Cloning, chromosomal localization and promoter analysis of the human transcription factor YY1

Ya-Li Yao, Barbara R. Dupont¹, Subir Ghosh+, Yong Fang, Robin J. Leach¹ and Edward Seto*

H. Lee Moffitt Cancer Center and Research Institute, Department of Microbiology and Immunology, Department of Biochemistry and Molecular Biology, College of Medicine, University of South Florida, 12902 Magnolia Drive, Tampa, FL 33612, USA and ¹Department of Cellular and Structural Biology, University of Texas Health Science Center, San Antonio, TX 78284, USA

Received April 9, 1998; Revised and Accepted July 2, 1998 DDBJ/EMBL/GenBank accession no. AF047455

ABSTRACT

Yin Yang 1 (YY1) is a protein that activates and represses transcription of a large number of cellular and viral genes. In addition, studies suggest that YY1 may play an important role in development and differentiation. Here, we report the isolation and analysis of a YY1 genomic clone from a lambda human liver library. Fluorescence in situ hybridization with the YY1 clone has localized the YY1 gene to chromosome 14 band q32. A major YY1 gene transcription initiation site has been mapped to 478 bp upstream of the ATG translation start site. The proximal promoter contains multiple Sp1 transcription factor binding sites but lacks a consensus TATA or CCAAT box. Transient transfections and detailed deletion analyses localized the promoter to no more than 277 bp upstream from the major transcription start site. Finally, we have found that overexpression of the adenovirus E1A protein represses expression of a reporter gene directed by the YY1 promoter.

INTRODUCTION

Yin Yang 1 (YY1, also known as δ, NF-E1, UCRBP and CF1) is a ubiquitously expressed 68 kDa zinc-finger transcription factor that may play an important role in the regulation of many cellular and viral genes through the consensus cis recognition sequence CCCTAAT (for comprehensive reviews, see 1 and 2). Shortly following the cloning of YY1, it was revealed that YY1 functions both as a transcriptional repressor and as a transcriptional activator. In addition, in the absence of the TATA-binding protein (TBP), YY1 together with TFIIB and RNA polymerase II was sufficient to correctly initiate transcription from a supercoiled DNA template in vitro (3). Several non-mutually exclusive models have been proposed to explain the multifunctional nature of YY1, and many of these models are based on the large number of cellular factors that have been reported to physically interact with YY1 (reviewed in 1).

Despite the large number of genes that have been found to be potentially regulated by YY1 and the increasingly large number of proteins that are claimed to interact with YY1, little is known concerning how YY1 itself is regulated. Although YY1 is generally regarded as a ubiquitous protein, since the YY1 mRNA and protein are present in many different tissues and cell types, there are situations where YY1 is clearly regulated. For example, expression of YY1 mRNA in NIH3T3 cells has been shown to be affected by cell density and growth factors such as IGF-1 (4). YY1 is also down-regulated in F9 cells following long-term treatment with retinoic acid (5). Levels of YY1 activity also change during myoblast differentiation and during aging (6,7). Therefore, understanding the mechanisms of regulation of YY1 may provide important clues to understanding the mechanisms of how YY1 activates and represses many genes. As a first step toward understanding how YY1 may be regulated, we report here the isolation of the human YY1 gene, chromosomal mapping of the human YY1 gene, and a systematic analysis of the human YY1 promoter.

MATERIALS AND METHODS

Isolation of genomic clones

Using a standard protocol (8), a human liver genomic library (ATCC 37333) in Charon 4A was screened with a 32P-labeled EcoRI/PstI human YY1 cDNA fragment (nt 1–315; 9). λ phage DNA was amplified and purified from two positive clones, digested with different sets of restriction enzymes, separated on agarose gels and analyzed by Southern blots (8). A 4.5 kb EcoRI fragment was subcloned into a pGEM7zf(+) vector (Promega) for subsequent analysis.

Fluorescent in situ hybridization (FISH)

FISH analysis was carried out with a 12 kb plasmid clone containing the human insert for YY1 using a modification of

*To whom correspondence should be addressed. Tel: +1 813 979 6754; Fax: +1 813 979 6700; Email: setoe@moffitt.usf.edu

+Present address: Geraldine Brush Cancer Research Institute, California Pacific Medical Center, San Francisco, CA 94115, USA
previously described methods (10). Metaphase spreads for FISH analysis were prepared by the standard methods with the addition of ethidium bromide to produce prometaphase spreads (11,12). The YY1 probe was labeled with biotin-14-dATP by nick translation (Gibco/BRL). The labeled probe (60 ng) was mixed with hybridization buffer and human Cot-1 DNA and was denatured for 5 min at 70°C. The probe was immediately added to denatured slides and incubated overnight. The slides were washed in 50% formamide, 2× SSC (pH 7.0) at 42°C, followed by washing in 2× SSC (pH 7.0). The biotin-labeled YY1 probe was fluorescently-tagged with avidin-FITC (Vector Laboratory). Signal amplification was performed using biotin-labeled anti-avidin (Vector Laboratory) followed by avidin-FITC. The chromosomes were counterstained with DAPI and were examined using a Zeiss Axioskop fluorescent microscope equipped with FITC, DAPI and triple band pass filter sets. Digital images were captured by computer using Applied Imaging Probevision software (Pittsburgh), and photographs were printed on a Kodak XL 7700 color image printer.

Plasmids

pGL2-Basic (Promega) contains a firefly luciferase gene as a reporter and lacks eukaryotic promoter and enhancer sequences. p–3600Luc was constructed by isolating an EcoRI/NcoI fragment from the 4.5 kb λ clone described above, treating the fragment with Klenow polymerase to create blunt ends, and then ligating the fragment to a pGL2-Basic vector cut with SmaI. To obtain plasmids containing 5′ progressive deletions of the YY1 promoter linked to the luciferase reporter gene, p–3600Luc was digested with NsiI/StuI and subjected to exonuclease III digestion. pRL-TK (Promega) contains a cDNA encoding Renilla luciferase under control of the herpes simplex virus thymidine kinase promoter. Plasmids pCMV12S, pCMV13S and pCMV-YY1 have been previously described (13,14). pCMV12S directs expression of the E1A 243R protein, and pCMV13S expresses the E1A 289R protein.

DNA sequencing

p–3600Luc and promoter deletion derivatives served as templates to obtain the complete sequence of one strand of the human YY1 promoter using the dideoxy chain-termination sequencing method with GLprimer 1 (Promega). Custom-designed oligodeoxynucleotide primers and GLprimer 2 (Promega) were used to obtain the sequence of the complementary strand.

Primer extension

Primer extension reactions were carried out essentially as previously described with minor modifications (8). An antisense oligonucleotide corresponding to the sequence from position +76 to +100 of the human YY1 promoter was synthesized and end-labeled with [γ-32P]ATP and T4 polynucleotide kinase. Total RNA from HeLa cells was isolated using the acid phenol–guanidinium thiocyanate method (15). The oligodeoxynucleotide primer (105 c.p.m.) was mixed with 10 µg of the HeLa total RNA and was ethanol-precipitated. The DNA–RNA mixture was then redissolved in 30 µl of hybridization buffer [40 mM PIPES (pH 6.4), 1 mM EDTA (pH 8.0), 0.4 M NaCl and 80% formamide], denatured at 85°C for 10 min, and annealed at 30°C overnight. The annealed hybridization mixture was then ethanol-precipitated, washed, and redissolved in 20 µl of reverse transcriptase buffer [50 mM Tris–Cl (pH 7.6), 60 mM KCl; 10 mM MgCl2, 1 mM each of dATP, dCTP, dGTP and dTTP; 1 mM dithiothreitol, 1 U/µl RNase inhibitor and 50 µg/ml actinomycin D]. Subsequently, 50 U of avian myeloblastosis virus reverse transcriptase was added to the mixture and incubated for 1 h at 37°C. The reaction mixture was then ethanol-precipitated, washed, and redissolved in 20 µl of hybridization buffer.

Figure 1. Chromosomal localization of the human YY1 gene. Metaphase chromosomes hybridized with a human genomic probe. Arrows indicate specific hybridization signals. Chromosomes are counterstained with DAPI.
Figure 2. Determination of the 5' end of human YY1 transcripts. Purified total RNA from HeLa cells (lane 5) or tRNA from yeast (lane 6) were used as templates. Primer extension analysis was done using a 25 bp 32P-labeled antisense oligonucleotide with AMV reverse transcriptase. A genomic sequencing ladder was created in parallel and is shown on the left with the appropriate radiolabeled nucleotides (lanes 1–4). The arrow indicates the most likely start site of transcription.

virus (AMV) reverse transcriptase was added, and the reactions were incubated for 2 h at 37°C. The reactions were then stopped with EDTA and treated with DNase-free RNase, and extracted with phenol–chloroform. Single-stranded DNA was then recovered by ethanol precipitation, washed, and dissolved in 4 µl of TE (pH 7.4) and 6 µl of formamide loading buffer [80% formamide, 10 mM EDTA (pH 8.0), 1 mg/ml xylene cyanol and 1 mg/ml bromophenol blue]. Samples were heated at 95°C for 5 min and resolved on a 6% polyacrylamide/7 M urea gel. The gel was then dried and images were obtained by autoradiography.

Cell line, transfection and luciferase assays

The HeLa human cervical carcinoma cell line was maintained in Dulbecco’s modified Eagle’s media supplemented with 10% fetal bovine serum and 100 mg/ml penicillin and streptomycin. For each transfection, HeLa cells (3 × 10^5) were seeded in a 60 mm tissue culture dish. After 16–18 h, 10 µg of pGL2-Basic or its derivatives were transfected into HeLa cells using the calcium phosphate precipitation method (16) along with 1 µg of the pRL-TK reporter plasmid, 10 µg of p–3600Luc or p–277Luc and 10 µg of pCMV12S, pCMV13S or pCMVYY1 were used. Forty-eight hours after transfection, cells were harvested and luciferase activity was determined using the Dual Luciferase Reporter Assay system (Promega).

Accession number

The nucleotide sequence data reported in this paper will appear in GenBank, EMBL and DDBJ Nucleotide Sequence Databases under the accession number AF047455.

RESULTS

Chromosomal location of the human YY1 gene

Purified YY1 genomic DNA was used as a probe for FISH localization on human metaphase chromosomes. The analysis of 20 metaphase spreads banded with DAPI localized the YY1 probe to the telomere region of human chromosome 14 at band q32 (Fig. 1). No hybridization was observed on any other chromosome.

Determination of the transcriptional start site of the human YY1 gene

To determine the transcriptional initiation site of the YY1 gene, we used a reverse transcriptase primer extension assay. Primers were designed to span the putative transcriptional start site and then used in extension reactions with total RNA isolated from HeLa cells. Identical reactions were carried out side by side with yeast tRNA as a negative control. The results from one primer revealed a consistently strong signal indicating a major transcriptional start site, and alignment with a dideoxynucleotide sequence ladder from the same primer revealed that the strong band corresponds to a G within a GC-rich region (Fig. 2). We suspect that transcription initiates at the preceding A and that the cap structure of the YY1 mRNA accounts for the staggered ends.

Transcriptional analysis of the human YY1 promoter

An EcoRI/NcoI fragment from one of the genomic clones corresponding to sequences immediately upstream of the ATG translation initiation codon in the YY1 cDNA was subcloned into a pGL2-Basic vector, which contains a luciferase gene without eukaryotic promoter or enhancer elements. This construct, p–3600Luc, when transiently transfected into HeLa cells, gave rise to a 180-fold increase in luciferase activity compared to the reporter vector (Fig. 3).

To determine the YY1 promoter sequence and to delineate the 5’ boundary of the human YY1 promoter, we generated a panel of promoter-luciferase 5’ serial deletion constructs for DNA sequencing and for transfection and luciferase assays. As shown in Figure 3, the p–3600Luc plasmid had a similar level of luciferase activity compared with deletion constructs p–2100Luc, p–1729Luc, p–1514Luc, p–1248Luc, p–1110Luc, p–917Luc, p–478Luc and p–277Luc, with <40% variation in the activities of these constructs. A significant increase in luciferase activity was seen, however, when comparing plasmid p–1514Luc to p–1248Luc, suggesting the existence of a negative regulatory element located between nt –1514 and –1248 of the YY1 promoter. In contrast, a drastic decrease in luciferase activity was observed when comparing plasmids p–277Luc to p+54Luc.
Figure 3. Expression of luciferase enzymatic activity driven by the human YY1 promoter in transient transfected cells. Left panel shows schematic drawings of various fragments of the human YY1 gene 5′ sequences subcloned upstream of the luciferase reporter plasmid pGL2-Basic. The bent arrow indicates the direction of transcription. Reporter constructs were transfected into different cell lines by calcium phosphate co-precipitation, harvested and assayed for luciferase activity. All relative luciferase activities, as shown in the right panel, are normalized with control Renilla luciferase expressions. Data shown represent the average of three independent experiments with standard deviations.

indicating the presence of a positive regulatory element located between positions −277 and +54. A further deletion of the YY1 promoter to position +379 reduced luciferase activity to background level, compared to the pGL2-Basic vector. These results, then, suggest that a minimal sequence of 54 bp, located downstream of the transcriptional start site of the YY1 gene, is sufficient for promoter activity, and this activity is enhanced by sequences between nt +54 and −277 relative to the start site.

Sequence analysis of the human YY1 promoter

The complete DNA sequence of the YY1 promoter is shown in Figure 4A. The human YY1 promoter is remarkably similar to the mouse YY1 promoter (Fig. 4B). It possesses a region rich in GC nucleotides (80%; from –427 to +96) and lacks a TATA box, a typical characteristic of 5′ untranslated sequences of many transcription factors. In addition to the GC-boxes located at positions –57, +29 and +231, which bind transcription factor Sp1 (17), an inspection of the human YY1 promoter sequence revealed several putative binding sites for ubiquitous and lineage-specific transcription factors. Just 20 bp upstream of the transcriptional start site is a binding site for CDP/CUT, which functions as a transcriptional repressor for many genes (18). Further upstream (position −142) is a consensus binding site for CREB/ATF (19) and three Myb binding sites (20) in tandem (positions −736 to −673). A binding site for the mouse tinman homeodomain factor Nkx-2, which has been shown to antagonize YY1’s negative regulatory effects in the cardiac α-actin promoter through competitive binding during myogenesis (21), is found far upstream at position −1522. Close to the Nkx-2 binding site at position −1640 is a binding site for MZF1, a myeloid zinc finger transcription factor (22). Interestingly, the three GC-boxes, the CREB/ATF binding site and the three Myb binding sites are all conserved between the mouse and the human YY1 promoter.

E1A mediated repression of the human YY1 promoter

It has been suggested that the adenovirus E1A protein can relieve YY1-mediated transcriptional repression either directly by associating with YY1 or indirectly through interactions with p300 in the YY1–p300 complex (23–25). Moreover, YY1 is an important regulator of the cardiac α-actin promoter, and E1A can block myogenesis and cardiac-specific gene transcription (26). To examine the possibility that E1A might affect expression of the YY1 gene, we introduced E1A expression plasmids into HeLa cells to examine the effects of overexpressed E1A protein on YY1 transcription. Either expression plasmid pCMV12S or pCMV13S was transiently transfected into HeLa cells together with the human YY1 promoter construct p–3600Luc and an internal control reporter plasmid. Human YY1 promoter activities were measured in luciferase assays and, as shown in Figure 5, YY1 promoter activity was significantly down-regulated by the 12S and 13S E1A proteins.

To examine whether E1A represses expression of the YY1 promoter through upstream sequences or through basal transcriptional machinery, we repeated the experiments with a minimal YY1 promoter, p–277Luc. Similar to p–3600Luc, luciferase activity was repressed by both 12S and 13S E1A with the p–277Luc reporter construct (Fig. 5B). These data suggest that the mechanism of YY1 repression by E1A most likely does not involve upstream sequence-specific DNA binding factors.
Rather, repression probably results from E1A interacting with proteins that modulate basal transcription elements surrounding the YY1 transcription initiation region.

YY1 does not autoregulate its own promoter

Many transcriptional factors are known to be capable of autoregulating their own transcription. Some transcriptional factors bind directly to cis DNA sequences located on their own promoters and activate or repress transcription, while other factors may activate or repress a separate set of transcriptional factors which in turn activate or repress expression of that factor’s own gene. An inspection of the human YY1 promoter did not reveal any YY1-binding consensus sequence. Therefore, we speculated that YY1 might alter expression of transcriptional factors which in turn bind to and regulate the YY1 promoter. As shown in Figure 6, overexpression of YY1 had no significant effect on expression of a reporter gene directed by the YY1 promoter. It is, of course, possible that the intracellular concentration of YY1 is sufficiently high in HeLa cells to negate the effect of an additional source of protein supplied by transfection, and that under other conditions the YY1 promoter might not be refractory to YY1 overexpression. However, at this time, our data suggest that YY1 probably does not autoregulate its own promoter.

**DISCUSSION**

YY1 is an important regulator of gene expression for many cellular and viral genes. In this report we describe the isolation and analysis of a genomic human YY1 clone. Using FISH, human YY1 was mapped to chromosome 14q32. This finding is in good agreement with the previous prediction that the human YY1 gene may be located on chromosomes 10 or 14 based on mouse–human chromosome homology and Southern blot analysis (27). A number of other genes have previously been mapped to this same region including the immunoglobulin heavy chain gene (28), the Usher syndrome 1A gene (29), the tumor necrosis factor-α-inducible gene (30) and the interferon-α-inducible gene (31). The significance of these co-localizations is unknown at this time. However, it is intriguing that, similar to the finding that translocations involving a breakpoint at 14q32 occur frequently in non-Hodgkin’s lymphomas, loss of heterozygosity at or around 14q32 has been found to be significant in ovarian carcinomas, renal cell carcinomas and lymphoblastic leukemias (32-34).

As is typical of many housekeeping genes, deletion analysis of the human YY1 promoter revealed that elements critical for basal transcription reside within a small area surrounding the transcription start site, and the gene segment from –277 to +475 is sufficient for full promoter activity. In addition, the human YY1 promoter does not contain a TATA box, another feature typical of many
Figure 4. (A) DNA sequences upstream of the translational start codon of the human YY1 gene. The major transcriptional initiation site (+1) is indicated by a bent arrow. Transcription factor binding sites are underlined and indicated below each binding site sequence. (B) Comparison of YY1 promoter DNA sequences between mouse and human. The YY1 promoter sequences of human [hYY1 (–741 to +476)] is aligned with the sequences of the mouse [mYY1 (–751 to +434)] (34). Identical nucleotide positions are indicated by double dots.

housekeeping genes. However, unlike many housekeeping genes, the human YY1 gene possesses only one single transcription start site. An especially notable feature of the human YY1 promoter is its similarity to the mouse YY1 promoter (35), with 70% identity in the proximal promoter region between positions –741 and +476 (human) and between positions –751 and +434 (mouse). Also, similar to the mouse YY1 promoter, which requires only a sequence as short as from positions –58 to +32 for activity, the human YY1 promoter is quite active with only the sequence from positions –277 to +475. Therefore, the remarkable sequence conservation between the mouse and human YY1 previously noted actually extends beyond the cDNA. As shown by sequence analysis, many transcription factor binding sites present in the human YY1 promoter are conserved in the mouse YY1 promoter, suggesting that similar conserved mechanisms are used to regulate the expression of the human and mouse YY1 genes. Finally, our observation that there are three Sp1 sites within the region critical for human YY1 promoter activity fits well with the previous finding that an Sp1 binding site plays an important role in the expression of mouse YY1 (35).

One unusual observation from our deletion analysis of the human YY1 promoter is that plasmid p+54Luc, which contains sequences downstream from the transcription start site, still possesses about one quarter of the luciferase activity of the wildtype promoter. Primer extension analysis thus far has not detected any transcription start site surrounding or downstream of position +54 (data not shown). It is possible that very weak transcription start sites may exist downstream of +1, but the assay we used was not sensitive enough to detect these minor start sites. Alternatively, it is reasonable to speculate that the YY1 promoter may contain ‘standby’ start sites that are utilized only when the natural start site is not present.

One of the most interesting observations that emerged from the studies of YY1 is the fact that the adenovirus E1A protein can convert YY1 from a repressor to an activator (9). Lee et al. (25) have elegantly demonstrated that the molecular mechanism underlying this activation and repression by YY1 results from a physical interaction between YY1 and p300. We have shown here that the E1A protein down-regulates the YY1 promoter and that a minimal YY1 promoter is sufficient for E1A responsiveness.
Figure 5. Adenovirus E1A proteins repress transcription by the YY1 promoter. Luciferase assays were done after transient transfections into HeLa cells with p–3600Luc or p–277Luc and pRL-TK reporter plasmids in the presence or absence of plasmids encoding the adenovirus 12S or 13S proteins. Results are presented as the mean of three independent transfections with standard deviations and normalized with control Renilla luciferase expressions.

Figure 6. The YY1 protein does not autoregulate the YY1 promoter. Luciferase assays were done after transient transfections into HeLa cells with p–3600Luc and pRL-TK reporter plasmids in the presence or absence of a plasmid encoding the YY1 protein. Results are presented as the mean of three independent transfections with standard deviations and normalized with control Renilla luciferase expressions.

Although we do not know at this time whether this down-regulation of YY1 by E1A is a direct or indirect phenomenon, our finding nevertheless raises another, although non-mutually exclusive, model of how E1A might regulate transcription through YY1. Perhaps under physiological concentrations of YY1, and depending on the promoters and cells being studied, YY1 represses transcription. However, in the presence of E1A, the expression of YY1 is decreased, and the abnormally low concentration of YY1 results in the activation of YY1-targeted genes. In this case, YY1 might closely resemble the Drosophila Kruppel transcription factor whose ability to activate or repress target genes is dependent on its cellular concentration (36). Unlike E1A, we found that YY1 did not affect transcription of the YY1 promoter. Work is now in progress to determine whether YY1 expression could be regulated by other viral and cellular transcription factors.

In summary, the isolation of a genomic human YY1 gene and the YY1 promoter provides us now with additional tools to study the mechanisms of how YY1 is regulated and eventually will contribute to our knowledge of how YY1 operates in the regulation of many viral and cellular genes.

ACKNOWLEDGEMENTS

We thank Julia Lee and Rosalind Jackson for discussion and critical reading of the manuscript. This work was supported by a grant from the National Science Foundation (MCB-9631067).

REFERENCES