Deregulated Direct Targets of the Hepatitis B Virus (HBV) Protein, HBx, Identified through Chromatin Immunoprecipitation and Expression Microarray Profiling*§

Wing-Kin Sung†§1, Yiwei Lu§1, Charlie W. H. Lee‡§1, Dongwei Zhang‡, Mostafa Ronaghi§§*, and Caroline G. L. Lee†§1‡‡2

From the Departments of †§1 Computer Science and §1 Biochemistry, National University of Singapore, Singapore 119077, Singapore, the §§1 Genome Institute of Singapore, Singapore 138672, Singapore, the ‡§1 Division of Medical Sciences, National Cancer Centre Singapore, Singapore 169610, Singapore, the ‡‡2 Department of Biochemistry, Stanford Genome Technology Center, Stanford University, Stanford, California 94305, and the ‡§‡1 Duke-NUS Graduate Medical School, Singapore 169547, Singapore

The hepatitis B-X (HBx) protein is strongly associated with hepatocellular carcinoma. It is implicated not to directly cause cancer but to play a role in hepatocellular carcinoma as a co-factor. The oncogenic potential of HBx primarily lies in its interaction with transcriptional regulators resulting in aberrant gene expression and deregulated cellular pathways. Utilizing ultraviolet irradiation to simulate a tumor-initiating event, we integrated chip-based chromatin immunoprecipitation (ChIP-chip) with expression microarray profiling and identified 184 gene targets directly deregulated by HBx. One-hundred forty-four transcription factors interacting with HBx were computationally inferred. We experimentally validated that HBx interacts with some of the predicted transcription factors (pTF) as well as the promoters of the deregulated target genes of these pTFs. Significantly, we demonstrated that the pTF interacts with the promoters of the deregulated HBx target genes and that deregulation by HBx of these HBx target genes carrying the pTF consensus sequences can be reversed using pTF small interfering RNAs. The roles of these deregulated direct HBx target genes and their relevance in cancer was inferred via querying against biogroup/cancer-related microarray databases using web-based NextBio™ software. Six pathways, including the Jak-STAT pathway, were predicted to be significantly deregulated when HBx binds indirectly to direct target gene promoters. In conclusion, this study represents the first ever demonstration of the utilization of ChIP-chip to identify deregulated direct gene targets from indirect protein-DNA binding as well as transcriptional factors directly interacting with HBx. Increased knowledge of the gene/transcriptional factor targets of HBx will enhance our understanding of the role of HBx in hepatocellular carcinogenesis and facilitate the design of better strategies in combating hepatitis B virus-associated hepatocellular carcinoma.

Hepatocellular carcinoma (HCC)3 is the 5th most common cancer and the 3rd leading cause of cancer deaths in the world, primarily due to late symptom manifestation and unresponsiveness to treatment (1). Chronic hepatitis B virus (HBV) infection is strongly associated with HCC, and the viral X-gene product (HBx) has been implicated to play a major role in the etiology of HCC.

The HBx protein is implicated not to directly cause cancer but to play a role in hepatocellular carcinogenesis as a co-factor or tumor promoter through its pleiotropic functions (see Ref. 2). The long latent period between the initial HBV infection and manifestation of HCC in chronically infected individuals provides ample opportunity for a tumor-initiating event to occur. HBx, which is resident in these chronic HBV carriers, then acts as a tumor promoter, modulating regulatory processes in these cells to facilitate the transformation process. One likely tumor initiating event is DNA damage. We previously observed that the HBx protein sensitizes cells to UV-induced DNA damage causing these cells to exhibit greater G2/M arrest, apoptosis, and reduced DNA repair (3). HBx was implicated to compromise host cell DNA repair mechanisms (4–6) by binding to DNA repair genes (e.g. the UV-damaged DNA-binding protein 1 or DDB1) and interfering with cell-cycle process including S-phase progression and chromosome segregation leading to chromosome instability, a hallmark of many cancers, including HCC (7).

HBx is an enigmatic 154-amino acid protein with pleiotropic functions. Although it was reported not to bind DNA directly, it has been implicated as a promiscuous transactivator acting through two modalities: in the cytoplasm, HBx was suggested to activate mitogenic signaling cascades, whereas in the nucleus, it was thought to interact with transcription factors to modulate gene expression (8).

The mechanism underlying HBx-mediated oncogenicity has been attributed to gene deregulation caused by the transcriptional activities of HBx (9, 10). This viral protein is reported to stimulate the activity of numerous transcription factors

*This work was supported by BioMedical Research Council (BMRC) of Singapore Grant BMRC06/1/21/19/449 and the Singapore Millennium Foundation (to C. G. L.) through the National Cancer Center, Block Grants from the National Cancer Centre, and the DUKE-NUS Graduate Medical School.

The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. S1 and Tables S1–S11.

1 These authors contributed equally to this work.

2 To whom correspondence should be addressed: Division of Medical Sciences, National Cancer Centre, Level 6, Lab S, 11 Hospital Dr., Singapore 169610, Singapore. Tel.: 65-6436-8353; Fax: 65-6372-0161; E-mail: bchlee@nus.edu.sg.

3 The abbreviations used are: HCC, hepatocellular carcinoma; HBV, hepatitis B virus; HBx, hepatitis B-X; ChIP, chromatin immunoprecipitation; PBS, putative binding site; TF, transcription factor; pTF, predicted transcription factors; CREB, cAMP-response element-binding protein; SIWTF, signal intensity weighted TF; siRNA, small interfering RNA; STAT, signal transducers and activators of transcription,
Deregulated Direct Target Genes of HBx

(e.g. NF-κB, NF-AT, AP1, CREB) and interact with various members of the general transcription machinery complex in the cell (e.g. TATA-binding protein, TFIIH, TFIIH, and RPBS) (11–14). As a consequence, HBx deregulates various cellular processes, including signaling pathways, proliferation, DNA repair mechanisms, and apoptosis (2, 15, 16). The modulation of host gene expression and subsequent deregulation of cellular pathways has been thought to be one of the major roles HBx plays in oncogenesis. Unfortunately, observations thus far from these studies offer only a fragmented picture of the mechanism of HBx-mediated oncogenesis as it remains inconclusive whether the observed changes are the direct effects of HBx or are merely indirect downstream “side effects” of HBx expression. To date, the direct gene targets of HBx remain unclear.

A systematic analysis of the genes targeted by HBx will be insightful in understanding the molecular mechanism of HBx-associated HCC. Various groups have embarked on this quest through gene expression profiling using serial analysis of gene expression and cDNA microarray techniques for large scale identification of genes modulated by HBx (10, 17, 18). Although such expression profiling offers invaluable insights into the role of HBx, it is difficult to differentiate if the observed change in gene expression is a result of a direct effect of HBx or indirect domino-like cascading effects whereby the observed altered expression of a particular gene is caused by HBx changing the cellular environment or due to HBx modulating the expression of another gene that in turn changes the expression of the observed gene.

The current state-of-the-art chromatin immunoprecipitation (ChIP) technology enables the identification of candidate genes directly regulated by a factor with critical consideration for the cellular milieu and chromatin environment (19). Major advancement in the study of mammalian gene regulation has been achieved with the application of ChIP profiling to the analysis of transcription factors, and this technique has been successfully employed to determine the target promoters of the E2F (20), c-MYC (21, 22), P53 (23), and RELA (p65) (24) transcription factors.

In addition to its stated advantage over microarray expression profiling, ChIP also offers advantages over in vitro assays such as the electrophoretic gel mobility shift analysis, which is not amenable to global screening. Although electrophoretic mobility shift assay is relatively fast and easy, it is also difficult to utilize the electrophoretic mobility shift assay to detect indirect protein-DNA binding (e.g. HBx binds to another protein that then binds to DNA) or binding by multiprotein complex (25).

As HBx was reported not to bind DNA directly but indirectly through its interaction with another transcription factor, ChIP would thus be a useful technique for the identification of genes directly targeted by HBx via indirect HBx-DNA binding.

Hence, we are interested in identifying both the gene and protein targets of HBx in cells exposed to UV irradiation to simulate a tumor-initiating event and understanding the cellular pathways deregulated by HBx under these conditions. Here, we report the use of a combination of chip-based chromatin immunoprecipitation (ChIP-chip) profiling and expression profiling to identify the direct gene targets that are deregulated by HBx in cells irradiated with UV. We experimentally validated that HBx interacts with some of the predicted transcription factors (pTF) as well as the promoters of the deregulated target genes of these pTFs. Significantly, we demonstrated that the pTF interacts with the promoters of the deregulated HBx target genes and that deregulation by HBx of these HBx target genes carrying the pTF consensus sequences can be reversed using pTF siRNAs. Finally, computational methods were also employed to delineate the cellular pathways deregulated through the modulation of the expression of the identified gene targets by HBx.

EXPERIMENTAL PROCEDURES

Preparation of Samples for ChIP-Chip and Expression Profiling

HepG2 cells were infected with either recombinant HBx adenoviruses (AdHBx) or control adenoviruses (AdEasy) as previously described (3). Seventy-two hours later, infected cells were exposed for 30 s to UVC (254 nm) irradiation with a germicidal lamp calibrated to deliver 8 or 16 J/m². Cells were then harvested for either ChIP-chip profiling or expression microarray profiling.

Chip-based Chromatin Immunoprecipitation (ChIP-Chip)

The ChIP assay kit (Upstate Cell Signaling Solutions/Millipore) was employed. Briefly, proteins and DNA in UV-treated HepG2 or control cells were cross-linked with 1% formaldehyde for 10 min at 37 °C. The cross-linked protein-DNA complexes were then sheared to an average size of 300 bp using the Bioruptor sonicator (Diagenode, Belgium) at medium settings for 12 cycles of 30 s on followed by 30 s off. The sonicated chromatin fragments were then immunoprecipitated using 1 μg of anti-HBX antibodies generated in our laboratory (3). The immunoprecipitated protein-DNA complexes were reverse cross-linked using 5 m NaCl, and the DNA was then recovered from the immunoprecipitate using phenol-chloroform extraction and ethanol precipitation (ChIP DNA). DNA (input DNA) extracted from the total lysate that has been cross-linked (but not immunoprecipitated) and reverse cross-linked was utilized to normalize for differences in the amount of DNA present in the different samples. The ends of the HBx-immunoprecipitated ChIP DNA and input DNA were polished using the End-It™ DNA end repair kit (EPICENTRE® Biotechnologies). Oligonucleotides with the following sequences (oligo 1, 5’-GCG-GTGACCGGAGATCTGATTAC-3’; oligo 2, 5’-GAATTCAGATC-3’) were then annealed to form a double-stranded DNA linker and then ligated to the end-polished ChIP DNA and input DNA. ChIP DNA and input DNA were then amplified using oligo 1 as primer. Conditions for the amplification were as follows. The linker-ligated ChIP DNA/input DNA were then annealed with oligo 1 primer at 55 °C for 2 min. A mixture of Taq and Pfu DNA polymerase was then added and the primer extended to 72 °C for 2 min. The reaction was then heated to 95 °C for 2 min to denature the DNA followed by 22 cycles of 95 °C for 1 min, 60 °C for 1 min, and 72 °C for 1 min. A final extension at 72 °C for 5 min was carried out.

The ChIP DNA and input DNA were end-labeled with Cy5 and Cy3, respectively, and co-hybridized to a Nimblegen ChIP-chip promoter tiling array (NimbleGen Systems, Inc.). This array contained 382,884 probes with a probe length of 50 bp and 21942 JOURNAL OF BIOLOGICAL CHEMISTRY
median probe space of 100 bp tiled 1200 bp upstream and 300 downstream 24,134 human promoters, present in triplicate. DNA end-labeling, hybridization, and scanning were performed at NimbleGen Systems, Inc., which provided the raw data file of foreground signal intensities.

**Expression Microarray Profiling**

The Agilent Whole Human Genome Oligo Microarray (G4112A), containing 44,000 60-mer oligonucleotide probes representing 41,000 unique human genes/transcripts, was employed for profiling of differential gene expression between AdHBx- and AdEasy-infected cells. Isolation of RNA, cRNA generation, probe labeling, and hybridization were performed by Miltenyi Biotech, an Agilent certified center, which provided us with normalized microarray data.

**Identification of Direct Gene Targets of HBx from the ChIP-Chip Data**

From the raw data file obtained from the Nimblegen ChIP-chip promoter tiling array, we searched for regions in which a minimum of five consecutive probes displayed more than 1.5-fold stronger signal intensity over background in HBx-expressing cells only. The same region does not show binding in vector control cells. Those regions are known as candidate HBx binding sites, whereas their corresponding genes are known as candidate HBx target genes.

**Computational Identification of Potential Transcription Factors That Interact with HBx from the ChIP-Chip Data**

From the list of candidate HBx binding sites (L) extracted, potential TF binding motifs were scanned using 500 human TF position weight matrices obtained from TRANSFAC 9.1. A segment within a candidate HBx binding region in L is defined to be a putative binding site (PBS) of some TF if the position weight matrix score is higher than the position weight matrix threshold score. The position weight matrix threshold score is defined as the score in which less than 5% of false positives will be accepted as PBSs of that TF in a set of 5000 randomly chosen sequences, each with a length 5000 bp.

To determine whether a particular pTF is statistically significantly associated with HBx in the list of L, we computed the p value as follows. Initially, we scanned random sequences from the human genome build 17 for PBSs and computed the chance p of the putative TF occurring in the random sequences. If we assume that PBS occurs in L randomly, the number of PBSs occurring in the list L will follow the binomial distribution Bin(m, p, 1 − p), where m is the total number of possible binding sites in the list L. The p value that the putative TF is significantly associated with HBx was calculated as $\sum_{j=k}^{\infty} C_i j(1-p)^{j-i}$, where k is the actual number of motif hits of the TF. A TF is deemed to be significantly over-represented in the list L if its p value is smaller than 0.05.

We also attempted to reduce Type I error by performing a signal intensity weighted TF (SIWTF) search. We reasoned that among the significantly over-represented TFs, those TFs that bind to HBX directly should be close to the probes with the peak probe signal intensity. To perform the SIWTF search, the weight on each position of each binding site was first computed based on the probe signals of that binding site. Positions that have probes are assigned a weight directly proportional to the signal intensity of the probes. Positions that do not have probes are assigned a weight based on the linear interpolation of the signal intensities of the nearest probes. The SIWTF score is then computed by averaging the sum of the positional weights over all the hits. To obtain a p value empirically, we generated 10,000 random hits (same length as the putative TF motifs in question) in the list of binding regions L and computed the SIWTF score for these random hits. The p value for the TF was then computed to be the proportion of the SIWTF scores of the random hits of 10,000 that exceed the SIWTF score of the putative TF. For each significant TF found based on the TRANSFAC scanning, we computed the SIWTF score and the corresponding p value.

**Computational Analyses of Genes That Are Differentially Expressed in AdHBx- versus Control-infected HepG2 Cells from the ChIP-Chip Data**

From the processed normalized data provided by the Agilent certified center, we selected genes that exhibit at least 2-fold differences in gene expression between the AdHBx- and control-infected cells.

**Computational Analyses of Target Genes Deregulated by HBx**

We computationally analyzed the deregulated target genes that displayed differential expression of at least 2-fold when HBx is bound indirectly to its promoter using the NextBio™ search engine. The list of target genes that are bound indirectly to HBx and their -fold difference in expression in HBx cells compared with vector control cells was uploaded into NextBio. To evaluate if these direct target genes deregulated by HBx were associated with HCC, this list of genes was queried against 192 oncology-related expression microarray studies that were available in NextBio (as of December 5, 2007). We also explored the putative functions of these HBx target genes by querying the same list against the major biogroups that were available in the NextBio (December 5, 2007), namely the Gene Ontology (GO) data base, KEGG pathway data base, Reactome Data base, and Broad MSigDB Regulatory Motif Data base.

**Computational Inference of Pathways Deregulated by the Indirect Binding of HBx to Promoters of Genes**

To delineate the pathways that are deregulated by HBx, data from both ChIP-chip and cDNA expression microarrays were analyzed as shown in Fig. 6. The pathways in which the deregulated direct HBx gene targets reside were mapped using the KEGG human pathway data base. These pathways were then ranked computationally taking in consideration not only the pathways where the target genes of HBx reside but also the pathways of the genes downstream of the HBx target genes that were also deregulated as a result of the deregulation of HBx target genes. Consistency in the relationships between genes in the pathway was also taken into consideration.

The pathways in which the deregulated direct HBx gene targets reside were mapped using the KEGG human pathway data base. Of the 175 pathways in KEGG, they were ranked compu-
Deregulated Direct Target Genes of HBx

tationally through the integration of ChIP-chip and expression microarray data as follows.

HBx Target Gene Score—HBx target gene score of a pathway is the ratio of the number of expressed HBx target genes in the pathway over the total number of expressed HBx target genes. Because pathways targeted by HBx are likely to have more deregulated direct HBx target genes, pathways with a higher HBx target gene score are more significant.

Expression Microarray Score—Expression microarray score is the ratio of the number of significantly expressed genes in the pathway over the total number of genes in the pathway. Pathways with higher microarray expression score are more significant.

Pathway Correctness Score—Pathway correctness score is the number of genes in the pathway whose microarray expressions are consistent with the differential expression of the HBx target genes in that pathway. For example, the relationship of a particular pathway shows that A up-regulates B and C. However, from the expression microarray data, A and B are up-regulated in HBx-expressing cells, whereas C is down-regulated. Hence the relationship between A and B and C is consistent, whereas the relationship between A and C or B and C is inconsistent. Pathways with more genes whose differential gene expressions are consistent with the predicted differential expression when HBx binds and deregulates the target genes will be ranked higher.

For each of the 175 pathways in the KEGG pathway database, the three scores (HBx Target Gene Score, Expression Microarray Score, and Pathway Correctness Score) were computed. For each type of score, the mean ± S.D. was determined. Approximating the distribution of each score to a normal distribution, the p values of the three scores of each candidate pathway were then computed. The combined p value of each pathway is the product of the p values of the three scores. Corrected p values represent p values that have been corrected for multiple testing. Pathways were then ranked according to the combined p values and corrected p values.

Experimental Validation of Transcription Factors Predicted to Interact with HBx to Modulate Expression of the Target Genes

Two different approaches were utilized to experimentally validate some of the transcription factors that were predicted computationally to interact with HBx to modulate gene expression.

In the first approach, we utilized the signal transduction AntibodyArray™ (Hypomatrix) to validate that pTF actually interacts with HBx. This array contains 400 immobilized antibodies against proteins important in various signal transduction pathways. Slightly less than 10% of these antibodies that were immobilized in this array hybridized to transcription factors listed in TRANSFAC 9.1. Whole cell extracts from UV-treated HepG2 cells infected with either control or AdHBx recombinant adenoviruses were prepared in RIPA buffer (150 mM NaCl, 1.0% Nonidet P-40, 0.5% deoxycholate, 0.1% SDS, and 50 mM Tris, pH 8.0) as recommended by the manufacturer. Cell debris was pelleted through centrifugation at maximum speed in a desktop microcentrifuge for 15 min at 4 °C. The supernatant containing the whole cell extract was then hybridized to the AntibodyArray™ according to the manufacturer’s instructions. Briefly the AntibodyArray was incubated in blocking solution comprising 5% dried milk in TBST (150 mM NaCl, 25 mM Tris, and 0.05% Tween 20, pH 7.5) for at least 1 h at room temperature with slow shaking. The AntibodyArray was then incubated with the whole cell extract for 2 h at room temperature with gentle shaking. Thereafter the AntibodyArray was washed 3 times with TBST for 15 min. The AntibodyArray was then incubated with horseradish peroxidase-conjugated anti-HBx antibodies at a concentration of 1 μg/ml in TBST for 2 h at room temperature. The AntibodyArray was again washed 3 times in TBST for 15 min, and peroxidase substrate was applied and the blot exposed to x-ray film.

In the second approach, we performed a series of experiments to validate 2 HBx-pTFs: (i) co-immunoprecipitation was performed to demonstrate that HBx interacts with the pTF (see Fig. 2A); (ii) ChIP with antibodies against HBx was utilized to demonstrate the interaction of HBx with the pTF target genes (see Fig. 3A); (iii) ChIP with antibodies against the pTF was utilized to demonstrate that the pTF interacts with the HBx target genes (see Fig. 4A); and (iv) specific inhibition of the pTF using siRNA was employed to demonstrate that HBx deregulates target genes through interacting with transcription factors that bind promoters of target genes (see Fig. 5A).

(i) Co-immunoprecipitation Was Performed to Demonstrate That HBx Interacts with the pTF—HBx co-immunoprecipitated with putative transcription factors, namely SMAD4, E2F1, and YY-1 following the protocol as described previously (26). Briefly, cells were washed twice in ice-cold PBS to remove remaining serum protein from the culture medium. RIPA (RIPA buffer with 1 tablet of protease inhibitor (Roche)) was then added to these cells and the cells were homogenized with a Type B pestle using ~10 repeated strokes. Cell debris was removed through centrifugation at 12,000 × g for 10 min at 4 °C. Fifty microliters of Protein G-agarose suspension (25-μl bed volume) was added and the mixture was incubated for 3 h at 4 °C to remove background caused by nonspecific absorption of irrelevant cellular proteins to Protein G beads. The beads were then pelleted by centrifugation at 12,000 × g for 20 s and the supernatant was transferred to a fresh tube. Two micrograms of HBx antibody were then added to the supernatant and gently rocked for 3 h at 4 °C. Fifty microliters of Protein G-agarose suspension (25-μl bed volume) was then added and the mixture incubated for 3 h at 4 °C on a rocking platform. The immunocomplexes on the beads were then centrifuged at 12,000 × g for 1 min at 4 °C. The proteins that were immunoprecipitated were then electrophoresed on an SDS-PAGE gel and Western blot analyses were performed as previously described (27) using antibodies against SMAD4, E2F1, and YY1 (Santa Cruz) at a dilution of 1 in 1,000.

(ii and iii) Experimental Validation to Demonstrate That HBx and the Predicted Transcription Factors Bind the Same Promoters of the Deregulated Target Genes—Chromatin immunoprecipitation was performed as described earlier on cells infected with either HBx or AdEasy (control) and immunoprecipitated with 1 μg of anti-HBx, anti-SMAD4, anti-E2F1, anti-YY1, or anti-SRF (nonspecific control for HBx ChIP) as well as anti-EGFP (nonspecific control for SMAD4, anti-E2F1, and anti-YY1 ChIPs) antibodies. PCR was then performed to detect
the promoters of deregulated HBx target genes, namely CYP17A and IL17B (targets of SMAD4), SOAT2 (target of E2F1), as well as AICDA and GRIN2D (targets of YY1) of the predicted transcription factors SMAD4, E2F1, and YY1. (Sequences of primers used for PCR are found in supplemental Table S1.)

(iv) Experimental Validation to Demonstrate That HBx Deregulates Target Genes through Interacting with Transcription Factors That Bind Promoters of Target Genes—HepG2 cells were transfected with control siRNA (ctrl siRNA) or siRNA against SMAD4, E2F1, or YY1 (Sigma–Proligo) through electroporation at 180 V for 100 ms. These cells were then infected 24 h later with either the control adenoviruses (CTL) or HBx adenoviruses (HBX). Twenty-four hours later, these infected cells were then UV-irradiated as previously described (3). Total RNA was then isolated from these cells 24 h later using the MirVana™ miRNA isolation kit (Ambion, Austin, TX) and reverse-transcribed using the SuperScript™II RT kit (Invitrogen). Expression of the target genes of SMAD4, E2F1, and YY1 transcription factors, namely CYP17A and IL17B (SMAD4), SOAT2 (E2F1), as well as AICDA and GRIN2D (YY1) were measured using real-time PCR as previously described (28) using the primers shown in supplemental Table S1. The PCR conditions are as follows: 95 °C for 10 min followed by 40 cycles at 95 °C for 1 min, 55 °C for 30 s, and 72 °C for 33 s.

RESULTS

HBx Antibodies Generated in the Laboratory Can Immunoprecipitate Known Targets of HBx—Antibodies against the HBx protein generated in our laboratory performed well on Western blots as well as immunoprecipitation (3). A preliminary proof-of-concept experiment was performed to evaluate if these same HBx antibodies are able to immunoprecipitate known targets of HBx through chromatin immunoprecipitation. As HBx was reported to interact directly with P53 and modulate the transcriptional activity of P53 (3), we evaluated if the HBx antibodies are useful for ChIP of MDM2, a well known gene target of P53. HepG2 cells either expressing or not expressing HBx were cross-linked and nuclear extracts isolated, sonicated, and then immunoprecipitated with either anti-HBX or anti-P53 antibodies. PCR amplification using primers spanning the P53 binding site in the MDM2 promoter was performed on these immunoprecipitated nuclear extracts. As shown in supplemental Fig. S1, as expected the MDM2 promoter can be amplified in control cells when the nuclear extracts were immunoprecipitated with p53 but not with HBx antibodies. However, in HBx-expressing cells, MDM2 promoters can be amplified in cells immunoprecipitated with either P53 or HBx antibodies, suggesting that binding of HBx to the MDM2 promoter is likely via the binding of P53. Thus these HBx antibodies are useful for chromatin immunoprecipitation to identify HBx target genes.

HBx Does Not Bind DNA Directly but Indirectly through Interaction with Various Transcription Factors—Previous isolated studies suggest that HBx does not bind specifically to DNA directly but through the interaction with various transcription factors (29) although HBx was reported to interact with single-stranded DNA nonspecifically (30). Here, we revisit this issue of whether HBx binds directly to promoters of genes using the global ChIP-chip that examines the promoter regions of ~40,000 genes in the human genome. From the ChIP-chip data (ncbi.nlm.nih.gov) (accession number GSE11108), the promoter region in which HBx shows strong binding (>1.5-fold stronger signal intensity over background) to a minimum five consecutive probes was selected and scanned for conserved elements (motifs) that are 6, 8, and 12 bp long using the Weeder algorithm (31). We found no single motif with a significantly high Weeder score, and none of the motifs were similar to any previously identified motifs in the TRANSFAC 10.2 data base (data not shown). Hence it is unlikely that HBx binds directly to DNA in a sequence-specific manner.

To identify the potential transcription factors that bind to DNA when HBx interacts with them, we extracted the binding regions of candidate direct HBx target genes (that display >1.5-fold stronger signal intensity over background to a minimum of 5 consecutive probes) and scanned for potential TF binding motifs. Of 331 unique vertebrate TF motifs examined, a total of 144 TF binding motifs were found to be significantly enriched (p < 0.05) in promoter regions of the HBx direct target genes suggesting that 144 TFs potentially interact with HBx and bind to the promoter regions of HBx target genes. The list of TFs is significantly enriched in direct HBx target genes as listed in supplemental Table S2. When we attempted to reduce Type I error by employing the SIWTF algorithm to identify TF motifs among the significantly over-represented TF motifs that occur close to the probes of the peak probe signal intensity, only 8 TF motifs, SMAD4, ELK1, CP2, MYOD, AHRHIF, GL1, AHR, and ZID, were found to be statistically significant (p < 0.05) (supplemental Table S2, highlighted in italicized bold).

To experimentally validate the binding between HBx and some of the putative TFs predicted computationally, two different approaches were utilized. Only the first approach will be explained in this section. In the first approach, we employed the signal transduction AntibodyArray™ (Hypomatrix) in which 400 antibodies against proteins important in various signal transduction pathways are immobilized. Of the 400 antibodies immobilized in this array, 33 of these are against TFs whose motifs are listed in the TRANSFAC 9.1 data base that we used to predict TFs interacting with HBx (supplemental Table S3).

Twenty-two and 2 of these immobilized antibodies were against TFs predicted to bind significantly to HBx before and after employing the SIWTF algorithm, respectively. As evident in supplemental Table S3, 100% (2/2) of the TFs predicted to significantly bind HBx after employing the SIWTF algorithm was found to interact with HBx in this array in three independent experiments. Interestingly, even without employing the SIWTF algorithm, >95% (21/22), ~82% (18/22), and ~60% (13/22) of the predicted HBx-interacting TFs were shown to bind HBx in this array in at least one, two, and all three independent experiments, respectively. These experimental data suggest that for TFs that were predicted to bind HBx, it may not be necessary to correct for Type I error as that may result in an increase in Type II errors.

Eleven of the 33 immobilized antibodies against TFs listed in the TRANSFAC 9.1 data base were predicted not to significantly bind to HBx (supplemental Table S3). Curiously, only ~55 (6/11), ~36 (4/11), and ~27% (3/11) of these TFs were...
found not to interact with HBx in this signal transduction AntibodyArray in at least 1, 2, and all 3 independent experiments, respectively. An unusually high percentage (~45% or 5/11) of these TF was found, in all 3 independent experiments, to interact with HBx using this array suggesting high Type II errors. A search through the literature also reveals that among the 144 TF predicted to significantly bind HBx, some of these were consistent with previous reports (e.g. ATF-2/CREB) (11) (supplementary Table S3). However, there were others including TATA-binding protein (13) and p53 (3, 32) that were previously reported to interact with HBx but were not predicted to significantly bind HBx in this study. Additionally, we observed that although both STAT5a and STAT5b were not predicted to bind significantly to HBx, the results from the signal transduction AntibodyArray showed that HBx interacts with STAT5b but not STAT5a (supplementary Table S3) in all 3 independent experiments, which is consistent with a previous observation that HBx activated STAT5b but not STAT5a (33). One possible explanation for the high false-negatives in this study may be due to the limitation of the ChIP-chip array that was utilized in this current study. The actual binding sites for these false-negative TFs may reside in a region that is beyond the ~1.5-kb region of the promoter examined in the current ChIP-chip array and hence not identified. For example, P53 binding sites in the P21 and MDM2 genes were found to be located upstream from the regions of promoters that were examined in this ChIP-chip array although the interaction between p53 and HBx was also evident in 2 of 3 experimental replicates on the signal transduction AntibodyArray (supplementary Table S3) and known P53 target genes, e.g. MDM2 can be amplified when nuclear extracts from HBx cells were immunoprecipitated with HBx antibodies (supplemental Fig. S1). This limitation may be overcome by using the ChIP-chip array that examines longer regions of the promoter of genes or using ChIP sequencing that examines the whole genome. Nonetheless, this study demonstrates that HBx significantly interacts with a wide variety of TF, which then binds to regions within 1.5 kb of the promoter of genes. Many of these TFs that we identified have never been reported to bind to HBx and represent novel proteins that interact with HBx.

**Direct Gene Targets That HBx Deregulate through Indirect HBx-DNA Binding**—We integrated data from ChIP-chip and the expression microarray (accession number GSE11108) to identify direct gene targets that are deregulated by HBx through indirect HBx-DNA binding. Of a total of 39,369 promoters of genes that were represented on the ChIP-chip array, 971 promoters (~2.5%) were found to show significant binding to HBx as evidenced by ≥5 consecutive probes showing >1.5-fold stronger signal intensity over background. Fig. 1 shows exam-
Deregulated Direct Target Genes of HBx

AUGUST 14, 2009• VOLUME 284 • NUMBER 33
JOURNAL OF BIOLOGICAL CHEMISTRY

FIGURE 2. HBx interacts with the pTFs. Co-immunoprecipitation (IP) using either anti-HBx or control antibodies (Ab) as described under “Experimental Procedures” was utilized to demonstrate the interaction between HBx and the following pTFs, namely E2F1 and SMAD4. A nonspecific rabbit IgG was used as the control antibody. A, schematic illustrating the various techniques employed to validate that the various pTFs interact with HBx and their respective predicted target genes to modulate the expression of these target genes. The circled information represents the technique employed to show the interaction between HBx and the pTF. B, Western blot (WB) of cells immunoprecipitated with anti-HBx and probed with anti-E2F1 antibody (first panel), anti-SMAD4 antibody (third panel), and anti-YY1 antibody (fifth panel). The second, forth, and sixth panels show the same blots probed with anti-HBx antibody.

Co-immunoprecipitation was employed to validate the interaction between HBx and the 3 above-mentioned pTFs (Fig. 2A). As shown in Fig. 2B, HBx was found to co-immunoprecipi-
Deregulated Direct Target Genes of HBx

**FIGURE 3.** HBx interacts with the promoters of the selected deregulated indirect gene targets. ChIP using anti-HBx antibodies as described under “Experimental Procedures” was employed to validate the interaction between HBx and the promoters of selected deregulated indirect gene targets. The gene targets selected represent the gene targets of the two pTFs that we demonstrated in Fig. 2 to interact with HBx. A, a schematic illustrating the various techniques employed to validate that the various pTFs interact with HBx and their respective predicted target genes to modulate the expression of these target genes. The circled information represents the technique employed to show the interaction between HBx and promoters of the deregulated indirect target gene. B–D, nuclear extracts of cells infected with control (Control) or HBx (HBx) adenoviruses were immunoprecipitated with HBx. Additional controls include HBx-infected cells that were also immunoprecipitated with no antibodies (No Ab) or a nonspecific antibody, anti-SRF (Ns). Deregulated indirect HBx target genes, SOAT2 (B), CYP17A1 (Ci), IL17B (Cii), AICDA (Di), and GRIN2D (Dii), that were co-immunoprecipitated were amplified using primers around the respective putative binding sites of the pTFs (E2F1, SMAD4, and YY1) that HBx was predicted to act through. The amplification product was quantitated using real-time PCR. Data are shown as mean ± S.E. from three independent experiments (* denotes p < 0.05).
Introduction of siRNA against E2F1, SMAD4, and YY1 significantly decreased the expression of the respective pTFs (Fig. 5B). As evident in Fig. 5, C–E, in the presence of ctrl siRNA, HBx-expressing cells showed significantly higher expression (p < 0.05) of SOAT2 and CYP17A1, and significantly lower expression (p < 0.05) of IL17B, AICDA, and GRIN2D compared with vector-infected cells, consistent with the expression microarray data. However, when siRNA against E2F1, SMAD4, and YY1 were introduced, there were no significant differences in the expression of SOAT2, CYP17A1, IL17B, AICDA, and GRIN2D in both the CTL and HBx-infected cells (Fig. 5, C and D). These series of data strongly suggest that the direct target genes of HBx are modulated by HBx through its interaction with transcription factors binding the promoters of these genes.

**Computational Characterization of the Direct Target Genes That HBx Deregulates through Indirect HBx-DNA Binding**—Of the 184 direct target genes deregulated by HBx, only WNT1 has previously been reported to activate Wnt/β-catenin signaling through HBx (34). None of the other genes were previously reported to be associated with HBx suggesting that most of the previous studies did not examine targets that are directly affected by HBx through indirect binding of HBx to their promoter region. We thus sought to gain a better understanding of these 184 HBx direct target genes by querying these target genes against several expression microarray studies as well as major biogroup databases using the web-based software develop-
Deregulated Direct Target Genes of HBx

To evaluate if these target genes were significantly differentially expressed in cancers, particularly HCC, 192 cancer-related expression microarrays available in NextBio software were queried. NextBio computes correlations between our study and the other studies using a rank-based algorithm based on an inverse log of a p value computed for the study versus our study. Only three of these studies examined hepatocellular carcinoma. Two of these studied HCC in mice and only one study examined HCV-associated HCC in humans, but none displayed statistical correlation in their expression profiles compared with HBx target gene expression profiles. Nonetheless, the expression profile of the sole HCC study in humans (35) was positively correlated with the expression profiles of HBx target genes (supplemental Table S7). When the livers from very advanced hepatocellular carcinoma patients were compared

FIGURE 5. Deregulation of target genes by pTFs can be reversed with siRNA against the pTFs. A, schematic illustrating the various techniques employed to validate that the various pTFs interact with HBx and their respective predicted target genes to modulate the expression of these target genes. The circled information represents the technique employed to validate the deregulation of target genes, which can be reversed using siRNA against their respective pTFs. B, Western blot showing E2F1, SMAD4, and YY1 protein expression in cells expressing siRNA against E2F1, SMAD4, and YY1, respectively (right lane), compared with cells expressing control siRNA (left lane). Relative expression was measured with reverse transcription real-time PCR of SOAT2 (C), CYP17A (Di), IL17B (Dii), AICDA (Ei), and GRIN2D (Eii) in cells infected with either control adenoviruses or HBx adenoviruses and transfected with either control siRNA (Ctrl) or siRNA against the pTFs. Data are shown as mean ± S.E. from three independent experiments (* denotes p < 0.05).
with normal liver, ~62% (41/66) of the common genes were positively correlated with the expression profiles of HBx target genes. The correlation was much weaker in cirrhotic liver without hepatocellular carcinoma versus normal liver or in dysplastic liver tissue versus normal liver of the same study. Although not statistically significant, these observations suggest that the list of HBxs direct target genes may be associated with very advanced HCC and is thus worth further investigation.

Using the NextBio search engine, the regulatory motifs at the promoters of these target genes deregulated by HBx were also explored by querying 173 regulatory motifs from the Broad MSigDB Regulatory Motif database. Eighteen regulatory motifs, of which 50% corresponded to unknown motifs, were found to be significantly enriched (p < 0.01) at the promoters of HBx target genes that were generally down-regulated (supplemental Table S8). Notably, nearly 80% (7/9) of the known transcription factor binding motifs represent transcription binding sites that were computationally predicted to be significantly bound by HBx from the ChIP-chip data (supplemental Tables S2 and S8).

To understand the functions of these direct target genes that are deregulated by HBx, the GO database within NextBio was queried, and HBx target genes seem to be enriched significantly (p < 0.001) in 55% (11/20) of the GO groups in the Biological Process category, 17.6% (3/17) of the GO groups in the Cellular Component category, and 22.2% (4/18) of the GO groups in the Molecular Function category (supplemental Table S9). The most significant GO group in the biological process was the cellular process in which HBx target genes were generally down-regulated in 26.8% (12/42) of the different GO subconcepts (p < 0.0001) (supplemental Table S9). Notably, HBx seems to generally deregulate target genes involved in cell communication; carboxylic, organic acid, and amino acid metabolism; and cell–cell signaling as well as signal transduction (supplemental Table S9). In the Cellular Component category, the most significant GO group in which the HBx target genes reside is the cell part. Interestingly HBx seems to deregulate target genes in 8 of 51 different GO subconcepts (p < 0.0001), all of which are associated with membrane or plasma membranes (supplemental Table S9). The most significant GO group in the Molecular Function category is binding, whereby the HBx gene seems to generally down-regulate target genes involved in calcium ion, cation, metal ion, as well as ion binding (p < 0.0001). Thus, it seems that the HBx gene directly deregulates cellular genes primarily involved in cell communication, metabolism (primarily amino acid), cell signaling, signal transduction, and ion (particularly calcium, cation, and metal) binding.

The REACTOME database was also queried using the NextBio search engine to identify core pathways and reactions that may be deregulated by HBx through its binding to the promoters of its target genes. Of 34 different categories in the REACTOME database that were examined, HBx target genes seem to be weakly enriched in five categories (p < 0.05) (supplemental Table S10). The two most significant categories in the REACTOME data base affected by HBx are Energy Metabolism and Metabolism of Nitrogenous Molecules, which is consistent with observations from the GO database.

Finally, the KEGG data base was also queried using the NextBio to identify cellular pathways in which these gene targets of HBx reside. Of the 202 KEGG pathways that were queried, HBx target genes reside only within 68 KEGG pathways. Direct target genes deregulated by HBx were significantly enriched (p < 0.0001) in six pathways, namely cell adhesion molecules, the Toll-like receptor signaling pathway, the Jak-STAT signaling pathway, D-arginine and D-ornithine metabolism, regulation of autophagy, and the calcium signaling pathway (supplemental Table S10).

Significant Pathways Affected by HBx Deregulation of Direct Target Genes—To computationally delineate pathways that are deregulated when HBx binds indirectly to the promoters of the target genes, data from ChIP-chip and expression microarray were integrated as shown in Fig. 6. This analysis takes into account not only the pathways in which HBx target genes reside but also pathways of genes downstream of the HBx target genes that were also deregulated as a result of the indirect domino-like cascading effects of the deregulation of HBx direct target genes.

Of the 184 direct gene targets of HBx, only 50 (~27%) could be mapped to the 68 pathways of the KEGG human pathway database. The list of the 50 direct target genes, their differential gene expression, the TFs that were predicted to significantly bind to them, as well as the pathway in which they reside are shown in supplemental Table S6. These 68 pathways were then computationally ranked through integrating data from the ChIP-chip and expression microarray taking into account pathways of the deregulated HBx gene targets as well as deregulated genes downstream of the HBx target genes and the consistency of the relationship between the genes in the pathway.

As shown in supplemental Table S11, six pathways seem to be significantly (p < 0.01) affected after correction for Type I error when HBx binds indirectly to the promoters of the direct target genes. These six pathways, the Jak-STAT signaling pathway, D-arginine and D-ornithine metabolism, neuroactive ligand-receptor interaction, cytokine-cytokine receptor interaction, cell adhesion molecules, and calcium signaling pathways, were among the top 17 pathways in which HBx target genes reside (p < 0.01) (supplemental Table S10), although ranking of the pathways changed when genes downstream of the HBx target genes that were deregulated and the consistency of the relationship between genes in the pathway were taken into consideration.

DISCUSSION

In this study, we employed a strategy, illustrated in Fig. 6, that facilitated the identification of TFs that interact with HBx as well as gene targets that were directly deregulated by HBx through indirect HBx-TF DNA binding. We also provided evidence to demonstrate that HBx interacted with the pTFs (Fig. 2), that HBx and the pTF interacted with the promoters of the deregulated target genes (Figs. 3 and 4), and that it is the interaction of HBx with the transcription factor that deregulates the expression of the direct HBx gene targets because inhibiting the expression of the pTFs using siRNAs resulted in the abolition of the deregulation of the gene expression by HBx of HBx direct target genes whose promoters contain the pTFs consensus
binding sequences (Fig. 5). As evident from our results, our strategy also has the potential to delineate significant pathways that are deregulated by HBx through its interaction with TFs and subsequent deregulation of its direct gene targets.

The HBx protein has long been touted as a promiscuous transcriptional regulator (36). Indeed, we found that HBx interacts with a large number of proteins/transcription factors and deregulates a large number of genes. Only a fraction of the transcription factors predicted to interact with HBx in this study has been previously studied, and all studies, except one, reported that these TFs either interacted with HBx or are activated by HBx, which was consistent with our observations. Curiously, MYOD was previously reported not to interact with HBx through in vitro protein-protein interaction assays with glutathione S-transferase and glutathione S-transferase-X fusion proteins (37). However, our data predicted that MYOD interacted with HBx both before \((p = 1.14 \times 10^{-2})\) and after \((p = 0.03351)\) employing the SIWTF algorithm (supplemental Tables S2 and S3). We also demonstrated the interaction between MYOD and HBx using the signal transduction AntibodyArray in three independent experiments (supplemental Table S3). A possible explanation is perhaps that MYOD interacts with HBx under specific conditions, e.g. stabilizing UV irradiation and the HBx protein. Hence, it would be interesting to revisit the association between HBx and MYOD. In addition to the list of TFs that are predicted to bind to HBx, there were other TFs that were predicted not to bind significantly to HBx but were found to bind to HBx using the signal transduction AntibodyArray or in previous publications (supplemental Table S3). This is likely due to the limitations of the ChIP-chip array, which we used that examined only \(\sim 1.5\) kb of promoter regions, and some of the TF binding sites may reside farther upstream on the promoters. This limitation can be readily overcome using chip arrays that examine longer regions of the promoters or genome-wide chip arrays or employing massively parallel sequence-based identification of the binding regions.

Our study identified 184 direct cellular gene targets that HBx deregulates through binding indirectly to the promoter of these gene targets. This observation that HBx can interact with such a large number of different transcription factors to deregulate so many different genes is interesting and further highlights the highly pleiotropic nature of HBx. Despite numerous reports about diverse genes in cells being affected by HBx, only one of these 184 direct gene targets of HBx that we identified has been previously associated with HBx (34), highlighting that most of those genes associated previously with HBx were likely due to indirect effects of HBx.

We thus utilized the NextBio search engine to obtain a glimpse into the function of these direct target genes as well as their potential roles in carcinogenesis. Although not statistically significant, it is interesting that these deregulated target genes of HBx were found to be enriched to a greater extent in very advanced HCC than in cirrhotic liver without HCC or in dysplastic liver tissue (supplemental Table S7).
Deregulated Direct Target Genes of HBx

Our present understanding of the functions of human genes and their relationships to each other remains incomplete. The current curated gene sets and pathways are fragmentary and do not comprehensively represent all functionally related gene cohorts in the human genome. A large number of genes in the human genome remain uncharacterized or poorly characterized. Of the 184 direct gene targets that HBx deregulates, only 50 (<30%) of these genes were sufficiently known to be mapped to the KEGG human pathway database. Thus to infer the potential functions of these HBx target genes, the GO database, which computationally predicts gene function based on studies of human orthologs and paralogs in model organisms, was also queried. More than 80% (149) of the deregulated target genes of HBx could be mapped to at least one GO term. The genes that HBx directly deregulates seem to be involved in cell communication, metabolism (primarily amino acids), cell signaling, signal transduction, and ion (particularly calcium, cation, and metal) binding. The 50 direct HBx gene targets that could be mapped to the KEGG pathway database was found to be significantly (p < 0.0001) associated with six pathways. Three of these pathways, cell adhesion molecule (38), the Jak-STAT signaling pathway (39), and the calcium signaling pathway (9), have been previously implicated as pathways that HBx deregulates.

To further delineate the pathway(s) deregulated by HBx when it binds to the transcription factors at the promoters of its direct target genes, data from both the ChIP-chip and expression microarray were integrated to take into account deregulated genes downstream of the deregulated direct target genes of HBx as well as the consistency of the relationships between genes within the pathways. This analyses also yielded six pathways that were significantly (p < 0.01) deregulated by HBx after Type I error correction, four of which were also among the top six significant (p < 0.0001) pathways in which the deregulated target genes of HBx reside. All three previously reported cell adhesion molecule, Jak-STAT, and calcium signaling pathways were also found to be significantly deregulated when genes downstream of the target genes as well as the relationship between the genes were taken into consideration. Interestingly, the cytokine-cytokine receptor interaction pathway, which was previously implicated to be deregulated by HBx (40) and was ranked the 10th pathway (based on p value significance) in which direct HBx target genes reside, is now ranked the 4th pathway when HBx deregulated downstream genes and the relationships between genes in a pathway are taken into consideration. These results reaffirm the validity of this strategy to identify significant pathways directly deregulated by HBx.

In summary, this present study represents the first ever to utilize an integration of ChIP-chip and expression microarray profiling to identify both the direct deregulated gene targets of a protein, which does not bind directly to DNA (HBx), and the transcription factors that this protein acts through to deregulate the direct gene targets. The interaction between HBx and some of these transcription factors has also been validated. The feasibility of identifying significantly meaningful pathways deregulated by HBx was also demonstrated. The complete and comprehensive identification of all transcription factors interacting with HBx is currently limited by our incomplete knowledge of all the transcription factors in the human genome and the consensus motifs on the promoter that they bind. Additionally this study may not identify all the pathways or the truly significant pathway affected by HBx due to our limited knowledge of the function of all the genes in the human genome as well as the pathways in which they reside. As we gain increased knowledge of the functions of more genes within the human genome as well as the inter-relationships between genes and the pathways in which these genes reside, our understanding of the significant pathways deregulated by HBx will be refined. Nonetheless, increased knowledge of the gene and transcriptional factor targets of HBx identified in this study will enhance our understanding of the role of HBx in hepatocellular carcinogenesis and facilitate the design of better strategies in combating HBV-dependent HCC.

Acknowledgments—We express our sincere gratitude to Wang Yu, Cheryl Chan, and He Huining for invaluable technical assistance.

REFERENCES

Downloaded from http://www.jbc.org/ by guest on August 16, 2017
Deregulated Direct Target Genes of HBx

Deregulated Direct Targets of the Hepatitis B Virus (HBV) Protein, HBx, Identified through Chromatin Immunoprecipitation and Expression Microarray Profiling

Wing-Kin Sung, Yiwei Lu, Charlie W. H. Lee, Dongwei Zhang, Mostafa Ronaghi and Caroline G. L. Lee

doi: 10.1074/jbc.M109.014563 originally published online May 13, 2009

Access the most updated version of this article at doi: 10.1074/jbc.M109.014563

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2009/06/04/M109.014563.DC1

This article cites 40 references, 19 of which can be accessed free at http://www.jbc.org/content/284/33/21941.full.html#ref-list-1