MOLECULAR AND CELLULAR BIOLOGY, Sept. 1991, p. 4561–4571
0270-7306/91/094561-11$02.00/0

Vol. 11, No. 9

Differential Ability of Proximal and Remote Element Pairs To Cooperate in Activating RNA Polymerase II Transcription

WEIDONG WANG AND JAY D. GRALLA*

Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90024

Received 4 March 1991/Accepted 24 June 1991

To investigate the synergism or cooperative interaction between transcription elements, we have designed and constructed a series of synthetic polymerase II promoters with different combinations of elements. These include three different CCAAT boxes, which correspond to the binding sites for CP1, CP2, and NFI, a GC box, a CACCC box, and an ATF/CREB-binding site. The synthetic promoters containing these elements in proximal positions were linked to a test gene (CAT). Tandem repeats of AP1- and AP2-binding sites, the simian virus 40 enhancer, and DNA-binding sites for GAL-estrogen receptor were cloned downstream of the test gene. The strength of these promoters was then tested in transient-expression assays in HeLa TK$^-$ cells. In the context of the adenovirus major late promoter TATA box, the promoters containing only certain combinations of elements are active in this assay. Some elements appear to cooperate nearly universally, but others exhibit strong selectivity. These results indicate strongly selective synergistic interactions between elements and suggest that levels of promoter strength may be determined by the extent of compatibility between factors bound to proximal and enhancer sites.

The DNA sequence elements controlling transcription by RNA polymerase II generally divide into two overlapping classes, proximal elements and remote elements. Proximal elements are those that reside near the start point of transcription, usually within approximately 100 bp. Remote elements, including enhancers and upstream activating sequences, are positioned at much greater distances, usually from a few hundred to a few thousand base pairs away. Both types of elements are generally involved in raising transcription to levels higher than that set by the basal transcription machinery, which generally includes TATA and sometimes initiator sequences. Most mammalian promoters require both proximal and remote elements for appropriate transcription. Both classes of elements are quite diverse, and many proteins have been identified that bind to each type of element (for reviews, see references 11 and 25). Some elements can function from both proximal and remote positions (for example, see reference 8).

Although many proximal transcription elements and enhancer elements have been identified so far, the relationships among these elements, or the interactions between the corresponding *trans*-acting factors, are not very well understood. Several studies have shown that certain factors appear to act universally in the sense that they stimulate expression in a wide variety of contexts. For example, studies on the glucocorticoid receptor element showed that in HeLa and certain other cell lines, this element, when placed upstream of a promoter, stimulated every proximal transcription element tested: binding sites for CP1, NFI, OTF, two SP1-binding sites, the CACCC box, estrogen receptor, VP16, and even itself (31, 32, 35, 36). The acidic activators such as GAL4 and VP16 may also fall into this category since it has been shown to work with many promoter-proximal regions and in diverse organisms (for a review, see reference 28). Other enhancer elements have not been tested as extensively, and the extent to which they might cooperate selectively with proximal elements is not known.

Little information is available concerning whether elements in proximal positions cooperate universally in the sense that many enhancers can activate them. The question of cooperation between upstream and proximal elements is now receiving more attention since recent studies have indicated that there may be adapter factors that mediate the interactions between the upstream and downstream elements (4, 20, 29). It is not known whether there are families of such factors that mediate gene-specific control by using the mix of elements present in promoters. Ultimately, the ability to describe the machinery responsible for promoter-specific activation will depend on learning the rules for cooperation between the various elements and isolating the molecules that allow this selective cooperation.

To allow a systematic approach to these questions, we have designed a test system in which combinations of elements may be tested for the ability to cooperate. Its basis is a promoterless vector containing a polylinker into which different combinations of synthetic DNA transcription elements may be inserted. Each combination is then tested for expression of the chloramphenicol acetyltransferase (CAT) gene in a transient-transfection assay. In these initial experiments, HeLa cells growing in 10% serum are used since prior studies have indicated that these cells contain active factors capable of using many proximal and enhancer elements.

We report here results obtained from studying 57 different synthetic promoters in this manner. The elements studied in proximal positions include CCAAT-box sequences that are bound differentially by factors CP1, CP2, and NFI in vitro (7), single and double GC boxes, a CACCC box (32), and an ATF/CREB-binding site (13, 14). Each of these elements is studied in combination with a synthetic TATA box and is also paired either with synthetic multicopy enhancer sites of the AP1 (1, 21), AP2 (15, 26), or estrogen receptor (ER) type (37) or with the simian virus 40 (SV40) enhancer that is known to contain sites for many transcription factors (for a

* Corresponding author.

review, see reference 17). These enhancer sites are present in a location that is far downstream. The results show that some enhancer and proximal elements appear to cooperate nearly universally but others exhibit strong selectivity.

MATERIALS AND METHODS

Plasmid constructions. The starting vectors UCAT and UCATX10 were obtained from the laboratory of R. Tjian (21). The SV40 early promoter upstream of the CAT gene in both vectors was removed by *Hin*dIII digestion to generate plasmids UOCAT and UOCATX10. The X10 designation refers to the presence of 10 synthetic AP1-binding sites downstream from the CAT gene. Two complementary synthetic oligonucleotides with *Hin*dIII ends were annealed, phosphorylated at their 5′ ends, and cloned into the *Hin*dIII site of the new vectors. This will be the matrix allowing construction of synthetic promoters. The sequence of each oligonucleotide is as follows:

5′-AGCTTCGAATCTAGATCTGCAGATCGATGATCAGAATTCTCGAGGCATGCGTCGACGAGCTTCCA,

and 5′-AGCTTGGAAGCTCGTCGACGCATGCCACGAGAATTCTGATCATCGATCTGCAGATCTAGATTCGA.

The restriction map of this matrix and the DNA sequences of the inserted elements are shown in Fig. 1. A synthetic TATA box from the adenovirus major late promoter sequence was cloned into the *Sph*I site of the matrix. Four kinds of CCAAT boxes, CP1, CP2 (7), an NFI consensus (16), and a natural NFI (24) binding site, were synthesized and cloned into the *Bcl*I sites in both orientations. One GC box that is a strong SP1-binding site (18) was also synthesized and cloned into the *Bcl*I site in both orientations. The same GC box was dimerized and cloned at the same site in both orientations. Another sequence was designed so that it contains one GC box and one mutated GC box, which was cloned directionally between the *Bcl*I and *Xho*I sites, to generate the promoter with the wild-type GC box at the distal position. A CACCC box was synthesized and cloned in both orientations at the *Bcl*I site. An ATF/CREB-binding site (13, 14) was synthesized and directionally cloned between the *Bcl*I and *Xho*I sites. The SV40 enhancer, which includes two copies of the 72-bp repeat, was cloned into the *Bam*HI site downstream of the gene. Two complementary oligonucleotides which correspond to the sequence of the AP2 site (15) or the sequence of the GAL4 site (19) were synthesized with a *Bam*HI-*Bgl*II site at each end. They were then multimerized with DNA ligase, and the 5-mers were selected to clone at the *Bam*HI site of the vector. The DNA sequence of each promoter construct was verified by primer extension sequencing by using W2 primer that hybridized to a sequence within the coding region of the CAT gene. Its sequence is 5′-CGGTGGTATATCCAGTGA. The GAL-ER plasmids were kindly provided by the laboratory of P. Chambon (37).

Cell culture and CAT assays. The HeLa TK^- cells were obtained from R. Tjian, University of California, Berkeley. The transfection experiments were done by the method of Lee et al. (21). In brief, the cells were maintained in Dulbecco modified Eagle medium supplemented with 10% fetal bovine serum. About 1.0×10^6 cells were seeded on a 10-cm plate 24 h prior to transfection. Three hours before transfection, the cells received fresh medium. They were then incubated with the calcium phosphate-DNA solution for 4 h. After glycerol shock, the cells received fresh medium and were harvested 24 h later. The conditions for cotransfection experiments in HeLa cells were the same except that the cells were assayed for the CAT activity 48 h posttransfection instead of 24 h.

Extracts were prepared and assayed for CAT activity. The plasmid pON260 (33), which expresses the *Escherichia coli* β-galactosidase gene, was used as a control in initial experiments to compare the efficiency of different transfection experiments. It was omitted in later experiments after it was found to affect the expression of some cotransfected constructs. There was considerable variation in the levels of CAT expression of some promoters in these experiments. Most data are the average of five or six different determinations, each done in duplicate. In no case were fewer than two sets of duplicate independent experiments done. At least two different preparations of plasmid were assayed for each construct. An expression level of less than twofold higher than background was scored as inactive. A level of 10 times background corresponds to approximately 3% CAT conversion in a 2-h incubation.

Mapping of the transcription start sites. The start sites of the synthetic promoters have been mapped both in vitro, by using HeLa nuclear extract, and in vivo, by using the above transient-transcription assays in HeLa cells. The in vitro transcription was performed as described previously (6, 22). Briefly, the reaction mixture contains 25 μl of HeLa extract, 1 μl of 0.33 M $MgCl_2$, 4 μl of 5 mM nucleoside triphosphates, and 100 ng of supercoiled templates. GAL-AH activator was added last to start the incubation. The mixture was incubated at 30°C for 1 h. The RNA products were then purified and analyzed by primer extension, using avian myeloblastosis virus reverse transcriptase and W2 primer that was also used for sequencing of promoter DNA. The extension products were fractionated on 8% denaturing polyacrylamide gels. Cellular RNA was extracted from transfected HeLa cells and purified by the methods of Auffray and Rougeon (2). They were then analyzed by primer extension as described above. A 30-μg portion of RNA was used for each extension reaction.

RESULTS

The organization of the synthetic polylinker constructed and inserted upstream from the CAT gene is shown in Fig. 1A. Most constructs contain a synthetic TATA box, which corresponds to the sequence of the element in the adenovirus major late promoter, inserted into the *Sph*I site. Most of the other proximal elements were inserted into the upstream *Bcl*I site. Four synthetic CCAAT box-type elements were inserted here, and their sequences are shown in Fig. 1A. These are designated according to transcription factors known to bind them in vitro: CP1 refers to a sequence

occurring in the long terminal repeat of murine sarcoma virus (7); CP2 refers to a sequence from the human gamma-fibrinogen gene promoter (7); NFI refers to a sequence present in the long terminal repeat of mouse mammary tumor virus (24); and NFI-C refers to the consensus sequence, obtained by comparing DNA sequence targets for NFI binding, that was used for purification of NFI by a DNA affinity column (16). Each synthetic element is approximately 25 bp long and thus includes both the central and peripheral sequences corresponding to the natural sites. All sites are known to bind HeLa cell transcription factors in vitro in the context of natural promoters. Their location, approximately 40 bp upstream from the TATA box, is common within the range of naturally occurring locations (see specific discussion below).

In some constructs, multiple tandem recognition sequences for enhancer factors are located downstream from the CAT gene. In the case of AP1, these sequences are known to be active in this location when a variety of natural promoters are present in the region occupied by the promoter insertion polylinker in this vector (21).

The stimulatory activity of these paired elements in proximal and remote positions will be assayed by measuring CAT activity, always in the context of the adenovirus major late promoter TATA box. The TATA element alone showed no detectable CAT activity (see below), implying that observed increases will be due to the effect of added elements. In addition, with the single exception of the ATF site, no single element plus TATA box yielded CAT activity, demonstrating that expression occurs in the context of the paired inserted elements. For several pairings, the transcription start sites were mapped from transfected HeLa cells or from an in vitro system by using HeLa nuclear extract (see examples in Fig. 1B). The start sites are at the position expected for TATA-dependent transcription. Two nearly adjacent start sites were detected in vitro (lanes 3 to 6), and only the stronger one (T) was found in vivo (lanes 1 and 2). The distance from the TATA box is only 1 bp different from the natural distance specified by this same TATA box in the adenovirus major late promoter, from which it was copied. Thus the increased activity shown below by certain paired elements represents joint stimulation of appropriate TATA-dependent transcription.

Proximal requirements for promoter activity with AP1 enhancer. At first, 10 synthetic promoter constructs containing CCAAT-box elements in a proximal position were transfected into HeLa TK^- cells, and 24 h later extracts were prepared and assayed for CAT activity. These constructs included four different CCAAT-box elements, alone and in combination with the AP1 enhancer element (Fig. 2A). All contained the TATA box in the usual position. Of the 10 constructs, 8 showed CAT activity that was scored as indistinguishable from the activity of the mock-transfected control (all less than twofold greater than this background signal). These include all the constructs that lacked the AP1 sites downstream. The data indicate that these promoters have, at most, extremely low promoter activity. Apparently the pairing of CCAAT-box and TATA elements is insufficient to create a minimal promoter in HeLa cells in the absence of an enhancer element.

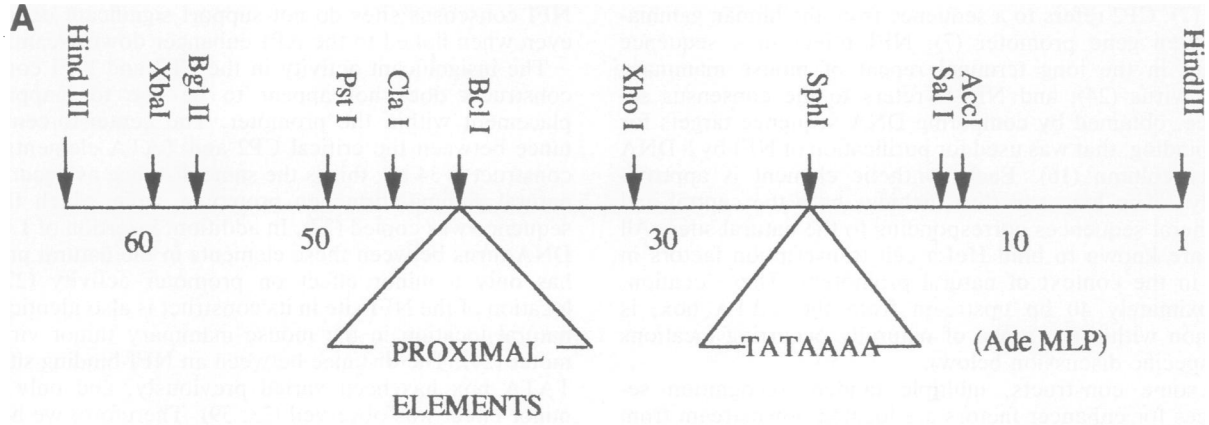
In the presence of the downstream, remote AP1 sites, certain of the proximal CCAAT-box elements show detectable expression. The signal from the promoter containing the synthetic CP1 site is approximately eight times background, indicating significant promoter strength. The synthetic NFI site supports weaker expression, yielding a signal four times that of the background control. By contrast, the CP2 and NFI consensus sites do not support significant expression, even when linked to the AP1 enhancer downstream.

The insignificant activity in the CP2 and NFI consensus constructs does not appear to be due to inappropriate placement within the promoter. The center-to-center distance between the critical CP2 and TATA elements in the construct is 34 bp; this is the same distance as occurs in the natural gamma-fibrinogen promoter, from which the CP2 sequence was copied (27). In addition, insertion of 1.0 or 1.5 DNA turns between these elements in the natural promoter has only a minor effect on promoter activity (27). The location of the NFI site in its construct is also identical to its natural location in the mouse mammary tumor virus promoter (24). The distance between an NFI-binding site and a TATA box has been varied previously, and only a very minor effect was observed (23, 39). Therefore, we have not attempted to activate the CP2 and NFI consensus synthetic promoters by changing the locations of the binding sites. The appropriateness of the placement is supported further by the ability of these constructs to be stimulated by other enhancers (see below).

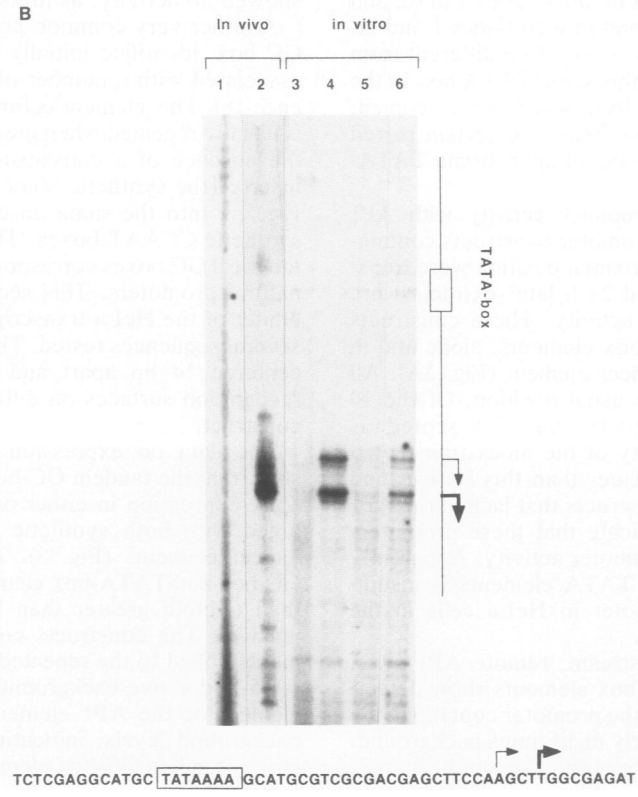
The orientations of these CCAAT-box constructs were chosen to correspond to the orientation in the large majority of natural promoters (5). In some promoters CCAAT boxes appear in a reverse orientation. To learn whether these reverse orientations were active in the context of AP1-dependent expression, we made a set of reverse constructs. These included reversed CP1, CP2, and NFI consensus elements. The results of CAT assays on these constructs (Fig. 2B) shows that the reversed CP1 site does not work as well as the more common forward orientation. Only about twofold residual activity was detected compared with eightfold in the forward orientation. The reversed NFI consensus site also showed very weak activity. The reversed CP2 site showed no activity, as in its common orientation.

Another very common proximal promoter element is the GC box, identified initially in SV40 but now known to be associated with a number of genes (for a review, see reference 18). This element is functional in both orientations and is usually repeated when present. It occasionally functions in the absence of a consensus TATA element. Initially, we inserted the synthetic tandem GC box sequences shown in Fig. 1A into the same upstream site used for insertion of synthetic CCAAT boxes. The sequence of each of the two identical GC boxes corresponds to DNA occurring in several natural promoters. This sequence is the strongest in vitro binder of the HeLa transcription factor SP1 among a set of several sequences tested. The two copies of the GC box are centered 14 bp apart and thus are expected to present recognition surfaces on different helical DNA faces in this construct.

The data on expression from the synthetic promoters show that the tandem GC-box element can direct significant CAT expression in either orientation, but only when combined with both synthetic TATA and synthetic AP1 enhancer elements (Fig. 3). The constructs containing only GC-box and TATA-box elements give CAT conversions less than twofold greater than background and are scored as inactive. The constructs containing the two proximal elements linked to the repeated AP1 enhancer give a signal 12- to 15-fold above background. Leaving out either the TATA element or the AP1 element reduces the signal to nearly background levels, indicating that in these synthetic constructs tandem GC-box elements cannot act alone to support expression.



CP1:	TTATTTGAACTAACCAATCAGTT	(MSV LTR)
CP2:	TGACCAGTTCCAGCCACTCTTTAT	(gamma-Fibrinogen)
NFI-C:	TTTTGGCTTGAAGCCAATATGAGT	(CONSENSUS)
NFI:	TTTTGGAATCTATCCAAGTCTTAT	(MMTV-LTR)
2GC:	GGGGCGGGGCGATCGGGGCGGGGC	(SP1-Consensus)
GC-up:	GGGGCGGGGCGATCTCCTGCAGCC	
GC-down:	GATCGGGGCGGGGC	
CACCC:	ATGAGAGCCACACCCAGCCTGTAT	(Tryptophan Oxygenase)
ATF/CREB:	CGGGCGGCTTTCGTCACAGGGTGCGGTC	(Ade E3)



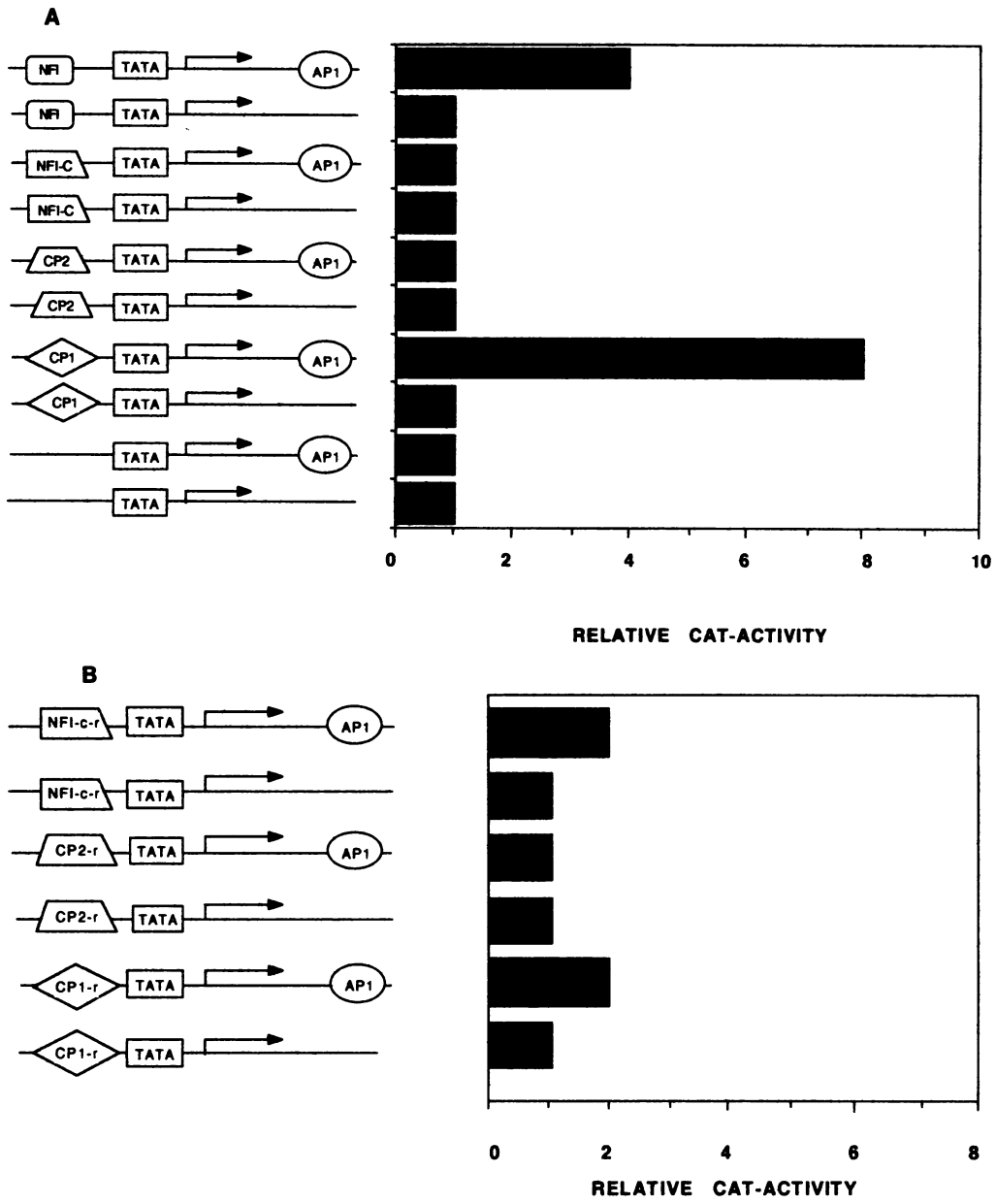


FIG. 2. Expression from various CCAAT-box-containing promoters. (A) The relative CAT activity of promoters containing binding sites for CP1, CP2, NFI, and NFI consensus, in combination with the TATA box and the AP1 enhancer. (B) The activity of promoters containing reversed CP1, CP2, and NFI consensus elements in combination with the TATA box and the AP1 enhancer.

FIG. 1. Restriction map, sequence, and start site mapping of the synthetic promoters. (A) The map of unique restriction sites of the synthetic promoter matrix (69 bp) and the sequences of the inserted promoter elements. The TATA-box sequence shown with *Sph*I sites on both ends was cloned into the *Sph*I site of the matrix. All other proximal elements, except for ATF/CREB, were cloned into the *Bcl*I site. The ATF/CREB site was directionally cloned between the *Bcl*I site and *Xho*I site. The DNA sequence and biological source of each element are shown. The sequence corresponds to the forward orientation. (B) Autoradiograph showing the mapping of RNA start sites. Lanes: 1, RNA from mock-transfected HeLa cells; 2, RNA from cells transfected from a synthetic promoter containing two GC boxes upstream of the TATA box and the SV40 enhancer downstream of the CAT gene; 3 to 6, RNA isolated from an in vitro transcription system of HeLa nuclear extract in the presence or absence of GAL-AH activator (6, 22). In lanes 3 and 4, the extract contains a template with five GAL4 sites upstream of the synthetic TATA box, in either the absence (lane 3) or presence (lane 4) of the activator. Lane 5 is the same as lane 4 except that the GAL4 sites on the template were located downstream of the CAT gene. Lane 6 is the same as lane 5 except that a proximal element, CP2, was present upstream of the TATA box. Similar results were obtained for promoters containing different proximal elements (data not shown). The positions of the two start sites are marked with arrows, with the heavy arrow indicating the in vivo start site. The location of the TATA box is marked with a open box. The sequence surrounding the start site DNA is shown below the figure.

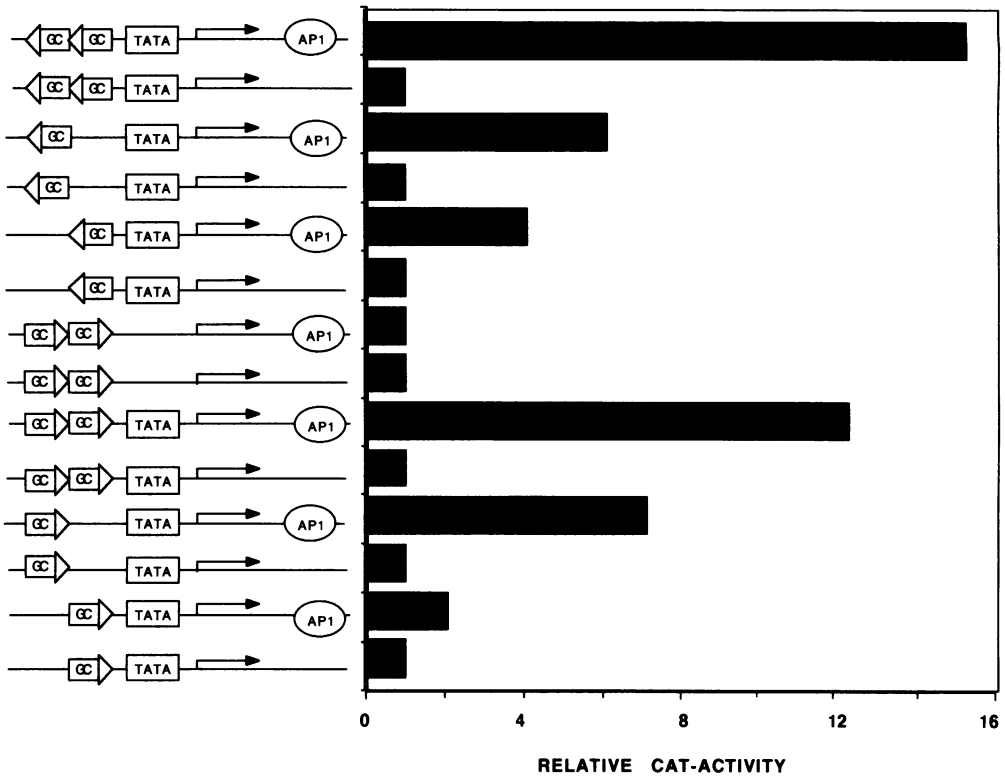


FIG. 3. Expression from various GC box-containing promoters. Different plasmids with either one or two GC boxes, and in either orientation, were transfected into HeLa TK$^-$ cells and analyzed for CAT activity. The arrows (◊) represent the orientation of the GC box as the sequence is shown in Fig. 1. The relationship of the GC box to the TATA box reflects the different positions of the GC box in various constructs (see GC-down and GC-up in Fig. 1A).

Next, we tested whether this promoter activity was due to one partner of the tandem GC-box sequences acting alone, to see whether the duplication of the element was important for its function. For this purpose, two additional synthetic elements were made, designated GC-up and GC-down in Fig. 1A. GC-up contains sequences identical to the upstream half of the tandem synthetic element, and GC-down has the sequences from the downstream half (Fig. 1). The upstream and downstream positions were also conserved compared with the tandem constructs, and they are shown as appropriately positioned single GC boxes in Fig. 3. Their reversed GC-box counterparts were also constructed and assayed.

The data (Fig. 3) show that both the upstream and downstream halves of the tandem element have lower activity than the corresponding double GC-box promoters. In one orientation (arrow pointing rightward in Fig. 3), the upstream half directs CAT expression at about 7 times background level and the downstream half directs expression at about 2 times background level (cf. 12 times background for the tandem GC-box construct). The difference between two- and sevenfold stimulation indicates that the position or phasing of the GC box is somewhat important in this orientation. In the other orientation (arrows pointing leftward in Fig. 3), the activity of the upstream GC-box promoter is more similar to that of the downstream one, yielding activities of sixfold and fourfold above background, respectively. This compares to a 15-fold-higher activity for the tandem construct containing elements in both positions. It is clear that the tandem constructs direct activity greater than that for each of the two parts. All constructs appear to require the remote AP1 enhancer element for significant activity, indicating that no proximal combinations alone contain sufficient information to direct significant assembly of active transcription complexes in HeLa cells. Only some combinations can even direct such assembly with the participation of the AP1 enhancer, as observed with the CCAAT box elements as well.

These results indicate that there is strong selectivity in the ability of certain proximal elements to cooperate with an AP1-type enhancer. Among the elements tested, only CP2 could not be activated. To learn the extent to which AP1 might cooperate with other types of proximal elements, we inserted and tested two less-common proximal elements.

The binding site for ATF/CREB is involved in the regulation of adenovirus E1A- and cellular cyclic AMP-inducible promoters (for a review, see reference 17). The synthetic sequence was chosen to match the adenovirus E3 promoter, since this promoter also has a linked AP1 enhancer site (13, 14). To preserve the distance between the ATF/CREB site and the TATA box as in this natural promoter, we cloned this element directionally between the *Bcl*I and *Xho*I sites in the synthetic promoter polylinker.

The results in Fig. 4 show that when the ATF/CREB site is combined with a TATA box the synthetic promoter has an activity that is about five times the background level. This is the only paired combination of elements in proximal positions that did not require linking to an enhancer site for significant activity. When this proximal promoter combination is linked to the AP1 enhancer, however, its activity increases to approximately 22-fold that of background.

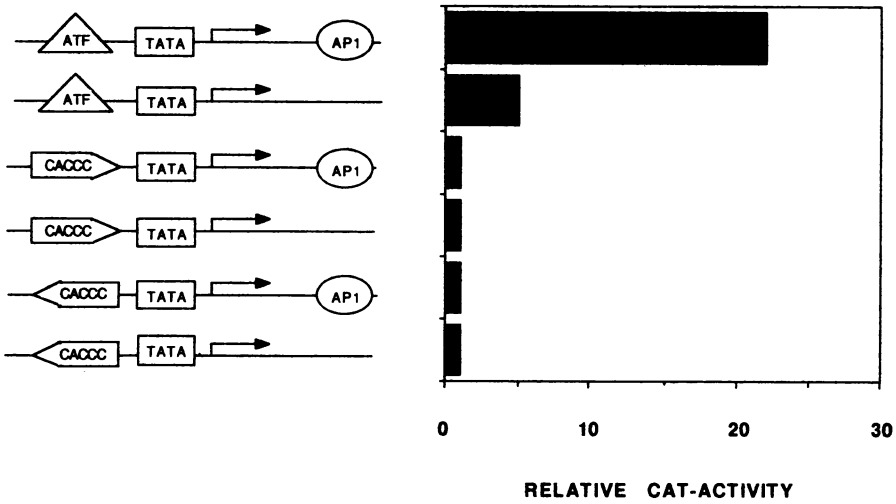


FIG. 4. Expression from plasmids containing the CACCC box and the ATF/CREB site in proximal positions. The arrows (◊) represent the orientation of the CACCC box as the sequence is shown in Fig. 1.

Lastly, we studied the CACCC box that has been found in a number of promoters, including various β-globin genes (10), rat tryptophan oxygenase (32), and rat tyrosine aminotransferase (3). It has been found in both orientations, and its position varies from within 100 bp of the start site of transcription to 2 kb away. We have chosen for study the CACCC-box region from the rat tryptophan oxygenase gene (Fig. 1A), which contains in reverse orientation a sequence of 11 nucleotides identical to the CACCC box of the β-globin promoter (32). This CACCC-box sequence was cloned in both orientations at the same restriction site used for the other proximal elements.

The results (Fig. 4) show that the CACCC box does not support any significant expression above the background when combined with the TATA box, regardless of its orientation. Linking the promoter to the AP1 enhancer does not appear to produce any higher activity. The distance between the two proximal sites is probably not the cause of inactivity, since one β-globin mutant promoter, which was deleted in its CCAAT box and therefore had a shortened distance between its TATA box and CACCC box, still had activity higher than the TATA box alone; the distance in that mutant is only 3 bp different from that in our synthetic promoter (12). The appropriateness of the construct is further supported by its ability to be activated by a different enhancer (see below).

Taken together, these results indicate that there is considerable selectivity in how AP1 acting as an enhancer cooperates with elements in proximal positions. Two of the six proximal elements could not be activated detectably in the context of the minimal promoter, and four other elements supported expression at 4-, 8-, 15-, and 22-fold above background levels. We now turn to other remote activation elements to learn whether they have related or unique patterns of selectivity. The proximal elements tested will be restricted to these six examples in their natural orientations.

The AP2 and SV40 enhancers show a different pattern of selectivity in cooperation with proximal elements. Multiple copies of AP2 sites were cloned downstream of proximal promoter elements. These AP2 sites were chosen from the human metallothionein IIA gene. They have been shown to be capable of mediating transcription as an enhancer (15).

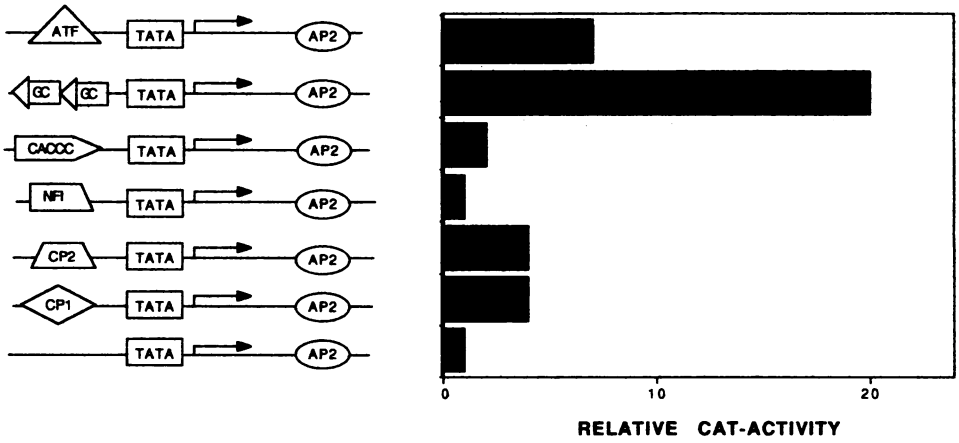


FIG. 5. Expression from plasmids containing various proximal elements when linked to the AP2 enhancer.

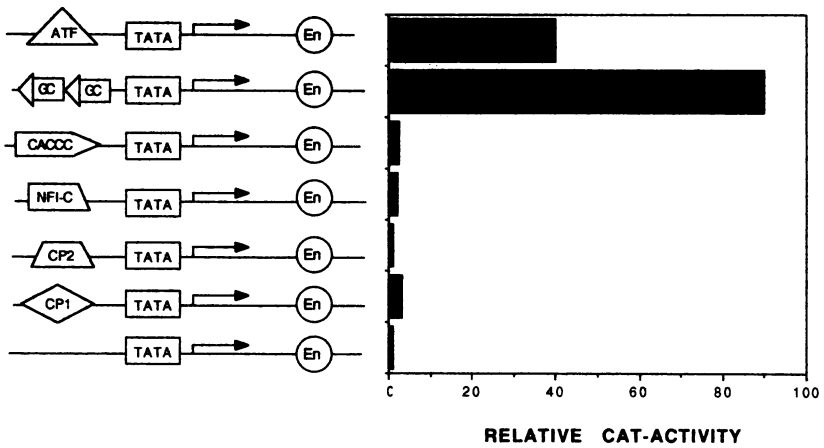


FIG. 6. Expression from plasmids containing various proximal elements when linked to the SV40 enhancer.

The activities of these constructs were determined, and the data are shown in Fig. 5.

The spectrum of AP2 activation is different from that of AP1. Two previously inactive proximal elements, CP2 and CACCC, can now be activated, although not very strongly. This refers to a fourfold activation of the CP2 site and a twofold activation of the CACCC site. The NF1 sequence that was fourfold activated by AP1 cannot be detectably activated by AP2. The ATF data are slightly complicated by its ability to be active without any enhancer (Fig. 4). The sevenfold background level of expression of the AP2-ATF-TATA construct is less than twofold higher than the fivefold for the ATF-TATA construct, and thus we score this as no AP2 stimulation. This contrasts with a signal 22-fold higher expression over background for the AP1-ATF-TATA construct.

There are similarities as well as these differences in the AP2 and AP1 activation patterns. In both cases the strongest stimulation was with the GC-box element (20- and 15-fold, respectively) and with the CP1 promoter (4- and 8-fold, respectively). Collectively, the results show that different proximal elements can be activated differently by different enhancers. Among the proximal elements, the best candidate for a universal-type cooperator is the GC-box element, which is stimulated strongly by both enhancers.

These results were obtained with multicopy synthetic enhancer elements. We used multiple copies on the basis of previous demonstrations that they are needed to generate a high signal strength, which could be particularly important in the context of these minimal promoters. Next, we replaced these synthetic sites with the natural SV40 enhancer. This sequence includes a number of different enhancer motifs, not all of which may bind factors in HeLa cells, a nonpermissive host for SV40 growth. We expected that if certain proximal elements were truly universally activable, the SV40 enhancer should activate them, even if a mixture of factors were bound to it. Conversely, for proximal elements with high selectivity in cooperation there may be insufficient binding of the appropriate type of enhancer proteins to achieve high activation.

A number of promoters with the linked SV40 enhancer were assayed for CAT expression, and the results are displayed in Fig. 6. The data show that by far the strongest activation (a 90-fold activation) is with the GC-box element. Recall that this element also cooperated best with the AP1 and AP2 enhancers, giving 15- and 20-fold stimulation, respectively. The only element that fails to be activated by the SV40 enhancer is CP2, which was also not responsive to AP1 activation but was responsive to AP2. These comparisons suggest that GC-box elements have the least selectivity and CP2 elements the most selectivity among the proximal elements tested. The remaining proximal elements were activated from two- to eightfold, consistent with the potential binding of a few copies of compatible factors over the enhancer.

Activation by the ER element. These experiments have relied on the activator proteins occurring naturally within HeLa cells to accomplish the activation. Although they are important in demonstrating selective interactions, they cannot, in the long run, identify the proteins involved. This is because there is probably a family of factors that recognizes each of the synthetic sequences and it is impossible to control or identify the factors bound. Therefore, in an initial attempt to investigate the selectivity of identified factors, we have used the GAL4 fusion system in which a cotransfected expression vector directs unique factor binding at introduced yeast GAL4 DNA-binding sites (19). When expressed in mammalian cells, these fusion proteins can function as enhancers when the GAL4 DNA sites are downstream of the test gene. Since there are apparently no endogenous proteins that activate from the unnatural yeast GAL4 sequence, the factor that activates is known.

Five copies of GAL4 DNA-binding sites were cloned downstream of these synthetic promoters at the same location as the synthetic enhancer sequences studied above. In this initial test, the fusion protein was a GAL4 ER (37); eventually we hope to test a variety of activation domains in this manner. Each promoter bearing a proximal element and GAL4 sites was cotransfected with the plasmid expressing the fusion protein. CAT assays were then done to determine how the ER cooperated with the various proximal domains. The ER has been shown to contain two activation domains, AB and EF. AB is a constitutive domain but is very weak in HeLa cells (37), and, consistent with this, there is no detectable stimulation on all synthetic promoters (data not shown), regardless of whether estrogen is present. This control demonstrates that estrogen itself does not change the signal above the background level. The domain EF has been shown to be a hormone-inducible activation domain and a strong activator in HeLa cells (37).

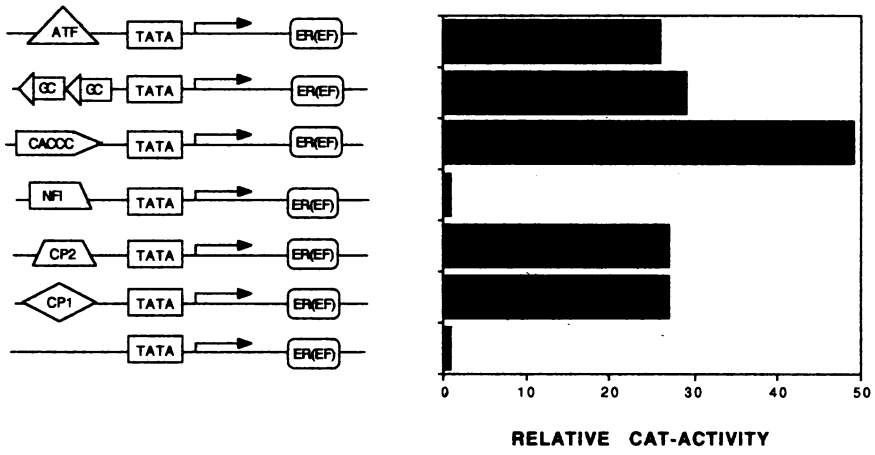


FIG. 7. Expression from plasmids containing various proximal elements when linked to the GAL4 DNA-binding sites. Each of them was cotransfected with the effector plasmid which expresses the GAL-ER(EF), the chimera of GAL4 DNA-binding domain and the EF domain of the ER (37). The CAT assays were performed 48 h posttransfection in these experiments.

When the ER domain is tested against the strongest and least selective proximal element, the tandem GC box, it gives a 29-fold stimulation (Fig. 7). This is consistent with the suggested universal response of this proximal element. Quantitatively, the 29-fold activation is typical of the response of synthetic elements directing endogenous stimulation, which was 15-, 20-, and 90-fold for the three enhancers tested. This indicates that the Gal4 and endogenous enhancer test systems are probably behaving similarly.

When tested against the other proximal elements, the ER activator stimulates all except the NF1 element (Fig. 7). This lack of stimulation of NF1 contrasts to the weak but detectable stimulation by the AP1 and SV40 enhancers, indicating some selectivity in ER activation. The CP2 element was greatly (27-fold) stimulated, which contrasts very strongly with the lack of activation by the AP2 and SV40 enhancers. The identical 27-fold stimulation of the CP1 element is the strongest stimulation of that element, and the 26-fold stimulation of the ATF element is also the strongest. With the exception of NF1, the quantitative stimulation of the proximal elements is very significant, suggesting that the ER is the least selective of the enhancers studied here.

DISCUSSION

These results indicate that there is very significant diversity in the mechanism by which enhancer and proximal elements cooperate to direct appropriate transcription levels. All elements tested could cooperate with at least one other element. Two of the six proximal elements could be stimulated by all of the enhancers. None of the enhancers stimulated all the tested proximal elements. About three-quarters of the enhancer-proximal combinations were active, and about one-quarter were inactive. Among the active combinations, about half were weak at 2- to 4-fold background levels and about half were strong at 15- to 90-fold background levels. These data are collected in Table 1, which also includes, for the sake of comparative discussion, data obtained by using the VP16 activation domain, since this appears to activate universally (38a).

We will attempt to break the data in Table 1 into categories. The expectation is that these categories will eventually be explained by the involvement of common remote-proximal adapters or coactivators or common protein-protein interactions of some sort. In the data set, the CP1 and tandem GC-box elements have in common that they cooperate in transcription with all five enhancers tested. The GC-box-dependent transcription is always strong, and the CP1-dependent expression ranges from strong to weak. To work, both require a TATA box. Thus, from these data, these elements appear to have very low selectivity in how they may be activated by other transcription elements. The protein sequence of factors known to bind these elements in vitro has been established. Although there is no common sequence, there is a common motif, a glutamine-rich transcription activation domain (8, 9, 38). One possibility is that the ability of glutamine to act as both a facile hydrogen bond donor and acceptor underlies the ability of these proteins to cooperate with the wide variety of other elements.

By contrast, two of the proximal elements, CP2 and CACCC, show very great selectivity in how they cooperate with enhancers. For CACCC the activation varies from

TABLE 1. Summary of the activation of different proximal elements by different enhancers[a]

Enhancer[b]	Activation (fold) of proximal element:						
	Only TATA	CP1 + TATA	CP2 + TATA	NFI + TATA	CACCC + TATA	2 GC + TATA	ATF + TATA
None	1	1	1	1	1	1	5
AP1	1	8	1	4	1	15	22
AP2	1	4	4	1	2	20	7
SV40 enhancer	1	3	1	2	3	90	40
GAL-ER(AB)	1	1	1	1	1	1	1
GAL-ER(EF)	1	27	27	1	49	29	26
GAL-VP16	3	34	45	4	31	103	116

[a] All promoters are in the context of the adenovirus major late promoter TATA box. All proximal elements are positioned at the *BclI* site upstream, and all enhancers are cloned at the *Bam*HI site which is about 1.7 kb downstream. Transfection in HeLa cells and CAT assays were described in the text.

[b] The AP1 enhancer contains 10 copies of its binding sites. The AP2 enhancer and the GAL4 reporter plasmids contain five copies of their DNA-binding sites. GAL-ER(AB) and GAL-ER(EF) refer to the chimera of GAL4 DNA-binding domain and the AB or EF activation domain of ER (37). GAL-VP16 refers to the fusion of the same DNA-binding domain with the VP16 acidic domain (30).

49-fold with the estrogen receptor to inactive with AP1. For CP2 the activation ranges from 27-fold with the estrogen receptor to inactive with AP1 and the SV40 enhancer. Except for the weak activation of CACCC by the SV40 enhancer, the patterns are quite similar for these two proximal elements. Since the activation domains of factors that might bind here have not been established, one cannot yet speculate on what they might have in common. The activation patterns for the two remaining elements in proximal positions, ATF and NF1, are different from all of the above and from each other, emphasizing the potential for diversity in the ability of enhancers and proximal elements to cooperate in directing differential transcription.

From the point of view of selectivity in remote interactions, none of the enhancers tested here can cooperate with all the six proximal elements. The least selective are the estrogen receptor and VP16, which fail to stimulate only NF1. NF1-dependent stimulation is quite weak with the AP1 and SV40 enhancers, and thus the lack of detectable stimulation by the other enhancers may not be an indication of inability to cooperate, since the source of general weak stimulation is not known. This element has been shown to cooperate well with the glucocorticoid receptor element in the same cells we used (31). The estrogen receptor has no common activation motifs and lacks the acidic motif of VP16, and thus one cannot yet say what determines the low selectivity of ER and VP16. Their common use of the GAL4 fusion system is not the source of their common strong response, since, for example, the AP1 and SV40 enhancers are roughly the same strength on ATF and GC-box constructs but fail to interact with other proximal elements. As a whole, the selectivity in the enhancer set is quite high.

This study was conducted by using one cell type growing in fixed media and using constructs with only one proximal element, one enhancer, and a TATA element. All the elements contributed to activation in at least one context, implying that at least one active transcription factor could bind them in this situation. Collectively, the data show that certain elements can cooperate with low selectivity and that others can cooperate with high selectivity. It will be a challenge to identify the functional domains on these factors and their hypothetical adapters or coactivators that are responsible for these differential effects. It is possible that the rules describing the extent of cooperation among elements will be different under different conditions; presumably the mix of active factors in different cells will alter the rules somewhat, accounting in significant part for differential transcription. It has been shown, for example, that an enhancer consisting of a steroid-responsive element coupled to an NFI site is functional in HeLa cells but not in some other cell lines (34). We hope that the application of this test system to other environments will facilitate the testing of these ideas and the identification of the classes of molecules whose ability to cooperate underlies differential transcription.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Science Foundation.

We thank R. Tjian, P. Chambon, and M. Ptashne for providing us with the plasmids and A. Courey for critical reading of the manuscript.

REFERENCES

1. **Angel, P., M. Imagawa, R. Chiu, B. Stein, R. J. Imbra, J. H. Rahmsdorf, C. Jonat, P. Herrlich, and M. Karin.** 1987. Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. Cell **49**:729–739.
2. **Auffray, C., and F. Rougeon.** 1980. Purification of mouse immunoglobulin heavy chain messenger RNAs from total myeloma tumor RNA. Eur. J. Biochem. **107**:303–314.
3. **Becker, P. B., B. Gloss, W. Schmid, U. Strähle, and G. Schütz.** 1986. *In vivo* protein-DNA interactions in a glucocorticoid response element require the presence of the hormone. Nature (London) **324**:686–688.
4. **Berger, S. L., W. D. Cress, A. Cress, S. J. Triezenberg, and L. Guarente.** 1990. Selective inhibition of activated but not basal transcription by the acidic activation domain of VP16: evidence for transcriptional adaptors. Cell **51**:1199–1208.
5. **Bucher, P., and E. N. Trifonov.** 1986. Compilation and analysis of eukaryotic POL II promoter sequences. Nucleic Acids Res. **14**:10008–10026.
6. **Carey, M., Y.-S. Lin, M. R. Green, and M. Ptashne.** 1990. A potent GAL4 derivative activates transcription at a distance *in vitro*. Science **247**:710–712.
7. **Chodosh, L. A., A. S. Baldwin, R. W. Carthew, and P. A. Sharp.** 1988. Human CCAAT-binding proteins have heterologous subunits. Cell **53**:11–24.
8. **Courey, A. J., D. A. Holtzman, S. P. Jackson, and R. Tjian.** 1989. Synergistic activation by the glutamine-rich domains of human transcription factor SP1. Cell **59**:827–836.
9. **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.
10. **Dierks, P., A. van Ooyen, M. D. Cochran, C. Dobkin, J. Reiser, and C. Weissmann.** 1983. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit beta-globin gene in mouse 3T6 cells. Cell **32**:695–706.
11. **Dynan, W.** 1989. Modularity in promoters and enhancers. Cell **58**:1–4.
12. **Efstratiadis, A., J. W. Posakony, T. Maniatis, R. M. Lawn, C. O'Connell, R. A. Spritz, J. K. Deriel, B. G. Forget, S. M. Weissman, J. L. Slightom, A. E. Blechl, O. Smithies, F. E. Baralle, C. C. Schoulders, and N. J. Proudfoot.** 1980. The structure and evolution of the human beta-globin gene family. Cell **21**:653–668.
13. **Garcia, J., F. Wu, and R. Gaynor.** 1987. Upstream regulatory regions required to stabilize binding to the TATA sequence in an adenovirus early promoter. Nucleic Acids Res. **15**:8367–8384.
14. **Hurst, H. C., and N. C. Jones.** 1987. Identification of factors that interact with the E1A-inducible adenovirus E3 promoter. Genes Dev. **1**:1132–1146.
15. **Imagawa, M., R. Chiu, and M. Karin.** 1987. Transcription factor AP2 mediate induction by two different signal-transduction pathways: protein kinase C and cAMP. Cell **51**:251–260.
16. **Jones, K. A., J. T. Kadonaga, P. J. Rosenfeld, T. J. Kelly, and R. Tjian.** 1987. A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell **48**:79–89.
17. **Jones, N. C., P. W. J. Rigby, and E. G. Ziff.** 1988. Trans-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses. Genes Dev. **2**:267–281.
18. **Kadonaga, J. T., K. A. Jones, and R. Tjian.** 1986. Promoter-specific activation of RNA polymerase II transcription by Sp1. Trends Biochem. Sci. **11**:20–23.
19. **Kakidani, H., and M. Ptashne.** 1988. GAL4 activates gene expression in mammalian cells. Cell **52**:161–167.
20. **Kelleher, R. J., P. M. Flanagan, and R. D. Kornberg.** 1990. A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. Cell **61**:1209–1215.
21. **Lee, W., P. Mitchell, and R. Tjian.** 1987. Purified transcription factor SP-1 interacts with TPA-inducible enhancer elements. Cell **49**:741–752.
22. **Lin, Y.-S., M. F. Carey, M. Ptashne, and M. R. Green.** 1988. GAL4 derivatives function alone and synergistically with mammalian activators *in vitro*. Cell **54**:659–664.
23. **McKnight, S. L.** 1982. Functional relationships between transcriptional control signals of the thymidine kinase gene of herpes simplex virus. Cell **31**:355–365.

24. **Miksicek, R., U. Borgmeyer, and J. Nowock.** 1987. Interaction of the TGGCA-binding protein with upstream sequences is required for efficient transcription of mouse mammary tumor virus. EMBO J. **6**:1355–1360.
25. **Mitchell, P. J., and R. Tjian.** 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245**:371–378.
26. **Mitchell, P. J., C. Wang, and R. Tjian.** 1987. Positive and negative regulation of transcription *in vitro*: enhancer-binding protein AP-2 is inhibited by SV40 T antigen. Cell **50**:847–861.
27. **Morgan, J. G., G. Courtois, G. Fourel, L. A. Chodosh, L. Campbell, E. Evans, and G. R. Crabtree.** 1988. Sp1, a CAAT-binding factor, and the adenovirus major late promoter transcription factor interact with functional regions of the gamma-fibrinogen promoter. Mol. Cell. Biol. **8**:2628–2637.
28. **Ptashne, M.** 1988. How eukaryotic transcriptional activators work. Nature (London) **335**:683–689.
29. **Pugh, B. F., and R. Tjian.** 1990. Mechanism of transcriptional activation by SP1: evidence for coactivators. Cell **61**:1187–1197.
30. **Sadowski, I., J. Ma, S. Triezenberg, and M. Ptashne.** 1988. Gal4-VP16 is an unusually potent transcriptional activator. Nature (London) **335**:563–564.
31. **Schüle, R., M. Müller, C. Kaltschmidt, and R. Renkawitz.** 1988. Many transcription factors interact synergistically with steroid receptors. Science **242**:1418–1420.
32. **Schüle, R., M. Müller, H. Otsuka-Murakami, and R. Renkawitz.** 1988. Cooperative of the glucocorticoid receptor and the CACCC-box binding factor. Nature (London) **332**:87–90.
33. **Spaete, R. R., and E. S. Mocarski.** 1985. Regulation of cytomegalovirus gene expression: α and β promoters are *trans*-activated by viral functions in permissive human fibroblasts. J. Virol. **56**:135–143.
34. **Strähle, U., M. Boshart, G. Klock, F. Stewart, and G. Schütz.** 1989. Glucocorticoid- and progesterone-specific effects are determined by differential expression of the respective hormone receptors. Nature (London) **339**:629–632.
35. **Strähle, U., W. Schmid, and G. Schütz.** 1988. Synergistic action of the glucocorticoid receptor with transcription factors. EMBO J. **7**:3389–3395.
36. **Tasset, D., L. Tora, C. Fromental, E. Scheer, and P. Chambon.** 1990. Distinct classes of transcriptional activating domains function by different mechanisms. Cell **62**:1177–1187.
37. **Tora, L., J. White, C. Brou, D. Tasset, N. Webster, E. Scheer, and P. Chambon.** 1989. The human estrogen receptor has two independent nonacidic transcriptional activation functions. Cell **59**:477–487.
38. **van Huijsduijnen, R. H., X. Y. Li, D. Black, H. Matthes, C. Benoist, and D. Mathis.** 1990. Co-evolution from yeast to mouse: cDNA cloning of the two NF-Y CP-1/CBF) subunits. EMBO J. **9**:3119–3127.

38a. **Wang, W., and J. D. Gralla.** Submitted for publication.

39. **Williams, G. T., T. K. McClanahan, and R. I. Morimoto.** 1989. E1a transactivation of the human hsp70 promoter is mediated through the basal transcription complex. Mol. Cell. Biol. **9**:2574–2587.