Transcription Repression of Human Hepatitis B Virus Genes by Negative Regulatory Element-binding Protein/SON*

Received for publication, February 12, 2001, and in revised form, April 2, 2001 Published, JBC Papers in Press, April 16, 2001, DOI 10.1074/jbc.M101330200

Chien-Tzang Sun‡, Wei-Yu Lo‡, In-Hsun Wang, Yu-Hsnn Lo, Sheng-Ru Shiou, Chao-Kuen Lai, and Ling-Pai Ting§

From the Institute of Microbiology and Immunology, School of Life Science, National Yang-Ming University, Shih-Pai, Taipei 112, Taiwan, Republic of China

A negative regulatory element (NRE) is located immediately upstream of the upstream regulatory sequence of core promoter and second enhancer of human hepatitis B virus (HBV). NRE represses the transcription activation function of the upstream regulatory sequence of core promoter and the second enhancer. In this study, we described the cloning and characterization of an NRE-binding protein (NREBP) through expression cloning. NREBP cDNA is 8266 nucleotides in size and encodes a protein of 2386 amino acids with a predicted molecular mass of 262 kDa. Three previously described cDNAs, DBP-5, SONB, and SONA, are partial sequence and/or alternatively spliced forms of NREBP. The genomic locus of the NREBP/SON gene is composed of 13 exons and 12 introns. The endogenous NREBP protein is localized in the nucleus of human hepatoma HuH-7 cells. Antibody against NREBP protein can specifically block the NRE-binding activity present in fractionated nuclear extracts in gel shifting assays, indicating that NREBP is the endogenous nuclear protein that binds to NRE sequence. By polymerase chain reaction-assisted binding site selection assay, we determined that the consensus sequence for NREBP binding is GA(G/T)AN(C/G)(A/G)CC. Overexpression of NREBP enhances the repression of the HBV core promoter activity via NRE. Overexpression of NREBP can also repress the transcription of HBV genes and the production of HBV virions in a transient transfection system that mimics the viral infection in vivo.

Infection of hepatitis B virus (HBV) causes acute and chronic hepatitis and is closely associated with the development of cirrhosis and hepatocellular carcinoma. HBV is a small enveloped DNA virus with a partially double-stranded 3.2-kb genome. The genome contains four partially overlapping open reading frames (ORFs) coding for the surface, core, polymerase, and X proteins. The transcription of these open reading frames is under the control of four promoters (two surface promoters, one core promoter, and one X promoter) and two enhancers (enhancer I and enhancer II). Core promoter produces two 3.5-kb RNAs: the precore and pregenomic RNAs. Precore RNA encodes precore protein and e antigen. Pregenomic RNA not only serves as the mRNA that encodes core and polymerase proteins but also can be packaged into nucleocapsids along with viral polymerase, serving as the template for reverse transcription. Regulated expression of pregenomic RNA plays a pivotal role in the control of the viral replication cycle. The core promoter can be divided into two elements: the basal core promoter (BCP) and the core upstream regulatory sequence (CURS). CURS can activate the adjacent downstream BCP activity in cis. Interestingly, the CURS is also colocalized with the second enhancer (ENII) in the HBV genome (1). The ENII can activate the surface and X promoters in a position- and orientation-independent manner (2). The CURS/ENII displays a differentiated liver cell specificity (3), which is the combined effect of several liver-enriched transcription factors, such as CCAAT/enhancer-binding protein (4–6), FTF (7–9), HNF4 (10, 11), HNF3 (12, 13), and HNF1 (14, 15).

We have previously identified a negative regulatory element (NRE) located upstream of CURS/ENII. NRE can effectively abolish the transcription stimulatory function of both CURS and ENII from a nearby upstream position (16). The minimal essential sequence required for the NRE function has been mapped (16, 17). A trans-acting factor present in the nuclear extracts derived from a human hepatoma cell line can specifically bind to this region (16).

In this paper, we describe the cloning and characterization of an NRE binding protein, NREBP. NREBP bears strong homology to previously described cDNAs: DBP-5, SONA, and SONB. Antiserum raised against NREBP recombinant protein can specifically abolish the nuclear extracts to form an NRE-protein complex in gel shifting assays. Recombinant NREBP can specifically interact with wild type but not mutated NRE sequence. PCR-assisted binding site selection assay reveals that the optimal sequence for binding to NREBP is a perfect match

* This study was supported by National Science Council Grants NSC-87-2315-B-010-002MH, NSC88-2315-B-010-012MH, and NSC89-2315-B-010-007MH, National Health Research Institute Grants NHRI-GT-EX89B906L and NHRI-EX90-8906BL, and Veterans General Hospital (Republic of China) Grant V89-383-4. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ The first two authors contributed equally to this work.

§ Recipient of an award from the Medical Research and Advancement Foundation in memory of Dr. Chi-Shuen Tsou. To whom correspondence should be addressed. Tel.: 886-2-28222400; Fax: 886-2-28212880; E-mail: lpting@ym.edu.tw.

# The abbreviations used are: HBV, human hepatitis B virus; CURS, HBV core promoter upstream regulatory sequence; BCP, basal core promoter region of HBV core promoter; ENII, second enhancer of HBV;
for the NRE sequence. Overexpression of NREBP enhances the transcription repression of core promoter mediated by NRE. Overexpression of NREBP also represses the viral replication and gene expression in a HBV replication system.

**EXPERIMENTAL PROCEDURES**

**Isolation of cDNA Clones—** A ZAPII cDNA library was made from messenger RNA isolated from human hepatoma HepG2 cells according to the manufacturer’s recommendation. This library was screened with concatenated double-stranded synthetic oligonucleotides of wild type NRE sequence. The oligonucleotides were derived from a 1636-bp EcoRI site in the unique EcoRI site, which is nt 1. Plasmids pSV2CAT, pNRE-CP-CAT, pCPCAT, and pPHV3.6 were described previously (15, 16). The recombinant GST-BP15 and pGST-BP15 and pET-BP15, were generated by cloning of a 1.6-kb EcoRI fragment containing the BP15 cDNA into the EcoRI site of pGEX-3X vector (Amersham Pharmacia Biotech), pET30a (Novagen), or pET28a (Novagen), respectively. The plasmid pCMV-BP15 was generated by cloning of the same 1.6-kb EcoRI fragment into the EcoRI site downstream of a CMV immediate early promoter of the pFLAG-CMV2 expression vector (Eastman Kodak Co.). The plasmid pCMV-f:A1 was generated by cloning of a 5-kb BamHI-XhoI fragment containing the A1 cDNA into BamHI and SalI sites downstream of a CMV immediate early promoter of the pFLAG-CMV2 vector. All of these constructs were verified by DNA sequencing using appropriate sequencing primers.

**Production of Bacterially Derived GST-BP15 Recombinant Protein and Induction of Rabbit Polyclonal Anti-BP15 Antibody—** GST-BP15 fusion protein was expressed and purified as previously described (18). Purified GST-BP15 protein was separated on 5% SDS-PAGE and eluted with electrophoresis (Bio-Rad). Protein eluent was concentrated with Centricon-10 concentrator (Amicon), left on ice for 2–3 h, centrifuged to remove SDS, and then dialyzed against renaturation buffer (10 mM HEPES (pH 7.9), 100 mM KCl, 0.1 mM EDTA, 0.1 mM ZnSO4, 1 mM dithiothreitol, 10 mM MgCl2, and 10% glycerol) at 4 °C for 24 h. Recombinant GST-BP15 after renaturation was then used as antigen to immunize rabbit.

**Affinity Purification of Anti-BP15 Polyclonal Antibody—** Expression of recombinant S.Tag-BP15 protein tagged with S.Tag was induced by isopropyl-1-thio-β-D-galactopyranoside and then purified by S-protein-agarose chromatography according to the manufacturer’s recommendation. Purified recombinant S.Tag-BP15 protein was separated on 10% SDS-PAGE and transferred to nitrocellulose membrane (Amersham Pharmacia Biotech). To identify the recombinant S.Tag-BP15 protein, a side chain was cut out of GST-Tag (19). The double-stranded degenerate template oligonucleotides in a DNA binding buffer at room temperature for 1 h. After wash with the binding buffer, bound DNA was eluted with elution buffer (0.5 mM ammonium acetate, 5 mM EDTA, and 0.5% SDS) at room temperature for 30 min. After phenol/chloroform extraction, bound DNA was purified by ethanol precipitation and then cloned by PCR. After five rounds of binding, elution, and amplification, PCR products were cloned. Individual clones were then isolated and sequenced.

**Transfection and CAT Assay—** For reporter gene expression, transfection of human hepatoma cell lines HepG2 and HuH-7 was performed with 2.7 μg of reporter plasmid, 0.5 μg of CH101, 0.91 μg of pFLAG-CMV2, or 2 μg of pCMV-f:A1 expression plasmid and carrier plasmid as previously described (19). The double-stranded degenerate template oligonucleotides in a DNA binding buffer at room temperature for 1 h. After wash with the binding buffer, bound DNA was eluted with elution buffer (0.5 mM ammonium acetate, 5 mM EDTA, and 0.5% SDS) at room temperature for 30 min. After phenol/chloroform extraction, bound DNA was purified by ethanol precipitation and then cloned by PCR. After five rounds of binding, elution, and amplification, PCR products were cloned. Individual clones were then isolated and sequenced.

**Northern Blotting Analysis—** Purified inserts containing BP9, BP15, 6-1, GAPDH, and β-actin probes. Each set of experiments was performed with two different preparations of plasmids and repeated two to three times for each preparation. The CAT activity was measured by a PhosphorImager and normalized against β-galactosidase activity.

For HBV replication experiment, 25 μg of the plasmid HBV 3.6, which contained more than a unit length of HBV viral genome (1), was transfected into HuH-7 cells in a 15-cm plate with either no plasmid, 15.9 μg of pFLAG-CMV2, or 35 μg of pCMV-f:A1 and carrier plasmid, respectively. At day 3 posttransfection, total RNAs were collected for Northern blotting analysis, while culture media were collected for endogenous DNA polymerase activity.

**Northern Blotting Analysis—** Purified inserts containing BP9, BP15, 6-1, GAPDH, and β-actin cDNAs were labeled with the random priming method and used as probes, respectively. Preparation of total cellular RNA from human hepatoma HepG2 and HuH-7 and Northern hybridization using BP9, BP15, and 6-1 probes were performed as previously described (19). Membranes containing 10 μg of total cellular RNA from different adult human tissues (CLONTECH) were used for Northern blotting analysis with BP9, BP15, and 6-1 as probes. The same blots were reprobed with β-actin probe.

In the HBV replication experiment, total cellular RNA was collected at day 3 posttransfection. Forty μg of total RNA of each sample was analyzed by Northern hybridization. The probe was a PstI fragment of pHV3.8 containing the HBV sequence from nt 25 to 1989. The same

HuH-7 cells cultured on slides were transiently transfected with pFLAG-CMV2, pCMV-f:B15, and pCMV-f:A1 plasmids, respectively. Immunofluorescence was performed as previously described (19). The endogenous NREBP protein was detected by purified anti-BP15 antibody, while overexpressed f:B15 and f:A1 proteins were detected by anti-FLAG antibody.

**Preparation of Crude and Fractionated Nuclear Extracts and Gel Shifting Assay—** Preparation of crude and fractionated nuclear extracts from cultured cells was performed as previously described (1). Gel shifting analysis was performed as previously described (6, 16). For blocking experiments, an appropriate amount of GST-anti-GST-BP15 antibody, respectively, was preincubated with the 0.4 μg at NcOD1 step-eluted nuclear extracts on ice for 25 min before the addition of labeled probe.

**Southwestern Analysis—** The probe was used to concentate double-stranded oligonucleotides containing the wild type or mutant NRE (5′-gactctTCTCTCAACGATGGCCCTACAG-3′ and its complement with mutation from nt 1613 to 1621) sequence of HBV. The probe was end-labeled with γ-32P-ATP and T4 polynucleotide kinase. Recombinant GST-BP15 protein or affinity-purified f:B15 protein with anti-FLAG M2 affinity gel (Sigma) used for the Southwestern assay was resolved by 8% SDS-PAGE and then transferred to nitrocellulose membrane. The membrane was renatured and blocked in renaturation buffer containing 5% nonfat milk for 8 h at 4°C. The membrane was then incubated in DNA binding buffer (20 mM HEPES, pH 7.9, 100 mM KCl, 0.5 mM dithiothreitol, 0.25 mM MgCl2, 0.5% nonfat milk, and 0.1 mg/ml salmon sperm DNA) containing radiolabeled probe at 106 cpm/ml at room temperature for 6 h. After washing three times with the binding buffer for 30 min, the membrane was subjected to autoradiography.

**PCR-assisted Binding Site Selection Assay—** Binding site selection was a modification of the assay described by Pollock and Treissman (20). The sequence of degenerate template is 5′-TCTCGAGTACTAGCANNNNNNNNNNNNNACTGAGCATGATGCTG-3′. Primers for PCR amplification are 5′-TCTCGAGTACTAGCAGCA-3′ and 5′-AGCATGAGCATGCTG-3′. Recombinant GST-BP15 protein was immobilized on nitrocellulose membrane, renatured in renaturation buffer containing 5% nonfat milk overnight at 4°C, and then incubated with 200 ng of concentrated double-stranded degenerate DNA template oligonucleotides in a DNA binding buffer at room temperature for 1 h. After wash with the binding buffer, bound DNA was eluted with elution buffer (0.5 mM ammonium acetate, 5 mM EDTA, and 0.5% SDS) at room temperature for 30 min. After phenol/chloroform extraction, bound DNA was purified by ethanol precipitation and then amplified by PCR. After five rounds of binding, elution, and amplification, PCR products were cloned. Individual clones were then isolated and sequenced.
Cloning of NREBP cDNAs—Three cDNAs (BP9, BP4, and BP15) were obtained by expression cloning with concatenators of NRE sequence. A1 cDNA was obtained by library screening with a probe derived from the 5¢ portion of the BP9 sequence. 6-1 cDNA was obtained by 5¢ rapid amplification of cDNA ends. These cDNAs assemble into NREBP cDNA totaling 8266 nt in length. NREBP cDNA encodes a protein of 2386 amino acids with the translation start codon and stop codon at nt 50 and 7208, respectively. Seven different repeats of this protein are shown on the NREBP ORF. The relationship of each cDNA to the combined NREBP cDNA sequence is shown with the position of nucleotides on the top and amino acids on the bottom.

RESULTS

Cloning of NREBP cDNA—NRE is a sequence element that represses the transcription activation function of the adjacent ENII and core promoter in the HBV genome. Mutational analysis reveals the minimal essential sequence, nt 1613–1621, within NRE that is required for its function (16, 17). This sequence is bound specifically by nuclear proteins derived from differentiated human hepatoma HepG2 cells (16). To search for NRE-binding protein(s), we performed expression cloning using labeled concatenators of oligonucleotides of NRE sequence to screen a cDNA library made from HepG2. Three overlapping cDNAs, BP4, BP9, and BP15, were obtained (Fig. 1). Sequence analysis reveals a long open reading frame (ORF) shared by these cDNAs. The gene encoded by these cDNAs is referred to as NREBP. The longest BP9 cDNA is 5145 nt in size.

Northern analysis using BP9 cDNA as probe showed that NREBP transcripts had apparent sizes of 8.3 kb and greater than 10 kb (see below). To obtain full-length NREBP cDNA, we used sequence from the 3¢-end of BP9 to screen an adult human liver cDNA library. A cDNA clone, A1, was obtained that was 5539 nt in size. A1 represented the 3¢ portion of the NREBP transcript, since it contained both translation stop codon and a poly(A) tract at its 3¢-end. 5¢ rapid amplification of cDNA ends was performed, and an additional cDNA clone, 6-1, was obtained that represented further 5¢ extension. The combined sequence information from BP9, A1, and 6-1 reveals a NREBP cDNA that is 8266 nt in size. The translation start codon is located at nt 50 within the sequence context of GCCAUGGCG, which conforms to the Kozak consensus sequence, (A/G)C-

Expression Pattern of NREBP Transcript—To examine whether NREBP was expressed in human hepatoma cell lines, total RNAs isolated from human hepatoma HuH-7 and HepG2 cells were used for Northern hybridization with BP9 (Fig. 2A), BP15, and 6–1 (data not shown) as probes, respectively. A 8.3-kb NREBP transcript was detected. Furthermore, poly(A)+ RNAs from a variety of human tissues were subject to Northern analysis with BP9 (Fig. 2B) and BP15 (data not shown) as probes, respectively. Two transcripts were found, one at 8.3 kb and the other at greater than 10 kb. The 8.3-kb transcript was the major one. NREBP transcripts could be detected in all tissues. The highest expression was seen in leukocyte and heart, followed by lymph node, spinal cord, ovary, testis, thymus, spleen, pancreas, placenta, and brain. Weaker expression was seen in colon, small intestine, prostate, kidney, skeletal muscle, liver, and lung.

Localization and Expression of NREBP Protein—To study the localization and expression of the NREBP protein, polyclonal anti-BP15 antiserum was generated after immunizing rabbits with recombinant GST-BP15 protein. This antiserum

ent in the amino acid (aa) sequence (Fig. 1). Repeat 1, which is composed of a 10-aa unit, LA(S/T)(N/S/G)/T/S/MDSQM, is repeated 13 times and occupies the position from aa 726 to 855. Repeat 2, present in 11 copies from aa 872 to 948, is composed of 7 amino acids, (D/R)PYR/I(L)A/G/Q/K/H/P. Nine copies of the repeat 3, PAYERMSM, is present from aa 973 through 1045. There are three copies of the repeat 4, PPLPPEEP/T/M/E(M/T/G), from aa 1107 to 1139. Four copies of repeat 5, VLES-SAVT, occupy the position from aa 1319 to 1350. Six copies of repeat 7, which is composed of PSRRRSKT, is located at the segment from aa 1913 through 1954. Two copies of repeat 7, which is composed of PSRRRSSVSVRRRSSFIS, are present from aa 1894 to 1912 and from 1955 to 1973, respectively. In addition, the region from aa 1090 to 1361 is rich in acidic amino acids, while the region from aa 1787 to 2009 is rich in basic amino acids.

A search of the data base reveals that the genomic region containing the NREBP gene has been completely sequenced (DDBJ accession number AP000046). The NREBP gene spans ~35 kb and contains 13 exons and 12 introns. The sizes of exons and introns and splicing junction sequences are shown in Table I. The combined NREBP cDNA contains exons 1–12, exon 13a, and intron 3. The NREBP gene has been mapped to chromosome 21q22.1 (21), which is a critical region for an autosomal dominant familial “aspirin-like” platelet disorder associated with development of acute myelogenous leukemia and in close proximity to the Down syndrome critical region (22–27).

Transcription Repression by NREBP/SON

24061

FIG. 1. Cloning of NREBP cDNAs. Three cDNAs (BP9, BP4, and BP15) were obtained by expression cloning with concatenators of NRE sequence. A1 cDNA was obtained by library screening with a probe derived from the 3¢ portion of the BP9 sequence. 6-1 cDNA was obtained by 5¢ rapid amplification of cDNA ends. These cDNAs assemble into NREBP cDNA totaling 8266 nt in length. NREBP cDNA encodes a protein of 2386 amino acids with the translation start codon and stop codon at nt 50 and 7208, respectively. Seven different repeats of this protein are shown on the NREBP ORF. The relationship of each cDNA to the combined NREBP cDNA sequence is shown with the position of nucleotides on the top and amino acids on the bottom.
Expression of NREBP mRNA in various human tissues. Poly(A) blotting. The membrane, after transfer, was probed with BP9 cDNA. Total RNA from each cell line as indicated was loaded for Northern mRNA in differentiated human hepatoma cell line HuH-7 and HepG2. In transfected HuH-7 cells were localized in the nucleus as which serves as negative control. Both f:BP15 and f:A1 proteins insertless, parental vector, pFLAG-CMV2, was performed, in HuH-7 cells by transient transfection. Transfection with an with a FLAG epitope at their N terminus, were overexpressed truncated forms of NREBP protein (f:BP15 and f:A1), tagged endogenous NREBP protein was localized in the nucleus. Two membranes were probed with BP9 cDNA. The tissue origins of the RNA samples examined are as indicated. The from various human tissues was used for Northern blot analysis. The identified anti-BP15 antibody. As shown in the Fig. 3, the top panel, expression of NREBP protein in HuH-7 cells was detected by purified anti-BP15 antibody and a fluorescein isothiocyanate-conjugated goat anti-rabbit IgG (top left panel). Plasmid pFLAG-CMV2 (negative control, data not shown), pCMV-f:BP15 (f:BP15), or pCMV-f:A1 (f:A1) was transiently transfected into HuH-7 cells. The localization of f:BP15 (middle left panel) and f:A1 (bottom left panel) proteins was examined by anti-FLAG monoclonal antibody and a fluorescein isothiocyanate-conjugated goat anti-mouse IgG. The nuclear DNA was stained with Hoechst 33258 (right panel). (Fig. 3, middle and bottom panels). HuH-7 cells transfected with a parental vector did not produce any signal (data not shown).

To identify the endogenous NREBP protein by Western blotting, crude nuclear extracts from HuH-7 cells were loaded onto a heparin-Sepharose column and step-eluted with NaCl at increasing concentrations. Each fraction was then collected and tested for binding to the NRE sequence. As previously observed, the NRE-binding protein was present in the 0.4 M fraction (16). The DNA-protein complex appeared to be specific for NRE, since it could be competitively abolished by wild type NRE sequence but not by mutant NRE or random sequence (16). When crude and 0.4 M NaCl fraction of fractionated nuclear extracts were examined by Western blotting with purified anti-BP15 antibody, several proteins were detected. The apparent molecular mass of the largest one was about 257 kDa, which was consistent with the expected molecular mass of protein encoded by the full-length NREBP, 262 kDa (Fig. 4). Proteins of smaller sizes may represent processed or degraded products, translation products from alternatively spliced transcripts, or products derived from internal initiation. No signal was seen with the preimmune serum (data not shown).

The expression of f:BP15 and f:A1 proteins was examined by Western blotting using purified anti-BP15 antibody. The expression of f:BP15 in HuH-7 transient transfectants could be detected, although the apparent molecular mass was larger than expected (138 and 130 kDa instead of 63.5 kDa; Fig. 5A). The same result was obtained in 293T cells after transient transfection (Fig. 5B). The expression of f:A1 in HuH-7 transient transfectants could not be detected by Western blotting, because the level of f:A1 expression was significantly lower than that of f:BP15 (Fig. 5A). However, the expression of f:A1 in 293T transient transfectants could be detected, although the apparent molecular mass of f:A1 was 217 and 197 kDa, which is larger than the expected molecular mass of 165.5 kDa (Fig. 5B). That the apparent molecular mass of f:BP15 and f:A1 was larger than expected is probably due to the high content of proline residues in NREBP. HuH-7 and 293T cells transfected with an insertless parental control vector did not produce any signal by Western blotting analysis.

NREBP Is Indeed the Endogenous NRE-binding Protein—To
test whether NREBP is the endogenous protein that binds to the NRE sequence, we performed gel shifting assays. When the 0.4 M fraction of fractionated nuclear extracts was preincubated with anti-BP15 antiserum, the formation of DNA-protein complex was abolished (Fig. 6, lane 4). Control anti-GST antiserum did not have any effect (lane 3). This result indicates that NREBP is indeed the endogenous protein that binds to NRE sequence in 0.4 M fractionated nuclear extracts.

Binding Specificity of NREBP Protein—To test whether NREBP protein can bind to the NRE sequence specifically, the f:BP15 protein produced by the transiently transfected HuH-7 cells was purified by affinity column chromatography coated with anti-FLAG antibody and then transferred to nitrocellulose membranes after SDS-PAGE. After denaturation and renaturation, Southwestern binding assays were performed using radioactively labeled wild type NRE sequence as probe. The amount of f:BP15 protein produced was verified with Western blotting (data not shown). As shown in Fig. 7, proteins derived from f:BP15 transfecants could interact with wild type NRE sequence but not mutant NRE sequence. Proteins derived from cells transfected with a parental control vector could not bind either wild type or mutant NRE sequence.

This study shows that the middle portion of NREBP, as represented by BP15, can interact with wild type NRE sequence specifically. Similar binding results were obtained using bacterially derived recombinant protein, GST-BP15, by Southwestern binding assays (data not shown). However, the binding of this GST-BP15 protein to the NRE sequence could not be detected in gel shifting assays (data not shown).

Recombinant GST-BP15 protein was subsequently used in a PCR-assisted binding site selection assay to examine the binding specificity of recombinant NREBP protein. Recombinant GST-BP15 protein was first immobilized on nitrocellulose filters and incubated in a DNA binding reaction with a pool of oligonucleotides. These oligonucleotides had random sequences in the middle flanked by primer binding sites suitable for PCR amplification. The binding experiment was performed in a condition identical to that of the Southwestern assay. After extensive wash, bound oligonucleotides were eluted and amplified by
Transcription Repression by NREBP/SON

FIG. 9. Activation of NRE function by overexpression of NREBP: a reporter gene assay. HuH-7 and HepG2 cells were transfected with pNRE-CP-CAT, pCP-CAT, or pBCP-CAT in the presence of a pCH101 and a pCMV-f:A1 or a pFLAG-CMV2 vector. This diagram shows the relative activity of core promoter achieved by overexpression of f:A1 for different core promoter constructs. The relative activity (percentage) was calculated by taking the normalized CAT activity of the transient transfecants co-transfected with pFLAG-CMV2 vector as 100%.

The NRE of HBV represses the activity of the nearby core promoter and ENII by 10–20-fold (16, 17). In this paper, we describe the cloning and characterization of a transcription factor, NREBP. NREBP is a 2386-aa protein. The middle portion of NREBP can interact with wild type NRE sequence specifically, but not with mutant NRE sequence. Furthermore, the consensus sequence for NREBP binding, GA(G/T)AN(C/G/A)G/C, is a perfect match for the NRE sequence in the HBV genome, GAGACCACC.

NREBP Mediates Transcription Repression by NRE—We have previously shown that the core promoter of HBV can be functionally dissected into two regions: a basal core promoter, BCP, and a core promoter upstream regulatory sequence, CURS. CURS can activate the BCP activity. This stimulatory activity can, however, be partially abolished by the nearby NRE sequence from an upstream position (16). To test the ability of NREBP to mediate this transcription repression effect, we performed co-transfection experiments with an NREBP expression construct and constructs containing a reporter gene driven by different core promoter elements. We chose to use f:A1 construct, since numerous attempts to engineer a full-length NREBP expression construct failed, probably because of the sequence repeats present in the ORF. An empty vector, pFLAG-CMV2, was used as a control. The reporter gene constructs we used were a CAT reporter gene driven by the BCP (pBCP-CAT), BCP plus CURS (pCP-CAT), and BCP plus CURS plus NRE (pNRE-CP-CAT). As shown in Fig. 9, the overexpression of f:A1 reduced the promoter activity of pNRE-CP-CAT to about 13.5 and 12.2% in HuH-7 and HepG2 cells, respectively. The promoter activity of neither pCP-CAT nor pBCP-CAT was significantly affected by f:A1. This result strongly suggests that NREBP can further augment the transcription repression effect of NRE on the core promoter of HBV.

NREBP Represses HBV Gene Expression and Replication—To test the effect of NREBP on the HBV gene expression and replication, we resorted to transient transfection with more than a unit length of HBV genome, pHBV3.6, in HuH-7 cells. Viral gene expression and production of mature virions that closely mimic viral infection in vivo have been seen after transfection (1). The effect of NREBP on the transcription and replication of HBV was tested by co-transfecting an NREBP expression plasmid pCMV-f:A1 or an empty vector pFLAG-CMV2 with pHBV3.6. To avoid the competition among promoters, no plasmid was co-transfected to normalize the transcription efficiency. To circumvent this, the experiment was repeated four times. Three days after transfection, the amount of the 2.4-kb large surface, 2.1-kb middle and major surface, and 3.5-kb precore and pregenomic transcripts was measured by Northern hybridization (Fig. 10A, top panel). The expression of GAPDH was used as an RNA loading control (Fig. 10A, bottom panel). Overexpression of NREBP reduced the expression of 3.5-kb RNAs to 20%, as well as 2.4- and 2.1-kb RNAs to 40%. The production of mature virions and core particles was quantified by an endogenous DNA polymerase activity assay. As shown in Fig. 10B, NREBP reduced the production of virions and core particles to 43.5%. These results strongly suggest that NREBP represses the gene expression and replication of HBV.

DISCUSSION

The NRE of HBV represses the activity of the nearby core promoter and ENII by 10–20-fold (16, 17). In this paper, we describe the cloning and characterization of a transcription factor, NREBP. NREBP is a 2386-aa protein. The middle portion of NREBP can interact with wild type NRE sequence specifically, but not with mutant NRE sequence. Furthermore, the consensus sequence for NREBP binding, GA(G/T)AN(C/G/A)G/C, is a perfect match for the NRE sequence in the HBV genome, GAGACCACC. Overexpression of an NREBP protein lacking the N-terminal 892 amino acids can further enhance this repression mediated by NRE. Overexpression of this truncated NREBP protein can also repress the transcription of HBV genes and the production of HBV virions in a transient transfection system that mimics the viral infection in vivo. The abolishment of the formation of DNA-protein complex between NRE and nuclear proteins by antibody against NRE demonstrates that NREBP is the endogenous NRE binding protein present in the cells. Taken together, NREBP appears to function as a transcription repressor at NRE.

It has been reported that the sequence of HBV from nt 1605 to 1625, named as NRE γ, can repress the core promoter activity by 2–3-fold in HuH-7 cells (28). This sequence can be bound by a transcription factor, RFX1. The consensus sequence for RFX1 binding is from nt 1605 to 1617. Overexpression of RFX1,
transcription may not be essential for the transcription repressor function of NREBP, since overexpression of A1 protein, which is devoid of the leucine zipper domain, can still repress transcription mediated through NRE.

Interestingly, the NREBP bears strong homology to previously described cDNAs: DBP-5, SONA, and SONB (EMBL accession numbers X63071, X63753, and X63751, respectively) (Fig. 11). DBP-5 cDNA, 4972 nt in size, was obtained through expression cloning from a human B cell cDNA library using a segment of the HLA-DR gene promoter as the probe (30). The sequence of DBP-5, from nt 1 to 4967, is almost identical (99.6%) to the NREBP cDNA sequence from nt 3290 to 8257 (Fig. 11A). The sequences of NREBP and DBP-5 were colinear except for the deletion of one single nucleotide after nt 3585 in DBP-5 (corresponding to nt 6875 in NREBP), which leads to a frameshift in the predicted ORF. The sequences of the corresponding genomic region and SONA cDNA (see below) do not carry this deletion. This discrepancy, therefore, is probably due to a sequencing error in DBP-5. Furthermore, the initially assigned start codon of DBP-5 (30) is probably only an internal ATG, since DBP-5 appears to be a 5′-truncated cDNA. DBP-5 cDNA therefore appears to encode a truncated 1306-aa protein instead of the originally assigned 1179 amino acids. The 1306-aa DBP-5 protein corresponds to an 1081–2386 of NREBP (Fig. 11B).

SONA and SONB cDNAs were isolated by hybridization screening of a human embryonic cDNA library probed with the rat gene K51, which was obtained initially based upon its cross-hybridization to the v-mos gene under a nonstringent condition (31–34). SONB appears to be a partial cDNA corresponding to the middle portion of NREBP cDNA. The sequence of SONB from nt 1 to 3373 is colinear and nearly identical (99.7%, not including the 240-nt insertion) to that of NREBP from nt 2174 to 5303 (Fig. 11A); however, there is a 240-nt insertion (from nt 112 to 351 of SONB) in SONB. The relative position of this insertion site is located between nt 2284 and 2285 (within intron 3) of the NREBP cDNA. The reported genomic sequence has a 120-nt insertion instead. This addition results in an increase in the number of repeats in the repeat 1 region; there are 13 repeats in NREBP and 21 repeats in the SONB sequence, while the reported genomic sequence has 17 repeats. SONB cDNA encodes a 1124-aa protein corresponding to the middle portion of NREBP from aa 709 to 1751 (Fig. 11B).

The sequence of SONA from nt 7 to 5676 is partially identical (99.7%, not including intron 3 and the 53-nt deletion in exon 13) to the sequence of NREBP from nt 1708 to 8266 (Fig. 11A). SONA, however, lacks the sequences corresponding to the segments from nt 2108 to 2945 (intron 3) and from 7151 to 7203 (5′-terminal 53 nt of exon 13a) in NREBP (Fig. 11A). Compared with the NREBP and genomic sequences, SONA appears to be a product of alternative splicing (Table I). The SONA cDNA starts from within exon 3, continues through exons 4–12 and ends in exon 13b. Exon 13b lacks the 5′-terminal 53 nt of exon 13a. Because of the absence of the third intron, there is a shift in the ORF. It is not clear if SONA represents a full-length or nearly full-length product of this alternatively spliced transcript. If SONA uses the same translation start codon as NREBP, SONA will yield a C-terminal truncated product as a stop codon (at nt 416 of SONA) is introduced from frameshift because of the splicing out of the third intron. If SONA uses a downstream translation start codon at nt 415 or a more distal location for internal initiation instead, it will yield an N-terminal truncated product of NREBP (Fig. 11B). If the latter is the case, the C-terminal portion of SONA protein will also be different from that of NREBP protein, since the omission of the 53 nt in exon 13 (corresponding to nt 7151–7203 of NREBP)
generates an additional frameshift of the ORF (Fig. 11B). In either case, SONA does not contain repeat 1 or 2. SONA may therefore encode a protein functionally distinct from NREBP. The sequence at the extreme 5′-end of SONA cDNA (nt 1–6) is not found in either the cDNA or genomic sequence of NREBP. Examination of the sequence from the corresponding region in the genomic locus of NREBP shows that the sequence difference is not likely to be the result of alternative splicing. The possibility that this sequence is introduced as a cloning artifact remains to be ruled out.

The sequence of the region at chromosome 21q22.1 containing the NREBP/SON gene is completely known. Based on our results, the NREBP/SON gene spans ~35 kb and contains 13 exons and 12 introns. Interestingly, comparison of the cDNA and genomic sequence of NREBP reveals the insertion of a segment of 120 nt in the genomic locus that is located between nt 2284 and 2285 of the NREBP cDNA sequence. SONB contains a 240-nt insertion at this location compared with the cDNA sequence of NREBP. These extra sequences represent an increase in the number of repeats in the repeat 1 region of NREBP. The protein encoded by our SONB cDNA, the sequenced genomic region of NREBP, and the SONB cDNA therefore will have 13, 17, and 21 repeats in the repeat 1 region, respectively. The increase in the number of repeats is not the result of alternative splicing and most likely represents genetic polymorphism. The effect of the number of repeats on the transcriptional regulation function of NREBP requires further study. Some other minor sequence difference and/or polymorphism among NREBP, DBP-5, SONB, and SONA cDNAs and the NREBP/SON gene are not discussed here.

Although the mechanism of transcription activation has been well described, relatively little is known about how the transcription repressor works. Transcription repressors can be passive or active repressors (35–38). Passive repressors can counteract the activator function by the following mechanisms: (a) direct competition of the same DNA-binding sites (39–45); (b) interference of overlapping or neighboring activator-binding sites (46–51); (c) prevention of the translocation of the transcription activators (52); (d) titration away of limiting protein factors required for transcription activator function (53, 54); (e) modification of the DNA-binding property of the transcription activators (55); (f) blocking of the DNA binding activity of the transcription activators through protein-protein interactions (50, 56, 57); or (g) masking or alteration of the function of the factors required for transcription activator function (53, 54).

Active repressors can directly inhibit the assembly or the activity of the general transcription machinery (60–66). Another mode of transcription repression is through the recruitment of histone deacetylase complexes by repressors or corepressors. Histone deacetylase complexes will alter the chromatin structure through increased histone deacetylation (reviewed in Refs. 67–69). We have previously shown that NRE can repress the transcription activation function of the ENII (16, 17). NRE is located 25 nt upstream of the ENII, and the repression function of NRE depends on its close proximity to the ENII. The spatial requirement of the function of NRE on the ENII favors the
Transcription Repression by NREBP/SON

notion that NREBP functions as a passive repressor. The exact repression mechanism of NREBP remains to be elucidated.

Acknowledgments—We are grateful to Dr. Shiu-Wen Luoh for stimulating discussions in the course of these experiments and for critical reading of the manuscript. We thank Dr. Shih-Feng Tsai for providing the human adult liver cDNA library. We thank Weber Chern for computer technology support.

REFERENCES

1. Yuh, C.-H., Chang, Y.-L., and Ting, L.-P. (1992) J. Virol. 66, 4073–4084
2. Lai, C.-K., and Ting, L.-P. (1999) J. Virol. 73, 114–124
3. Dong, G., Broker, T. R., and Chow, L.-T. (1994) J. Virol. 68, 1115–1127
4. Lehming, N., Tanos, D., Brickman, J. M., Ma, J., Maniatis, T., and Ptashne, M. (1994) Nature 371, 175–179
5. Gray, S., Szymanski, P., and Levine, M. (1994) Genes Dev. 8, 1829–1838
6. Kessling, A. M., Farrer, M., Carritt, B., Kamalati, T., and Buluwela, L. (1997) Hum. Mol. Genet. 6, 203–210
7. Fuentes, J.-J., Pritchard, M. A., Planas, A. M., Bosch, A., Ferrer, I., and Estivill, X. (1995) Eur. J. Hum. Genet. 3, 1935–1944
8. Delabar, J. M., Theophile, D., Rahmani, Z., Chettouh, Z., Blouin, J. L., Prieur, M., and Antinorakis, S. (1999) Genomics 5, 325–331
9. McBride, A. A., Romanczuk, H., and Howley, P. M. (1991) J. Biol. Chem. 266, 2060–2064
10. Olave, L., Reinberg, D., and Bayordo, E. J. (1992) EMBO J. 11, 3941–3945
11. Eilers, M., Yeung, K., Cao, J., Inostroza, J. A., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (1996) Genes Dev. 10, 1035–1048
12. Merleinstein, F., Yeung, K., Cao, J., Inostroza, J., Erdjument-Bromage, H., Engelsohn, K., Landsman, D., Levitt, P., Tempest, P., and Reinberg, D. (1996) Genes Dev. 10, 1035–1048
13. Kang, P., Halle, J.-P., Stelzer, G., Meisterernst, M., and Kawakami, K. (1998) Mol. Cell. Biol. 18, 10–18
14. Li, C., and Manley, J. L. (1998) Mol. Cell. Biol. 18, 3771–3781
15. Ollave, L., Reinberg, D., and Mayes, L. D. (1998) Genes Dev. 12, 1621–1637
16. Ross, J. F., Liu, X., and Dynlacht, B. D. (1999) Mol. Cell. Biol. 19, 855–865
17. Pazin, M. J., and Kadonaga, J. T. (1997) Cell 99, 325–328
18. Wolke, A. P. (1997) Nature 387, 16–17
19. Struhl, K. (1998) Genes Dev. 12, 599–606