity of amino acid sequences in these proteins were found to be 28% between HCC1.4 and U2AF<sup>65</sup>, and 22% between HCC1.4 and YCL11c by the GCG program BESTFIT (gap weight 3.0, length weight 0.1). Interestingly, the 18-nucleotide deletion in HCC1.3 was located in the RNP-1-like region of the putative third RNP-CS domain.

Within the RS domain, six contiguous direct repeats of an octapeptide sequence were found (Fig. 2B, Fig. 3B). This octapeptide was designated RS-ERK motif and consisted of two repeats of a basic residue (arginine or lysine), a serine followed by a basic residue (arginine, lysine, or histidine), an acidic residue (glutamic acid or aspartic acid), and two basic residues (arginine or lysine). The RS-ERK motif was determined in this analysis to occur three times in the RS region of U2AF<sup>65</sup> (16), once in both the U2AF 35-kD protein (U2AF<sup>35</sup>) (24) and the U1 snRNP 70-kD protein (25) (Fig. 3B).

**Characterization of HCC1 proteins.** Recombinant proteins derived from the two longest clones, HCC1.3 and HCC1.4, were used to affinity-purify antibodies from NK serum. In Western blotting, NK serum reacted with several bands, while affinity-purified antibodies recognized only the 64-kD band (Fig. 4A). The predicted molecular mass from the deduced



**Figure 3.** RNP-CS domains and RS-ERK motif in HCC1. (A) Alignment of HCC1.4, U2AF<sup>65</sup>, and YCL11c with respect to RNP-CS domains. Alignment of sequences was initially made with the GCG programs BESTFIT and PILEUP. Further refinements in alignments were performed manually. Black boxes represent identical matches and stippled boxes represent conservative changes (V=L=I, F=Y=W, R=K, Q=N, D=E, S=T). RNP-1 and RNP-2 motifs in the first two RNP-CS domains are boxed. The broken line in the putative third RNP-CS indicates a weaker match for the RNP-1 and -2 motifs. The third RNP-2 position for U2AF<sup>65</sup> (residues 378–383) is indicated based on the data reported previously (23). (B) Six octapeptide repeats in HCC1. The consensus sequence for the octapeptide consists of Arg-Ser-Arg-Ser-Arg(Lys)-Glu(Asp)-Arg-Lys(Arg) called the RS-ERK motif. Using the GCG program WORDSEARCH, octapeptide sequences homologous with the RS-ERK motif were found in U2AF<sup>65</sup>, U2AF<sup>35</sup>, and U1 snRNP 70-kD protein (U1 70K).



**Figure 4.** Electrophoretic properties of HCC1. (A) In Western blotting analysis using whole nuclear extracts of HEP-2 cells, several bands were recognized by NK serum. Affinity-purified antibodies from HCC1.3 or HCC1.4  $\lambda$ Zap phage plaques recognized exclusively a 64-kD band. (B) Comparison of the *in vitro* transcription and translation product of HCC1.4 with the 64-kD cellular protein in HEP-2 cells. The *in vitro* translated products from a HCC1.4 cDNA template and nuclear extracts from HEP-2 cells were fractionated on 10% SDS-polyacrylamide gel and transferred to nitrocellulose. The strips of nitrocellulose containing HEP-2 antigens were subsequently processed for Western blotting using NK serum (lane 2) and rabbit antibody to pHCC1 recombinant protein (lane 3). *In vitro* translated products were detected by autoradiography (lane 1).

amino acid sequences of HCC1.3 and HCC1.4 were 58.6 and 59.3 kD, respectively. *In vitro* transcribed and translated products of HCC1.3 and HCC1.4 using rabbit reticulocyte lysate showed that the largest polypeptide migrated at 64 kD (Fig. 4 B, lane 1). Similar products were observed with wheat germ extract (data not shown). Furthermore, the 64-kD *in vitro* translated product was shown to comigrate with the 64-kD antigen in HEP-2 cells detected by Western blotting using NK serum (Fig. 4 B, lane 2) and by rabbit antisera raised against pHCC1 recombinant protein (Fig. 4 B, lane 3).

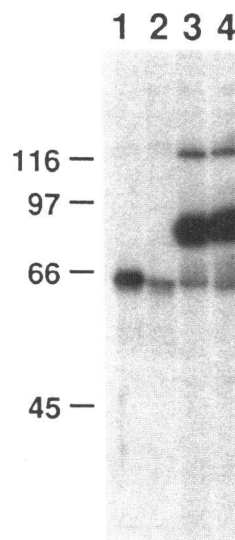
Depletion experiments were performed to determine whether the antigen recognized by rabbit antiserum was identical to that recognized by NK serum. The NK serum (at time point depicted in Fig. 1 B, lane *iv*), which contained antibody to the 64-kD antigen and NK serum obtained before detection of HCC and lacking this antibody (Fig. 1 B, lane *ii*) were used to deplete HEP-2 cell extracts. The resultant depleted extracts were analyzed in Western blots with rabbit antiserum as shown in Fig. 5, lanes 1 and 2. There was marked reduction of the 64-kD antigen recognized by rabbit antiserum in extracts depleted with NK serum containing the cognate antibody (lane 2). The converse experiment in which HEP-2 cell extracts were depleted with rabbit antiserum and the depleted extracts subsequently immunoblotted with NK serum is shown in Fig. 5, lanes 3 and 4. Rabbit antiserum was ineffective in depleting the 64-kD protein, a feature that was predictable on the basis of

other data which suggested that the rabbit antibody was highly reactive with antigen in Western blots but poorly reactive with the antigen in solution, possibly related to the fact that rabbit antibody was more reactive with denatured than native antigen. In further experiments, affinity-purified antibodies from the 64-kD band not only showed specific reactivity to the 64-kD band alone in HEP-2 nuclear extracts but also recognized the recombinant HCC1 protein in Western blotting (data not shown). These results further support the authenticity of HCC1 protein.

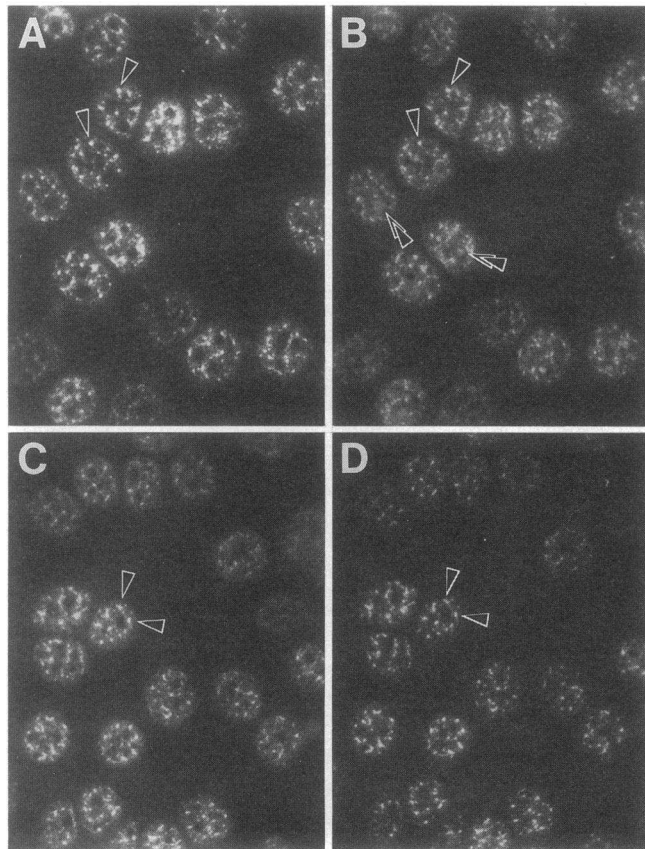
In immunofluorescence microscopy, NK serum used for cDNA screening showed staining of nuclear speckles and nucleoli (Fig. 1 D), whereas affinity-purified antibodies showed staining of nuclear speckles and interconnecting strands in the nucleoplasm, but no nucleolar staining (Fig. 6, A and C). Since the 64-kD antigen contained motifs characteristic of splicing factors, double immunofluorescence was performed using affinity-purified antibodies and mAb to  $m_3G$ , a common 5' cap structure of U-snRNAs (13), and mAb to the non-snRNP splicing factor SC35 (14). All three proteins colocalized in speckled regions in the nucleoplasm of HeLa cells (Fig. 6). The mAb to  $m_3G$  also showed weak nucleolar staining (Fig. 6 B, indicated by double arrowheads). Of these three antigens, SC35 showed greater preferential localization to discrete regions (Fig. 6, D). When HeLa cells were digested with either RNase A or DNase I before immunolocalization of the 64-kD antigen with affinity-purified antibodies, RNase A digestion greatly reduced fluorescence intensity of nuclear speckles, but DNase I treatment did not affect the staining (data not shown). These observations suggest that HCC1 protein may be associated with RNA in a similar manner to U-snRNPs but dissimilar from SC35, since the immunolocalization of the SC35 is reported to be resistant to RNase A digestion (26).

**HCC1 protein and mRNA expression in cell lines and tissues.** Rabbit antibody raised against pHCC1 recombinant protein detected a 64-kD band in Western blots with extracts from three human epithelial cancer cell lines, a human T cell leukemia, a rat hepatoma, and a mouse fibroblast cell line (Fig. 7 A). The signal intensities varied although equal number of cells were used in each lane.

Using full-length HCC1.4 cDNA as a probe, two major messages of 2.3 and 3.2 kb, and minor messages of 4.3 and 5.0



**Figure 5.** Depletion of the 64-kD antigen from HEP-2 cell extracts. Extracts were depleted with NK serum with (lane 2) or without (lane 1) antibody to the 64-kD antigen followed by analysis of the resultant depleted extracts in Western blots with rabbit antiserum. Conversely, the extracts were depleted with preimmune rabbit serum (lane 3) or postimmune rabbit antiserum (lane 4), and the depleted extracts were immunoblotted with NK serum which contained antibody to the 64-kD antigen.



**Figure 6.** Colocalization of HCC1 with snRNPs, and non-snRNP splicing factor SC35. HeLa cells were double labeled with affinity-purified antibodies to HCC1.4 (A and C) and mAb to m<sub>3</sub>G (B), or to SC35 (D). Arrowheads point to some colocalized speckles. Double arrowheads indicate nucleolar staining with mAb to m<sub>3</sub>G.  $\times 500$ .

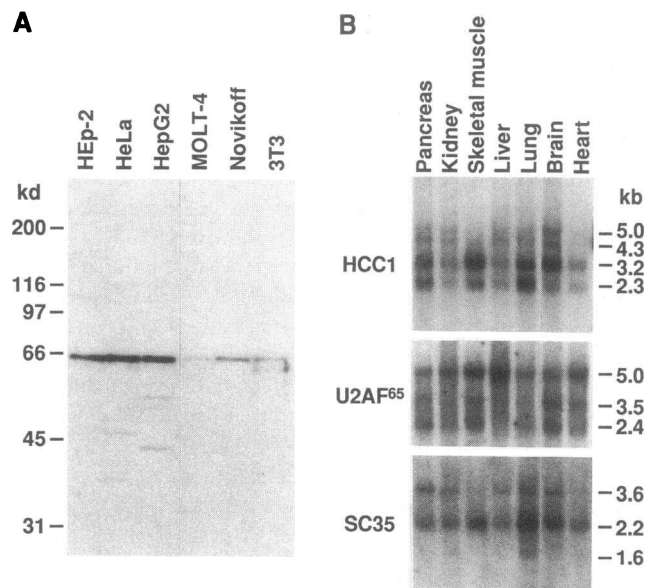
kb were detected in Northern blotting using either total RNA extracted from human cancer cell lines (data not shown) or poly(A)<sup>+</sup> RNA extracted from normal human tissues (Fig. 7B). To compare with the expression of splicing factors, the same RNA blots were subsequently reprobbed with cDNA probes of SC35 and U2AF<sup>65</sup> (Fig. 7B). Consistent with reported findings in cell lines (17), three mRNA species for SC35 were detected in normal human tissues. Three mRNA species for U2AF<sup>65</sup> were also detected in normal human tissues, although a 5-kb mRNA was not reported previously in HeLa cells (16). Expression of HCC1 mRNA appeared to be substantially higher in pancreas, skeletal muscle, lung, and brain than in kidney, liver, and heart. The 2.4- and 3.5-kb mRNAs of U2AF<sup>65</sup> also appeared to be heterogeneous in expression level with strong signals detected in skeletal muscle and brain, and weaker signals in kidney, liver, and lung. The 5-kb signal for U2AF<sup>65</sup> and 2.2-kb signal for SC35 showed less variable intensity among different tissue types.

## Discussion

HCC1 cDNA sequences encoded two known protein motifs, the RS motif and the RNP-CS domain. Proteins containing these motifs belong to a recently described family of non-snRNP splicing factors that are usually not complexed to snRNAs. The RS motif is present in U1 snRNP 70-kD protein

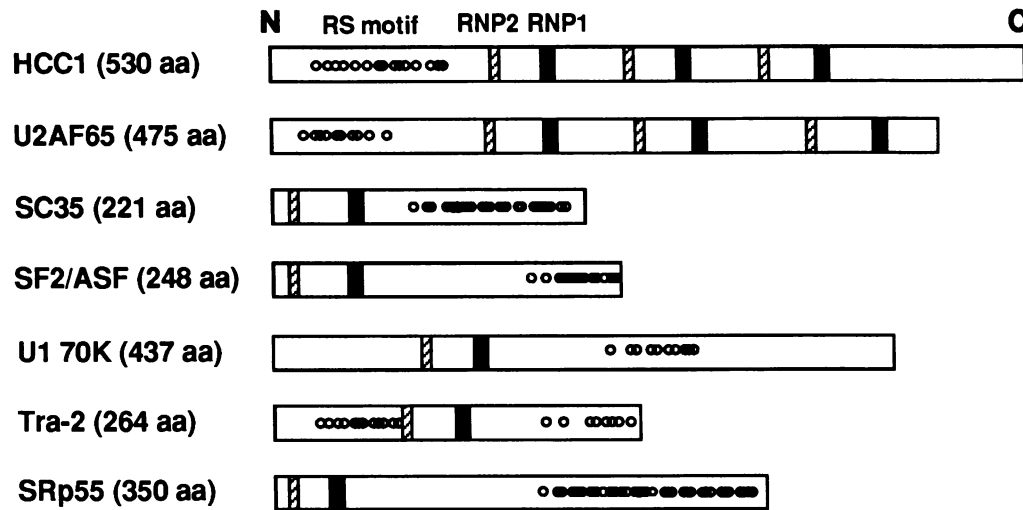
(25, 27), the mammalian non-snRNP splicing factors SF2/ASF (28, 29), SC35 (17), and U2AF (16, 24), and the *Drosophila* splicing regulators *transformer* (30), *transformer-2* (31, 32), and *suppressor-of-white-apricot* (33). Recently, additional family members of pre-mRNA splicing factors containing the RS motif were identified (34, 35). It has been suggested that the RS motif might be involved in nuclear targeting signals (36). The RS motif was necessary for *in vitro* splicing in U2AF<sup>65</sup> and has also been suggested to have RNA binding activities (16). In the RS region of HCC1, an octapeptide sequence called the RS-ERK motif is repeated six times between residues 41 and 90, and is also found in U2AF<sup>65</sup>, U2AF<sup>35</sup>, and U1 snRNP 70-kD protein (Fig. 3B), suggesting that this putative motif may be of biologic significance.

RNP-CS domains have been shown to be responsible for sequence specific RNA binding (25, 37), are important for protein-protein interactions (38), and are found in many proteins involved in RNA processing (22, 25, 39). U2AF<sup>65</sup> and HCC1 contain an RS domain in the NH<sub>2</sub>-terminal region and multiple RNP-CS domains in the central to COOH-terminal region (Fig. 8). The six-amino acid deletion in HCC1.3 results in the loss of the RNP-1-like motif in a possible third RNP-CS domain. If this region does have RNA-binding properties, deletion of the six amino acids observed in HCC1.3 could have functional consequences. At the present time, it is not known whether the deletion in clones HCC1.1, HCC1.2, and HCC1.3 might have resulted from alternative pre-mRNA splicing or whether these highly related mRNAs might have been derived



**Figure 7.** Expression of HCC1 mRNAs and protein. (A) Whole cell extracts from HEp-2, HeLa, HepG2, MOLT-4, rat Novikoff hepatoma, and mouse 3T3 cells ( $\sim 10^5$  cells in each lane) were fractionated on 10% SDS-polyacrylamide gel and transferred to nitrocellulose. The membrane was probed with rabbit antibody to pHCC1 recombinant protein and bound antibodies were detected using <sup>125</sup>I-labeled protein A. A 64-kD band was detected in all cell lines. Preimmune rabbit sera did not show reactivity with the 64-kD antigen. (B) Radiolabeled probe derived from HCC1.4 cDNA (2.5 kb) was used to detect HCC1 mRNA using poly(A)<sup>+</sup> RNA extracted from normal human tissues. The same RNA blot was stripped and reprobbed with SC35 and U2AF<sup>65</sup> cDNAs.





**Figure 8.** Domain distribution of HCC1.4 protein and splicing factors with RS and RNP-CS motifs. RNP-2 and RNP-1 regions in such RNP-CS motif are in hatched and black boxes, respectively. The clustered Arg-Ser or Ser-Arg dipeptides are indicated by open circles. The left and right side of the open boxes correspond to the NH<sub>2</sub> and COOH termini, respectively, and the length of the boxes reflect the number of amino acids in the peptide. *Tra-2* stands for *transformer-2*. References are given in the text. SRp55 is in references 34, 35, and 40.

from different genes. Features including a polypyrimidine tract followed by an AG dinucleotide within the 18-nucleotide deletion (boxed in Fig. 2 B) suggest that the two classes of HCC1 clones are likely to be generated by alternative 3' splice sites (41).

The HCC1 transcripts were also alternatively polyadenylated. Two shorter clones HCC1.1 and HCC1.6, and the other four clones HCC1.2 to HCC1.5 with longer 3' untranslated sequences may correspond to the 2.3- and 3.2-kb mRNA detected in Northern analysis respectively. The presence of two major mRNAs at 2.3 and 3.2 kb is consistent with the usage of two polyadenylation sites, one of which is located 100 nucleotides from the stop codon and the other at 800 nucleotides further downstream. Minor messages detected at 4.3 and 5.0 kb might be caused by alternative or incomplete pre-mRNA splicing or to cross-hybridization. High molecular weight mRNA species were also detected in U2AF<sup>65</sup> and SC35, as well as a 5-kb mRNA of U2AF<sup>65</sup> found in normal human tissues. It appeared that the level of HCC1 mRNA and the 2.4- and 3.5-kb U2AF<sup>65</sup> mRNAs varied among tissues, but whether these levels correspond with protein expression in these tissues remain to be investigated.

From a clinical point of view, it is of interest that antibody to HCC1 was not detectable in the serum of patient NK until several months before the clinical diagnosis of liver malignancy, but was readily demonstrable as one of several novel antibody responses temporally related to malignant transformation (Fig. 1). Also of interest is the current finding that the HCC1 antigen is a nuclear protein with features characteristic of an expanding group of proteins associated with pre-mRNA splicing. These data support previous observations that human autoantibody responses against intracellular antigens are often directed at molecules involved in cellular biosynthetic or proliferative functions (12). It has been proposed that autoantibody responses are antigen driven and not solely caused by disorders of immune regulation such as polyclonal T or B lymphocyte activation. In HCC, the autoantibody responses are quite heterogeneous (5, 11), and in patient NK only one of several antigens has been characterized. One approach towards a better understanding of cellular processes associated with malignant transformation is based on the notion that autoantibodies

might reflect immune responses to cellular components in which functions are exaggerated or aberrantly activated during transformation (11). An important step would be to determine the nature of other antigens that are driving the neo-antibody responses in patients such as NK to ascertain whether the antigens might belong to some related structural or functional families. Molecular cloning of antigens as shown in this study could provide the means to obtain this type of information.

## Acknowledgments

We thank Dr. Michael R. Green for making U2AF<sup>65</sup> cDNA available. We are grateful to John C. Hamel for excellent technical assistance, and to Ying Guo for DNA sequencing.

This is publication 7746-MEM from The Scripps Research Institute. This work was supported by National Institutes of Health grants CA56956, AR32063, AR41803, and M01RR00833. X.-D. Fu was supported by a grant from American Cancer Society (IRG93U).

## References

1. Tan, E. M. 1989. Antinuclear antibodies: diagnostic markers for autoimmune diseases and probes for cell biology. *Adv. Immunol.* 44:93-151.
2. Prelich, G., M. Kostura, D. R. Marshak, M. B. Mathews, and B. Stillman. 1987. The cell-cycle regulated proliferating cell nuclear antigen is required for SV40 DNA replication in vitro. *Nature (Lond.)* 326:471-475.
3. Toschi, L., and R. Bravo. 1988. Changes in cyclin/proliferating cell nuclear antigen distribution during DNA repair synthesis. *J. Cell Biol.* 107:1623-1628.
4. Gottlieb, T. M., and S. P. Jackson. 1993. The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen. *Cell* 72:131-142.
5. Imai, H., R. L. Ochs, K. Kiyosawa, S. Furuta, R. M. Nakamura, and E. M. Tan. 1992. Nucleolar antigens and autoantibodies in hepatocellular carcinoma and other malignancies. *Am. J. Pathol.* 140:859-870.
6. Kass, S., K. Tyc, J. A. Steitz, and B. Sollner-Webb. 1990. The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. *Cell* 60:897-908.
7. Jantzen, H.-M., A. Admon, S. P. Bell, and R. Tjian. 1990. Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature (Lond.)* 344:830-836.
8. Chan, E. K. L., H. Imai, J. C. Hamel, and E. M. Tan. 1991. Human autoantibody to RNA polymerase I transcription factor hUBF: molecular identity of nucleolar organizer region autoantigen NOR-90 and ribosomal RNA transcription upstream binding factor. *J. Exp. Med.* 174:1239-1244.
9. Feuerstein, N., P. K. Chan, and J. J. Mond. 1988. Identification of numatin, the nuclear matrix protein associated with induction of mitogenesis, as the nucleolar protein B23: implication for the role of the nucleolus in early transduction of mitogenic signals. *J. Biol. Chem.* 263:10608-10612.

10. Wands, J. R., and H. E. Blum. 1991. Primary hepatocellular carcinoma. *N. Engl. J. Med.* 325:729-731.
11. Imai, H., Y. Nakano, K. Kiyosawa, and E. M. Tan. 1993. Increasing titers and changing specificities of antinuclear antibodies in patients with chronic liver disease who develop hepatocellular carcinoma. *Cancer (Phila.)*. 71:26-35.
12. Tan, E. M. 1991. Autoantibodies in pathology and cell biology. *Cell*. 67:841-842.
13. Krainer, A. R. 1988. Pre-mRNA splicing by complementation with purified human U1, U2, U4/U6 and U5 snRNPs. *Nucl. Acids Res.* 16:9415-9426.
14. Fu, X.-D., and T. Maniatis. 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature (Lond.)*. 343:437-441.
15. Chan, E. K. L., J. C. Hamel, J. P. Buyon, and E. M. Tan. 1991. Molecular definition and sequence motifs of the 52-kD component of human SS-A/Ro autoantigen. *J. Clin. Invest.* 87:68-76.
16. Zamore, P. D., J. G. Patton, and M. R. Green. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature (Lond.)*. 355:609-614.
17. Fu, X.-D., and T. Maniatis. 1992. Isolation of a complementary DNA that encodes the mammalian splicing factor SC35. *Science (Wash. DC)*. 256:535-538.
18. Adam, S. A., T. Nakagawa, M. S. Swanson, T. K. Woodruff, and G. Dreyfuss. 1986. mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Mol. Cell. Biol.* 6:2932-2943.
19. Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156-159.
20. Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6-13.
21. Kozak, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucl. Acids Res.* 15:8125-8148.
22. Dreyfuss, G., M. S. Swanson, and S. Pinòl-Roma. 1988. Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. *Trends Biochem. Sci.* 13:86-91.
23. Birney, E., S. Kumar, and A. R. Krainer. 1992. A putative homolog of U2AF<sup>65</sup> in *S. cerevisiae*. *Nucl. Acids Res.* 20:4663.
24. Zhang, M., P. D. Zamore, M. Carmo-Fonseca, A. I. Lamond, and M. R. Green. 1992. Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. *Proc. Natl. Acad. Sci. USA*. 89:8769-8773.
25. Query, C. C., R. C. Bentley, and J. D. Keene. 1989. A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein. *Cell*. 57:89-101.
26. Spector, D. L., X.-D. Fu, and T. Maniatis. 1991. Associations between distinct pre-mRNA splicing components and the cell nucleus. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:3467-3481.
27. Mancebo, R., P. C. H. Lo, and S. M. Mount. 1990. Structure and expression of the *Drosophila melanogaster* gene for the U1 small nuclear ribonucleoprotein particle 70K protein. *Mol. Cell. Biol.* 10:2492-2502.
28. Krainer, A. R., A. Mayeda, D. Kozak, and G. Binns. 1991. Functional expression of cloned human splicing factor SF2: homology to RNA-binding proteins, U1 70K, and *Drosophila* splicing regulators. *Cell*. 66:383-394.
29. Ge, H., P. Zuo, and J. L. Manley. 1991. Primary structure of the human splicing factor ASF reveals similarities with *Drosophila* regulators. *Cell*. 66:373-382.
30. Boggs, R. T., P. Gregor, S. Idriss, J. M. Belote, and M. McKeown. 1987. Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from the transformer gene. *Cell*. 50:739-747.
31. Amrein, H., M. Gorman, and R. Nöthiger. 1988. The sex-determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein. *Cell*. 55:1025-1035.
32. Goralski, T. J., J.-E. Edström, and B. S. Baker. 1989. The sex determination locus transformer-2 of *Drosophila* encodes a polypeptide with similarity to RNA binding proteins. *Cell*. 56:1011-1018.
33. Chou, T.-B., Z. Zachar, and P. M. Bingham. 1987. Developmental expression of a regulatory gene is programmed at the level of splicing. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:4095-4104.
34. Mayeda, A., A. M. Zahler, A. R. Krainer, and M. B. Roth. 1992. Two members of a conserved family of nuclear phosphoproteins are involved in pre-mRNA splicing. *Proc. Natl. Acad. Sci. USA*. 89:1301-1304.
35. Zahler, A. M., W. S. Lane, J. A. Stolk, and M. B. Roth. 1992. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes & Dev.* 6:837-847.
36. Li, H., and P. M. Bingham. 1991. Arginine/serine-rich domains of the su(w<sup>+</sup>) and tra RNA processing regulators target proteins to a subnuclear compartment implicated in splicing. *Cell*. 67:335-342.
37. Nagai, K., C. Oubridge, T. H. Jessen, J. Li, and P. R. Evans. 1990. Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature (Lond.)*. 348:515-520.
38. Scherly, D., N. A. Dathan, W. Boelens, W. J. van Venrooij, and I. W. Mattaj. 1990. The U2B<sup>+</sup> RNP motif as a site of protein-protein interaction. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:3675-3681.
39. Kenan, D. J., C. C. Query, and J. D. Keene. 1991. RNA recognition: towards identifying determinants of specificity. *Trends Biochem. Sci.* 16:214-220.
40. Roth, M. B., A. M. Zahler, and J. A. Stolk. 1991. A conserved family of nuclear phosphoproteins localized to sites of polymerase II transcription. *J. Cell Biol.* 115:587-596.
41. Green, M. R. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu. Rev. Cell Biol.* 7:559-599.