MOLECULAR AND CELLULAR BIOLOGY, May 1996, p. 2274–2282 Vol. 16, No. 5
0270-7306/96/$04.00+0

A Unique Transactivation Sequence Motif Is Found in the Carboxyl-Terminal Domain of the Single-Strand-Binding Protein FBP

ROBERT DUNCAN,† IRENE COLLINS, TAKESHI TOMONAGA, TIAN ZHANG, AND DAVID LEVENS*

Laboratory of Pathology, National Cancer Institute, Bethesda, Maryland 20892

Received 4 October 1995/Returned for modification 14 November 1995/Accepted 26 February 1996

The far-upstream element-binding protein (FBP) is one of several recently described factors which bind to a single strand of DNA in the 5′ region of the c-*myc* gene. Although cotransfection of FBP increases expression from a far-upstream element-bearing c-*myc* promoter reporter, the mechanism of this stimulation is heretofore unknown. Can a single-strand-binding protein function as a classical transactivator, or are these proteins restricted to stabilizing or altering the conformation of DNA in an architectural role? Using chimeric GAL4-FBP fusion proteins we have shown that the carboxyl-terminal region (residues 448 to 644) is a potent transcriptional activation domain. This region contains three copies of a unique amino acid sequence motif containing tyrosine diads. Analysis of deletion mutants demonstrated that a single tyrosine motif alone (residues 609 to 644) was capable of activating transcription. The activation property of the C-terminal domain is repressed by the N-terminal 107 amino acids of FBP. These results show that FBP contains a transactivation domain which can function alone, suggesting that FBP contributes directly to c-*myc* transcription while bound to a single-strand site. Furthermore, activation is mediated by a new motif which can be negatively regulated by a repression domain of FBP.

The level of expression of the c-*myc* proto-oncogene is influenced by a multitude of signals, many of which have been shown to act through nuclear proteins binding to specific sequences within the *myc* DNA (19). The far-upstream element (FUSE) is one such sequence required for maximal transcription of the c-*myc* gene (3). A FUSE-binding protein (FBP) interacts specifically with this site and stimulates expression in a FUSE-dependent manner (11). The level of the FBP mRNA and FUSE DNA binding activity declines sharply with differentiation of the leukemia cell line HL60 concurrent with the decrease in c-*myc* transcription, suggesting that FBP is important to maintain active c-*myc* transcription in the undifferentiated state. In peripheral blood lymphocytes the expression of FBP and that of c-*myc* are highly correlated, rising and falling in parallel (4).

The DNA in the FUSE region of an actively transcribed c-*myc* gene is sensitive to permanganate modification, in vivo, suggesting that this region is single stranded (11, 22). Significantly, FBP binds stably to a single DNA strand of the FUSE site but does not form a stable complex with linear or relaxed B-DNA. Remarkably, FBP specifically recognizes its target sequence as a duplex when FUSE is embedded in a plasmid under negative superhelical strain (5, 22). FBP is a 644-amino-acid protein in which three domains have been defined from primary structure elements. The central domain of the protein employs a novel repeated structure to destabilize the double helix and to bind sequence specifically to one strand. The functions of the N-terminal and C-terminal domains have heretofore been unknown.

The in vitro DNA binding properties of FBP and the peculiar secondary structure of FUSE when the c-*myc* gene is activated, in vivo, suggest that FBP is bound to FUSE in cells. Furthermore, increasing the level of FBP by transient transfection stimulates expression from a c-*myc* promoter-driven reporter; however, the mechanism of FBP-mediated activation has yet to be solved. One means to modify c-*myc* expression may utilize FBP to pry open FUSE, facilitating strand separation and creating a focus of melted DNA at a strategic site; this nidus of melted DNA might dramatically reduce torsional strain while simultaneously augmenting flexibility to DNA bending (17). FBP would act indirectly to augment c-*myc* expression by facilitating interactions between bona fide transcription factors. Creating a single-stranded loop, FBP would eliminate the stereochemical barriers imposed by helical phasing and the intrinsic rigidity of duplex B-DNA. In this model, FBP action is driven by its peculiar DNA binding properties and the protein itself need not possess effector domains capable of modifying transcription directly. Thus, FBP would be a transcriptional coactivator functioning in a manner conceptually similar to architectural transcription factors (31, 33) such as HMG I/Y and HMG-1/2 (28). Further evidence to suggest that FBP functions as a coactivator comes from the observation that a FUSE site does not stimulate expression when placed in front of a heterologous promoter (2). The action of FBP binding may have no effect in the absence of adjoining activators in the *myc* regulatory region.

Alternatively, FBP might employ its DNA binding domain to recognize a single-stranded or negatively supercoiled FUSE. By tethering an activation domain to this platform, FBP function could be restricted to those conditions in which FUSE adopts a single-stranded conformation or is under negative superhelical torsion. In this model, FBP behaves as a modular gene regulator possessing both DNA binding and activator domains similar to conventional transcription factors. A key prediction of this model requires that FBP bear an activation domain which is functionally dissociable from FUSE binding. Outside of the DNA binding domain, FBP has short segments rich in glycine, proline, or glutamine as well as an unusual tyrosine motif (YM) repeated thrice (YM1, YM2, and YM3) in the carboxyl-terminal domain. However, none of these segments was sufficiently homologous to known activator domains to indicate effector function. To address this point, a series of

* Corresponding author. Phone: (301) 496-2176. Fax: (301) 402-0043. Electronic mail address: levens@helix.nih.gov.

† Present address: Laboratory of Molecular Pharmacology, Center for Biologics Evaluation and Research, Food and Drug Administration, Bethesda, MD 20892.

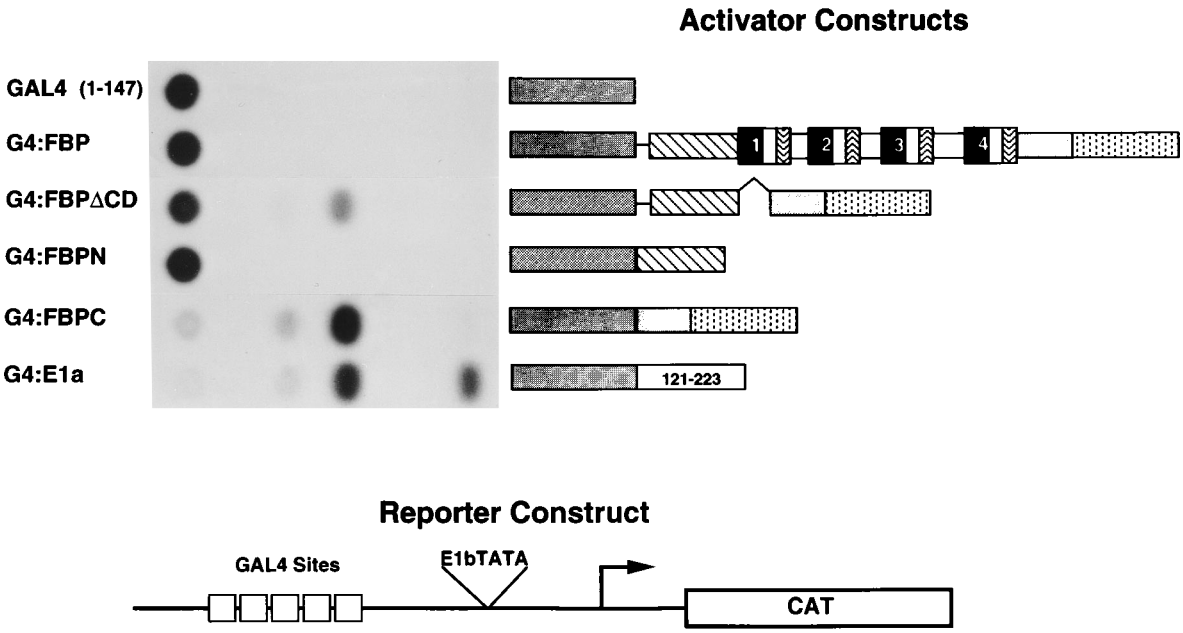

FIG. 1. The carboxyl-terminal region of FBP is a potent activation domain. The GAL4 DNA binding domain was fused to full-length or separate domains of FBP and cotransfected with the reporter diagrammed in the lower part of the figure. A representative CAT assay is shown to the left of the box diagram of each activator construct. Darkly shaded boxes, the GAL4 DNA binding domain; diagonally striped boxes, FBP N-terminal domain; lightly shaded boxes, C-terminal region rich in proline and glycine; dotted boxes, C-terminal region containing tyrosine dyad repeats. The central domain has four repeated units made up of a KH motif (black boxes), a spacer, and an α-helix (wavy-lined boxes). The name of each construct is shown at the far left. The activation domain of E1a (residues 121 to 223) fused to a GAL4 DNA binding domain was used as a positive control. The reporter plasmid (2.5 μg) and the expression plasmid (1 μg) were transfected into HeLa cells. At least four independent transfections gave comparable results.

chimeric molecules were generated by fusing FBP or portions thereof to the DNA binding domain of GAL4. These constructs were assayed for their ability to transactivate reporters driven by GAL4 binding sites.

MATERIALS AND METHODS

Cell culture and transfection. HeLa cells were cultured in Dulbecco's modified Eagle's medium (GIBCO-BRL) with 10% fetal calf serum. Transfections were performed by electroporation as described elsewhere (32). After 48 h of culture, cell extracts were prepared, normalized for protein content, and assayed for chloramphenicol acetyltransferase (CAT) activity as described previously (13).

Plasmid constructs. GAL4 fusion protein expression plasmids were generated for full-length FBP and FBPΔCD by inserting a restriction fragment from the cDNA clone into the cloning site of pSG424 (26). The central domain deletion was generated by PCR and removes amino acids 107 to 450. FBP N-terminal, C-terminal, and C-terminal deletion constructs were created by using PCR primers which anneal at the proper position in the FBP cDNA and carried restriction sites that allowed PCR fragments to be fused in frame to the 3′ end of the GAL4(1-147) coding sequence. The G4:FBP2C and G4:FBP3C plasmids were similarly constructed by using the corresponding cDNA clones as the template and primers which encompass amino acid residues 439 to the C terminus at 651 for FBP2 and 458 to the C terminus at 600 for FBP3 (10). The G4:FBPCΔYM plasmid was made by deleting a *Bam*HI-*Cvn*I fragment from G4:FBPC and blunting and religating the plasmid. The G4:FBPC YM3 mutants were formed by synthesizing two complementary oligonucleotides which, after annealing, encoded the residues shown in Fig. 5; these fragments were substituted for a restriction fragment in the G4:FBPC YM3 plasmid. The G4:YM3 construct was formed by ligating a restriction fragment which encodes the first residue of YM3 (Pro-609) to the C terminus of FBP into pSG424, thereby fusing YM3 in frame directly to the GAL4 DNA binding domain. The G4:VP16 and G4:E1a plasmids were previously described (references 25 and 18, respectively).

The reporter plasmids contain either one or five GAL4 binding sites upstream of a minimal E1b TATA sequence and the CAT coding sequence described previously (18).

The reporter pMP *Spe-Acc*-CAT is derived from pMP CAT (11) by deletion of a *Spe-Acc* restriction fragment from the 5′ *myc* flanking region of that plasmid and has no effect on FUSE-mediated activation (unpublished data).

Immunoblot analysis. COS cells were used to quantitate protein levels, because higher levels of protein were expressed in them. Extracts were made 48 h after transfection, and 50 μg protein was separated on a sodium dodecyl sulfate–12.5% polyacrylamide gel electrophoresis gel, transferred to Immobilon-P (Millipore), and probed with a polyclonal rabbit antibody raised against GAL4(1-147) (Upstate Biotechnology Inc.). The blot was developed by incubation with peroxidase-coupled antibody to rabbit immunoglobulin followed by enhanced chemiluminescence (Amersham).

RESULTS

Transactivation by the C-terminal domain of FBP. Plasmids encoding GAL4-FBP fusion proteins (Fig. 1) were transfected into HeLa cells along with a reporter plasmid which had five GAL4 binding sites placed upstream of the adenovirus E1b TATA sequence to drive expression of the CAT gene [$(G4)_5$CAT]. The full-length FBP (G4:FBP) showed no trace of transactivating capacity in multiple transfections. Why was the full-length protein devoid of transactivating properties? First, the presence of the FBP DNA binding domain might simply overpower the GAL4 DNA binding domain, thereby redirecting the chimeric transactivator away from the GAL4 upstream activating sequences (UASs) on the reporter plasmid to chromosomal sites. Second, the activating properties of FBP might be latent or masked, thereby requiring protein modification, conformational changes, or dissociation from an inhibitor in order to function. Precedent exists for each of these mechanisms in well-studied transcription factors such as HSF (34), NF-κB (29), SP1 (23), and STATs (9, 15). Lastly, FBP might simply not possess a transcription activation domain.

To test the hypothesis that the FBP DNA binding domain prevented GAL4-FBP from transactivating plasmid-directed CAT synthesis, the central domain was deleted and the new chimera was assayed for GAL4 UAS-dependent gene stimulation. This FBP derivative, G4:FBPΔCD, weakly but clearly transactivated reporter expression (Fig. 1). In contrast, the GAL4(1-147) polypeptide was completely inactive, as expected. Therefore, deletion of the DNA binding domain either

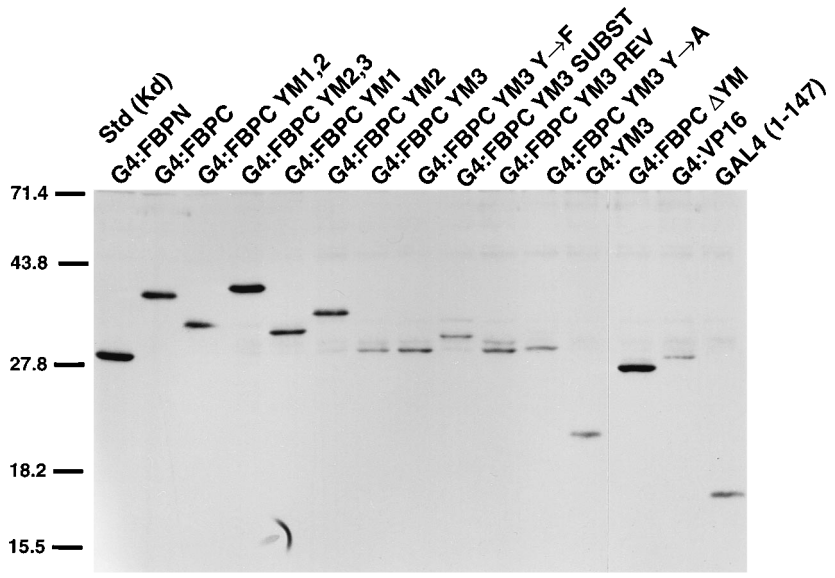

FIG. 2. GAL4-FBP fusion proteins are expressed appropriately in cells. Shown is an immunoblot of 50-μg protein samples of extracts from COS cells transiently transfected with plasmid constructs listed above each lane. The primary antibody was polyclonal anti-GAL4. The sequence compositions of the FBP constructs are diagrammed in Fig. 1, 4, 5, and 6. Std, molecular mass standards.

exposed transactivating properties inherent within FBP or generated an artificial transactivating segment. To localize better the portion of FBP responsible for transactivation, the amino and carboxyl termini of FBP were separately fused to the GAL4 DNA binding domain and assayed by transient transfection. Although the N-terminal 107-amino-acid segment alone was completely incapable of transactivation (Fig. 1, G4:FBPN), the C-terminal domain, composed of residues 451 to 644 (G4:FBPC), strongly activated CAT expression. Estimation of transactivation potency from a dose response of transfected G4:FBPC plasmid suggested that the C-terminal domain of FBP approached the activity of the adenovirus activator E1a in its ability to drive CAT expression. Notably, the activity of G4:FBPC greatly exceeded that of G4:FBPΔCD. All of the constructs expressed proteins of the predicted molecular weight, as indicated by Western blots (immunoblots) of whole-cell extract from transfected cells using an anti-GAL4 antibody (Fig. 2 and data not shown). The minor variation in the level of chimeric protein expression did not correlate with the degree of CAT gene activation.

Unlike other activators tested in the GAL4 system (7, 18), the C-terminal activation domain of FBP did not rely on multiple copies of the GAL4 site to stimulate CAT expression. Cotransfection of the G4:FBPC-encoding plasmid with a reporter construct containing a single GAL4 binding site upstream of an E1b TATA and a CAT gene resulted in CAT expression at least comparable to, if not greater than, that seen with $(G4)_5$CAT (Fig. 3).

Tyrosines are essential for activation by the AW(A/E)(A/E)YY sequence motif. The amino acid sequence of the C-terminal region of FBP resembles no known activation domains, and in fact no homologous regions were found after screening several databases. Several peculiar features indicated that the C-terminal domain is itself a patchwork of shorter segments, perhaps defining functional units. First, the C-terminal domain begins with a 63-amino-acid segment rich in proline, glycine, and histidine including the sequence PHGP repeated three times. A striking feature of the remaining segment is the presence of five sets of tyrosine dyads within or closely associated with one of three copies of a novel sequence motif, AW(A/E)(A/E)YY. Between these repeated elements the protein is rich in proline, glutamine, glycine, and threonine. Bulky hydrophobic and ionic residues are almost absent except within or adjacent to a tyrosine-rich motif.

To test the hypothesis that some of these structures play a functional role in activation, a series of deletions were made in the G4:FBP C-terminal domain construct (Fig. 4A). Each of these constructs was cotransfected with $(G4)_5$CAT. CAT assays from these transfections (Fig. 4B) revealed that the FBP activation domain is composed of functionally redundant elements, because nonoverlapping segments expressed separately

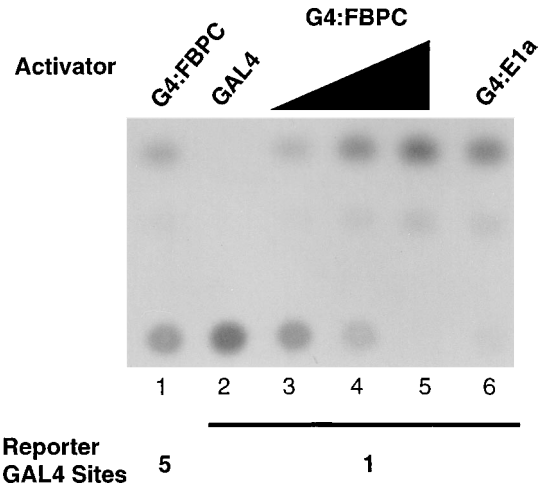

FIG. 3. The C-terminal domain of FBP can activate transcription when anchored to a single GAL4 site. The GAL4 DNA binding domain/FBP C-terminal domain-expressing plasmid (G4:FBPC) was cotransfected with a CAT reporter containing a single GAL4 UAS. Five micrograms of reporter was cotransfected into HeLa cells along with 200 ng of expression plasmid in lanes 1, 2, and 6. A dose-response curve for G4:FBPC used 40 ng, 200 ng, and 1 μg of this plasmid in lanes 3, 4, and 5 respectively.

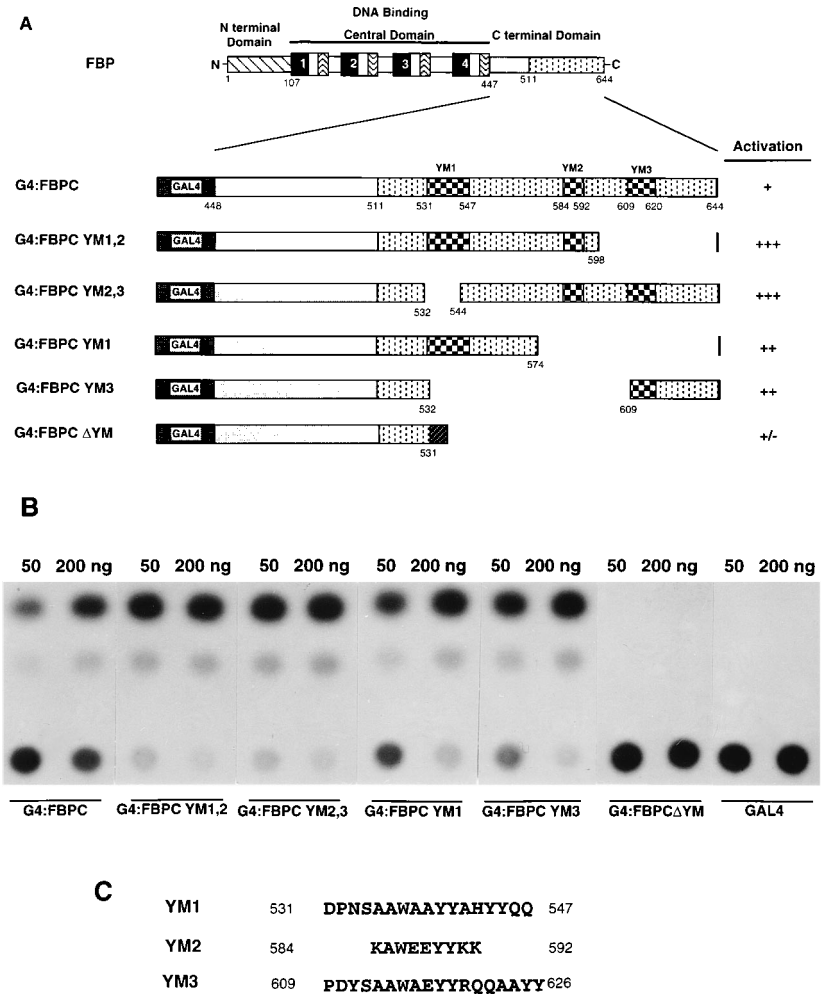

FIG. 4. The *trans*-activating derivatives of the FBP C terminus have the tyrosine motif (YM) as a common feature. (A) Box diagrams of FBP C-terminal domain–GAL4 fusion proteins with segments deleted as shown; numbers below each diagram indicate amino acid positions of structures. Shading symbolisms as described in the legend to Fig. 1; checkered boxes indicate tyrosine motifs (YM); heavy diagonal stripes in ΔYM indicate 8 residues added as a result of subcloning. The relative level of CAT expression activated by each construct is shown at the right. (B) CAT assays of transfections which contained 2.5 μg of the $(G4)_5$CAT reporter and either 50 or 200 ng of the expression plasmid as indicated. (C) The conservation of the tyrosine motifs is shown by alignment of residues between the amino acid position numbers indicated.

each drove reporter expression. The transcription activation appeared to be associated with the tyrosine motif (YM1, YM2, and YM3; Fig. 4C), because every transactivating protein contained at least one such element. Deletion of all the tyrosine motifs yielded a construct expressing only the proline-glycine-rich segment of the carboxyl terminus (G4:FBPCΔYM) which was severely crippled for activation.

The multiple activating elements within the carboxyl terminus of FBP did not synergize with each other to augment CAT activity. A truncated construct terminating at residue 598 (FBPC YM1,2) lacks the third tyrosine motif yet activated at least as well as the full C-terminal domain. Similarly, a small deletion from residues 532 to 544 which eliminated the first tyrosine motif (FBPC YM2,3) also yielded a potent activator. In fact, molecules possessing only a single tyrosine motif also proved to be effective transactivators. Constructs deleted beyond residue 574 expressing only the first tyrosine motif and constructs containing a fusion of just the last 35 residues of FBP containing the third tyrosine motif to the end of the Pro-Gly segment both activated. Another construct containing residues 544 to 596 encompassing only the second tyrosine motif activates as effectively as the other single-motif constructs (data not shown). In summary, the only feature shared between the different segments of the FBP carboxyl terminus which confer activator activity upon the GAL4 DNA binding domain is the presence of one or more tyrosine motifs.

Is the tyrosine motif itself necessary to mediate activation by the FBP carboxyl terminus? To further examine the sequence and structural restrictions on the activation motif, plasmids encoding mutant proteins possessing alterations within the last 35 amino acids of FBP fused to the Pro-Gly segment joined to the GAL4 DNA binding domain were generated (Fig. 5A). Four mutants were constructed; the first, FBPC YM3:Y→F, conservatively replaced each of the five tyrosines in the last 35 residues of FBP with phenylalanine. The second mutant, YM3:Y→A, more severely altered the primary sequence by substituting alanine for the same five tyrosines. The third mutant, YM3:SUBST, contained nonconservative substitutions throughout YM3, including a pair of glycines which could disrupt the secondary structure. The fourth mutant, YM3:REV, reversed amino acid residues 614 to 620, thereby retaining the amino acid composition but altering primary structure com-

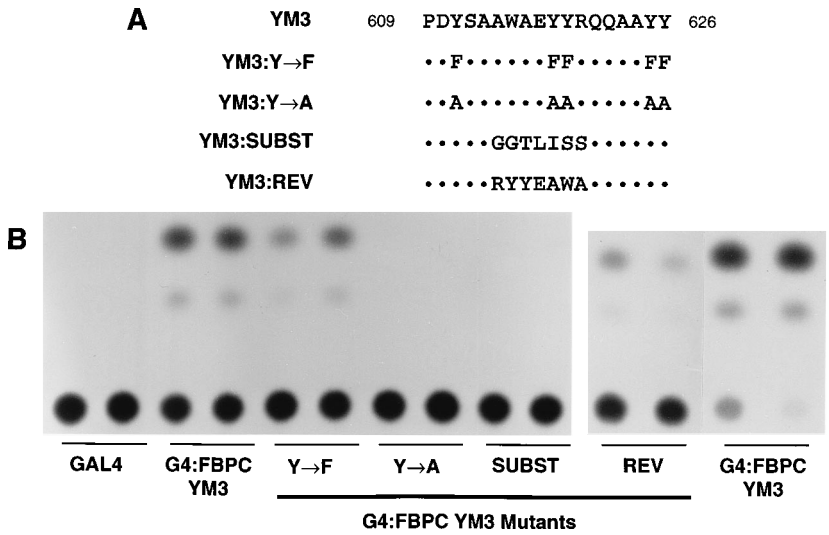

FIG. 5. Nonconservative substitution of the tyrosines abrogates activity of YM3. (A) The amino acid sequence of YM3 found between residues 609 and 626 is shown with the changes introduced in the YM3 mutants aligned below. The mutations were made in a G4:FBPC YM3 construct as diagrammed in Fig. 4A. Dots in the mutant sequences indicate an unchanged residue; all constructs continued with the native sequence from position 626 to the C terminus at 644. (B) CAT assays from duplicate transfections of the $(G4)_5$CAT reporter (1 μg) and the activator construct (20 ng) shown below for the transfections with GAL4, G4:FBPC YM3, and the YM3 mutants Y→F, Y→A, and SUBST. CAT assays shown for YM3 mutant REV and G4:FBPC YM3 at the far right were from a separate transfection in which 2 μg of reporter and 100 ng of activator plasmid were used.

pared to FBPC YM3. Each construct was cotransfected with the $(G4)_5$CAT reporter, and the cell extracts were assayed for CAT activity. Though some of the mutant proteins retained the ability to transactivate through GAL4 UAS sites (Fig. 4B), nonconservative substitution of the tyrosines abrogated activation. Therefore, it is apparent that at least one intact YM segment is necessary for FBP action.

Although the data to this point support the notion that YM motifs are required to augment gene expression, because all of the GAL4 FBP expression plasmids employed retained residues 511 to 532, there remained the formal possibility that this proline-glycine-histidine-rich segment, essentially inert alone, was necessary for YM motif function. To eliminate this possibility, a 36-amino-acid segment encompassing YM3 was fused directly to the GAL4 DNA binding domain and shown to upregulate reporter expression (Fig. 6). Together the above data prove that the YM units alone or in combination can function as a self-contained, transcription-activating protein module.

Repression of C-terminal activation by the N-terminal domain of FBP. The failure of the full-length FBP to activate as a GAL4 fusion and the weak activation supported by G4:FBPΔCD contrast with the dramatic activity supported by the C-terminal domain and derivatives thereof. Assigning a negative transcription effector function to the FBP amino terminus is a simple model which reconciles this apparent discrepancy. To test whether the N-terminal domain could in fact play such a repressive role, the highly activating G4:FBPC plasmid and the $(G4)_5$CAT reporter were transfected into HeLa cells together with G4:FBPN or GAL4(1-147) or a mixture of the last two plasmids. Since the GAL4 DNA binding domain functions as a dimer (6), the formation of heterodimers together with binding at adjacent GAL4 UASs ensures that FBPN and C-terminal domains should be brought within distances resembling those in the native protein. Transactivation by the G4:FBPC construct was drastically reduced by cotransfection with G4:FBPN in a dose-responsive fashion (Fig. 7A). In contrast, GAL4(1-147) did not perturb activation, arguing that the repression was unlikely to be due either to competition for binding to the tandem GAL4 sites or to disruption of G4:FBPC homodimers. Cotransfection of the same amounts of plasmids expressing other nonactivating GAL4 constructs (data not shown) also has no effect on G4:FBPC-mediated activation.

To ascertain the specificity of N-terminal repression, G4:FBPN was cotransfected with G4:FBPC or the equivalent chimeras constructed with the activating domains excised from

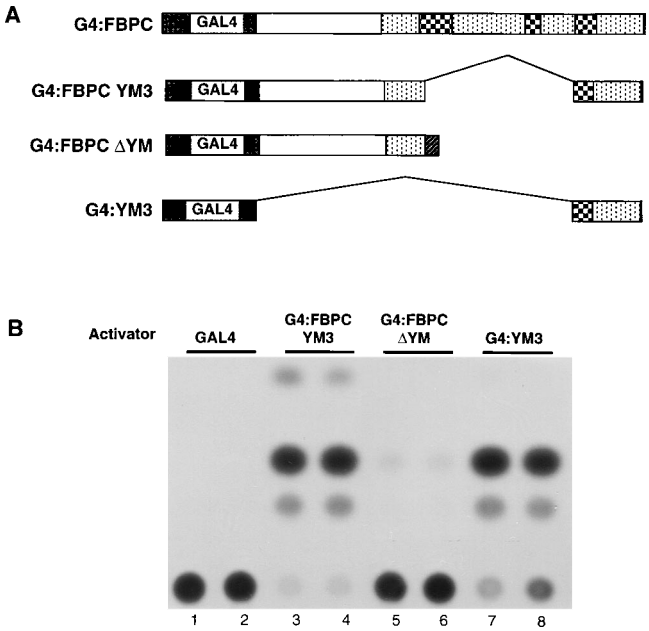

FIG. 6. YM3 alone is sufficient for activation. (A) Box diagrams of activator constructs; symbols are as in the legend to Fig. 4. (B) CAT assays of duplicate independent transfections of the $(G4)_5$CAT reporter (2.5 μg) and 200 ng of the activator indicated.

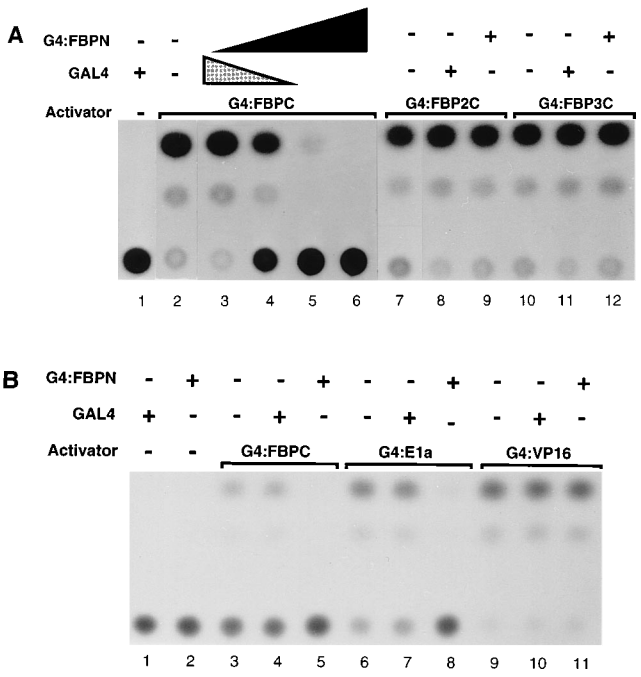

FIG. 7. The N-terminal domain of FBP represses activation by the C-terminal domain. (A) HeLa cells were transfected with reporter plasmid $(G4)_5$CAT (2.5 μg), 200 ng of the activator plasmids indicated (see the text for a description of FBP2 and FBP3), and the following amounts of repressor plasmids: lane 3, 500 ng of GAL4; lane 4, 250 ng GAL4 and 250 ng of G4:FBPN; lane 5, 500 ng G4:FBPN; lane 6, 1 μg G4:FBPN; lanes 8, 9, 11, and 12, 500 ng of the indicated repressor plasmid. (B) FBPN also represses E1a activation but not VP16 activation. HeLa cells were transfected with $(G4)_5$CAT (1 μg), 20 ng of activator plasmids, and 20 ng of G4:FBPN or GAL4 where indicated.

proteins highly related to FBP, FBP2, and FBP3 (10). Each of these FBP siblings possesses a C-terminal domain with a primary structure very similar to that of FBP (FBP2 has four YMs, while FBP3 bears two YMs). Plasmids encoding fusions of the GAL4 DNA binding domain to the COOH terminus of either FBP2 or FBP3 drive high levels of CAT activity when cotransfected with $(G4)_5$CAT. Remarkably, neither of these transactivators was repressed by cotransfecting G4:FBPN, notwithstanding the similarity of these molecules to G4:FBPC, seemingly indicating that FBPN is only an autologous repressor. Yet cotransfection of G4:FBPN with the well-studied activation domain of E1a or VP16 fused to GAL4 DNA binding domains revealed a broader capacity for repression. FBPN was able to repress E1a-mediated activation but not VP16-mediated activation when the same quantity of activator and repressor plasmid was transfected (Fig. 7B). From this result we conclude the following. First, the amino terminus of FBP is neither a general nor a nonspecific repressor of gene expression but rather antagonizes activation mediated by a restricted set of proteins, perhaps sharing some characteristic, as yet undefined. Second, FBPN does not directly antagonize tyrosine motif function (or else the FBP2 and FBP3 fusions should have been inactivated).

How does the FBP amino terminus antagonize activation by the carboxyl terminus? One model requires interactions between determinants at both ends of FBP to generate an inactive or repressed conformation, perhaps with the participation of additional cellular components. Such interaction between N and C termini is a common feature of many protein structures (16, 27). To test this explanation, several forms of FBP, untethered to the GAL4 DNA binding domain, were coexpressed with the G4:FBC activator and the effect on $(G4)_5$CAT was monitored. Supplying full-length FBP in *trans* had no effect on reporter activity (Fig. 8); presumably, the intramolecular interaction between the amino and carboxyl termini of FBP would be greatly entropically favored over the intermolecular interaction required in order for the amino terminus of FBP to repress G4:FBPC. However, by deleting the carboxyl end of FBP, the amino end would be deprived of its partner and become available to block G4:FBPC action in *trans*. Exactly this repression in *trans* by FBPΔCterm was observed. In contrast, FBP missing its repressing amino terminus (FBPΔN) and thus, lacking the molecular tools to block G4:FBPC, is relatively inert.

If the C-terminal domain of FBP acts as an activator in regulating c-*myc* expression, then a *myc* promoter-driven reporter plasmid previously shown to be stimulated by an FBP expression construct (11) should not be responsive to a C-terminally deleted FBP. In such a transfection, the N-terminal domain-deleted construct was not impaired for activation, but the C-terminal domain-deleted molecule scarcely increased expression above the basal level seen with a vector-alone control (Fig. 9).

DISCUSSION

The amino-terminal, central, and carboxyl-terminal domains of FBP function as separable repression, DNA binding, and activation domains, respectively. Although the modular assembly of FBP from functional protein segments conforms to the paradigm which generally describes transcription factors, several important properties distinguish FBP from other gene regulators. The activating segments primarily responsible for G4:FBPC-mediated transcription all share a conserved sequence, AW(A/E)(A/E)YY (tyrosine motif, YM). Amino acid substitutions prove that this sequence constitutes a motif required for activation. Activation mediated by YM3 is lost when its five tyrosines are nonconservatively replaced by alanine. The transcription-activating domains of other transcription factors have often been classified according to the predominant amino acids therein contained. For example, acidic domain (glutamic and aspartic) (24, 30), glutamine-rich (8), and proline-rich (21) as well as glycine-rich (20) or isoleucine-rich (1) activating segments have all been identified. In the case of FBP's YMs, a conserved and reiterated motif preserving tyrosine dyads is required for activation. If the smallest segment of the FBP carboxyl terminus proven to possess strong transactivating properties assumes an α-helical configuration (this 21-residue sequence is devoid of helix-destabilizing residues), then all five tyrosines as well as tryptophan form a single aromatic surface (Fig. 10). An alanine-rich helical surface is also formed. Similarly, displaying the residues comprising YM1 and YM2 as putative α-helices reveals that in each case all of the aromatic residues fall upon one face. Both the Y→F and REV mutants of YM3 preserve an aromatic face and still retain some capacity to activate transcription, thus indicating that neither the phenolic OH of tyrosine nor the exact arrangement of the aromatic groups is crucial to activation. A tryptophan contained within a glutamine-rich domain of SP1 has been shown to be a critical residue for activation, supporting the notion that aromatic residues are critical for activation (12). A very different tyrosine-rich segment of the pou-homeobox protein pit-1 has been shown to synergize with the estrogen receptor to stimulate prolactin expression. Unlike the tyrosine motifs in FBP, the pit-1 tyrosine-rich segment does not activate transcription alone and synergizes only with the

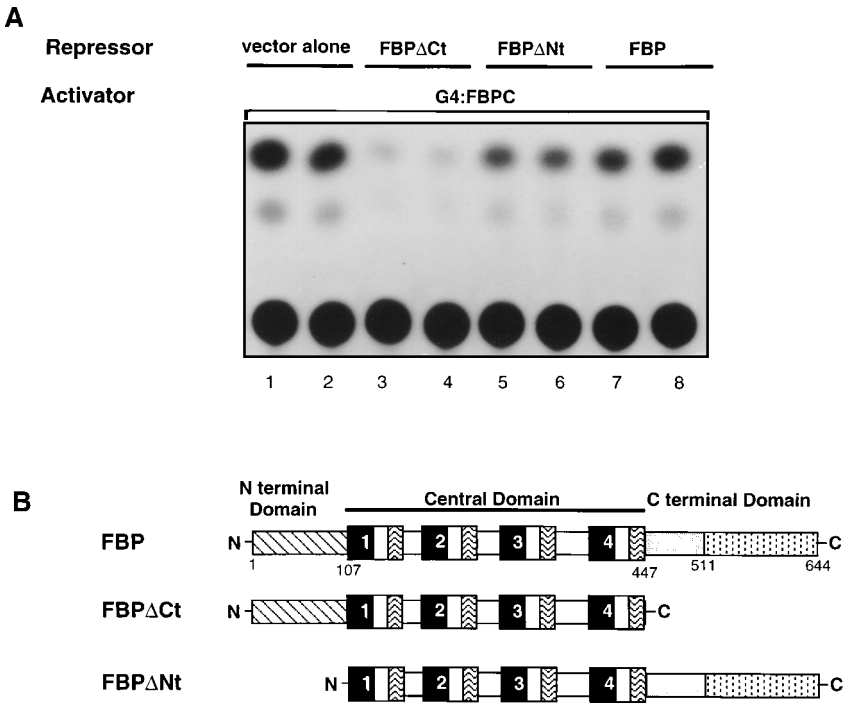

FIG. 8. Transactivation by an FBP C-terminal domain–GAL4 fusion can be repressed by an N-terminal domain of FBP which is not tethered to a GAL4 DNA binding domain. (A) CAT assays of transfections which received 2 μg of $(G4)_5$CAT reporter and 20 ng of G4:FBPC activator plasmid in all lanes; 100 ng of repressor plasmid in lanes 1, 3, 5, and 7; and 200 ng of repressor plasmid in lanes 2, 4, 6, and 8. The repressor plasmid was as follows: lanes 1 and 2, pCDNA1/AMP (Invitrogen) with no insert; lanes 3 and 4, the same vector with an FBP cDNA insert lacking the carboxyl-terminal domain; lanes 5 and 6, FBP cDNA lacking the amino-terminal domain; lanes 7 and 8, full-length FBP. (B) Box diagram of the FBP protein sequence encoded by the repressor plasmids described for panel A. See the legend to Fig. 1 for definitions of symbols.

proper arrangement of binding sites for pit-1 and estrogen receptor (14).

Two aspects of FBP activation may indicate an atypical mode of action. First, multimerization of a *cis* element is generally a very effective means of amplifying its influence on an adjacent promoter. In the case of FBP, no synergy is supported by tandemly repeating the GAL4 UAS to which G4:FBPC binds. Although the details of *cis*-element synergy have not been elucidated in most cases, one favored explanation involves the simultaneous interaction of multiple activating surfaces on *trans* factors with complementary surfaces on the TBP-TAF/basal transcription factor assembly at the promoter. Such multipoint attachment would be expected to facilitate promoter selection and transcription initiation. Second, embedding multiple activation domains within a single protein usually drives significantly greater transcription than is supported by a solitary activating sequence; however, in the case of FBP, no clear correlation exists between the number of YMs and the activity of the GAL4 chimera. The motifs singly, in pairs, or three together in the native configuration of FBP are not qualitatively different in level of activation, and FBP-2 and FBP-3, closely related homologs with four and two YMs, respectively, activate at the same level, both higher than FBP. Although it is possible that a fine-resolution genetic analysis would define specific protein determinants capable of modifying or limiting YM activity, it is also possible that the mechanism of YM action is intrinsically insensitive to either the number of *cis* elements from which it operates or the number of activating peptide segments. Perhaps the FBP activating sequence stimulates transcription after formation of the preinitiation complex, when the ability of multimerized *cis* elements and activation domains to recruit components of the transcription machinery might become irrelevant to later steps such as promoter clearance and elongation. If each successive step in the transcription process can be stimulated independently, then even a very low rate of preinitiation complex formation in the absence of powerful activation domains might be driven to increase expression through stimulation of rapid promoter clearance and elongation.

If tyrosine motifs do not synergize with each other, why are they repeated in each member of the FBP family? One rationalization for this observation invokes independent regulation

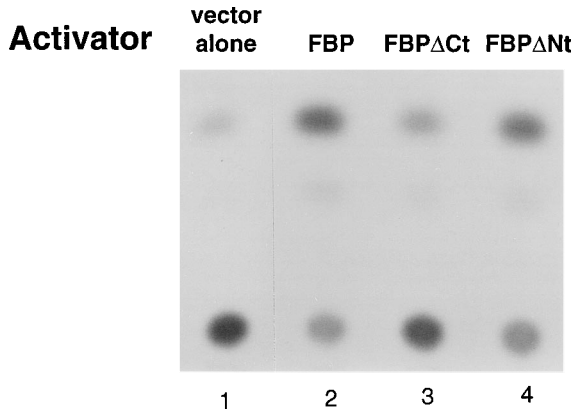

FIG. 9. The C-terminal domain is required for FBP to activate expression driven by the native *myc* regulatory sequence. Shown are results of CAT assays of HeLa cell extracts from cells cotransfected with an expression plasmid described in the legend to Fig. 8 as indicated above each lane and a *myc* regulatory region/CAT reporter, pMP *Spe-Acc*-CAT (see Materials and Methods).

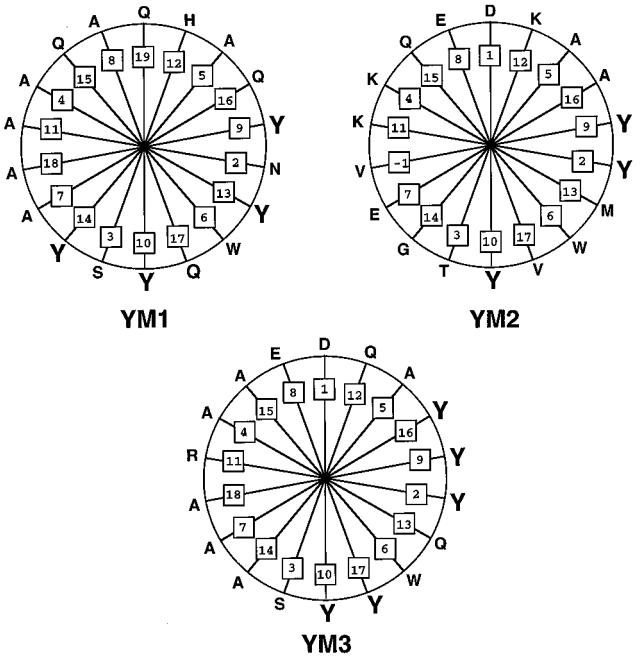

FIG. 10. The tyrosine motifs have the potential to form α-helices which isolate the tyrosine and tryptophan residues to one face and small side chain and charged residues to the opposite face. These projections place the tryptophan of each motif at position 6 of the helix.

of YM1, YM2, and YM3. The resemblance of YMs to tyrosine kinase recognition motifs evokes the idea that phosphorylation might regulate the function of this activation domain. If so, then the activity of each motif could be independently modulated, depending on the accessibility of the domain and the specificity of the presumptive kinase(s). The fact that phenylalanine can substitute for tyrosine in the G4:FBPC YM3 constructs does not exclude a role for tyrosine phosphorylation in the regulation of the activation potential of native FBP. In this manner, FBP could help link myc transcriptional regulation to signal transduction pathways.

In the full-length FBP, the activation domain is masked. This blockage employs the amino terminus of FBP in a specific manner. If FBP is to function as a gene regulator, then the interaction between its N and C termini must be regulated. Several mechanisms might couple activator and repressor function. First, covalent modification of one end or the other might promote or inhibit N-C interaction to determine FBP activity. Second, allosteric changes upon binding nucleic acid might alter FBP conformation. Since FBP binding to duplex DNA requires negative superhelicity and since transcription generates negative supercoils in upstream sequences, a conformational change of FBP's effector domains upon DNA binding would create a *cis*-acting feedback loop coupling FBP action to promoter activity.

The general functional properties of each domain of FBP have been assigned through the study of chimeric molecules and deletional analysis. Protein domains are generally fairly robust, and their individual functional properties often withstand genetic and biochemical manipulation. In contrast, the mechanisms coordinating functional domain activity often utilize weak, sometimes allosteric interactions which are easily disturbed and may not survive coarse dissection. The elucidation of the role of FBP in c-*myc* transcription may be expected to require approaches designed to probe subtle coupling of signal transduction, DNA binding, and effector domain activity. Despite its unusual recognition of single-stranded or supercoiled DNA, FBP appears to be more than an architectural transcription factor because it possesses a transactivation domain which can function tethered to an alternate DNA binding domain on a heterologous promoter-enhancer.

ACKNOWLEDGMENTS

We thank Suzanne Sanford for construction of plasmids, Charles Robinson for technical assistance, and Lance Liotta for critical review of the manuscript.

REFERENCES

1. **Attardi, L. D., and R. Tjian.** 1993. Drosophila tissue specific transcription factor NTF-1 contains a novel isoleucine-rich activation motif. Genes Dev. **7:**1341–1353.
2. **Avigan, M., and D. Levens.** Unpublished results.
3. **Avigan, M. I., B. Strober, and D. Levens.** 1990. A far upstream element stimulates c-*myc* expression in undifferentiated leukemia cells. J. Biol. Chem. **265:**18538–18545.
4. **Bazar, L., V. Harris, I. Sunitha, D. Hartmann, and M. A. Avigan.** 1995. A transactivator of c-*myc* is coordinately regulated with the proto-oncogene during cellular growth. Oncogene **10:**2229–2238.
5. **Bazar, L., D. Meighen, V. Harris, R. Duncan, D. Levens, and M. Avigan.** 1995. Targeted melting and binding of a DNA regulatory element by a transactivator of c-*myc*. J. Biol. Chem. **270:**8241–8248.
6. **Carey, M., H. Kakidani, J. Leatherwood, F. Mostashari, and M. Ptashne.** 1989. An amino-terminal fragment of GAL4 binds DNA as a dimer. J. Mol. Biol. **209:**423–432.
7. **Carey, M., Y.-S. 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.
8. **Courney, A., and R. Tjian.** 1988. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell **55:**887–898.
9. **Darnell, J. E., Jr., I. M. Kerr, and G. R. Stark.** 1994. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. Science **264:**1415–1421.
10. **Davis-Smyth, T., R. Duncan, T. Zhang, G. Michelotti, and D. Levens.** The human FBPs are members of an evolutionarily conserved family of sequence specific, single strand DNA binding transactivators. Submitted for publication.
11. **Duncan, R., L. Bazar, G. Michelotti, T. Tomonaga, H. Krutzsch, M. Avigan, and D. Levens.** 1994. A sequence-specific, single-strand binding protein activates the far upstream element of c-*myc* and defines a new DNA-binding motif. Genes Dev. **8:**465–480.
12. **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.
13. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2:**1044–1051.
14. **Holloway, J. M., D. P. Szeto, K. M. Scully, C. K. Glass, and M. G. Rosenfeld.** 1995. Pit-1 binding to specific DNA sites as a monomer or dimer determines gene-specific use of a tyrosine dependent synergy domain. Genes Dev. **9:**1992–2006.
15. **Ihle, J. N., and I. M. Kerr.** 1995. Jaks and Stats in signaling by the cytokine receptor superfamily. Trends Genet. **11:**69–74.
16. **Johansen, F.-E., and R. Prywes.** 1993. Identification of transcriptional activation and inhibitory domains in serum response factor (SRF) by using GAL4-SRF constructs. Mol. Cell. Biol. **13:**4640–4647.
17. **Kahn, J. D., E. Yun, and D. M. Crothers.** 1994. Detection of localized DNA flexibility. Nature (London) **368:**163–166.
18. **Lillie, J. W., and M. R. Green.** 1989. Transcription activation by the adenovirus E1a protein. Nature (London) **338:**39–44.
19. **Marcu, K. B., S. A. Bossone, and A. J. Patel.** 1992. *myc* function and regulation. Annu. Rev. Biochem. **61:**809–860.
20. **Meijer, D., A. Graus, and G. Grosveld.** 1992. Mapping the transactivation domain of the Oct-6 POU transcription factor. Nucleic Acids Res. **20:**2241–2247.
21. **Mermod, N., E. O'Neill, T. Kelly, and R. Tjian.** 1989. The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. Cell **58:**741–753.
22. **Michelotti, G. A., E. F. Michelotti, A. Pullner, R. C. Duncan, D. Eick, and D. Levens.** Multiple single-stranded *cis* elements are associated with activated chromatin of the human c-*myc* gene in vivo. Mol. Cell. Biol., in press.
23. **Murata, Y., H. G. Kim, K. T. Rogers, A. J. Udavia, and J. M. Horowitz.** 1994. Negative regulation of Sp1 trans-activation is correlated with the binding of

cellular proteins to the amino terminus of the Sp1 trans-activation domain. J. Biol. Chem. **269:**20674–20681.
24. **Ptashne, M.** 1988. How eukaryotic transcriptional activators work. Nature (London) **335:**683–689.
25. **Sadowski, I., J. Ma, S. Triezenberg, and M. Ptashne.** 1988. GAL4-VP16 is an unusually potent transcriptional activator. Nature (London) **335:**563–564.
26. **Sadowski, I., and M. Ptashne.** 1989. A vector for expressing GAL4(1-147) fusions in mammalian cells. Nucleic Acids Res. **17:**7539.
27. **Shi, Y., P. E. Kroeger, and R. I. Morimoto.** 1995. The carboxyl-terminal transactivation domain of heat shock factor 1 is negatively regulated and stress responsive. Mol. Cell. Biol. **15:**4309–4318.
28. **Shykind, B. M., J. Kim, and P. A. Sharp.** 1995. Activation of the TFIID-TFIIA complex with HMG-2. Genes Dev. **9:**1354–1365.
29. **Siebenlist, U., G. Franzoso, and K. Brown.** 1994. Structure, regulation and function of NF-kappa B. Annu. Rev. Cell Biol. **10:**405–455.
30. **Struhl, K.** 1987. Promoters, activator proteins and the mechanism of transcriptional initiation in yeast. Cell **49:**295–297.
31. **Tjian, R., and T. Maniatis.** 1994. Transcriptional activation: a complex puzzle with few easy pieces. Cell **77:**5–8.
32. **Tomonaga, T., and D. Levens.** 1995. Heterogeneous nuclear ribonucleoprotein K is a DNA-binding transactivator. J. Biol. Chem. **270:**4875–4881.
33. **Wolffe, A. P.** 1994. Architectural transcription factors. Science **264:**1100–1101. (Comment.)
34. **Wu, C.** 1995. Heat shock transcriptional factors. Structure and regulation. Annu. Rev. Cell. Dev. Biol. **11:**491–469.