MOLECULAR AND CELLULAR BIOLOGY, Nov. 1995, p. 5906–5916 Vol. 15, No. 11
0270-7306/95/$04.00+0

Core Promoter Specificities of the Sp1 and VP16 Transcriptional Activation Domains

KATAYOON H. EMAMI, WILLIAM W. NAVARRE, AND STEPHEN T. SMALE*

Howard Hughes Medical Institute and Department of Microbiology and Immunology, University of California, Los Angeles, School of Medicine, Los Angeles, California 90095-1662

Received 12 July 1995/Accepted 21 July 1995

The core promoter compositions of mammalian protein-coding genes are highly variable; some contain TATA boxes, some contain initiator (Inr) elements, and others contain both or neither of these basal elements. The underlying reason for this heterogeneity remains a mystery, as recent studies have suggested that TATA-containing and Inr-containing core promoters direct transcription initiation by similar mechanisms and respond similarly to a wide variety of upstream activators. To analyze in greater detail the influence of core promoter structure on transcriptional activation, we compared activation by GAL4-VP16 and Sp1 through synthetic core promoters containing a TATA box, an Inr, or both TATA and Inr. Striking differences were found between the two activators, most notably in the relative strengths of the TATA/Inr and Inr core promoters: the TATA/Inr promoter was much stronger than the Inr promoter when transcription was activated by GAL4-VP16, but the strengths of the two promoters were more comparable when transcription was activated by Sp1. To define the domains of Sp1 responsible for efficient activation through an Inr, several Sp1 deletion mutants were tested as GAL4 fusion proteins. The results reveal that the glutamine-rich activation domains, which previously were found to interact with *Drosophila* TAF110, preferentially stimulate Inr-containing core promoters. In contrast, efficient activation through TATA appears to require additional domains of Sp1. These results demonstrate that activation domains differ in their abilities to function with specific core promoters, suggesting that the core promoter structure found in a given gene may reflect a preference of the regulators of that gene. Furthermore, the core promoter preference of an activation domain may be related to a specific mechanism of action, which may provide a functional criterion for grouping activation domains into distinct classes.

Appropriate expression of a eukaryotic gene is influenced by several regulatory layers, ranging from alterations in chromatin structure to direct interactions between transcriptional activator proteins and the general transcription machinery assembled at the core promoter (26, 65). To approach these complex mechanisms, one strategy has been to elucidate the fundamental events occurring at the core promoter and then to define the mechanisms through which regulatory proteins influence these events. In mammals, two control elements are commonly found in the core promoters for protein-coding genes: the TATA box, located 25 to 30 bp upstream of the transcription start site, and the initiator (Inr), which overlaps the start site (2, 22, 28, 35, 57). These elements are functionally similar in that each is sufficient for directing accurate basal transcription initiation by RNA polymerase II. Although many core promoters contain one or both of these elements, several genes that contain neither element have been described.

During transcription initiation, recognition of both TATA and Inr elements appears to be carried out by the TFIID complex, which contains the TATA-binding protein (TBP) and several TBP-associated factors (TAFs) (24); the TBP subunit is responsible for TATA recognition, and one of the TAFs recognizes the Inr (32, 51). Inr activity also benefits from a direct interaction between RNA polymerase II and functional Inr elements (3). In some promoters, Inr recognition appears to be facilitated by other DNA-binding proteins, including YY1, TFII-I, USF, and E2F (15, 41, 53, 54, 56).

Following template recognition, TATA-mediated transcription and Inr-mediated transcription appear to proceed via similar mechanisms. With both types of promoters, a similar set of general transcription factors appear to assemble into a preinitiation complex that is competent for initiation of RNA synthesis (3, 33, 39, 50). Furthermore, a variety of biochemical parameters are identical for the two types of promoters (29, 39, 66, 67), which is consistent with the idea that TATA and Inr elements direct transcription initiation through similar pathways.

If, indeed, the mechanisms of TATA- and Inr-mediated transcription are similar, it is necessary to consider why some core promoters have evolved to contain TATA boxes and others Inr elements. In addition, since TATA and Inr appear to carry out similar functions, it is not clear why some core promoters contain both of these elements. One suggestion was that TATA-less promoters may be necessary for appropriate regulation of "housekeeping" genes (for a review, see reference 57). Other suggestions were that TATA-less promoters may be needed for genes expressed at low levels or for genes requiring strict downregulation during development. Although there may be some merit to these suggestions, numerous exceptions to each model exist.

An alternative model to explain the need for core promoters containing TATA, Inr, or both TATA and Inr is that transcriptional regulatory proteins may prefer to modulate transcription through core promoters with specific structures. A core promoter preference could occur if a specific regulator targets a basal factor that is rate limiting only with a specific core promoter structure. Indeed, recent studies have suggested that different activation domains interact with different general factors (7, 19, 25, 43). Possibly, some of these interactions lead to

* Corresponding author. Phone: (310) 206-4777. Fax: (310) 206-3800. Electronic mail address: steves@hhmi.ucla.edu.

preferential regulation of TATA-containing or Inr-containing core promoters.

Recently, the above model has received support from studies of three proteins, p53, topoisomerase I, and a novel protein, which appear to repress transcription from core promoters that contain a TATA box but not from core promoters that contain an Inr (1, 38, 42). Although these proteins appear to function by directly influencing the TBP-TATA interaction, the results demonstrate that core promoter structure can indeed influence transcriptional regulation.

The structure of a core promoter also appears to influence the activities of typical transcriptional activators. For example, in the adenovirus major late (AdML) and β-globin promoters, which contain both TATA and Inr elements, TATA mutations strongly reduce transcription, whereas Inr mutations have little effect (27, 45). These results suggest that activators which bind to the AdML and β-globin promoters function only in the presence of a TATA box. More recently, a study of transcriptional activation by c-Fos revealed that a TBP binding domain mediates activation only through TATA-containing core promoters (43); deletion of this c-Fos domain resulted in similar activation from core promoters containing either TATA or Inr elements. This result is consistent with the idea that some activators prefer TATA boxes for transcriptional activation. Finally, the E2 transactivator of bovine papillomavirus 1 appears to prefer a core promoter containing both TATA and Inr elements, since efficient activation was detected only from a promoter containing both of these elements (23).

Despite the growing evidence in support of core promoter specificities for transcriptional activators, only one study has directly compared the abilities of various activation domains to stimulate transcription through core promoters containing TATA or Inr elements (5). In that study, a wide variety of activation domains, including those found in VP16, Sp1, p53, CTF, and C/EBP, were tested as GAL4 fusion proteins. The results revealed that all activation domains function similarly with the two core promoters, suggesting that, in general, the presence of a TATA box or Inr element in a given core promoter has no relevance to the transcriptional activation process.

To address this issue in more detail, we compared transcriptional activation through core promoters containing TATA, Inr, or TATA and Inr with two well-characterized activators: Sp1, which contains a zinc finger DNA-binding domain and strong glutamine-rich activation domains (11, 30), and GAL4-VP16(N), which contains an acidic activation domain from the herpes simplex virus type 1 (HSV-1) VP16 protein fused to the yeast GAL4 DNA-binding domain (10, 16, 18, 55). Our studies reveal that these activators differ in their abilities to activate transcription through specific core promoters. These results support a model in which the VP16 and Sp1 transcriptional activation domains function through distinct mechanisms and suggest that the core promoter preference of an activation domain can be linked to its mechanism of action.

MATERIALS AND METHODS

Plasmid DNAs. The reporter plasmids containing GAL4 binding sites were constructed by first inserting a *Cla*I linker into the *Hin*dIII site of pG5E4T, which contains five tandem GAL4 binding sites (4). The GAL4 binding sites were then excised with *Cla*I and *Bam*HI and inserted into the *Cla*I and *Bgl*II sites of the pSP72 vector (Promega), producing GAL4-SP72. The GAL4-TATA plasmid used for the in vitro transcription assays was constructed by inserting an oligonucleotide containing the AdML TATA box (58) into the *Eco*RI and *Sac*I sites of the GAL4-SP72 plasmid. The GAL4-TATA/Inr plasmid was constructed by inserting an oligonucleotide containing the terminal deoxynucleotidyltransferase (TdT)Inr (58) into the *Sac*I and *Bam*HI sites of the GAL4-TATA plasmid. The GAL4-weak TATA/Inr and GAL4-weak TATA plasmids were constructed by inserting the following oligonucleotide and its complement into the *Eco*RI and *Sac*I sites of GAL4-TATA/Inr and GAL4-SP72, respectively: 5′-CGGGTGATA TCATGGGGTGGGGGG-3′. The GAL4-Inr and GAL4-no TATA plasmids were constructed by inserting the following oligonucleotide and its complement into the *Cla*I and *Sac*I sites of GAL4-TATA/Inr and GAL4-SP72, respectively: 5′-ATGCGTCGCCTGCGATGAATTCG-3′.

For the transient-transfection studies, the promoters containing GAL4 binding sites described above were inserted upstream of the HSV thymidine kinase (TK) coding sequences in plasmid pSVPyTK (58). DNA fragments containing the promoter regions were obtained by PCR with primers with the following sequences: 5′-GGCCGATTCAGGATCCCAGGTTAACCTGGCTTATCG-3′ and 5′-GCCTGCAGGTCGACTCTAGAGGATCC-3′. The first primer is complementary to sequences upstream of the GAL4 binding sites, and the second primer is complementary to sequences downstream of the core promoter. Both primers contain a *Bam*HI restriction site, which allowed the amplified products to be cleaved with *Bam*HI and inserted into the *Bgl*II site of pSVPyTK.

The plasmids containing 12 Sp1 binding sites were described previously. For the in vitro experiments, Sp1-TATA, Sp1-TATA/Inr, Sp1-weak TATA/Inr, Sp1-Inr, Sp1-weak TATA, and Sp1-no TATA are plasmids III-b, VI-b, VI-e, VI-c, III-e, and III-c, respectively (66). For the transient-transfection experiments, the above plasmids containing Sp1 binding sites were inserted into pSVPyTK as described previously (66) (the resulting plasmids contain only six Sp1 binding sites). All plasmids were purified by column chromatography (Qiagen, Inc.), and their sequences were confirmed by DNA sequencing with a Sequenase kit (U.S. Biochemical).

The expression vector encoding the GAL4 DNA-binding domain (amino acids 1 to 147 [18, 55]) fused to the N-terminal activation domain of VP16 (amino acids 413 to 454 [12, 40]) was reported previously (16). Most of the GAL4-Sp1 expression plasmids were reported previously (17); all plasmids express a fusion protein containing the GAL4 DNA-binding domain (amino acids 1 to 147) and a portion of Sp1. GAL4-Sp1 contains Sp1 amino acids 83 to 778 and is the same as GAL4-Sp1WT from Gill et al. (17). GAL4-Sp1(A+B) contains Sp1 amino acids 83 to 621 and is the same as GAL4-Sp1(N) (17). GAL4-Sp1(B) contains Sp1 amino acids 263 to 542. GAL4-Sp1(Bc) contains amino acids 425 to 542 and is the same as GAL4-Sp1(D23) (17). GAL4-Sp1(Bc-M37) was described previously and contains a 10-amino-acid substitution mutation in the Bc domain (17). GAL4-Sp1(A), containing Sp1 amino acids 83 to 262, was constructed by inserting a fragment from GAL4-Sp1(A+B) into expression vector pBXG1 (16). The GAL4-Sp1(A) fragment was amplified by PCR with the following primers complementary to the 5′ end of the GAL4 fragment and to the 3′ end of the Sp1 A domain, respectively: 5′-AAGCTTCCTGAAAGATGAAGCTA-3′ and 5′-GAT TTCTAGAAGATCTTCGACAGGTAGCAA-3′. The amplified product was cleaved with *Hin*dIII and *Xba*I and inserted into the same sites of pBXG1.

In vitro transcription and primer extension. In vitro transcription reactions in nuclear extracts derived from HeLa cells were performed as described previously (58, 67). GAL4-VP16(N), containing VP16 amino acids 413 to 454, was expressed in *Escherichia coli* and purified with the assistance of Michael Carey's laboratory (University of California, Los Angeles) according to a protocol published previously (46). Reaction mixes contained 100 μg of HeLa nuclear extract, 600 ng of GAL4-VP16 or GAL4(1–94) protein, and a total of 300 ng of DNA (10 ng of template and 290 ng of pSP72 vector as carrier DNA; various template amounts were used for titration experiments). Primer extension analyses were carried out as described previously (58), with an SP6 promoter primer (Promega). The cDNA products were visualized by electrophoresis on 8% denaturing polyacrylamide gels, followed by autoradiography. Quantitation was performed by PhosphorImager analysis (Molecular Dynamics, Inc.).

Transient-transfection assays. Human 293 cells were grown to approximately 80% confluency in Dulbecco's modified Eagle's medium supplemented with 10% bovine calf serum (HyClone) and transfected by the calcium phosphate coprecipitation method of Graham and van der Eb (20) as modified by Parker and Stark (48). In all experiments, the total amount of transfected DNA was 20 μg, including 5 μg of reporter plasmid and 5 μg of expression vector. RNAs were isolated 2 days after transfection by the Nonidet P-40 lysis method (58). RNA (30 μg) was analyzed by primer extension analysis with a 20-nucleotide primer complementary to HSV TK sequences (66). The products were analyzed by electrophoresis as described above.

RESULTS

In vitro, the ratio of activated transcription through TATA/Inr versus Inr core promoters is much greater for GAL4-VP16 than for Sp1. To compare the abilities of Sp1 and GAL4-VP16 to activate transcription through core promoters with different structures, four synthetic core promoters were inserted downstream of multiple binding sites for either Sp1 or GAL4 (Fig. 1A; see Materials and Methods). One core promoter contains both a strong, consensus TATA box (with the sequence TATA AAA from the AdML promoter) and a strong Inr (CCTCAT TCT from the murine TdT promoter), with the two elements

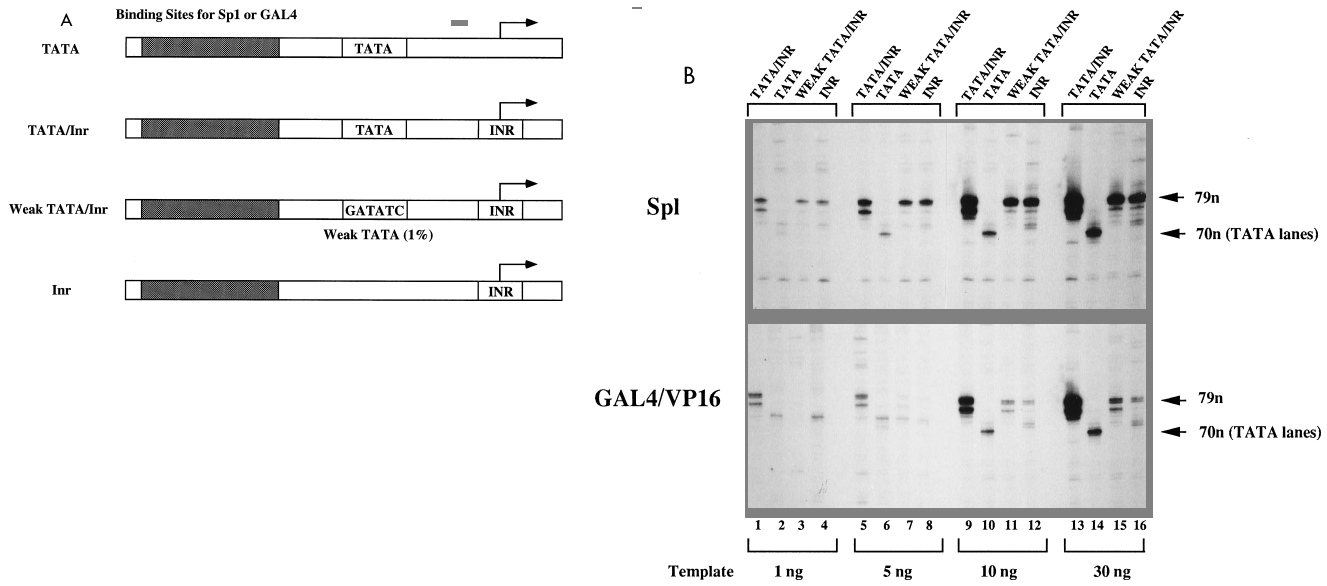

FIG. 1. Comparison of in vitro transcriptional activation by Sp1 and GAL4-VP16 through four different core promoters. (A) Promoters used in the in vitro and in vivo studies (see Materials and Methods). The TATA core promoter contains the AdML TATA box but no Inr. The TATA/Inr core promoter contains the AdML TATA box and the TdT Inr. The weak TATA/Inr core promoter contains the TdT Inr, with a very weak TATA box (GATATC [66]) located at −25. The Inr core promoter contains the TdT Inr with a highly GC-rich sequence at −25 (67). Inserted upstream of the core promoters were either five binding sites for GAL4 or 6 (in vivo experiments) or 12 (in vitro experiments) binding sites for Sp1. No difference was found in previous experiments between the strengths of promoters containing 6 or 12 Sp1 binding sites. (B) In vitro transcription reactions were performed in nuclear extracts derived from HeLa cells, with various amounts of template DNAs. Plasmids containing binding sites for Sp1 and GAL4 were tested (upper and lower panels, respectively). The reaction mixes presented in the lower panel were supplemented with 600 ng of purified GAL4-VP16 protein expressed in *E. coli*. RNA products were analyzed by primer extension followed by electrophoresis on an 8% denaturing polyacrylamide gel. The template concentrations are indicated at the bottom, and the core promoters tested are indicated at the top. The cDNA product corresponding to accurate initiation from the TATA core promoter is 70 nucleotides (n) (lanes 2, 6, 10, and 14), and the product corresponding to initiation from the TATA/Inr, weak TATA/Inr, and Inr core promoters is 79 nucleotides (lanes 1, 3 to 5, 7 to 9, 11 to 13, and 15 to 16), as indicated by arrows to the right.

separated by the spacing found in the AdML promoter. The other three promoters contain either the TATA box with no Inr, the Inr with no TATA box, or the Inr with a very weak TATA box (with the sequence GATATC). We previously showed that this weak TATA box directs transcription about 1% as efficiently as the consensus AdML TATA and, by all criteria measured, behaves very similarly to the promoter with a GC-rich −30 sequence (66).

The plasmids described above, with binding sites for either Sp1 or GAL4, were tested for promoter activity in an in vitro transcription assay in nuclear extracts derived from HeLa cells, with the resulting RNA transcripts analyzed by primer extension analysis. Because HeLa cells contain endogenous Sp1 but not GAL4-VP16, the extracts were supplemented with purified, recombinant GAL4-VP16 (containing VP16 amino acids 413 to 454; see Materials and Methods). The cDNA product corresponding to accurate initiation from the TATA core promoter is 70 nucleotides in length, and the product corresponding to initiation from the TATA/Inr, weak TATA/Inr, and Inr core promoters is 79 nucleotides.

Figure 1B shows the results of an experiment in which the amount of promoter-containing template introduced into the reaction mixes was varied from 1 to 30 ng. This titration experiment revealed that both Sp1 and GAL4-VP16 direct high levels of activated transcription from the TATA/Inr core promoter, with the activated signals increasing with increasing template amount (Fig. 1B, lanes 1, 5, 9, and 13). The activated transcription found with the TATA core promoter exhibited a similar pattern (lanes 2, 6, 10, and 14), but this promoter was considerably weaker than the TATA/Inr promoter. (The presence of two initiation sites with the TATA/Inr promoter was observed previously [47] and has been shown to result from a TATA-Inr spacing that is 1 to 3 nucleotides less than optimal for initiation from a single site. This spacing is used for our experiments because it corresponds to the TATA-Inr spacing found in the strong AdML core promoter and does not influence promoter strength [see reference 47].)

In contrast to the results described above, the activated signals observed with the weak TATA/Inr and Inr core promoters were much different for the two activators (Fig. 1B, lanes 3, 4, 7, 8, 11, 12, 15, and 16) when compared with the respective TATA and TATA/Inr signals. The relative signals vary to some degree with template amount, but in general, the Inr signal was 2- to 6-fold weaker than the TATA/Inr signal when transcription was activated by Sp1 and at least 20-fold weaker when transcription was activated by GAL4-VP16 (Fig. 1B; compare upper and lower panels). Similar results were found when the amount of purified GAL4-VP16 added to the reaction mixes was varied (data not shown).

The in vitro transcription experiment shown in Fig. 2 confirms and extends the above results. In this experiment, the activated signal detected with 10 ng of each plasmid was compared with the basal signal detected in the absence of activator. For the Sp1 panel (Fig. 2A, top), the plasmids containing binding sites for endogenous Sp1 (Fig. 2A, top, lanes 2, 4, 6, and 8) were compared with plasmids lacking Sp1 binding sites (Fig. 2A, top, lanes 1, 3, 5, and 7). For the GAL4-VP16 panel (Fig. 2A, bottom), the plasmids containing binding sites for GAL4 were tested in the absence (lanes 1, 3, 5, and 7) and presence (lanes 2, 4, 6, and 8) of added GAL4-VP16 protein. The data confirm that, with all four promoters, both Sp1 and GAL4-VP16 activate transcription above the basal levels. Moreover, the data confirm the results in Fig. 1B, in that the strengths of the Inr promoters were much different with the

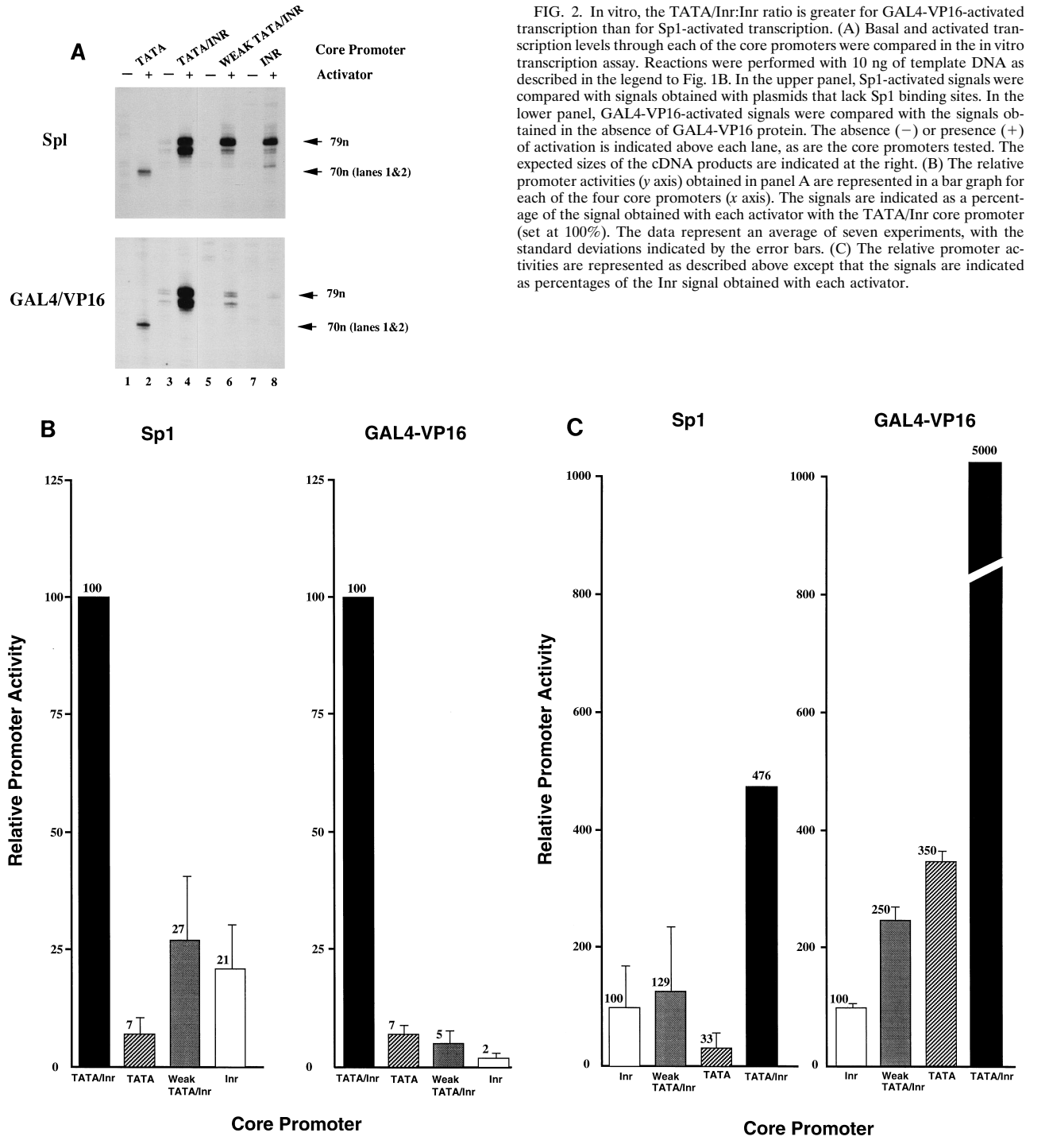

FIG. 2. In vitro, the TATA/Inr:Inr ratio is greater for GAL4-VP16-activated transcription than for Sp1-activated transcription. (A) Basal and activated transcription levels through each of the core promoters were compared in the in vitro transcription assay. Reactions were performed with 10 ng of template DNA as described in the legend to Fig. 1B. In the upper panel, Sp1-activated signals were compared with signals obtained with plasmids that lack Sp1 binding sites. In the lower panel, GAL4-VP16-activated signals were compared with the signals obtained in the absence of GAL4-VP16 protein. The absence (−) or presence (+) of activation is indicated above each lane, as are the core promoters tested. The expected sizes of the cDNA products are indicated at the right. (B) The relative promoter activities (*y* axis) obtained in panel A are represented in a bar graph for each of the four core promoters (*x* axis). The signals are indicated as a percentage of the signal obtained with each activator with the TATA/Inr core promoter (set at 100%). The data represent an average of seven experiments, with the standard deviations indicated by the error bars. (C) The relative promoter activities are represented as described above except that the signals are indicated as percentages of the Inr signal obtained with each activator.

two activators when normalized to the strengths of the TATA and TATA/Inr promoters.

Figures 2B and C present the results of the experiment in Fig. 2A in two different formats. In Fig. 2B, the promoter strengths are normalized to the strengths of the TATA/Inr promoters (set at 100% for each activator). With this interpretation, GAL4-VP16 and Sp1 appear to function equally well with the TATA/Inr and TATA promoters, but Sp1 appears to activate transcription much more effectively through the Inr promoters than does GAL4-VP16. In Fig. 2C, the promoter strengths are normalized to the strengths of the Inr promoters (set at 100% for each activator). With this interpretation, GAL4-VP16 and Sp1 appear to function equally well with the Inr promoter, but GAL4-VP16 appears to activate transcription more effectively with the TATA promoter and also with the TATA/Inr promoter, in which the TATA and Inr elements strongly synergize with each other. A direct comparison of the absolute signals obtained in Fig. 2A suggests that the representation in Fig. 2B is more accurate. However, we cannot directly compare the absolute signals obtained with the two

different activators because, for example, they may bind to DNA with different affinities and may differ in their abilities to function in the in vitro assay. Therefore, it is not yet clear if Fig. 2B or 2C more accurately represents the differences between Sp1 and GAL4-VP16 (see below).

The Inr and the VP16 activation domain are required for the in vitro activities detected. The results described above suggest that Sp1 and GAL4-VP16 differ in their relative abilities to activate transcription through the various core promoters. To confirm that the Inr elements within the weak TATA/Inr and Inr core promoters are responsible for the differences observed, the Inr was deleted from these plasmids. The resulting plasmids (referred to as weak TATA and no TATA) were then compared with the TATA/Inr and TATA plasmids in the in vitro transcription assay (Fig. 3A; also included is a plasmid containing a functional TATA box of intermediate strength [rTATA]). In the absence of the Inr, transcription could not be detected from either of the plasmids lacking a functional TATA box (Fig. 3A, lanes 8 and 10). These data demonstrate that the Inr is critical for the activated transcription directed by the weak TATA/Inr and Inr plasmids in Fig. 1 and 2.

With the weak TATA/Inr and Inr plasmids, GAL4-VP16 appears to activate transcription relatively poorly. This observation raises the possibility that the signals detected with these plasmids might not result from "true" VP16-dependent activation. An alternative possibility was that the GAL4 DNA-binding domain might bind to the template and prevent histones and nonspecific DNA-binding proteins in the extract from inhibiting basal transcription (36, 67). In other words, the VP16 activation domain might actually be incapable of direct activation from the weak TATA/Inr and Inr core promoters. To address this issue, an in vitro transcription experiment was performed, comparing activation by GAL4-VP16 with activation by the isolated, recombinant GAL4 DNA-binding domain, GAL4(1–94). [Gel shift experiments demonstrated that the purified GAL4-VP16 and GAL4(1–94) proteins were equally active for DNA binding (data not shown).] The data shown in Fig. 3B reveal that the VP16 domain is needed for the activation detected with all of the promoters tested, including the weak TATA/Inr and Inr promoters (lanes 10 to 15). This result suggests that GAL4-VP16 directly activates transcription from all four core promoters. However, we cannot rule out the possibility that the signals detected result from a VP16-dependent antirepression mechanism (13).

Relative promoter activities found with Sp1 and GAL4-VP16 are similar in vivo and in vitro. To study the core promoter specificities of Sp1 and GAL4-VP16 in vivo, transient-transfection assays were performed. The experiments described in this section compared activation by endogenous Sp1 with activation by GAL4-VP16(N) expressed from a cotransfected plasmid (16). For these experiments, the TATA, TATA/Inr, weak TATA/Inr, and Inr promoters were inserted upstream of a reporter gene containing coding sequences for the HSV TK gene. The resulting plasmids were transfected into the human embryonic kidney cell line 293, which was chosen because of its high transfection efficiency. Promoter activity was measured by primer extension analysis with a primer complementary to an HSV TK sequence. The cDNA product expected for accurate initiation from the TATA core promoter is 75 nucleotides in length, and the product expected for initiation from the TATA/Inr, weak TATA/Inr, and Inr core promoters is 84 nucleotides.

Figure 4A shows the promoter strengths obtained following activation by GAL4-VP16 (upper panel) and Sp1 (lower panel). Consistent with the in vitro data, the relative strengths of the TATA/Inr and Inr promoters were much different for the two activators (compare Fig. 4A, lanes 2 to 4, with Fig. 2A, lanes 4, 6, and 8): with GAL4-VP16, the TATA/Inr promoter was typically 10- to 20-fold stronger than the Inr promoter, whereas with Sp1-activated transcription, the TATA/Inr promoter was typically only 2- to 3-fold stronger than the Inr promoter. One inconsistency between the in vitro and in vivo assays was observed, however; with Sp1-activated transcription in the in vivo assay, the strengths of the Inr promoter and the TATA promoter were similar (Fig. 4A, lanes 1 and 4), whereas in the in vitro assay, the Inr promoter was about threefold tronger than the TATA promoter (see Fig. 2A). The results presented below (Fig. 5 and 6) might help to explain why the relative strengths of the Sp1-TATA and Sp1-Inr promoters vary.

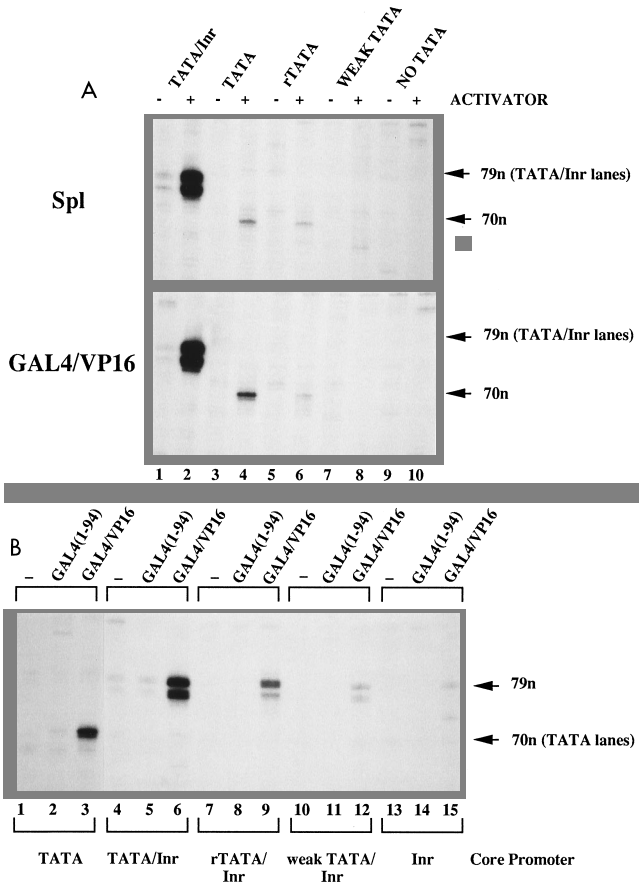

FIG. 3. The Inr and the VP16 activation domain are required for the in vitro activities detected. (A) In vitro transcription experiments were performed as described in the legend to Fig. 1B with core promoters containing the −30 sequences found in the weak TATA/Inr and Inr plasmids but after deletion of the TdT Inr. The resulting plasmids are referred to as weak TATA and no TATA, respectively. Plasmids containing a TATA box of intermediate strength (rTATA) were also tested. Sp1-activated transcription is shown in the upper panel, and GAL4-VP16-activated transcription is shown in the lower panel. The core promoter tested is indicated at the top, as is the absence (−) or presence (+) of activator. The sizes of the expected cDNA products are indicated at the right. The basal signals were determined by the method described in the legend to Fig. 2A. (B) In vitro transcriptional activation by GAL4-VP16 was compared with activation by the isolated GAL4 DNA-binding domain, GAL4(1–94). The DNA-binding activities of the purified GAL4(1–94) and GAL4-VP16 proteins were similar (data not shown). In vitro transcription reaction mixes were supplemented with either no activator (lanes 1, 4, 7, 10, and 13), 600 ng of GAL4(1–94) (lanes 2, 5, 8, 11, and 14), or 600 ng of GAL4-VP16 (lanes 3, 6, 9, 12, and 15). The core promoter located downstream of the GAL4 binding sites is indicated at the bottom, and the expected cDNA sizes are indicated at the right (in nucleotides [n]).

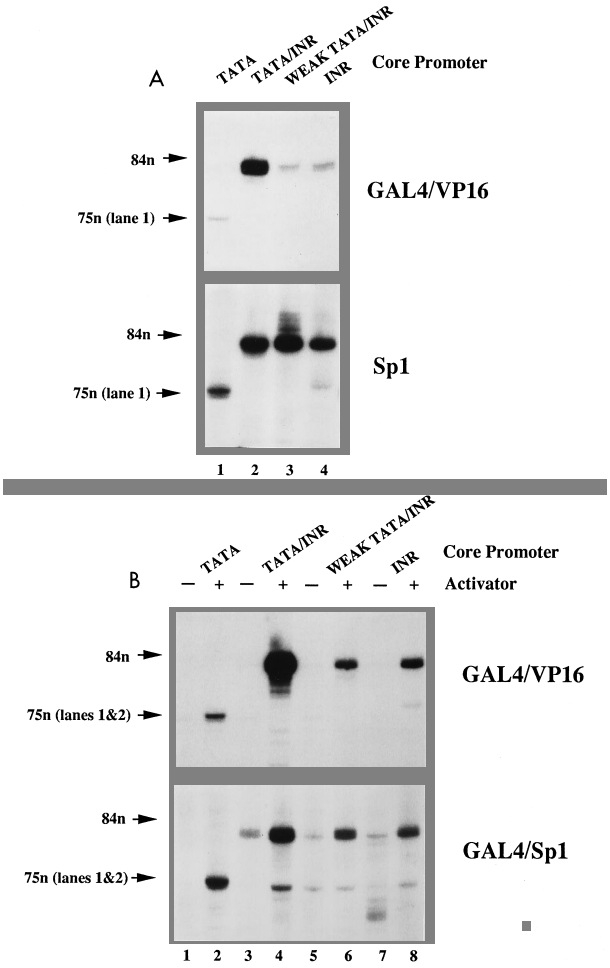

FIG. 4. Comparison of transcriptional activation in vivo by Sp1, GAL4-VP16(N), and GAL4-Sp1 through four different core promoters. (A) The promoters containing binding sites for Sp1 or GAL4 were inserted upstream of HSV TK coding sequences in the vector pSVPyTK (see Materials and Methods). Various amounts of the resulting plasmids were transfected into 293 cells. The mRNA products were isolated 48 h after transfection and analyzed by primer extension with a primer complementary to HSV TK sequences. In the upper panel, plasmids containing GAL4-binding sites were tested, with activation by GAL4-VP16(N) expressed from 5 μg of a cotransfected expression vector. In the lower panel, plasmids containing Sp1-binding sites were tested, with activation by the endogenous Sp1. The total DNA transfected was kept constant (20 μg) by addition of carrier plasmid DNA. cDNA products of 75 nucleotides (n) were expected with the TATA plasmid, and products of 84 nucleotides were expected with the other three plasmids. The arrows at the right indicate the locations of the expected cDNAs. (B) Transient-transfection experiments were performed in 293 cells to compare activation by transiently expressed GAL4-VP16 (upper panel) and GAL4-Sp1 (lower panel). Either the GAL4-VP16 or GAL4-Sp1 expression plasmid (5 μg) was cotransfected with 5 μg of the reporter plasmids. Basal signals were determined by transfection of the reporter plasmids in the absence of a cotransfected expression plasmid. The total concentration of transfected DNA was kept constant (20 μg) by addition of carrier plasmid DNA. The absence (−) or presence (+) of activator and the core promoter tested are indicated at the top. The sizes of expected cDNA products are indicated by arrows at the left. The lower panel was derived from a longer autoradiographic exposure than the upper panel.

Together, the in vitro and in vivo results suggest that Sp1 and GAL4-VP16 differ in their abilities to activate transcription from the core promoters tested. The primary difference appears to be that with GAL4-VP16, the strength of the TATA/Inr promoter is much greater than that of the Inr promoter, whereas with Sp1, the strengths of the TATA/Inr and Inr promoters are more similar.

In vivo, GAL4-Sp1 functions similarly to endogenous Sp1 with the four core promoters. The results described above suggest that activator proteins differ in their abilities to activate transcription through core promoters with different structures. However, these results do not permit a direct comparison between the VP16 and Sp1 activation domains for three reasons. First, multiple proteins in mammalian cells are capable of binding to the Sp1 sites (e.g., reference 34), making it likely that other proteins contributed to the results obtained. Second, the results may be influenced by the fact that an endogenous activator was compared with a recombinant fusion protein. Third, Sp1 and GAL4 bind to different recognition sequences with different affinities. The different DNA-binding parameters prevent us from directly comparing the relative strengths of the Sp1 and VP16 transcriptional activation domains. To establish whether the differences observed between Sp1 and GAL4-VP16 reflect the properties of their transcriptional activation domains, we compared GAL4-VP16 and GAL4-Sp1 fusion proteins in transient-transfection assays.

For these experiments, a eukaryotic expression vector for GAL4-Sp1 (containing Sp1 amino acids 83 to 778; a gift of Grace Gill, Harvard Medical School) was obtained. Figure 4B shows the results obtained with GAL4-VP16 (upper panel) and GAL4-Sp1 (lower panel) cotransfected with the four reporter plasmids containing GAL4 binding sites. These results are very similar to those shown in Fig. 4A; with both endogenous Sp1 and GAL4-Sp1, the TATA/Inr promoter was two- to threefold stronger than either the TATA or Inr promoter. In contrast, with GAL4-VP16, the TATA/Inr promoter was consistently at least 12-fold stronger than either the TATA or Inr promoter. Since the same reporter plasmids and same DNA-binding domain were used in comparing GAL4-VP16 and GAL4-Sp1, these results demonstrate that the VP16 and Sp1 transcriptional activation domains are primarily responsible for the characteristic differences between GAL4-VP16 and Sp1.

A GAL4-Sp1 deletion mutant activates transcription primarily from Inr-containing core promoters. The GAL4-Sp1 fusion protein tested in Fig. 4B functioned extremely poorly in the transient-cotransfection assay (the GAL4-Sp1 panel in Fig. 4B is derived from a relatively long exposure of the film). The weak signals may result from the large size (93 kDa) of the fusion protein and from the presence of both Sp1 and GAL4 DNA-binding domains. Therefore, to provide a more accurate comparison between the VP16 and Sp1 activation domains, we employed an expression plasmid, GAL4-Sp1(A+B) (17), which contains the GAL4 DNA-binding domain fused to the N-terminal transcriptional activation domains of Sp1 (see Fig. 6A). This N-terminal fragment, extending from amino acids 83 to 621, lacks the zinc finger DNA-binding domain but contains both of the strong glutamine-rich activation domains of Sp1 as well as the adjacent serine/threonine-rich domains (11). By immunoblot analysis, similar amounts of GAL4-VP16 and GAL4-Sp1(A+B) were expressed following transfection into 293 cells (data not shown).

A direct comparison of GAL4-VP16 and GAL4-Sp1(A+B) in a cotransfection experiment with plasmids containing the four distinct core promoters yielded two critical results (Fig. 5A). First, GAL4-VP16 activated transcription through all four core promoters more strongly than did GAL4-Sp1(A+B). With the TATA/Inr core promoter, for example, GAL4-VP16 was 44 times stronger than GAL4-Sp1(A+B) (Fig. 5A, lanes 5 and 6). With the Inr core promoter, GAL4-VP16 was six times stronger than GAL4-Sp1(A+B) (Fig. 5A, lanes 11 and 12). This result suggests that the model proposed in Fig. 2C rather than that proposed in Fig. 2B more accurately describes the

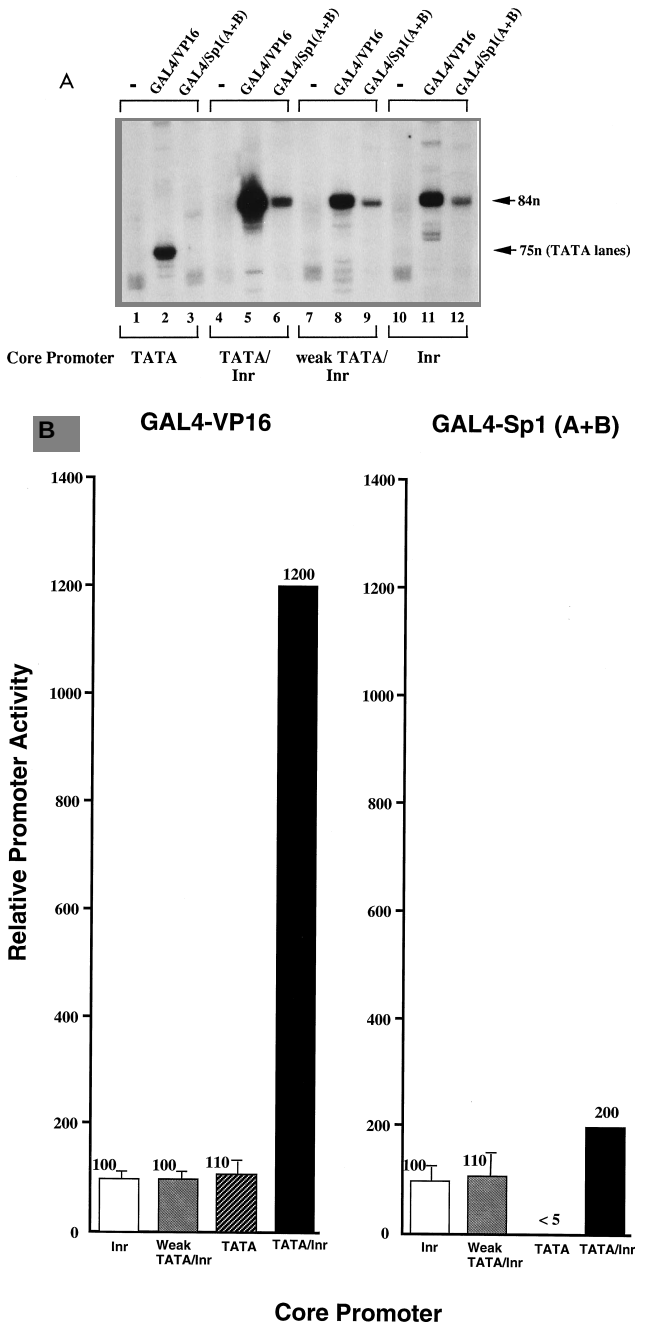

FIG. 5. GAL4-Sp1(A+B) prefers a core promoter containing an Inr for transcriptional activation. (A) Transient-transfection experiments were performed in 293 cells as described in the legend to Fig. 4. Plasmids containing GAL4 binding sites with the four core promoters (indicated at the bottom) were cotransfected with the GAL4-VP16 (lanes 2, 5, 8, and 11) or GAL4-Sp1(A+B) (lanes 3, 6, 9, and 12) expression vector or with no expression vector (lanes 1, 4, 7, and 10). The expected sizes of the cDNA products are indicated (in nucleotides) by the arrows at the right. (B) The relative promoter strengths (*y* axis) from the experiment in panel A are presented as a bar graph for each core promoter (*x* axis). The promoter strengths are indicated as a percentage of the strengths of the Inr promoters (set at 100%). Separate graphs are presented for GAL4-VP16 and for GAL4-Sp1(A+B). The data represent the averages of five independent experiments, with standard deviations indicated by error bars. For the TATA core promoter with GAL4-Sp1(A+B), signals could not be detected above background in any of the experiments, even with long exposures of the gels.

relative transcriptional activation properties of GAL4-VP16 and Sp1. As suggested in Fig. 2C, both activators function quite well with the Inr core promoter, but the TATA and Inr elements strongly synergize with each other when activated by GAL4-VP16, leading to extremely strong transcription through the TATA/Inr core promoter (Fig. 5A, lane 5). In contrast, with GAL4-Sp1 and GAL4-Sp1(A+B) (Fig. 4B and 5A), activated transcription through TATA/Inr is only two- to threefold stronger than through the weak TATA/Inr or Inr core promoters.

The second critical result in Fig. 5A is that GAL4-Sp1(A+B) could not activate transcription through the TATA core promoter to detectable levels (lane 3). In more than 20 repetitions of this experiment in 293 and HeLa cells, with multiple preparations of reporter and expression plasmids, we have been unable to detect an activated signal from the TATA promoter with GAL4-Sp1(A+B). In a recent study, Das et al. (14) also failed to detect activation by GAL4-Sp1(A+B) from the c-*fos* core promoter, which contains only a TATA box. Although GAL4-Sp1(A+B) may indeed activate transcription from these promoters at a level that is below the detection limits, these results clearly demonstrate that GAL4-Sp1(A+B) activates transcription much less efficiently through TATA than do endogenous Sp1 and GAL4-Sp1 (compare Fig. 4 and 5). This surprising result reveals that the A and B domains of Sp1 preferentially activate transcription through a core promoter containing an Inr. Moreover, this Inr preference of the A and B domains of Sp1 is strikingly different from the preference obtained with the VP16 activation domain, which results in equivalent activation through the TATA and Inr core promoters.

A 117-amino-acid glutamine-rich fragment of Sp1 is sufficient for transcriptional activation through core promoters containing an Inr. To localize the domains of Sp1 involved in the Inr preference for transcriptional activation, various N-terminal Sp1 subfragments were tested as GAL4 fusion proteins. Previous studies of Sp1 in cultured *Drosophila* cells revealed that the N terminus contains two distinct domains, A and B, each containing a glutamine-rich stretch and an adjacent region rich in serines and threonines (11) (Fig. 6A). In mammalian cells, each of the glutamine-rich domains was sufficient for transcriptional activation (17). Moreover, the C-terminal half of the B domain (called the Bc domain) was found to activate transcription as well as the entire B domain, but the N-terminal half (called the Bn domain) was inactive (17).

Figure 6B shows the results obtained with plasmids that express the GAL4 DNA-binding domain fused to the A, B, or Bc subdomain of Sp1 or to a Bc mutant called M37. This mutant, containing a 10-amino-acid substitution, reduces an interaction between the Sp1 Bc domain and *Drosophila* TAF110 (17). By immunoblot analysis, similar amounts of each fusion protein were expressed following transfection into 293 cells (data not shown). The primer extension analysis reveals that the A, B, and Bc fragments activate transcription from the TATA/Inr, weak TATA/Inr, and Inr core promoters (Fig. 6B, lanes 6 to 20). However, none of these fusion proteins activates transcription from the TATA core promoter (lanes 1 to 5). Interestingly, the activated signals obtained with GAL4-Sp1(Bc-M37) (Fig. 6B, lanes 10, 15, and 20) were reduced by 50 to 90% relative to those with GAL4-Sp1(Bc). This result suggests that a direct interaction between Sp1 and a human TAF110 homolog may be responsible for the Inr preference of the Sp1 glutamine-rich activation domains.

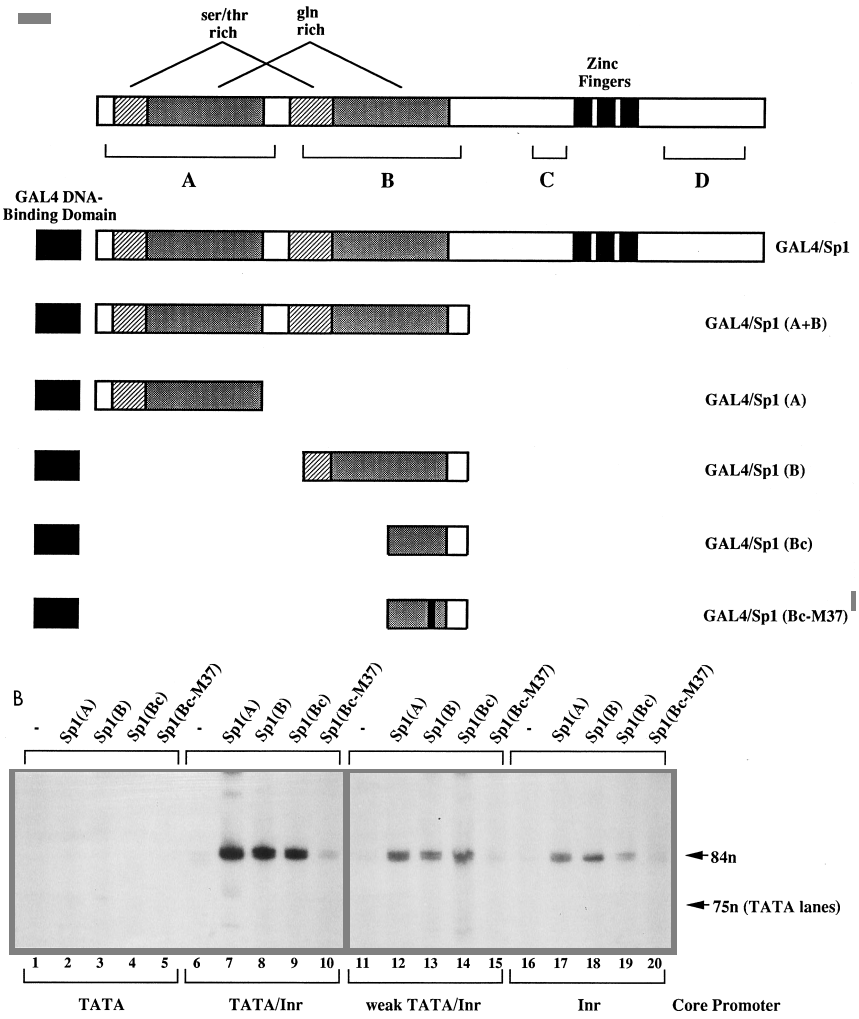

FIG. 6. The 117-amino-acid Bc domain of Sp1 is sufficient for transcriptional activation through core promoters containing an Inr. (A) Schematic representation of the GAL4-Sp1 fusion proteins tested in Fig. 4 through 6. The four previously defined activation domains, A through D, are indicated, as are the zinc fingers and the glutamine-rich and serine/threonine-rich regions. All fusion proteins contain the GAL4 DNA-binding domain (amino acids 1 to 147) fused to an Sp1 fragment that begins with Sp1 amino acid 83. (B) Transient-cotransfection assays were performed in 293 cells with the four reporter plasmids containing GAL4 binding sites (5 μg; core promoter indicated at the bottom) and with no expression plasmid or with the expression plasmids (5 μg) GAL4-Sp1(A), GAL4-Sp1(B), GAL4-Sp1(Bc), and GAL4-Sp1(Bc-M37) (indicated at the top). The expected locations of the cDNA products are indicated (in nucleotides) by the arrows at the right.

DISCUSSION

This study demonstrates that the VP16 and Sp1 activation domains have intrinsic functional preferences for specific core promoter structures during transcriptional activation. VP16, although capable of efficient activation from any core promoter, possesses a strong preference for a core promoter containing both TATA and Inr elements to achieve an optimal signal. In contrast, Sp1-activated transcription through TATA/Inr is only moderately stronger than through TATA or Inr alone. Dissection of the Sp1 protein revealed that it contains activation domains with different specificities. The glutamine-rich domains function preferentially with core promoters containing Inr or TATA/Inr, but full-length Sp1 also activates transcription efficiently from the TATA core promoter. These results suggest that specific core promoter structures may have evolved in protein-coding genes to match the preferences of the gene-specific activators.

The finding that typical transcriptional activation domains possess distinct core promoter specificities may provide a simple method for classifying transcriptional activation domains. Recent studies have suggested that activation domains stimulate transcription by interacting with specific components of the basal transcription apparatus (7, 19, 25, 43). Distinct classes of activation domains are likely to exist, based on their physical targets or mechanisms of action. Activation domains currently are classified by amino acid content (44) (e.g., glutamine-rich activators and acidic activators). This method of classification might not accurately reflect the mechanism of activation. In contrast, the pattern of activation through core promoters with different structures may provide a method for rapidly determining which activation domains are likely to function similarly.

To elucidate the mechanisms underlying the core promoter specificities observed, it first is helpful to consider the relationship between the activated and basal transcription signals. In these experiments, we were unable to accurately measure the weak basal transcription signals. However, our previous studies have demonstrated that TATA and Inr elements strongly synergize with each other when tested in the absence of upstream activators (47, 59). Interestingly, GAL4-VP16-activated tran-

scription reveals a similar degree of synergy between TATA and Inr elements. Thus, GAL4-VP16 appears to activate transcription to the same extent from all core promoters. In other words, the activated signal for each core promoter may be directly proportional to the basal signal. In contrast, the Sp1-activated signals are not proportional to the basal signals. Apparently, the extent of Sp1 activation is greater for promoters containing only an Inr than it is for the TATA/Inr promoter.

One possible mechanism to explain the above observations is that GAL4-VP16 may act primarily on steps that follow TFIID binding, whereas Sp1 may play a role in recruiting TFIID to the core promoter. If GAL4-VP16 does not influence the binding of TFIID to the core promoter, it might be expected to activate transcription in proportion to the affinity of TFIID for the core promoter. In contrast, if Sp1 stabilizes TFIID binding, it could activate transcription to a similar level from all core promoters by rendering irrelevant the intrinsic affinity of TFIID for the core promoter.

The above model is consistent with previous studies of the mechanisms of activation by GAL4-VP16 and Sp1. GAL4-VP16 contains an acidic domain that has been shown to interact with TBP, TFIIB, TFIIH, and a TAF (7, 19, 37, 60, 63, 64). In the experiments of Green and colleagues, GAL4-VP16 was found to recruit TFIIB to the preinitiation complex but to have little effect on TFIID binding (7, 37, 52). Sp1 has been shown to interact with *Drosophila* TAF110 and human TAF55, providing routes through which it could recruit TFIID to a core promoter (6, 17, 25). In fact, we previously showed that Sp1, possibly through an Sp1-TAF interaction, directly recruits TFIID to an Inr element in the absence of a TATA box (32).

The models described above are sufficient to explain why activation domains may differ in their abilities to stimulate transcription from TATA/Inr relative to TATA alone or Inr alone. However, these models do not address the mechanism underlying the observation that the glutamine-rich activation domains of Sp1 prefer Inr-containing core promoters. The Inr preference indicates that a fundamental difference must exist between the mechanisms of TATA-mediated and Inr-mediated transcription. The glutamine-rich domain may target a factor, possibly a human TAF110 homolog, that plays a role only in Inr-mediated transcription, whereas full-length Sp1 may interact with a second protein that leads to activation through a TATA box. One candidate for this second protein is TAF55, which recently was reported to interact with the C terminus of Sp1 (6). Alternatively, identical protein-protein interactions may be needed for activation through both TATA-containing and Inr-containing promoters, but the fundamental difference may simply be that the two core promoters depend on different rate-limiting steps. This possibility is supported by both in vivo and in vitro studies, which have implicated different rate-limiting steps in transcription initiation from the two types of promoters (8, 67).

Previous studies of Sp1 have revealed that at least four domains, A, B, C, and D, are required for optimal transcriptional activation (11) (Fig. 6A). The A and B domains contain the glutamine-rich regions that are sufficient for transcriptional activation. The C domain weakly activates transcription by itself, whereas the D domain has no independent activity but enhances activation by the A and B domains. Further dissection of the B domain in *Drosophila* and human cells localized a subdomain called Bc, which is sufficient for activation (11, 17, 49) and for the interaction with *Drosophila* TAF110 (25). Most of these studies were carried out in *Drosophila* cultured cells (11, 17), but the only studies measuring the activities of the isolated A and B domains were performed in mammalian cells with reporter plasmids containing both TATA and Inr elements (17). Thus, there are no inconsistencies between our results and these previous results.

Our results reveal that the A and B domains and the Bc subfragment stimulate transcription from the core promoters containing an Inr but that the entire Sp1 protein was needed for TATA activation. We have not localized the domains needed for efficient transcription through TATA, but from the previous mutagenesis, TATA activation is likely to require multiple domains. The D domain does not function by itself on a TATA/Inr-containing promoter, and the C domain functions poorly (11). Since TATA activation is quite strong in the presence of the entire Sp1 molecule, it is likely to require one or both of the glutamine-rich domains in combination with the C or D domain. It is interesting that the D domain has been linked to the formation of higher-order multimers (49), suggesting that multimerization may be needed for activation through a core promoter containing only a TATA box.

The results of this analysis are consistent with several other studies that have revealed core promoter specificities during transcriptional regulation (see the introduction). In addition, our results are somewhat consistent with a recent study by Colgan and Manley (9), which revealed that in *Drosophila* cells, Sp1 activated transcription more strongly through an Inr-containing promoter than through a TATA-containing promoter. However, our results are inconsistent with a study by Chang and Gralla (5), which reported that GAL4-Sp1(A+B) strongly activated transcription from a core promoter containing a TATA box. This inconsistency can easily be explained by the fact that the previous study used as a TATA-containing promoter the adenovirus E4 promoter, which contains an Inr element at its start site (TTTACACT) in addition to a TATA box at the −30 region. Thus, it is not surprising that GAL4-Sp1(A+B) activated transcription from this promoter.

The results of our study reveal interesting similarities to recent studies comparing a TATA box with the proximal sequence element (PSE) found in the core promoters of small nuclear RNA genes (14, 62). In one study, the GAL4-Sp1(A+B) fusion protein strongly activated a core promoter containing the PSE but was unable to activate a core promoter containing a TATA box (14). In contrast, GAL4-VP16 preferentially activated transcription through the TATA box. The finding that the Sp1 glutamine-rich domains preferentially activate transcription from core promoters containing either a PSE or an Inr but not a TATA box suggests that the mechanisms of PSE- and Inr-mediated transcription may be more closely related than that of TATA-mediated transcription.

In addition to the studies addressing functional differences between TATA elements and either Inr or PSE elements, numerous other studies have compared activation through distinct TATA sequences (57, 61). A well-characterized example of specificity among TATA sequences is found in the *Saccharomyces cerevisiae HIS3* gene. Struhl and colleagues have shown that this gene contains two functional TATA elements, but GAL4 and GCN4 can activate transcription only through one of the two elements (61). Although these experiments focused on different TATA sequences rather than different core promoter elements, they are consistent with the idea that activator proteins are influenced by the precise interactions between general transcription factors and the core promoter.

Most of the studies of core promoter specificity carried out by us and others have focused on the core promoter elements required by specific activators. However, appropriate regulation of most mammalian genes is likely to depend on the combined action of numerous proteins bound to promoter and enhancer sequences. Since several activation domains with different properties are likely to act on a given gene, one might

expect that at least one activation domain will always be present that is compatible with any core promoter structure. Thus, altering the core promoter structure of a complex gene might not inactivate the gene but may instead play a more subtle role in transcriptional regulation. To approach these issues, a more detailed understanding of the biochemical basis of transcriptional activation and of core promoter specificity is needed.

ACKNOWLEDGMENTS

We are extremely grateful to Grace Gill for providing GAL4-Sp1 expression plasmids and to Michael Carey for helpful suggestions, for providing GAL4(1–94) protein, and for assistance with the purification of GAL4-VP16. We also thank Michael Carey, Albert Courey, Patricia Ernst, Grace Gill, Jörg Kaufmann, and Kiersten Lo for critical reading of the manuscript.

This work was supported by the Howard Hughes Medical Institute.

REFERENCES

1. **Aso, T., H. Serizawa, R. C. Conaway, and J. W. Conaway.** 1994. A TATA sequence-dependent transcriptional repressor activity associated with mammalian transcription factor IIA. EMBO J. **13:**435–445.
2. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50:**349–383.
3. **Carcamo, J., L. Buckbinder, and D. Reinberg.** 1991. The initiator directs the assembly of a transcription factor IID-dependent transcription complex. Proc. Natl. Acad. Sci. USA **88:**8052–8056.
4. **Carey, M., Y. Lin, M. R. Green, and M. Ptashne.** 1990. A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. Nature (London) **345:**361–364.
5. **Chang, C., and J. D. Gralla.** 1993. Properties of initiator-associated transcription mediated by GAL4/VP16. Mol. Cell. Biol. **13:**7469–7475.
6. **Chiang, C. M., and R. G. Roeder.** 1995. Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. Science **267:** 531–536.
7. **Choy, B., and M. R. Green.** 1993. Eukaryotic activators function during multiple steps of preinitiation complex assembly. Nature (London) **366:**531–536.
8. **Colgan, J., and J. L. Manley.** 1992. TFIID can be rate limiting in vivo for TATA-containing, but not TATA-lacking, RNA polymerase II promoters. Genes Dev. **6:**304–315.
9. **Colgan, J., and J. L. Manley.** 1995. Cooperation between core promoter elements influences transcriptional activity in vivo. Proc. Natl. Acad. Sci. USA **92:**1955–1959.
10. **Conaway, J. W., J. P. Hanley, K. P. Garrett, and R. C. Conaway.** 1991. Transcription initiated by RNA polymerase II and transcription factors from liver: structure and action of transcription factors ε and τ. J. Biol. Chem. **266:**7804–7811.
11. **Courey, A. J., and R. Tjian.** 1988. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell **55:**887–898.
12. **Cress, W. D., and S. J. Triezenberg.** 1991. Critical structural elements of the VP16 transcriptional activation domain. Science **251:**87–90.
13. **Croston, G. E., P. J. Laybourn, S. M. Paranjape, and J. T. Kadonaga.** 1992. Mechanism of transcriptional antirepression by GAL4-VP16. Genes Dev. **6:**2270–2281.
14. **Das, G., C. S. Hinkley, and W. Herr.** 1995. Basal promoter elements as a selective determinant of transcriptional activator function. Nature (London) **374:**657–659.
15. **Du, H., A. L. Roy, and R. G. Roeder.** 1993. Human transcription factor USF stimulates transcription through initiator elements of the HIV-1 and the Ad-ML promoters. EMBO J. **12:**501–511.
16. **Emami, K. H., and M. Carey.** 1992. A synergistic increase in potency of a multimerized VP16 transcriptional activation domain. EMBO J. **11:**5005–5012.
17. **Gill, G., E. Pascal, Z. H. Tseng, and R. Tjian.** 1994. A glutamine-rich hydrophobic patch in transcription factor Sp1 contacts the $dTAF_{II}110$ component of the Drosophila TFIID complex and mediates transcriptional activation. Proc. Natl. Acad. Sci. USA **91:**192–196.
18. **Giniger, E., S., M. Varnum, and M. Ptashne.** 1985. Specific DNA binding of GAL4, a positive regulatory protein of yeast. Cell **40:**767–774.
19. **Goodrich, J. A., T. Hoey, C. J. Thut, A. Admon, and R. Tjian.** 1993. Drosophila $TAF_{II}40$ interacts with both a VP16 activation domain and the basal transcription factor TFIIB. Cell **75:**519–530.
20. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus type 5 DNA. Virology **52:**456–467.
21. **Greenblatt, J.** 1991. Roles of TFIID in transcriptional initiation by RNA polymerase II. Cell **66:**1067–1070.
22. **Grosschedl, R., and M. L. Birnstiel.** 1980. Identification of regulatory sequences in the prelude sequences of an H2A histone gene by the study of specific deletion mutants in vitro. Proc. Natl. Acad. Sci. USA **77:**1432–1436.
23. **Ham, J., G. Steger, and M. Yaniv.** 1994. Cooperativity in vivo between the E2 transactivator and the TATA box binding protein depends on core promoter structure. EMBO J. **13:**147–157.
24. **Hernandez, N.** 1993. TBP, a universal eukaryotic transcription factor? Genes Dev. **7:**1291–1308.
25. **Hoey, T., R. O. J. Weinzierl, G. Gill, J. L. Chen, B. D. Dynlacht, and R. Tjian.** 1993. Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. Cell **72:**247–260.
26. **Hori, R., and M. Carey.** 1994. The role of activators in assembly of RNA polymerase II transcription complexes. Curr. Opin. Genet. Dev. **4:**236–244.
27. **Hu, S. L., and J. L. Manley.** 1981. DNA sequence required for initiation of transcription in vitro from the major late promoter of adenovirus 2. Proc. Natl. Acad. Sci. USA **78:**820–824.
28. **Javahery, R., A. Khachi, K. Lo, B. Zenzie-Gregory, and S. T. Smale.** 1994. DNA sequence requirements for transcriptional initiator activity in mammalian cells. Mol. Cell. Biol. **14:**116–127.
29. **Jiang, Y., S. T. Smale, and J. D. Gralla.** 1993. A common ATP requirement for open complex formation and transcription at promoters containing initiator or TATA elements. J. Biol. Chem. **268:**6535–6540.
30. **Kadonaga, J. T., K. R. Carner, F. R. Masiarz, and R. Tjian.** 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell **51:**1079–1090.
31. **Kamakaka, R. T., M. Bulger, and J. T. Kadonaga.** 1993. Potentiation of RNA polymerase II transcription by GAL4/VP16 during but not after DNA replication and chromatin assembly. Genes Dev. **7:**1779–1795.
32. **Kaufmann, J., and S. T. Smale.** 1994. Direct recognition of initiator elements by a component of the transcription factor IID complex. Genes Dev. **8:**821–829.
33. **Kaufmann, J., and S. T. Smale.** Unpublished results.
34. **Kingsley, C., and A. Winoto.** 1992. Cloning of GT box-binding proteins: a novel Sp1 multigene family regulating T-cell receptor gene expression. Mol. Cell. Biol. **12:**4251–4261.
35. **Kollmar, R., and P. J. Farnham.** 1993. Site-specific initiation of transcription by RNA polymerase II. Proc. Soc. Exp. Biol. Med. **203:**127–139.
36. **Laybourn, P. J., and J. T. Kadonaga.** 1991. Role of nucleosomal cores and histones H1 in regulation of transcription by RNA polymerase II. Science **254:**238–245.
37. **Lin, Y. S., and M. R. Green.** 1991. Mechanism of action of an acidic transcriptional activator in vitro. Cell **64:**971–981.
38. **Mack, D. H., J. Vartikar, J. M. Pipas, and L. A. Laimins.** 1993. Specific repression of TATA-mediated but not initiator-mediated transcription by wild-type p53. Nature (London) **363:**281–283.
39. **Martinez, E., C. Chiang, H. Ge, and R. G. Roeder.** 1994. TATA-binding protein-associated factor(s) in TFIID function through the initiator to direct basal transcription from a TATA-less class II promoter. EMBO J. **13:**3115–3126.
40. **McGeoch, D. J., A. Dolan, S. Donald, and D. H. K. Brauer.** 1986. Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. Nucleic Acids Res. **14:**1727–1745.
41. **Means, A. L., and P. J. Farnham.** 1990. Transcription initiation from the dihydrofolate reductase promoter is positioned by HIP1 binding at the initiation site. Mol. Cell. Biol. **10:**653–661.
42. **Merino, A., K. R. Madden, W. S. Lane, J. J. Champoux, and D. Reinberg.** 1993. DNA topoisomerase I is involved in both repression and activation of transcription. Nature (London) **365:**227–232.
43. **Metz, R., A. J. Bannister, J. A. Sutherland, C. Hagemeier, E. C. O'Rourke, A. Cook, R. Bravo, and T. Kouzarides.** 1994. c-Fos-induced activation of a TATA-box-containing promoter involves direct contact with TATA-box-binding protein. Mol. Cell. Biol. **14:**6021–6029.
44. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245:**371–378.
45. **Myers, R. M., K. Tilly, and T. Maniatis.** 1986. Fine structure genetic analysis of a β-globin promoter. Science **232:**613–618.
46. **Ohashi, Y., J. M. Brickman, E. Furman, B. Middleton, and M. Carey.** 1994. Modulating the potency of an activator in a yeast in vitro transcription system. Mol. Cell. Biol. **14:**2731–2739.
47. **O'Shea-Greenfield, A., and S. T. Smale.** 1992. Roles of TATA and initiator elements in determining the start site location and direction of RNA polymerase II transcription. J. Biol. Chem. **267:**1391–1402.
48. **Parker, B. A., and G. R. Stark.** 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by the virus or viral DNA. J. Virol. **31:**360–369.
49. **Pascal, E., and R. Tjian.** 1991. Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism. Genes Dev. **5:**1646–1656.
50. **Pugh, B. F., and R. Tjian.** 1990. Mechanism of transcriptional activation by Sp1: evidence for coactivators. Cell **61:**1187–1197.
51. **Purnell, B. A., P. A. Emanuel, and D. S. Gilmour.** 1994. TFIID sequence-recognition of the initiator and sequences further downstream in Drosophila

class II genes. Genes Dev. **8:**830–842.
52. **Roberts, S. G. E., I. Ha, E. Maldonado, D. Reinberg, and M. R. Green.** 1993. Interaction between an acidic activator and transcription factor TFIIB is required for transcriptional activation. Nature (London) **363:**741–744.
53. **Roy, A. L., S. Malik, M. Meisterernst, and R. G. Roeder.** 1993. An alternative pathway for transcription initiation involving TFII-I. Nature (London) **365:** 355–359.
54. **Roy, A. L., M. Meisterernst, P. Pognonec, and R. G. Roeder.** 1991. Cooperative interaction of an initiator-binding transcription factor and the helix-loop-helix activator USF. Nature (London) **354:**245–248.
55. **Sadowski, I., J. Ma, S. Triezenberg, and M. Ptashne.** 1988. Gal4-VP16 is an unusually potent transcriptional activator. Nature (London) **335:**563–564.
56. **Seto, E., Y. Shi, and T. Shenk.** 1991. YY1 is an initiator sequence-binding protein that directs and activates transcription in vitro. Nature (London) **354:**241–245.
57. **Smale, S. T.** 1994. Core promoter architecture for eukaryotic protein-coding genes, p. 63–81. *In* R. C. Conaway and J. W. Conaway (ed.), Transcription: mechanisms and regulation. Raven Press, New York.
58. **Smale, S. T., and D. Baltimore.** 1989. The "initiator" as a transcription control element. Cell **57:**103–113.
59. **Smale, S. T., M. C. Schmidt, A. J. Berk, and D. Baltimore.** 1990. Transcriptional activation by Sp1 as directed through TATA or initiator: specific requirement for mammalian transcription factor IID. Proc. Natl. Acad. Sci. USA **87:**4509–4513.
60. **Stringer, K. F., C. J. Ingles, and J. Greenblatt.** 1990. Direct and selective binding of an acidic transcriptional activation to the TATA-box factor TFIID. Nature (London) **345:**783–786.
61. **Struhl, K.** 1989. Molecular mechanisms of transcriptional regulation in yeast. Annu. Rev. Biochem. **58:**1051–1077.
62. **Tanaka, M., J. S. Lai, and W. Herr.** 1992. Promoter-selective activation domains in Oct-1 and Oct-2 direct differential activation of an snRNA promoter and mRNA promoter. Cell **68:**755–767.
63. **White, J., C. Brou, J. Wu, Y. Lutz, V. Moncollin, and P. Chambon.** 1992. The acidic transcriptional activator GAL4-VP16 acts on preformed template-committed complexes. EMBO J. **11:**2229–2240.
64. **Xiao, H., A. Pearson, B. Coulombe, R. Truant, S. Zhang, J. L. Regier, S. J. Triezenberg, D. Reinberg, O. Flores, C. J. Ingles, and J. Greenblatt.** 1994. Binding of basal transcription factor TFIIH to the acidic activation domains of VP16 and p53. Mol. Cell. Biol. **14:**7013–7024.
65. **Zawel, L., and D. Reinberg.** 1992. Advances in RNA polymerase II transcription. Curr. Opin. Cell Biol. **4:**488–495.
66. **Zenzie-Gregory, B., A. Khachi, I. P. Garraway, and S. T. Smale.** 1993. Mechanism of initiator-mediated transcription: evidence for a functional interaction between the TATA-binding protein and DNA in the absence of a specific recognition sequence. Mol. Cell. Biol. **13:**3841–3849.
67. **Zenzie-Gregory, B., A. O'Shea-Greenfield, and S. T. Smale.** 1992. Similar mechanisms for transcription initiation mediated through a TATA box or an initiator element. J. Biol. Chem. **267:**2823–2830.