Cloning and Characterization of the 5'-Flanking Region of the Human Transcription Factor Sp1 Gene*

Received for publication, November 28, 2000, and in revised form, March 16, 2001
Published, JBC Papers in Press, April 9, 2001, DOI 10.1074/jbc.M010740200

Marta Nicolás‡§, Véronique Noé‡, Kirk B. Jensen†, and Carlos J. Ciudad‡
From the ‡Department of Biochemistry and Molecular Biology, School of Pharmacy, University of Barcelona, 08028 Barcelona, Spain and the ¶Laboratory of Molecular Neuro-Oncology, Rockefeller University, New York, New York 10021

The 5'-flanking region of the human Sp1 gene was cloned and characterized. Sequence analysis of this region showed the absence of both CAAT and TATA boxes and an initiator element. The proximal promoter of the Sp1 gene is a GC-rich region that contains multiple GC boxes and Ap2 binding sites. The major transcription start site is located 63 base pairs upstream of the translation start site. Transfection experiments demonstrate that all the elements necessary to achieve significant basal transcription activity are located between positions −443 and −20 relative to the translational start. Sp1 and Sp3 proteins bind to the downstream GC box located in the proximal promoter of Sp1. Furthermore, we demonstrate that the Sp1 protein activates Sp1 transcription activity; thus the Sp1 gene is autoregulated.

Transcription factor Sp1 was originally identified by its binding to the multiple GC boxes in the simian virus 40 (SV40) early promoter (1, 2, 3) and the thymidine kinase promoter (4). Sp1 belongs to a small protein family, which is presently composed of Sp2, Sp3, and Sp4. This family contains a highly conserved DNA-binding domain composed of three zinc fingers close to the C terminus and serine-, threonine-, and glutamine-rich domains in their N-terminal regions. The 81-amino acid C2H2-type zinc finger region, which comprises the DNA-binding domain, is the most highly conserved part of the protein. The two glutamine-rich regions (termed A and B) can act as strong activation domains (5, 6). Synergistic activation of promoters by Sp1 through multiple GC boxes also requires a short C-terminal domain (termed D) (7). More recently, an inhibitory domain has been mapped to the N terminus (8). Sp1 can be phosphorylated, a modification that affects its binding to the CCAAT-binding transcription factor NF-YA (22). In addition, Sp1 can interact with cell cycle regulators like p107 (23), E2F (24), and the retinoblastoma protein (25). In the latter case the protein association increases Sp1 transcriptional activity.

The GC-rich boxes bound by Sp1 are also recognized by Sp3. Thus, these two Sp family members compete for DNA binding, and the resulting Sp transcriptional activity depends on a given ratio of Sp1/Sp3. This ratio varies by cell type and is subject to changes in the cell cycle or cellular conditions (26, 27). It is also not clear whether Sp3 acts as an activator or as a repressor of Sp1-mediated transcription. Recent studies suggest that the GC box located in the promoter region of p21 is responsible for trichostatin A-induced promoter activation and that both Sp1 and Sp3 are the mediators of the activation of this GC box (28). Also, the linkage between histone acetylation and Sp1-mediated transcription is not yet clear. It is speculated that Sp1 and Sp3 may form complexes with histone acetylase or deacetylase directly or indirectly. In fact, there is evidence that p300 is required for trichostatin A-induced, Sp1-mediated p21 transcription (29).

The chromosomal localization of Sp1 is 12q13 (30, 31). A partial human Sp1 cDNA was described in 1987 by Tjian and co-workers (Ref. 3; GenBankTM accession number J03133); it encodes the 696 C-terminal amino acids of the human Sp1. More recently, highly conserved full-length Sp1 cDNA from rat (GenBankTM accession number D12768; Ref. 32) and mouse (GenBankTM accession numbers AF022363 and AF062566; Ref. 33) have been cloned. (For reviews, see Refs. 34, 35).

Sp1 is a ubiquitous transcription factor that has been implicated in the activation of a large number of genes. Although the activity of Sp1 is generally believed to be constitutive, it has been shown recently to participate in several cases of regulated gene transcription. There is now evidence that its activity can be modulated, as in the case in differentiation (36, 37), cell cycle (38, 39), and development (40). The function of Sp1 appears to be essential for differentiated cells after day 10 of development. According to results obtained in Sp1−/− mice, thymidine kinase and the methyl-CpG-binding protein 2 genes are key targets of Sp1 (41).

Our interest in this work was to study the regulation of the Sp1 gene itself. With this aim in mind, we cloned the 5' region of the human Sp1 gene, analyzed the putative transcription factor binding sites, and investigated the regulation of this promoter. Most notably, we observed that the Sp1 promoter is positively regulated by its own gene product, the Sp1 protein.
MATERIALS AND METHODS

Cell Culture—293T cells, adenovirus-transformed human embryonic kidney cells expressing SV40 large T antigen originally referred to as 239tsA1609neo, were obtained from the ATCC, and grown in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum (Life Technologies, Inc.). The HeLa human cervical carcinoma and HepG2 human hepatoma cell lines were grown in Ham’s F-12 medium supplemented with 7% fetal calf serum (Life Technologies, Inc.). All cultures were maintained at 37 °C in a humidified 5% CO₂-containing atmosphere.

DNA Cloning and Constructs—Human genomic DNA from five libraries of the GenomeWalker kit (CLONTech) was used to isolate a clone containing a partial genomic DNA sequence of 1006 bp corresponding to Sp1. Each library was made by cutting the genomic DNA with either EcoRV, Scal, DraI, PvuII, or Sphi and ligating each pool of the digested genomic DNA to the GenomeWalker Adaptor. Two PCR amplifications/library were performed. The primary PCR was made with the adaptor primer (AP1) provided in the kit and sp1mRNARev (5’-ACCATTCGTCCTGACTTTGGGC-3’) as the gene-specific primer located at 333 bp from the 5’-end of the known human cDNA sequence J03133. The primary PCR mixture was diluted 1:50, and 1 μl of this dilution was used as the template for a secondary PCR. This secondary PCR was carried out using the nested adaptor primer (AP2) and a nested gene-specific primer Sp1rev2 (5’-GATCTGCGGATGGTGACGCC-3’) located at 68 bp from the 5’-end of the known human cDNA sequence J03133 as shown in Fig. 1B. We isolated two specific bands from agarose gels libraries of approximately 500 and 1000 bp respectively. The 1000-bp fragment was gel-purified and cloned into the pTarget vector (Promega). This new construct was named pSP1PRM.

The insert was sequenced using the T7 Sequencing Kit (Amersham Pharmacia Biotech). The numbers indicated after the primer sequences correspond to the nucleotide position (underlined). The reverse primer followed a similar structure but contained a XhoI restriction site (underlined) in the reverse sequence. The numbers indicated after the primer sequences correspond to the distance in nt from the 5’-end of the sequence in uppercase to the translational start site.

Fwd 1 5’-gcagttcccgggggagctgtgtgtgggggtttcctc-3’ (146 nt)
Fwd 2 5’-gcagttcccgggggagctgtgtgtggggggctgctg-3’ (281 nt)
Fwd 3 5’-gcagttcccgggggagctgtgtgtgggggtttcctc-3’ (357 nt)
Fwd 4 5’-gcagttcccgggggagctgtgtgtgggggttccctc-3’ (757 nt)
Fwd 5 5’-gcagttcccgggggagctgtgtgtgggggtttcctc-3’ (1612 nt)
Rev 5’-gctgccggatacagccgggggggggagctgtgtgtgggggtttcctc-3’ (20 nt)

Database Submission and Searching—The sequences reported in this work were submitted to GenBank obtaining the corresponding accession numbers that are mentioned under “Results”. For sequence comparison, the BLAST 2.0 program (42) was used using the web server at the NCBI (Bethesda, MD). Locus AC021103, a 166,697-bp piece of genomic DNA corresponding to human clone RP11–147A18, was reported by R. H. Waterson from the Washington University Genome Sequencing Center (TIGR) (43). This sequence is formed by ordered pieces (contig) whose sequencing is in progress. This clone is part of the RPCL-11 human BAC library that was made from the blood of one male donor as described by Osoegawa, et al. (43) and was obtained from the Roswell Park Cancer Institute (Buffalo, NY). The putative transcription factor binding sites in the DNA were analyzed using the Transfection Factor Sites (TF-Sites) program (part of the GCG Software package), using the online computer facilities at the EMBO-Net (Madrid, Spain), and using the Mathinspector software using Transfac 4.0

Total RNA and Genomic DNA Preparation—Total RNA was extracted from HeLa, 293T, or HepG2 cells using the Ultraspread™ RNA reagent (Biotecx) in accordance with the manufacturer’s instructions.

Primer Extension—One hundred ng of the 47-mer REV3 primer (5’-CCTCTCATGGTGGCAGTCGAGGGGGGGGCTCTGTCCGGG-3’) complementary to nt 40 to 7 with respect to the ATG was end-labeled with 20 μCi of [32P]ATP (3000 Ci/mmole, Amersham Pharmacia Biotech) using T4 polynucleotide kinase (New England Biolabs). Thirty μg of total RNA either from HeLa, 293T, or HepG2 cells was hybridized with 106 cpm of the labeled primer in a total volume of 10 μl containing 150 mM KCl, 1 mM EDTA, and 10 mM Tris/HCl, pH 8.0. The same amount of yeast RNA was used as a negative control. The hybridization was performed at 80 °C for 5 min followed by 65 °C for 90 min. After hybridization, the reverse transcription reaction was carried out at 37 °C for 60 min upon the addition of 500 μM dNTPs, 10 mM dithiothreitol, 20 units of RNAsin (Promega), 10 mM MgCl2, and 200 units of murine leukemia virus reverse transcriptase (Life Technologies, Inc.) in 50 mM Tris/HCl, pH 8.0. Three μl of the reaction mixture were run on a 5% polyacrylamide 8 M urea sequencing gel in 1x TBE, (0.1 M Tris/ HCl, 90 mM boric acid, 1 mM EDTA, pH 8.3), dried, and exposed to x-ray film with an intensifying screen.

Electrophoretic Mobility Shift Assays (EMSA)—Nuclear extracts were prepared from exponentially growing 293T cells and HeLa cells as described (44) with the following modification: the amount of Triton used for lysing the cells was 0.02% Triton for 293T cells and 0.05% Triton for HeLa cells.

Electrophoretic mobility shift assays were performed as described (38) but using herring sperm as the nonspecific competitor. The probes FOR2, FOR2U, Umut, FOR2D, and Dmut were prepared by PCR using as template the pGL3FOR2 luciferase construct, which contains the 5’-flanking region necessary for the amplifications. The forward primer for FOR2, FOR2U, and Umut was 5’-CGCAACTTGTCTTACACGCCC-3’; for FOR2D it was 5’-AGAGAGGCGCGGCTCTAGG-3’; and for Dmut it was 5’-AGAGAGGCGCGGCTCTAGG-3’. The reverse primer for FOR2, FOR2D and Dmut was 5’-GTCGAAAGGGGTCTCTGCGG-3’. The PCR fragments were gel-purified, end-labeled with T4 polynucleotide kinase (New England Biolabs) using [γ-32P]ATP (3000 Ci/mmole, Amersham Pharmacia Biotech), and used as the probe in gel shift experiments.

In the competition experiments, increasing amounts of unlabeled FOR2U or FOR2D were added to the reaction with the nuclear extract for 15 min before the addition of the FOR2 as the radiolabeled probe. In the supershift experiments, 0.5 μg of rabbit polyclonal antibody (Santa Cruz) against Sp1, which does not cross-react with Sp2, Sp3, or Sp4, or 1 μg of rabbit polyclonal antibody (Santa Cruz) against Sp3 were added to the reaction mixture and incubated on ice for 15 min before the addition of each probe.

Transfections, Co-transfections, and Luciferase Assay—293T cells were seeded into 6-well plates the day before transfection at a density of 1.5 × 104 cells/well in Dulbecco’s modified Eagle’s medium containing 10% fetal calf serum. The medium was renewed 2 h before transfection. Transfection was performed using FuGENE™ 6 (Roche Molecular Biochemicals) for each well, 3 μl of FuGENE™ 6 in 100 μl of serum free Dulbecco’s modified Eagle’s medium was incubated at room temperature for 5 min. The mixture was added to 1.5 μg of each of the promoter deletion constructs. In co-transfection experiments, 500 ng of pGL3, pGL3FOR1, or pGL3FOR2 plus 1.5 μg of Sp1 expression vector was used. The DNA-lipid mixture was incubated at room temperature for 15 min. The mixture was added to the 293T cells for 24 h.

Luciferase activity was assayed 24 h after transfection of the constructs. Cell extracts were prepared by lysing the cells with 500 μl of freshly diluted 1× reporter lysis buffer (Promega). The lysate was centrifuged at 13,000 rpm for 2 min to pellet the cell debris. The supernatants were transferred to a fresh tube and their protein concentration was determined using the Bio-Rad protein assay reagent according to the manufacturer’s protocol. A 10-μl aliquot of the extract was added to 25 μl of the luciferase assay substrate (Promega) and the luminescence of the samples were read immediately on a TD-20/20 Luminometer, in which light production (relative light units) was measured for 10 s. Each transfection was performed in duplicate. Luciferase activity was normalized to cellular protein concentration.

RESULTS

Cloning of the 5’-Flanking Region of the Human Transcription Factor Sp1—Until now, the known sequence of the human Sp1 mRNA lacked the 276 nt downstream of the translational start, taking the rat cDNA sequence for comparison (Fig. 1B). To complete the cDNA sequence and to clone the 5’ region of

1 The abbreviations used are: bp(s), base pair(s); PCR, polymerase chain reaction; contig, group of overlapping clones; nt, nucleotide(s); CREE, cAMP-response element-binding protein.

2 R. H. Waterson, unpublished material.
the gene, antisense primers (sp1mRNArev and Sp1rev2) directed to the known 5′ region of Sp1 were designed as shown in Fig. 1B to walk toward the 5′ region of human Sp1 using the GenomeWalker kit from CLONTECH (Fig. 1A). This method generated a 1006-bp fragment that was cloned and sequenced (GenBank™ accession number AF255681). Comparing this sequence to the known rat Sp1 cDNA, which is 98% identical to the existing human cDNA, we were able to identify the human exon 2 and a section of exon 3. The remaining sequence of our cloned fragment corresponded to a section of intron 1 and the full intron 2 (Fig. 1C). The sequence of intron 2 was identified by comparing the PCR products obtained with either the genomic DNA or RNA from HeLa cells using primers hybridizing to exons 2 and 3 (data not shown). However, our 1006-bp sequence did not contain the 5′-untranslated region or the Sp1 promoter.

Next, we used our fragment as a virtual probe to perform a BLAST comparison using the unfinished High Throughput Genomic Sequences database at the NCBI (Bethesda), finding contig AC021103 corresponding to the Homo sapiens clone RP11–147A18, which contained additional sequence in the 5′-region. The sequence in this contig helped us to complete the sequence of intron 1 and to confirm the exon-intron organization of the 1006-bp fragment of Sp1 reported here. The comparison of our AF255681sequence, the published Sp1 mRNA 3′-end (J03133), and the contig AC021103 allowed us to deduce the genomic structure of Sp1 shown in Fig. 1D. Primers designed from contig sequence AC021103 were used to isolate a DNA fragment containing 1593 bp, which was sequenced, submitted to GenBank™ (number AF261690), and corresponds to the human Sp1 promoter region (Fig. 2).

Mapping of the Putative Cis-acting Elements in the Sp1 Promoter—Sequence AF261690 was subjected to computational analysis (TF sites within the GCG package, Univ. of Wisconsin and Mathinspector software using Transfac 4.0). Putative binding sites for p53, E2F, Sp1, C/EBP (CCAAT-enhanced binding protein), NF-Y (nuclear factor-Y), CREB, Ap2, and Ap1 were found in the proximal region of the promoter. The sequence and the DNA boxes are shown as in Fig. 2.

Identification of the Transcriptional Start Site—The transcriptional start site was determined by primer extension analysis carried out with RNA preparations from HeLa, 293T, and HepG2 cells using a 47-nt labeled oligonucleotide (Fig. 3). In all cases a major transcriptional start site was observed at nucleotide –63 from the translational start. A negative control was performed using yeast RNA. DNA sequencing analysis revealed that the Sp1 gene lacks basal elements such as a TATA box, a CAAT box, or an initiator sequence (45).

Transcriptional Activity of Sp1 Promoter Deletions—To localize the DNA elements that are important for basal and stimulated transcription of the Sp1 promoter, a series of 5′-deletion constructs were generated by PCR and cloned into the promoterless pGL3-basic, a luciferase reporter vector. The resulting constructs, containing 5′-flanking regions from –20 to –1612 relative to the translational start codon (ATG), were transiently transfected into human 293T cells. After 24 h, cell extracts were prepared and luciferase activity was measured (Fig. 4). The index of promoter strength was normalized to the concentration and was compared with the activity of the promoterless pGL3-basic vector.

Transfections of 293T cells with pGL3FOR5, which contained the longest 5′ sequence, yielded a 75-fold increase in promoter activity relative to pGL3 alone. Deletion from nucleotide –1612 to –443 (pGL3FOR3) reduced the Sp1 promoter activity by 50%. Further deletion to nucleotide –281 (pGL3FOR2) was still 25 times higher than the activity of the reporter vector alone. However, deletion to nucleotide –146 (pGL3FOR1) abolished the promoter activity altogether. These data suggest that the sequence between –1612 and –281 contains positive elements that enhance the basal promoter activity and that the sequence between –281 and –146 contains elements that contribute to the basal transcription activity of the Sp1 gene; thus the region from –281 to –146 could be referred as the proximal promoter.

Binding Analysis of Transcription Factors to the Proximal Promoter—Given that the sequence contained in the
pGL3FOR2 construct conferred the basal activity of the promoter, we used this sequence to perform electrophoretic mobility shift assays to characterize and identify the functional DNA-protein binding sites among the putative ones found by computational analysis. Using probe FOR2 (spanning nucleotides 222 to 2920 relative to ATG), which contains two Sp1, three AP2, one C/EBP, and one E2F potential binding sites (Fig. 2), four major bands could be observed using HeLa or 293T nuclear extracts (Fig. 5A).

To better define the protein binding within the major broad band we performed gel shift experiments with FOR2D and FOR2U probes as shown at the top of Fig. 5, C and D. FOR2D contains the downstream GC box, and FOR2U contains the upstream-Sp1 GC box. The FOR2D probe gave rise to one broad upper band (presumably composed of several bands) and a narrow lower band. The broad upper band could be resolved into three separate bands upon addition of anti-Sp1 and anti-Sp3 antibodies (Fig. 5C, lanes 6 and 7); the two lower bands corresponded to binding by Sp1, and the upper band corresponded to that of Sp3. The band with the fastest mobility in the gel shift generated by FOR2D probe also corresponded to Sp3 as it was supershifted by anti-Sp3 antibody (Fig. 5C, lanes 6 and 7). In addition, recombinant Sp1 protein gave rise to a shifted band with the same mobility as the broad upper band observed with nuclear extracts from HeLa cells (Fig. 5C, lane 3).

In contrast, the FOR2U probe produced two bands in the gel shift using nuclear extracts (Fig. 5D, lanes 2 and 5). In the presence of anti-Sp1 or anti-Sp3, no supershift could be observed (Fig. 5D, lanes 6 and 7). However, FOR2U probe could be bound by recombinant Sp1 protein (Fig. 5D, lane 3).

Given that the GC boxes contained in FOR2D and FOR2U.
probes differed only in a G → T change in the 5' region of each probe and that they show different patterns of binding by Sp1 in nuclear extracts, we constructed the Umut probe, which differs from FOR2 by a single T → G change. No changes in the binding patterns in the absence or in the presence of antibodies against Sp1 and Sp3 were observed with Umut compared with FOR2 (Fig. 5D, lanes 9–11). In addition we also constructed the Dmut probe, which is the analogous G → T change in FOR2; this probably did not show any difference in the gel shift or supershift patterns either (data not shown).

To better characterize the contribution of each one of the two GC boxes present in FOR2 probe, we performed gel shift experiments using FOR2 as the labeled probe and increasing amounts of either unlabeled FOR2 or FOR2D. FOR2U could compete only one of the lowest bands, but FOR2D competed away all of the other bands (Fig. 5E). Therefore, the downstream GC box is the binding site for both Sp1 and Sp3.

**Fig. 5. Characterization of Sp binding to Sp1 proximal promoter.** A gel mobility shift assay is shown. In the top left diagram the sequences used as the probes in these assays and the putative binding sites for transcription factors are depicted. The asterisks in the Umut and Dmut probes indicate the positions of the point mutations introduced in FOR2 and FOR2D, respectively. A, binding reactions were performed with 20,000 cpm of probe FOR2, 2 μg of nuclear extracts (NE) from either exponentially growing HeLa or 293T cells, and 2 μg of herring sperm DNA as the nonspecific competitor. Twenty-five ng of recombinant-purified Sp1 protein, rSp1, (Promega) was used as a control. B, supershift mobility assay with probe FOR2 in the presence of specific antibodies against Sp1 or Sp3. C, gel shift with probe FOR2D (20,000 cpm) using either nuclear extracts or recombinant Sp1 protein and supershift performed in the presence of specific antibodies against Sp1 or Sp3. D, in lanes 1–7, gel shift with probe FOR2U (20,000 cpm) using either nuclear extracts or recombinant Sp1 protein and supershift performed in the presence of specific antibodies against Sp1 or Sp3. In lanes 8–11, gel shift with probe Umut (20,000 cpm) and supershift performed in the presence of specific antibodies against Sp1 or Sp3. E, gel shift performed with probe FOR2 and competitions with either 10- or 40-fold of unlabeled FOR2U or FOR2D probes.

**Fig. 4. Deletion analysis of Sp1 promoter activity in 293T cells.** Depicted on the left side of the figure are the Sp1 promoter-luciferase deletion constructs. The location of putative nuclear protein binding sites for Ap1, Sp1, Ap1, Creb, and E2F are indicated. The 5' end point of each deletion mutant is indicated, and all constructs start at position −20 from the ATG. Each construct was transiently transfected into 293T cells and assayed for luciferase (Luc) activity. Transfections were performed in duplicate, and the results are the mean of two experiments ± S.E. Luciferase activity is normalized to micrograms of protein for each sample and is expressed relative to that of pGL3.
ther examine the role of Sp1 in its regulation, transient co-transfections in 293T cells were performed with pGL3, pGL3FOR1, or pGL3FOR2 together with an expression vector for Sp1. As shown in Fig. 6, minimal activity was detectable in 293T cells transfected with pGL3 or pGL3FOR1 plus Sp1. However, a marked increase in activity was induced by co-transfection of pGL3FOR2 with Sp1. The level of promoter activity accomplished by using the pGL3FOR2 construct co-transfected with Sp1 was 4-fold compared with that of pGL3FOR2 alone and 90-fold when referred to pGL3 activity. These results demonstrate that the Sp1 promoter is activated by its gene product.

**DISCUSSION**

In the present study, we isolated and identified 1593 bp of the 5′-flanking region of the human Sp1 gene (Fig. 1). This sequence allowed us to characterize the human Sp1 promoter region and identify cis-element regions involved in the regulation of the human Sp1 gene.

The mapping of the transcriptional start using the primer extension assay demonstrated the presence of a start site located at −63 bp relative to the translation initiation start that was observed for the three cell lines analyzed: HeLa, 293T, and HepG2 codon (Fig. 3). This result is in contrast with that reported by Takahara et al. (45) where the transcriptional start site was mapped at −171 relative to the ATG by primer extension in HepG2 cells. This difference could be due to the fact that we used a 47-mer in the primer extension reaction thus allowing a high specificity in the hybridization, whereas the group of Takahara used a 22-mer. The 5′-flanking region of the human transcription factor Sp1 lacks TATA and CAAT boxes (54). The TATA box is typically located 30 bp upstream from the transcription initiation site and helps to specify the transcription initiation site. There is also no initiator element (PyPyA, N/TAPyPy) in the human Sp1 promoter that in the absence of a TATA box commonly drives transcriptional initiation (46). In contrast, this region is GC-rich and contains potential binding sites for several well characterized transcription factors (Fig. 2). The absence of TATA or CAAT boxes and the presence of several GC boxes have been found primarily in housekeeping genes (47) and might have been expected to result in initiation of transcription at several locations (1, 48). However, a major single initiation site was consistently found in human 293T, HeLa, and HepG2 cells.

We found that the 5′-flanking region including up to −281 relative to the translational initiation codon contains all the elements necessary to achieve basal promoter activity. Of particular note is the high GC content of this region; in fact computer analysis indicates the presence of two putative Sp1-binding sites between nucleotides −281 to −20 and two more between −281 and −443. Our findings are consistent with previous studies showing that the transcription of other TATA-less promoters frequently involves the action of a proximal Sp1 site (49).

By EMSA, we demonstrated that several DNA-protein complexes were produced with a probe including the proximal promoter (up to base −281 relative to the translational start) that contains 2 GC boxes. The same qualitative binding pattern was observed when using two different human cell lines (293T and HeLa cells). The broad band in Fig. 5A is composed of three bands when using the shorter probe FOR2D as clarified in Fig. 5C. In Fig. 5C, anti-Sp1 supershifted the two lower bands, and anti-Sp3 supershifted the upper one. These three bands within the broad band are due to the binding to the most downstream GC box in FOR2 probe as shown by the results obtained in Fig. 5C and the competition assays between labeled FOR2 and unlabeled FOR2D (Fig. 5E). Sp3 contributed also to the formation of two more bands in the gel shift performed with FOR2 probe, the uppermost (Fig. 5B) and the lowermost bands (Fig. 5, B and C).

On the other hand, the FOR2U probe gives rise to two bands in the gel shift with nuclear extracts that are not supershifted with either Sp1 or Sp3 antibodies. The inability of Sp1 in the nuclear extract to bind to the FOR2U probe could be due to the binding of other transcription factors to adjacent sites that would interfere with Sp1 binding or induce a DNA conformation at this level of the promoter. Because the FOR2U probe can be bound by recombinant Sp1 (Fig. 5D, lane 3) it is likely that there is competition between Sp1 and other transcription factors present in the nuclear extract. The computational analysis of this promoter region revealed the presence of putative binding sites for C/EBP, NF-Y, CBP (C/CAAT-binding protein), and AP2. Therefore, it is possible that a complex of transcription factors binds to this sequence. These results demonstrate that Sp1 and Sp3 bind to the GC box located in the proximal promoter of Sp1 and that the downstream FOR2 GC box (the one contained in the FOR2D probe) is responsible for the major gel shift bands seen with the larger FOR2 probe.

An interesting conclusion of this work is the fact that the Sp1 gene is positively autoregulated. Indeed, co-transfection experiments of Sp1 proximal promoter constructs (pGL3FOR1 and pGL3FOR2) and a Sp1 expression vector in 293T cells demonstrate that Sp1 can activate its own transcription activity. The mechanism of positive autoregulation has already been described for other transcription factors such as NF-κB (50), the human junD (51) and CREB (52). For all these transcription factors, the actual binding of these proteins to their gene promoters has been shown, whereas for the human retinoblastoma susceptibility gene promoter, the positive autoregulation is mediated through an ATF-2 binding site in the Rb promoter (53). Autoregulation of Sp1 may be an important mechanism of modulating Sp1 activity. It is known that Sp1 protein is subjected to cell cycle regulation (38, 39) and the rapid synthesis and degradation of a cell cycle-regulated gene may be facilitated by an autoregulation mechanism of the corresponding gene promoter. Future work will analyze the possible regulation of the Sp1 expression achieved by Sp3 alone or at different ratios with Sp1, given that we have observed that Sp3 also binds to the Sp1 promoter as well as the effect of other transcription factors like Ap2 and E2F whose binding sites are also present in the Sp1 promoter.
Acknowledgments—Plasmid encoding Sp1 (CMV-Sp1) protein was generously provided by Dr. R. Tjian. Dr. R. Darnell is especially thanked for helpful suggestions and for hosting M. Nicolas in his laboratory at Rockefeller University. We also thank Dr. L. A. Chasin (Columbia University) for critical reading of the manuscript.

REFERENCES

1. Dynan, W. S., and Tjian, R. (1983) Cell 35, 79–87
2. Gidoni, D., Dynan, W. S., and Tjian, R. (1984) Nature 312, 409–413
3. Kadenaga, J. T., Carner, K. R., Masiarz, F. R., and Tjian, R. (1987) Cell 51, 1079–1090
4. Jones, R. A., Yamamoto, K. R., and Tjian, R. (1985) Science 228, 559–572
5. Courey, A. J., and Tjian, R. (1988) Cell 55, 877–898
6. Courey, A. J., Holtzman, A., Jackson, S. P., and Tjian, R. (1989) Cell 59, 827–836
7. Pascal, E., and Tjian, R. (1991) Genes Dev. 5, 1646–1656
8. Murata, Y. H. G. K., Rogers, K. T. Udvadia, A. J., and Horowitz, J. M. (1994) J. Biol. Chem. 269, 20674–20681
9. Jackson, S. P., MacDonald, J. J., Lees-Miller, S., and Tjian, R. (1990) Cell 63, 155–165
10. Armstrong, S. A., Barry, D. A., Leggett, R. W., and Mueller, C. R. (1997) J. Biol. Chem. 272, 13489–13495
11. Merchant, J. L., Du, M., and Todisco, A. (1999) Biochem. Biophys. Res. Commun. 254, 454–461
12. Jackson, S. P., and Tjian, R. (1998) Cell 93, 125–133
13. Han, I., and Kudlow, J. E. (1997) Mol. Cell. Biol. 17, 2550–2558
14. Mastrangelo, I. A., Courey, A. J., Wall, J. S., Jackson, S. P., and Hough, P. V. C. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 5670–5674
15. Emili, A., Greenblatt, J., and Ingles, C. J., (1994) Mol. Cell. Biol. 14, 1582–1593
16. Hoey, T., Weinzierl, R. O. J., Gill, G., Chen, J.-L., Dynlacht, B. D., and Tjian, R. (1993) Cell 72, 247–260
17. Chiang, C. M., and Roeder, R. G. (1995) Science 267, 531–536
18. Javahery, R., Khachi, A., Zenzie-Gregory, B., and Smale, S. T. (1994) Mol. Cell. Biol. 14, 116–127
19. Seto, E., Lewis, B., and Shenk, T. (1993) Nature 365, 462–464
20. Roder, K., Wolf, S. S., Larkin, K. J., and Schweizer, M. (1999) Gene 234, 461–469
21. Saltzman, A. G., and Lipman, D. J. (1997) Crit. Rev. Eucaryot. Gene Expr. 3, 229–254
22. Liptay, S., Schmid R. M., Nabel, E. G., and Nabel G. J. (1994) Mol. Cell. Biol. 14, 7695–7703
23. Meyer, T. E., Waeber, G., Lin, J., Beckmann, W., and Habener, J. F. (1993) Endocrinology 132, 770–780
24. Saltzman, A. G., and Weinmann, R. (1989) FASEB J. 3, 1723–1733