Review

Activation domains of transcriptional regulatory proteins

Peter F. Johnson, Esta Sterneck, and Simon C. Williams*

ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, MD USA

Introduction

Changes in gene expression are central to nearly every developmental or adaptive response that eukaryotic cells are known to undergo. Gene regulation usually involves changes in the rate of transcription, and this has led to intense interest in the regulatory proteins that control the transcription of specific genes. The past decade has seen striking progress in the identification and characterization of these factors.[1-4] It is now well established that regulated transcription requires the binding of one or more proteins to a set of DNA sequence motifs (*cis*-regulatory sequences) in the flanking regions of the gene. These sequences collectively specify a program of transcriptional activity that is elicited by developmental, physiological, and other signals. The factors that recognize *cis*-regulatory sequences usually function positively to activate transcription, in which case they are termed activator proteins. The DNA-binding domains of activator proteins have been studied extensively, primarily because straightforward assays to measure DNA:protein interactions have been available, and they have been the subject of numerous reviews.[1,5]

In the past few years, increased attention has been focused on other functional domains within transcriptional regulatory proteins, especially the sequences known as activation domains or activating regions that are required for an activator protein to stimulate transcription of a target gene. In this article we review the discovery, properties, and classification of activation domains. We use specific examples to illustrate how these elements are used to modulate the activity of regulatory proteins, and to distinguish the regulatory functions of related activator proteins that possess identical DNA-binding specificities.

*Simon C. Williams is a Special Fellow of the Leukemia Society of America.
Research sponsored by the National Cancer Institute, DHHS, under contract No. N01-C0-74101 with ABL.
The contents of this publication do not necessarily reflect the views or policies of the Department of Health & Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.
Address reprint requests to Dr. Peter F. Johnson at the NCI-Frederick Cancer Research and Development Center, P.O. Box B, Bldg, 539, Frederick, MD 21702-1201 USA.
Received January 22, 1993; accepted February 19, 1993.

Transcriptional activation

A discussion of activation domains requires a rudimentary introduction to the concept of transcriptional activation. Therefore, we begin by briefly summarizing the principles and mechanisms of activated transcription. These topics have been reviewed in greater detail elsewhere.[6,7]

Basal versus activated transcription

Activator proteins stimulate transcription by elevating the initiation rate from the associated core (or basal) promoter. The core promoters of most genes transcribed by RNA polymerase II (RNAPII) consist of a TATA box and an initiator sequence, or INR, as depicted in *Figure 1*. These two sequence elements serve as a scaffold on which the transcriptional initiation complex is assembled and thereby determine the start-site of transcription.[4] The initiation complex is composed of at least seven defined factors in addition to RNAPII[8] (Figure 1). In the absence of additional regulatory sequences, the core promoter generates relatively low (basal) levels of transcription, both in vivo and in cell-free transcription reactions. Basal levels of transcription can also be observed in vitro with promoters that include upstream regulatory sequences, provided that the transcription extracts consist of purified preparations of initiation factors and RNAPII, and consequently lack activator proteins.

A significantly higher level of transcription (activated transcription) is achieved when activator proteins and their cognate binding sites are both present. The binding of an activator can increase the transcription rate by as much as several hundred fold in vivo. The stimulatory effects of proximal regulatory sequences located near the core promoter are apparent in vitro as well, but the degree of activation is usually significantly less than that observed in vivo. The effects of distantly located regulatory sequences (enhancers) are often quite dramatic in

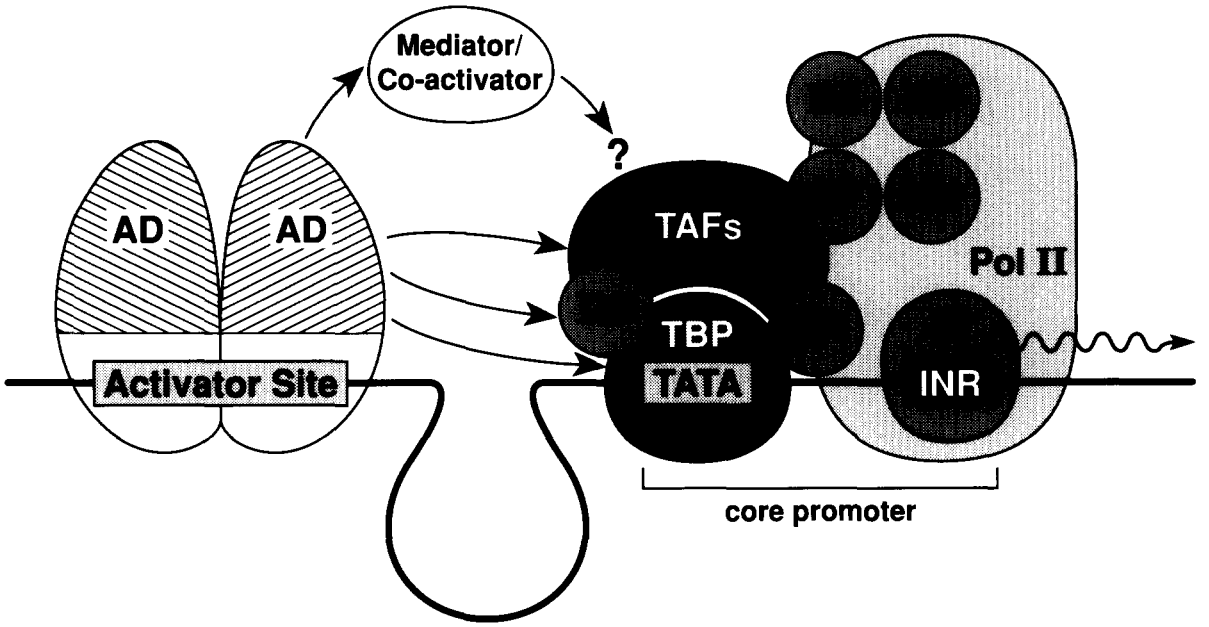


Figure 1 Potential targets for activators in the transcriptional initiation complex. A complex of general initiation factors (TFIIA, TFIIB, etc.) and RNA polymerase II assembles on the core promoter, which for most mRNA genes consists of a TATA box and an initiator (INR) element. An activator protein (shown here as a dimer) bound to a *cis*-regulatory site can stimulate the transcription initiation rate from an associated core promoter by an ill-defined mechanism. Arrows indicate direct binding interactions that have been observed between activation domains and initiation factors or accessory proteins.

vivo but are rarely evident in cell-free transcription assays.

Mechanisms of transcriptional activation

How does an activator protein tethered to a *cis*-regulatory sequence increase the rate of transcription from its associated core promoter? The basic underlying mechanism of this process, much less the details of the reaction, have not been definitively determined. However, the prevailing view is that activators either recruit or stabilize the assembly of basal transcription factors on the core promoter, or possibly catalyze a rate-limiting step in the maturation of the initiation complex to a fully active form.[7] There is now evidence to suggest that activators directly bind to components of the initiation complex. Proteins in the basal complex that have been implicated as targets for activators include the TATA binding protein (TBP) subunit of TFIID,[9] TBP-associated proteins (TAFs),[10] and TFIIB.[11] Other studies suggest the existence of intermediary proteins that serve as bridges between activators and initiation factors, and which have been called co-activators, adaptors, or mediators.[12–15] Co-activators are defined as proteins that are neither tightly associated with any of the basal factors nor required for basal transcription and do not bind DNA directly, but are necessary for transcriptional stimulation by certain activators or classes of activators. At present, co-activators are defined primarily as biochemical activities and only one potential co-activator, the yeast ADA2 gene, has been cloned.[16] As described in a later section, co-activators may in some cases play critical roles in controlling cell-specific gene activation.

An additional complication in understanding the mechanism of activation arises from the observation that two different effects of activator proteins are apparent in cell-free transcription experiments depending on whether the template DNA is free of histone proteins or is assembled into a chromatin-like structure. These processes have been called "true activation" and "anti-repression."[17] True activation refers to the enhanced initiation rate generated by the presence of activator proteins when naked DNA is used as the transcriptional template. This is distinct from anti-repression, which can be observed when the DNA template is assembled into nucleosomes[18] or complexed with histone H1.[17] Without an activator present, nucleosomes and H1 repress transcription below the basal level that occurs with naked DNA and initiation factors alone. However, if chromatin assembly occurs after binding of the activator protein, this histone-mediated repression can be alleviated. Both true activation and anti-repression require the presence of an activation domain, and therefore involve more than the mere binding of the activator to its binding site. It remains to be determined whether both of these experimentally defined effects of activator proteins operate in vivo.

Discovery and classification of activation domains

Activator proteins are bifunctional and modular

Eukaryotic activator proteins are bifunctional, being composed of distinct domains mediating DNA binding and transcriptional activation. The initial demonstration of this principle came from studies of the yeast GAL4

protein by Brent and Ptashne.[19] GAL4 is a sequence-specific regulatory protein from *Saccharomyces cerevisiae* that activates genes involved in galactose catabolism. A hybrid protein in which the DNA-binding domain of GAL4 was replaced by that of a repressor protein from *Escherichia coli,* LexA, was found to retain transcriptional activation function in yeast cells. Activation was dependent on the presence of a LexA binding site upstream of the target gene and could not be elicited by the LexA protein itself. These and similar studies[20] pointed to the presence of a transcriptional activation domain that is distinct from the DNA-binding region of the protein. Furthermore, these experiments demonstrated that the DNA-binding and activation regions exist as structurally autonomous domains (modules) that can be exchanged between proteins without loss of function. This domain-swapping approach has since been widely exploited to characterize functional segments of eukaryotic activator proteins. These studies have overwhelmingly supported the notion that activator proteins contain functionally independent DNA-binding and activation domains.[21]

Classes of activation domains

The identification and sequence analysis of activation domains in many activator proteins has revealed the existence of several distinct classes of these elements. These groupings are based primarily on similarities in amino acid content or, to a lesser extent, amino acid sequence. At this point, the following major groups of activators have been classified.

Acidic domains The analysis of deletion mutants within the GAL4 activating region revealed that most of the protein was dispensable for activation function.[22] However, two sequences designated region I (48 aa) and region II (113 aa), were found to be required for activation. While each domain alone displayed activating potential, both were necessary to generate full GAL4 activity. It was noted that these sequences are coincident with highly acidic regions of the protein, and in this respect were similar to the activation domain of another yeast regulatory protein, GCN4.[23] The activating region of GCN4 contains approximately 25% acidic residues and displays a net charge of -10.[24] Mutations affecting activation region I of GAL4 showed that negative charge was generally correlated with activation efficiency.[25] However, not all missense mutations that decreased transcriptional activation affected acidic residues, and amino acid substitutions involving different acidic residues could generate greatly differing effects on activity.[25] Therefore, net negative charge is a necessary, but not sufficient, feature of acidic activation domains.

Acidic activators are by no means restricted to fungal organisms. One of the most potent and best-studied acidic activation domains is derived from the Herpes Simplex Virus (HSV) transactivator protein, VP16. VP16 is not a direct DNA-binding protein, but is able to activate transcription of HSV immediate early (IE) genes by complexing with cellular proteins that themselves bind to *cis*-regulatory sequences in the promoters of these genes.[26] Deletion mutagenesis of VP16 revealed that transactivation function was abolished when an acidic region at the carboxyl terminus was removed.[27] Fusion of this 78 aa C-terminal segment to the DNA-binding domain of GAL4 creates a sequence-specific binding protein (GAL4-VP16) that activates transcription efficiently in both yeast and mammalian cells. This rather unlikely chimera between yeast and HSV proteins has become the prototypical acidic activator, and has been used extensively in studies of the mechanisms of transcriptional activation.[28]

A systematic mutagenesis study of the core activation domain within VP16 (residues 427 to 451) has been carried out by Cress and Triezenberg.[29] As in the case of GAL4, net negative charge was found to be an important feature but not the sole determinant of activation function. Acidic activator domains have been proposed to fold as amphipathic α-helical structures, with negatively charged residues located on one face of the helix.[24, 30, 31] However, the consequences of amino acid substitution mutations designed to test this possibility in VP16 were not consistent with a requirement for α-helicity.[29] Whether acidic activation domains are obligatorily α-helical remains an open question at the present time.

The single most critical amino acid was found to be Phe^{442}, because substitution by Ala or Ser at this position completely inactivated the protein. This residue is imbedded in a cluster of acidic (aspartate) residues that were also found to be important for transcriptional activation, and even conservative glutamate substitutions for these aspartate residues reduced VP16 activity. It was proposed that Phe^{442} makes a critical hydrophobic contact, either in tertiary folding interactions or through interactions with the target protein that mediates VP16 activation.[29] Other neighboring amino acids with aliphatic side chains might also contribute to these van der Waals contacts.

Because physical structures of acidic activation domains have not yet been reported, one is limited to speculation about their three-dimensional folded states. The facts that most single amino acid substitutions do not abolish activation and that functional acidic activating elements can be readily obtained from random fragments of *E. coli* genomic DNA in genetic selections for activation domains[22, 30] indicate that these domains do not have rigid conformational or sequence requirements. This view is supported by structural analyses of VP16[104,105], which suggest that the activation domain exists predominantly as a random coil. It is possible that an ordered conformation is achieved only when the domain is associated with its target protein. If this is the case, then a determination of the structure will probably require studies of the activator: target protein complex once the identity of the molecular target has been definitively established.

Glutamine-rich domains A second class of activator elements became evident from the analysis of functional

domains in the mammalian regulatory protein Spl. Spl is a ubiquitously expressed protein that was first identified as a DNA-binding protein in HeLa cell extracts that recognized a G:C-rich DNA motif.[32] Four activation domains were later identified in Spl, the two most potent of which contain approximately 25% glutamine residues.[33] The functional importance of the high Gln content in these activation domains was suggested by the presence of similar Gln-rich sequences in several other regulatory proteins.[2] One of these sequences, from the *Drosophila* Antennapedia homeobox protein, indeed functions as an activation domain when fused to the DNA-binding domain of Spl.[2] Other Gln-rich activating sequences have been identified in the NF-YA subunit of the ubiquitous CCAAT box binding protein NF-Y,[34] in the yeast protein MCM1,[35] and in the mammalian activators Oct-1 and Oct-2 (see below).

As in the case of the acidic activators, the Gln-rich activating elements from different proteins do not exhibit obvious similarities at the primary sequence level. However, Cress and Triezenberg[29] noted subtle sequence similarities between the Gln-rich domains found in Spl and the acidic activation domains in VP16 and GAL4. Particularly evident are the similar patterns of hydrophobic residues such as Leu, Val, Ile, and Phe in the two kinds of domains, and the fact that these hydrophobic amino acids are surrounded by residues that contain carbonyl groups in their side chains (glutamate or aspartate in acidic domains, and glutamine in Gln-rich domains). These observations raise the possibility that the two classes of activating sequences possess related folded structures. In light of these putative structural similarities, it will be interesting to determine whether the substitution of acidic amino acids for glutamine residues can transform a Gln-rich domain into a functional acidic activator, or vice versa.

Proline-rich domains Another class of activating sequence was discovered in the CTF/NFI protein, which is present in nearly all cells and has been implicated in the regulation of both transcription and DNA replication.[36] Inspection of the activation domain of CTF/NFI revealed that it contained a high frequency of proline residues, approximately 25%.[37] There are now several other examples of proteins containing proline-rich activation domains, including the lymphoid cell-specific activator Oct-2[38] (also characterized as a Ser/Thr-rich domain; see below) and activator protein 2 (AP-2).[39] In addition, regulatory proteins such as hepatocyte nuclear factor 1 (HNF1) and the plant transcription factor GT-2 feature regions rich in both proline and glutamine residues that are likely to function as activator sequences.[40,41] Whether these Pro/Gln-rich elements represent a class distinct from the Pro-rich and Gln-rich activators is not yet clear.

Serine/threonine-rich domains A small group of activators contain activation domains that have been characterized as serine- and threonine-rich. These include the C-terminal activation domain of Oct-2, which also possesses high proline content,[38,42] and the immunoglobulin enhancer binding proteins ITF-1, ITF-2,[43] and TFE3.[44] So far, the Ser/Thr-rich activators have not been extensively characterized by mutagenesis to determine their sequence requirements and, as is the case for all classes of activation domains, their structural nature is unknown. The presence of many serine and threonine residues raises the possibility that phosphorylation might contribute to their function or regulation. Although there is no evidence that these proteins are hyper-phosphorylated, the prospect of limited modifications within Ser/Thr-rich domains that are critical for function cannot be ruled out.

HOB1/HOB2 domains Kouzarides' group[45] recently noted two short regions within the A1 activation domain of the Jun oncoprotein[44] that share weak homology with a segment of Fos, a member of the same basic region: leucine zipper (bZIP)[50] DNA-binding protein family. These two adjacent sequences were designated homology box 1 and 2, or HOB1 and HOB2. HOB1 and HOB2 are spaced an identical distance apart in Jun and Fos, and both sequences occur within potential α-helical regions. The two elements function cooperatively to activate transcription, single copies of each motif being essentially inactive. All binary combinations of HOB1 and HOB2, including HOB1:HOB1 and HOB2:HOB2 duplications, generate functional activation domains, demonstrating the modularity of these elements. HOB1 and HOB2 are unrelated at the primary sequence level, however, and appear to be distinct types of activation modules. A sequence related to HOB2 was identified in an activating region of another bZIP regulatory protein, C/EBP (CCAAT/enhancer binding protein). The C/EBP sequence can substitute for the corresponding HOB2 elements in Jun and Fos and is therefore a functional homologue.

It is likely that the activity of HOB1 is regulated by phosphorylation. The Jun HOB1 element contains a serine residue (S-73) that undergoes phosphorylation in response to H-ras, and this modification appears to be important for H-ras–induced stimulation of the A1 activation domain.[46,47] S-73 is an in vitro substrate for MAP kinase,[48] and therefore the HOB1/HOB2 activation domain in Jun (and in Fos as well) may be a target for the H-ras/MAP kinase pathway that transduces extracellular mitogenic signals.[45]

Many activation domains have yet to be classified on the basis of protein sequence, amino acid content, or functional similarities with other activators. The identification of the conserved HOB1/HOB2 structure in Jun, Fos, and C/EBP raises the expectation that future low-stringency and structure-based homology searches will disclose other homologous motifs that define new classes of activation domains.

Functional differences among classes of activators One activator can often inhibit the activity of another in a phenomenon known as "squelching,"[51] which has been explained as a competition between the two activation domains for a common target molecule that is limiting in the cell. In principle, squelching can be used as an

assay to define functional classes of activation domains that interact with the same target factors. Tasset et al.[52] used the transcriptional interference test to demonstrate distinctions among the multiple activation domains present in the estrogen and glucocorticoid receptor proteins. Their data also indicate that the receptor activation domains are functionally distinguishable from acidic activators.

A study by Seipel et al.[53] tested the abilities of several different activation domains, fused to the GAL4 DNA-binding domain, to activate transcription in mammalian cells when bound either to remote enhancer sites or to proximal promoter sites. Acidic, Gln-rich, Pro-rich, and Ser/Thr-rich domains all activated efficiently from a location near the core promoter. However, Gln-rich domains were unable to stimulate transcription from a remote site, and Pro-rich sequences exhibited weak activity in this context. In contrast, acidic domains and Ser/Thr-rich domains were potent long-range activators. The ability to activate from a distal site was not correlated with the strength of the activator, indicating that different functional classes of domains are distinguished by qualitative and not quantitative properties. The authors propose that activation from enhancers may involve a fundamentally different mechanism (such as chromatin restructuring) from that of the proximal activators, which are believed to function by interacting directly with transcriptional initiation factors.

Activation domains influence the biological properties of activator proteins

Activation domains can modulate activator function in response to extracellular signals

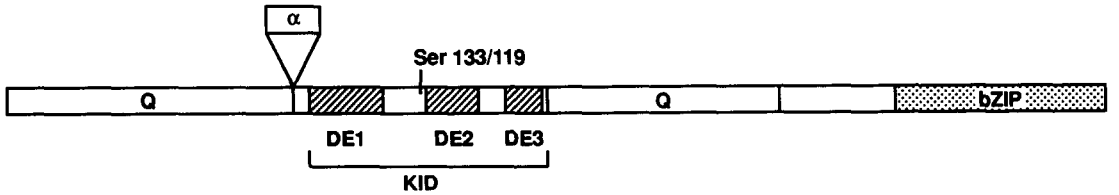
To initiate rapid changes in gene expression in response to environmental stress or extracellular cues, cells frequently utilize mechanisms that activate pre-existing, "dormant" regulatory proteins. Extracellular cues may alter either the DNA-binding activity or the activation functions of the transcription factor that is targeted by the incoming signal.[54] Two proteins in which the activation properties are stimulated either by amino acid modification (CREB) or by the binding of a ligand (thyroid hormone receptor) are presented below.

Modulation of CREB activation function by phosphorylation A number of genes are transcriptionally activated by an increase in intracellular cAMP concentration. A subfamily of the G-protein–coupled cell surface receptors activates adenylyl cyclase as an initial step in transduction of the ligand-induced signal, leading to an elevation in cAMP concentration. A sequence motif in the promoters of genes regulated by cAMP, termed the cAMP response element (CRE), is responsible for transcriptional induction by cAMP. The CRE-binding protein (CREB) recognizes the CRE motif and can mediate activation of gene expression in response to cAMP. CREB is a member of the CREB/ATF/CREM family of transcriptional regulators that belong to the basic region:leucine zipper (bZIP) class of DNA-binding proteins. Two forms of CREB have been identified (CREB341 and CREB327) that differ by an additional sequence in the N-terminal portion of CREB341, termed the alpha domain, that probably arises by an alternative splicing mechanism[55] (see *Figure 2A*).

Both recombinant and cellular CREB proteins bind constitutively to the CRE motif in vitro. However, in vitro and in vivo stimulation of transcription by CREB requires phosphorylation of the protein.[56] Although a recent study suggests that phosphorylation may also modulate the DNA-binding activity of CREB to low affinity binding sites,[57] the major effect of modification is to regulate activation function. The critical phosphorylated residue is Ser133 in CREB341 (Ser119 in CREB327). This residue lies within a region of the protein, termed the KID domain (kinase inducibility domain), which is flanked by two glutamine-rich regions (Q-regions, *Figure 2A*). Ser133/119 can be phosphorylated by the cAMP-activated Protein Kinase A (PKA),[56] and also by the Ca^{2+}/calmodulin-dependent protein kinases I and II (CaMK).[58, 59]

To identify and characterize functional domains of CREB, Lee et al.[60] fused segments of CREB327 lacking the bZIP domain to the GAL4 DNA-binding domain and tested these proteins in transactivation assays in vivo. The C-terminal bZIP domain of CREB was not required for regulated transactivation, but a 59 aa deletion encompassing the KID domain (residues 92–150) eliminated the 200-fold transcriptional activation that is observed for the intact protein under conditions of cAMP induction. The 59 aa segment alone, when fused to GAL4, functions as a weak (15-fold) cAMP-regulated activator. The isolated Q-domains did not function as activation domains in this study, and deletion of the N-terminal Q-region had no effect on regulated transactivation. However, when Gonzales et al.[61] deleted the N-terminal 87 amino acids from CREB341, creating a protein with the domain composition alpha-KID-Q-bZIP, transcriptional activation was abolished. Possibly, the discrepancy in these results points to a structural role for the N-terminal Q-domain, which can be functionally substituted by the GAL4 DNA-binding domain in a GAL4-KID-Q chimera. However, it cannot be excluded that the different results are due to the alpha domain present in the protein used by Gonzales et al.,[61] which is absent in the CREB327 protein analyzed by Lee et al.[60]

Mutational analysis showed that two deletions (DE1 and DE2; *Figure 2A*) within the KID domain, which do not interfere with phosphorylation of Ser133/119 by PKA, abolish transcriptional activation.[60] DE1 (aa 92–108) encompasses a region that is phosphorylated in vivo under cAMP stimulating conditions, and in vitro can be phosphorylated by casein kinase II. When Ser133/119 is mutated to Ala, phosphorylations within DE1 can no longer occur. Removal of DE2 (aa 121–131) does not affect phosphorylation of CREB. These results show that a phosphorylation event is apparently not required for the function of the DE2 region, and that phosphorylation of Ser133/119 and sites within DE1 are necessary, although not sufficient,



B. THRα

Figure 2 Schematic representation of the structural and functional domains within CREB and THRα. A. cAMP response element binding protein (CREB). Q, Gln-rich regions; α, alternative exon present in CREB341; KID, kinase inducibility domain; Ser133/119, serine phosphorylation site; DE1-3, deletion mutations; bZIP, basic region-leucine zipper domain. See text for detailed descriptions of domains and references. B. Thyroid hormone α (THRα). DBD, DNA-binding domain; HBD, hormone binding domain. See text for additional information and references.

for activation of transcription. Gonzales et al.[61] identified the sequence Asp-Leu-Ser-Ser-Asp (DLSSD) and the first Asp within this motif as the most critical sequence and amino acid, respectively, within the DE2 region. Spacing changes between the DLSSD motif and Ser133/119 also affected activation, suggesting a possible structural coupling between these elements. The authors argue that the DLSSD motif may be involved in salt bridges important for the active conformation of CREB.

Taken together, these data suggest that phosphorylation of CREB at Ser133/119 leads to structural changes that expose DE1 as a target for further phosphorylation events. Protein folding transitions that result from these modifications, perhaps stabilized by sequences within DE2, generate an active protein conformation in which the transcriptional activation surface becomes accessible for interactions with its target(s) in the transcriptional machinery. This model is supported by the finding that partial proteolysis of phosphorylated and unphosphorylated CREB yield different patterns of proteolytic fragments that are consistent with a conformational change in the protein near the PKA modification site.[61] As yet, it has not been possible to dissect a constitutive activation domain that is separate from inhibitory sequences that silence its function, as has been observed for the c-Jun activator protein (see below). Thus, it appears likely that the phosphorylated regions are either required for the structural integrity of the activation domain or serve as functional components of the activation domain itself.

Regulation of an activation domain in the thyroid hormone receptor by ligand binding Thyroid and steroid hormone signaling molecules can cross through cell membranes and bind to intracellular proteins that serve as receptor molecules. This family of structurally related receptors, in contrast to cell surface receptors, do not initiate a signal transduction pathway but rather exert transcriptional regulation directly. Thus, intracellular receptors must fulfill at least three different functions: hormone binding, DNA binding, and transcriptional activation. Regions that are critical for each of the respective functions have been identified in these proteins.

A well-conserved 70 aa domain that forms zinc finger structures mediates sequence-specific DNA binding. A region of approximately 200 amino acids C-terminal to the DNA-binding domain constitutes the hormone binding domain (HBD) and also contains a dimerization domain (*Figure 2B*). These sequences show some degree of conservation among the different receptors, whereas the remaining segments of the proteins are not related. Activation of transcription is mediated by the HBD. However, depending on the hormone receptor, various other regions of the protein can also contribute to transcriptional activation.[62-64]

The thyroid hormone (T_3) receptor-alpha (THRα or c-erbA-α), the cellular homolog of the transforming oncogene v-erbA of the avian erythroblastosis virus, is localized in the nucleus and exhibits constitutive DNA-binding activity. However, transcription is only activated in the presence of the ligand.[65]

The oncoprotein v-erbA cannot bind T_3 and acts as a dominant negative inhibitor of transcription, as does THRα without its ligand.[66, 67] Inhibition is not simply the result of competition for the binding site, preventing binding of activated THRα, because down-regulation of even basal levels of transcription is observed. The mechanism of this "active" negative regulation is not known.

One alteration in v-erbA that is critical for its oncogenicity is the deletion of the C-terminal 13 amino acids. Introduction of the same truncation in THRα creates a dominant negative inhibitor of wild-type THRα. However, a further deletion of 22 amino acids from the C

terminus, including sequences implicated in receptor dimerization, render the protein functionally inert. This result suggests that dimerization is important in regulation of THRα activity by the ligand and is required to exert the dominant negative phenotype of mutants.[65] The C-terminal dimerization function, however, is not required for DNA-binding or transcriptional activation by THRα. Removing approximately half of the HBD by C-terminal truncation of either THRα or the estrogen receptor (ER) leads to constitutive transcriptional activation.[68] This latter result shows that hormone binding releases a negative regulatory effect conferred by the HBD.

Thus, the structure of the THRα without its ligand allows specific binding to DNA, but the activation domain may be masked such that it cannot interact with components of the initiation complex (directly or indirectly). Binding of ligand may then induce a conformational change that allows the activation domain to function. The notion that conformational changes may be generally involved in activating nuclear receptors is supported by the finding that binding of progesterone to its receptor leads to resistance of the HBD to proteolysis. As detected by limited proteolysis, binding of antihormone also leads to structural alterations; however, these differ from the changes induced by binding of hormone.[69]

This model for hormone-receptor regulation is complicated by the findings of Privalsky et al.,[70] who showed that in yeast, THRα is a transcriptional activator in the absence of ligand, and is only modestly stimulated further by ligand (the hormone derivative TRIAC). Also, v-erbA expressed in yeast is almost as potent an activator as THRα, and activation is significantly enhanced by the ligand to levels even higher than seen with THRα.[70] These data may indicate that stabilization of the inactive conformation of THRα requires a cellular factor not present in yeast cells.

Recent results suggest that the DNA-binding domain can also influence the activity of the activation domain of hormone receptors.[63] THRα and ER are able to bind to each other's response element with high affinity. The response elements for THR and ER are very similar, both consisting of identical consensus palindromic half-sites. They differ in that the estrogen response element (ERE) half-sites are separated by a 3-nucleotide insertion. Despite the cross-recognition of binding sites, hormone-dependent transcription is only activated when the receptor is complexed with its cognate response element. However, receptors that are constitutive activators due to C-terminal deletions (see above) function on both types of elements. Domain swap experiments revealed that the binding site-specific activation function is mediated by the C-terminal region of the proteins. The intact T_3-HBD can only activate transcription when bound to the thyroid hormone response element (TRE), regardless of the DNA-binding domain (ER or THR), and vice versa. This suggests that a ligand-induced conformational change involving the HBD is not sufficient for the receptor to activate transcription. Rather, the cognate DNA sequence also seems to be required, perhaps as a second allosteric ligand that augments the effect of hormone binding in unmasking the activation domain.[68] Thus, the transcriptional activation functions of steroid hormone receptors are regulated by a complex set of intra- and intermolecular interactions that are only beginning to be understood.

Activation domains can mediate cell-specificity or gene-specificity of an activator protein

Many activator proteins are expressed in tissue-restricted patterns, and the cell-specific expression of these proteins is an important factor in determining the tissue-specific regulation of subordinate sets of genes. However, it has recently become apparent that certain activator proteins, as a result of their activation domains, also function in a cell-dependent fashion and that this may represent another mechanism by which tissue-specific gene expression is controlled. Moreover, activator proteins have now been shown to exhibit selectivity for different classes of core promoters, a property that can be termed gene specificity. Gene specificity is determined by the activation domain and appears to be one means to differentiate the regulatory properties of a set of related activator proteins that possess identical DNA-binding specificities. In the following two sections we summarize experiments that led to the discovery of these novel properties of activation domains.

Cell-specific inhibition of an activation domain in c-Jun

c-Jun is a member of the AP-1 family of transcription factors, which control the stimulation of a variety of cellular genes in response to growth factors and other mitogens.[71] The AP-1 family is composed of the Jun and Fos proteins and their relatives, all of which bind specific sequences via their bZIP DNA-binding domains. Jun homodimers and Jun-Fos heterodimers bind DNA and activate transcription from promoters containing AP-1 binding sites, although with differing efficiencies.[72] Deletion mutagenesis of the c-Jun protein led to the identification of two independent activation domains in the N-terminal half of the protein.[49, 73]

c-Jun is the cellular homolog of v-Jun, originally identified as the transforming oncogene of the avian sarcoma virus ASV-17.[74] The c-Jun and v-Jun proteins are highly related at the amino acid sequence level, v-Jun having sustained a deletion of 27 amino acids and 3 amino acid substitutions with respect to c-Jun. A comparison of the transcriptional activation properties of c-Jun and v-Jun in HeLa cells demonstrated that v-Jun is the more potent activator. The difference in transactivation properties of the two proteins was mapped to the 27 amino acid segment that is absent in v-Jun, called the δ region.[73] The δ region is located immediately upstream of the more N-terminal activation domain of c-Jun, designated A1.[73, 75] Removal of the δ region from c-Jun generated a protein that was nearly equivalent to v-Jun in transactivation assays. In addition to its effect on transactivation, removal of the δ region from c-Jun increased its ability to transform chicken embryo fibroblasts (CEF) to a

level similar to that of v-Jun. However, despite the apparent correlation between transactivation and transformation of c-Jun proteins lacking the δ region, it is not yet clear whether the increased transforming ability of the deleted form of c-Jun is due to its increased transactivation potential.[76-79]

The observation that deletion of the δ region in c-Jun increased its transcriptional activation potential identified an element in this region that negatively regulates the biological functions of c-Jun. The mechanism of action of this inhibitory domain was first characterized by Baichwal and Tjian[49] using an in vivo competition assay. A chimeric protein containing the activation domains of c-Jun linked to the DNA binding domain of GAL4 was cotransfected into HeLa cells with a reporter gene containing GAL4 binding sites, together with increasing amounts of a plasmid expressing intact c-Jun. Expression from the reporter gene was elevated with increasing amounts of intact c-Jun protein, indicating that the c-Jun protein was titrating a limiting cellular factor that inhibits the activity of the GAL4/c-Jun hybrid. The putative inhibitory factor was found to be cell specific, because c-Jun dependent transactivation was not inhibited (i.e., could not be enhanced by over expression of excess c-Jun) in HepG2 hepatocarcinoma cells, F9 embryonal carcinoma cells, or SL2 *Drosophila Schneider* cells, but was inhibited in Hela cells, mouse fibroblast (L) cells, CEF, and NIH 3T3 cells.[49,80] However, the raison d'etre for an inhibitory system that modulates c-Jun activity in a cell-specific manner is not readily apparent.

The δ region was shown to play a part in mediating negative regulation, as competing proteins that lacked the δ region were less efficient than intact c-Jun in stimulating transactivation by GAL4/c-Jun. However, interaction with the inhibitor also requires sequences outside the δ region. For example, although v-Jun lacks the δ region, it relieved inhibition of the GAL4/c-Jun hybrid, although not as efficiently as c-Jun.[49] In addition, transactivation by a v-Jun/E2 hybrid protein, consisting of the activation domains of v-Jun joined to the DNA-binding domain of the adenovirus E2 protein, could be enhanced by coexpression of c-Jun.[76] Deletion analysis of the c-Jun protein revealed a second region of c-Jun, termed ϵ, which was also involved in interactions with the inhibitor. The ϵ region is located just C-terminal to the A1 activation domain, such that the A1 domain is bounded by the δ and ϵ regions. In the latest model, the cell-specific inhibitory factor interacts with sequences within the δ and ϵ regions, thus blocking access of the activator domain to targets in the basal transcriptional machinery. The effect of the inhibitor is not specific to the A1 activation domain, as replacement of the A1 region with the VP16 activation domain also leads to inhibition of activation. However, the inhibitor does appear to exhibit activator protein specificity, at least within the AP-1 family, because the related JunB protein cannot compete for the inhibitor.[81] Moreover, Jun/Fos heterodimers are not susceptible to inhibition, suggesting that the inhibitor must bind to both subunits in a c-Jun homodimer to block activation.[49]

c-Jun, as well as other members of the AP-1 family, participate in the regulation of a set of cellular immediate-early (IE) genes whose expression is rapidly induced in response to mitogens. The induction of IE gene expression occurs without prior protein synthesis and therefore must utilize regulatory factors that pre-exist in the cell, perhaps in an inactive form.[82] The c-Jun protein could be held in an inactive form by the inhibitor molecule, and exposure of cells to mitogens could initiate signal transduction pathways that lead to the release of c-Jun from inhibition and subsequently, the activation of IE genes. The relationship between c-Jun phosphorylation and transactivation is somewhat controversial.[83,84] However, it is possible that phosphorylation of the c-Jun inhibitor is also an important factor in regulating the activity of c-jun. Baichwal et al.[75] showed that coexpression of either the Ras or Src oncogenes relieved c-Jun inhibition, suggesting that the putative c-Jun:inhibitor interaction can be regulated by signal transduction pathways in vivo. Future efforts to purify and/or clone the inhibitor and to characterize the effect of mitogens on c-Jun:inhibitor associations will undoubtedly reveal important insights into the regulation of c-Jun activity.

Another family of transcription factors in which the activity of different family members may be under cell type control is the C/EBP related protein (CRP) family.[85,86] At least two of the family members, C/EBP and CRP2 (also called NF-IL6, Il-6DBP, AGP/EBP, LAP, C/EBPβ) transactivate a target promoter in a cell-specific fashion. Both C/EBP and CRP2 activate transcription efficiently in HepG2 cells but are much weaker activators in HeLa cells or L cells,[85,87,88] whereas GAL4-VP16 activates robustly in all three cell types (S. Williams and P. Johnson, unpublished results). The cell specificity of CRP2 has been examined in some detail and appears to involve an inhibitory mechanism, because removal of a segment adjacent to the DNA-binding domain generates a protein that activates transcription in all cell lines tested (S. Williams and P. Johnson, manuscript in preparation). It is not yet known whether this cell type regulation is due to the binding of a cell specific inhibitor, as in the case of c-Jun, or arises from a cell-specific modification of the CRP2 protein.

Octamer binding proteins and immunoglobulin gene regulation Immunoglobulin (Ig) genes are expressed exclusively in B lymphoid cells, and analysis of Ig gene promoters revealed a sequence element known as the octamer motif (consensus ATTTGCAT) that is critical for high level B cell-specific transcription.[89] Subsequent studies showed that insertion of an octamer element into a heterologous, non-lymphoid promoter resulted in efficient B cell-specific expression that could be abolished by mutations that prevent factor binding to the octamer element.[90,91] Tissue-specific transcription of Ig genes was therefore proposed to result from the binding of a B cell-restricted factor to the octamer sequence. This hypothesis was brought into question by the identification of a large number of non-lymphoid genes that

contain functional octamer elements in their promoters. These include the ubiquitously expressed small nuclear RNA (snRNA) genes such as U2 and U6, the cell cycle regulated histone H2B gene, and the VP16-dependent HSV immediate early genes.[92] A potential explanation for the differential regulation of octamer-containing promoters was provided by the identification and cloning of two proteins that recognize the octamer motif, Oct-1 and Oct-2.[93, 94] Oct-1 is expressed ubiquitously, whereas Oct-2 is found only in B cells. The restricted expression of Oct-2 led to the supposition that Oct-2 in some way selectively activates B cell-specific transcription of Ig genes, while the other octamer-containing promoters are controlled by Oct-1.

Oct-1 and Oct-2 are members of the POU domain class of DNA-binding proteins[95] and contact DNA via a highly related POU domain located in the central portion of each protein. Oct-1 purified from HeLa or B cells and Oct-2 from B cells exhibit essentially identical DNA-binding properties.[96, 97] Both proteins possess a Gln-rich segment N-terminal to the POU domain, whereas their C-terminal regions share little similarity with each other. Oct-2 contains a Ser-, Thr- and Pro-rich C-terminal sequence that, together with the Gln-rich N-terminal element, is necessary for activation of octamer-containing mRNA promoters. Oct-1 does not have a similar C-terminal sequence and consequently, is incapable of transactivating many octamer-containing mRNA promoters that respond to Oct-2.[42] Other mRNA promoters, however, can be activated by both Oct-1 and Oct-2, or by an Oct-2 variant that lacks the C-terminal region.[98, 99]

In contrast, Herr et al.[100] showed that Oct-1 and Oct-2 have the opposite capabilities for activation of the U2 snRNA gene. The U2 gene is transcribed by RNA polymerase II, and its regulatory sequences include an octamer motif, but its core promoter lacks an identifiable TATA element and instead has a motif known as the proximal sequence element (PSE) that is common to many snRNA genes. In co-transfection experiments, Oct-1 efficiently activates a construct containing multiple octamer motifs upstream of the U2 snRNA promoter, whereas Oct-2 over-expression has little or no effect in a parallel experiment. The reciprocal result was obtained when a TATA-containing core promoter (from the non-lymphoid β globin gene) was used as the target in a similar transactivation assay; that is, Oct-2 enhanced transcription but Oct-1 did not. These experiments were performed using Oct derivatives in which the POU domains were replaced by the POU region from the Pit-1 protein, which has a different DNA-binding specificity, to circumvent the effects of endogenous Oct-1 activity in the host HeLa cells. (The octamer sites in the target promoters were likewise replaced by Pit-1 binding sites.) Because the Oct DNA-binding domains were absent in these chimeras, the different core promoter specificities of Oct-1 and Oct-2 must be attributable to dissimilarities in their activation domains and not their DNA-binding domains. Additional domain swapping and deletion mutagenesis experiments revealed that Oct-1 contains several redundant U2-specific activating regions, most of which are located in the C-terminal segment of the protein. Interestingly, an alternative splicing variant of Oct-2 (Oct-2B) carries extra sequences at the C terminus that exhibit homology to the Oct-1 C-terminal domain and, as a result of this addition, Oct-2B gains the capacity to activate the U2 promoter.

These results are most easily explained by a model in which different initiation complexes are formed on the U2 (PSE) and β globin (TATA) core promoters. Oct-1 would then have the ability to selectively interact with a component of the U2 initiation complex, while Oct-2 would find a specific target among the basal factors assembled on the β globin core promoter.[100] Under this scheme, the B cell-specific activation of the TATA-containing Ig genes would result from the ability of Oct-2, but not Oct-1, to enhance transcription from these core promoters. However, the activation of ubiquitously expressed mRNA promoters such as histone H2B by Oct-1 is not satisfactorily rationalized by this model. One must propose that the core promoters of some TATA-containing mRNA genes such as H2B can respond to the Oct-1 activation domain, whereas those of Ig genes and β globin are refractory to Oct-1.

A series of studies from Roeder's laboratory provides an alternative explanation for the B cell-specific activation of Ig genes.[97] This group used in vitro transcription assays to investigate the requirements for B cell-specific transcription of the Igκ gene promoter. Reconstitution experiments were performed using partially or highly purified preparations of Oct-1 and Oct-2 added to Oct-depleted cell extracts.[97] The results of these experiments suggest that (1) both Oct-1 and Oct-2 are in fact capable of activating the Igκ promoter in vitro, and (2) an additional B cell factor is required for high level activation to occur.

In a further characterization of this B cell factor, Luo et al.[101] succeeded in chromatographically separating the B cell activity, which they designated OCA-B (*oc*tamer *co-a*ctivator from *B* cells), from the Oct-1 and Oct-2 proteins. This fraction, when added to HeLa cell nuclear extracts, enhances both Oct-1– and Oct-2–stimulated transcription of the IgH promoter to the level seen in unfractionated B cell extracts. No similar activity could be isolated from HeLa cells, demonstrating the cell type restriction of OCA-B. OCA-B was also shown to be promoter selective, as it enhanced transcription of the IgH promoter but not the H2B promoter. In addition, OCA-B was found to be activator-protein specific,[101] because this factor had no effect on transcriptional stimulation by the regulatory proteins Spl and USF.[102] Surprisingly, OCA-B was found to associate preferentially with Oct-1, not Oct-2, both in transcription assays and in gel mobility shift assays where the addition of OCA-B caused a supershift of the Oct:DNA complex.

These findings suggest that the cell-specific activation of Ig promoters is primarily determined by the presence of OCA-B, and may involve an Oct-1:OCA-B complex. Presently it is difficult to reconcile this conclusion with the in vivo studies by Herr's group and others demon-

strating an intrinsic capability of Oct-2 in activating Ig promoters, even in non-lymphoid cells. However, Annweiler et al.[103] recently reported that Oct-2 can activate transcription when bound to a remote enhancer site in B cells, but not in fibroblasts or HeLa cells. Oct-1, on the other hand, was unable to stimulate transcription from a distantly placed binding site in any cell line, including B cells. It is conceivable, then that OCA-B, as well as the unique potential of Oct-2 to activate transcription from a distal site, are essential for high level expression of Ig genes in lymphoid cells. In this view Oct-1, facilitated by OCA-B, would be responsible for the effects of octamer sites that are located just upstream of the core promoter in Ig genes, while Oct-2 accounts for activation via octamer motifs present in Ig gene enhancers. The long-range effects of Oct-2 also apparently require a B cell-specific function, which could be the presumptive OCA-B co-activator or some other factor.

The precise roles that Oct-1 and Oct-2 play in governing B cell-specific activation of Ig genes remain controversial. Nevertheless, the study of these proteins demonstrates that activation domains with different properties can engender functional distinctions between otherwise related activators that, on the basis of their DNA-binding characteristics, would be predicted to exert similar regulatory effects.

Perspectives for the future

Despite the advances in our understanding of the nature and regulation of activation domains in recent years, many of the most interesting questions concerning these elements remain unanswered. Perhaps the most important issue for the future, which has not been reviewed in detail here, will be to identify the molecular targets for the various classes of activation domains. This may prove to be a challenging task, as activator:target associations are likely to entail relatively low affinity binding interactions. The difficulties inherent in these experiments are emphasized by studies to define the target of the VP16 activating region. This domain has been reported to directly interact with three different proteins: the TBP component of TFIID,[7] TFIIB,[9] and co-activator/mediator factor.[13] Whether one or all of these interactions are relevant remains to be conclusively determined.

One might predict that each class of activator domain will interact specifically with a different protein, or at least a different protein surface, in the transcriptional apparatus. These associations may occur directly or via a co-activator or bridging factor. The identification of targets will not only serve to explain the functional distinctions between different classes of activation domains, but will also be critical for elucidating the mechanism(s) of transcriptional activation. For example, it should become possible to determine whether activator:target associations are merely binding interactions that tether proteins to the promoter, or whether the target protein is covalently or allosterically altered by the activator so as to stimulate its activity in the initiation pathway.

Many activating regions have only been roughly delineated by deletion mutagenesis. It will be important to define the minimal functional sequence for each element, which in the case of HOB1/HOB2 proved to be relatively short.[45] In fact, other activation domains appear to contain small subdomains that function efficiently only when dimerized, such as an 18 aa segment of Oct-2 and an 11 aa sequence of VP16.[53] This multimerization approach may prove useful in defining minimal functional subdomains that can then be analyzed further using amino acid substitution mutagenesis (W. Herr, personal communication). Such fine-structure analysis may be particularly important for investigating the mechanisms by which activation domains are regulated by modifications or ligand binding, because this approach should allow one to determine whether a constitutive activating sequence is imbedded within a larger regulated activation domain. This strategy will also be useful for dissecting activation domains that are controlled by masking or intrapeptide folding mechanisms (e.g., THRα) and by interactions with inhibitory molecules (e.g., c-Jun). Investigations of the structural and mechanistic properties of regulatory domains that can modulate the activity of an associated activation domain also promise to be an increasingly active area of research.

Another major goal will be to elucidate the physical structures of activation domains. All crystallographic structures of eukaryotic activator proteins that have presently been determined were obtained with truncated proteins that contained only the DNA-binding domains, and consequently, no ultrastructural information regarding activating regions has been reported. With the exception of some potential α-helical regions, most activation domains have not been predicted to form specific secondary structures. This is due in part to the unusual amino acid compositions of these regions, which often contain a preponderance of one or two specific amino acids. As was discussed earlier, it is conceivable that a defined folded conformation is a consequence of, or is stabilized by the interaction of the activator with its target protein(s), in which case it may be necessary to study the activator:target protein complex. Nevertheless, we anticipate that studies to determine the structures of intact activator proteins will represent a productive area of investigation in the future.

The advances made thus far in investigations on the structure, regulation, and targets of activation domains in regulatory proteins have opened new and exciting areas of inquiry. The continued pursuit of these problems should soon lead to a better understanding of the mechanisms of transcriptional activation and regulated gene expression.

References

1 Johnson, P.F. and McKnight, S.L. (1989). Eukaryotic transcriptional regulatory proteins. *Annu. Rev. Biochem.* **58,** 799–839

2 Mitchell, P.J. and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245,** 371–378

3 Lewin, B. (1991). Oncogenic conversion by regulatory changes in transcription factors. *Cell* **64,** 303–312

4 Krajewska, W.M. (1992). Regulation of transcription in eukaryotes by DNA-binding proteins. *Int. J. Biochem.* **24,** 1885–1898

5 Pabo, C.O. and Sauer, R.T. (1992). Transcription factors: Structural families and principles of DNA recognition. *Annu. Rev. Biochem.* **61,** 1053–1095

6 Roeder, R.G. (1991). The complexities of eukaryotic transcription initiation: regulation of preinitiation complex assembly. *TIBS* **16,** 402–408

7 Ham, J., Steger, G., and Yaniv, M. (1992). How do eukaryotic activator proteins stimulate the rate of transcription by RNA polymerase II? *FEBS Lett.* **307,** 81–86

8 Zawel, L. and Reinberg, D. (1992). Advances in RNA polymerase II transcription. *Curr. Opin. Cell Biol.* **4,** 488–495

9 Stringer, K.F., Ingles, C.J., and Greenblatt, J. (1990). Direct and selective binding of an acidic transcriptional activation domain to the TATA-box factor TFIID. *Nature* **345,** 783–786

10 Dynlacht, B.D., Hoey, T., and Tjian, R. (1991). Isolation of coactivators associated with the TATA-binding protein that mediate transcriptional activation. *Cell* **66,** 563–576

11 Lin, Y.-S. and Green, M.R. (1991). Mechanism of action of an acidic transcriptional activator in vitro. *Cell* **64,** 971–981

12 Pugh, B.F. and Tjian, R. (1990). Mechanism of transcriptional activation by Spl: Evidence for coactivators. *Cell* **61,** 1187–1197

13 Kelleher, R.J., III, Flanagan, P.M., and Kornberg, R.D. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* **61,** 1209–1215

14 Berger, S.L., Cress, W.D., Cress, A., Triezenberg, S.J., and Guarente, L. (1990). Selective inhibition of activated but not basal transcription by the acidic activation domain of VP16: Evidence for transcriptional adaptors. *Cell* **61,** 1199–1208

15 Meisterernst, M., Roy, A.L., Lieu, H.M., and Roeder, R.G. (1991). Activation of class I gene transcription by regulatory factors is potentiated by a novel activity. *Cell* **66,** 981–993

16 Berger, S.L., Pina, B., Silverman, N., Marcus, G.A., Agapite, J., Regier, J.L., Triezenberg, S.J., and Guarente, L. (1992). Genetic isolation of ADA2: A potential transcriptional adaptor required for function of certain acidic activation domains. *Cell* **70,** 251–265

17 Laybourn, P.J. and Kadonaga, J.T. (1991). Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. *Science* **254,** 238–245

18 Workman, J.L., Taylor, I.C.A., and Kingston, R.E. (1991). Activation domains of stably bound GAL4 derivatives alleviate repression of promoters by nucleosomes. *Cell* **64,** 533–544

19 Brent, R. and Ptashne, M. (1985). A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43,** 729–736

20 Keegan, L., Gill, G., and Ptashne, M. (1986). Separation of DNA binding from the transcription-activating function of a eukaryotic transcriptional activator protein. *Science* **231,** 699–704

21 Frankel, A.D. and Kim, P.S. (1991). Modular structure of transcription factors: Implications for gene regulation. *Cell* **65,** 717–719

22 Ma, J. and Ptashne, M. (1987). Deletion analysis of GAL4 defines two transcriptional activating segments. *Cell* **48,** 847–853

23 Hope, I.A. and Struhl, K. (1986). Functional dissection of a eukaryotic transcriptional activator protein. *Cell* **46,** 885–894

24 Hope, I.A., Mahadevan, S., and Struhl, K. (1988). Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. *Nature* **333,** 635–637

25 Gill, G. and Ptashne, M. (1987). Mutants of GAL4 protein altered in an activation function. *Cell* **51,** 121–126

26 Triezenberg, S.J., LaMarco, K.L., and McKnight, S.L. (1988a). Evidence of DNA: protein interactions that mediate HSV-1 immediate early gene activation by VP16. *Genes Dev.* **2,** 730–742

27 Triezenberg, S.J., Kingsbury, R., and McKnight, S.L. (1988b). Functional dissection of VP16, the transactivator of herpes simplex virus immediate early gene expression. *Genes Dev.* **2,** 718–729

28 Ptashne, M. and Gann, A.A.F. (1990). Activators and targets. *Nature* **346,** 329–331

29 Cress, W.D. and Triezenberg, S.J. (1991). Critical structural elements of the VP16 transcriptional activation domain. *Science* **251,** 87–90

30 Giniger, E. and Ptashne, M. (1987). Transcription in yeast activated by a putative amphipathic α helix linked to a DNA binding unit. *Nature* **330,** 670–672

31 Cousens, D.J., Greaves, R., Goding, C.R., and O'Hare, P. (1989). The C-terminal 79 amino acids of the herpes simplex virus regulatory protein, Vmw65, efficiently activate transcription in yeast and mammalian cells in chimeric DNA-binding proteins. *EMBO J.* **8,** 2337–2342

32 Dynan, W.S. and Tjian, R. (1983). Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. *Cell* **32,** 669–680

33 Courey, A.J. and Tjian, R. (1988). Analysis of Spl in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* **55,** 887–898

34 Li, X.-Y., Hoof van Huijsduijnen, R., Mantovani, R., Benoist, C., and Mathis, D. (1992). Intron-exon organization of the NF-Y genes. *J. Biol. Chem.* **267,** 8984–8990

35 Bruhn, L., Hwang-Shum, J.-J., and Sprague, G.F., Jr. (1992). The N-terminal 96 residues of MCM1, a regulator of cell type-specific genes in *Saccharomyces cerevisiae,* are sufficient for DNA binding, transcription activation, and interaction with α1. *Mol. Cell. Biol.* **12,** 3563–3572

36 Santoro, C., Mermod, N., Andrews, P.C., and Tjian, R. (1988). A family of human CCAAT-box-binding proteins active in transcription and DNA replication: cloning and expression of multiple cDNAs. *Nature* **334,** 218–224

37 Mermod, N., O'Neill, E.A., Kelly, T.J., and Tjian, R. (1989). The proline-rich transcriptional activator of CTF/NF-1 is distinct from the replication and DNA binding domain. *Cell* **58,** 741–753

38 Gerster, T., Balmaceda, C.-G., and Roeder, R.G. (1990). The cell type-specific octamer transcription factor OTF-2 has two domains required for the activation of transcription. *EMBO J.* **9,** 1635–1643

39 Williams, T. and Tjian, R. (1991). Characterization of a dimerization motif in AP-2 and its function in heterologous DNA-binding proteins. *Science* **251,** 1067–1071

40 Raney, A.K., Easton, A.J., Milich, D.R., and McLachlan, A. (1991). Promoter-specific transactivation of hepatitis B virus transcription by a glutamine- and proline-rich domain of hepatocyte nuclear factor 1. *J. Virol.* **65,** 5774–5781

41 Dehesh, K., Hung, H., Tepperman, J.M., and Quail, P.H. (1992). GT-2: a transcription factor with twin autonomous DNA-binding domains of closely related but different target sequence specificity. *EMBO J.* **11,** 4131–4144

42 Tanaka, M. and Herr, W. (1990). Differential transcriptional activation by Oct-1 and Oct-2: Interdependent activation domains induct Oct-2 phosphorylation. *Cell* **60,** 375–386

43 Henthorn, P., Kiledjian, M., and Kadesch, T. (1990). Two distinct transcription factors that bind the immunoglobulin enhancer μE5/κE2 motif. *Science* **247,** 467–470

44 Beckman, H., Su, L.-K., and Kadesch, T. (1990). TFE3: A helix-loop-helix protein that activates transcription through the immunoglobulin enhancer μE3 motif. *Genes Dev.* **4,** 167–179

45 Sutherland, J.A., Cook, A., Bannister, A.J., and Kouzarides, T. (1992). Conserved motifs in Fos and Jun define a new class of activation domain. *Genes Dev.* **6,** 1810–1819

46 Benitruy, B., Smeal, T., and Karin, M. (1991). Ha-ras augments c-Jun activity and stimulates phosphorylation of its activation domain. *Nature* **351,** 122–127

47 Pulverer, B.J., Kyriakis, J.M., Arruch, J., Nikolakaki, E.,

and Woodgett, J.R. (1991). Phosphorylation of c-Jun mediated by MAP kinases. *Nature* **353,** 670–674
48 Smeal, T., Binetruy, B., Mercola, D.A., Birrer, M., and Karin, M. (1991). Oncogenic and transcriptional cooperation with Ha-ras requires phosphorylation of C-Jun on Serines 63 and 73. *Nature* **354,** 494–496
49 Baichwal, V.R. and Tjian, R. (1990). Control of c-Jun activity by interaction of a cell specific inhibitor with regulatory domain δ: differences between v-Jun and c-Jun. *Cell* **63,** 815–825
50 Landschulz, W.H., Johnson, P.F., and McKnight, S.L. (1988). The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* **240,** 1759–1764
51 Gill, G. and Ptashne, M. (1988). Negative effect of the transcriptional activator GAL4. *Nature* **334,** 721–724
52 Tasset, D., Tora, L., Fromental, C., Scheer, E., and Chambon, P. (1990). Distinct classes of transcriptional activating domains function by different mechanisms. *Cell* **62,** 1177–1187
53 Seipel, K., Georgiev, O., and Schaffner, W. (1992). Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. *EMBO J.* **11,** 4961–4968
54 Jackson, S.P. (1992). Regulation of transcription factor activity by phosphorylation. *Trends Cell Biol.* **2,** 104–108
55 Borrelli, E., Montmayeur, J.-P., Foulkes, N.S., and Sassone-Corsi, P. (1992). Signal transduction and gene control: The cAMP pathway. *Crit. Rev. Onc.* **3,** 321–338
56 Brindle, P.K. and Montminy, M.R. (1992). The CREB family of transcription activators. *Curr. Opin. Gen. Dev.* **2,** 199–204
57 Nichols, M., Weih, F., Schmid, W., DeVack, C., Kowenz-Leutz, E., Luckow, B., Boshart, M., and Schutz, G. (1992). Phosphorylation of CREB affects its binding to high and low affinity sites: implications for cAMP induced gene transcription. *EMBO J.* **11,** 3337–3346
58 Dash, P.K., Karl, K.A., Colicos, M.A., Prywes, R., and Kandel, E.R. (1991). cAMP response element-binding protein is activated by Ca^{2+}/calmodulin-as well as cAMP-dependent protein kinase. *Proc. Natl. Acad. Sci. USA* **88,** 5061–5065
59 Sheng, M., Thompson, M.A., and Greenberg, M.E. (1991). CREB: A Ca^{2+}-regulated transcription factor phosphorylated by calmodulin-dependent kinases. *Science* **252,** 1427–1430
60 Lee, C.Q., Yun, Y., Hoeffler, J.P., and Habener, J.F. (1990). Cyclic-AMP-responsive transcriptional activation of CREB-327 involves interdependent phosphorylated subdomains. *EMBO J.* **9,** 4455–4465
61 Gonzalez, G.A., Menzel, P., Leonard, J., Fischer, W.H., and Montminy, M.R. (1991). Characterization of motifs which are critical for activity of the cyclic AMP-responsive transcription factor CREB. *Mol. Cell. Biol* **11,** 1306–1312
62 Beato, M. (1989). Gene regulation by steroid hormones. *Cell* **56,** 335–344
63 Danielian, P.S., White, R., Lees, J.A., and Parker, M.G. (1992). Identification of a conserved region required for hormone dependent transcriptional activation by steroid hormone receptors. *EMBO J.* **11,** 1025–1033
64 Bocquel, M.T., Kumar, V., Stricker, C., Chambon, P., and Gronemeyer, H. (1989). The contribution of the N- and C-terminal regions of steroid receptors to activation of transcription is both receptor and cell-specific. *Nucl. Acids Res.* **17,** 2581–2595
65 Damm, K. (1992). c-erbA: Protooncogene or growth suppressor gene? *Adv. Cancer Res.* **59,** 89–113
66 Damm, K., Thompson, C.C., and Evans, R.M. (1989). Protein encoded by v-erbA functions as a thyroid-hormone receptor antagonist. *Nature* **339,** 593–597
67 Sap, J., Munoz, A., Schmitt, J., Stunnenberg, H., and Vennstrom, B. (1989). Repression of transcription mediated by a thyroid hormone response element by the v-erbA oncogene product. *Nature* **340,** 242–244
68 Holloway, J.M., Glass, C.K., Adler, S., Nelson, C.A., and Rosenfeld, M.G. (1990). The C′-terminal interaction domain of the thyroid hormone receptor confers the ability of the DNA site to dictate positive or negative transcriptional activity. *Proc. Natl. Acad. Sci. USA* **87,** 8160–8164
69 Allan, G.F., Leng, X., Tsai, S.Y., Weigel, N.L., Edwards, D.P., Tsai, M.-J., and O'Malley, B.W. (1992). Hormone and antihormone induce distinct conformational changes which are central to steroid receptor activation. *J. Biol. Chem.* **267,** 19513–19520
70 Privalsky, M.L., Sharif, M., and Yamamoto, K.R. (1990). The viral erbA oncogene protein, a constitutive repressor in animal cells, is a hormone-regulated activator in yeast. *Cell* **63,** 1277–1286
71 Curran, T. and Franza, B.R., Jr. (1988). Fos and Jun: The AP-1 Connection. *Cell* **55,** 395–397
72 Ryseck, R.-P. and Bravo, R. (1991). c-JUN, JUN B, and JUN D differ in their affinities to AP-1 and CRE consensus sequences: effect of fos proteins. *Oncogene* **6,** 533–542
73 Bohmann, D. and Tjian, R. (1989). Biochemical analysis of transcriptional activation by Jun: differential activity of c- and v-Jun. *Cell* **59,** 709–717
74 Vogt, P.K. and Tjian, R. (1988). jun: transcriptional regulator turned oncogenic. *Oncogene* **3,** 3–7
75 Bos, T.J., Monteclaro, F.S., Mitsunobu, F., Ball, Jr., A.R., Chang, C.H.W., Nishimura, T., and Vogt, P.K. (1990). Efficient transformation of chicken embryo fibroblasts by c-Jun requires structural modification in coding and noncoding sequences. *Genes Dev.* **4,** 1677–1687
76 Havarstein, L.S., Morgan, I.A., Wong, W.Y., and Vogt, P.K. (1992). Mutations in the Jun Delta region suggest an inverse correlation between transformation and transcriptional activation. *Proc. Natl. Acad. Sci. USA* **89,** 618–622
77 Hartl, M. and Vogt, P.K. (1992). Oncogenic transformation by Jun: Role of transactivation and homodimerization. *Cell Growth & Diff.* **3,** 899–908
78 Alani, R.P., Brown, P., Binetruy, B., Dosaka, H., Rosenberg, R.K., Angel, P., Karin, M., and Birrer, M.J. (1991). The transactivating domain of the c-jun proto-oncoprotein is required for cotransformation of rat embryo cells. *Mol. Cell. Biol* **11,** 6286–6295
79 Lloyd, A.N., Yancheva, N., and Wasylyk, B. (1991). Transformation suppressor activity of a jun transcription factor lacking its activation domain. *Nature* **352,** 635–638
80 Baichwal, V.R., Park, A., and Tjian, R. (1991). v-Src and EJ Ras alleviate repression of c-Jun by a cell-specific inhibitor. *Nature* **352,** 165–168
81 Baichwal, V.R., Park, A., and Tjian, R. (1992). The cell-type-specific activator region of c-Jun juxtaposes constitutive and negatively regulated domains. *Genes Dev.* **6,** 1493–1502
82 Karin, M. (1991). The AP-1 complex and its role in transcriptional control by protein kinase C. In *Molecular Aspects of Cellular Regulation* (P. Cohen and J.G. Foulkes, eds.), p 235–253, Elsevier, Oxford, UK
83 Baker, S.J., Kerppola, T.K., Luk, D., Vandenberg, M.T., Marshak, D.R., Curran, T., and Abate, C. (1992). Jun Is phosphorylated by several protein kinases at the same sites that are modified in serum-stimulated fibroblasts. *Mol. Cell. Biol* **12,** 4694–4705
84 Boyle, W.J., Smeal, T., Defize, L.H.K., Angel, P., Woodgett, J.R., Karin, M., and Hunter T. (1991). Activation of protein kinase C decreases phosphorylation of c-Jun at sites that negatively regulate its DNA-binding activity. *Cell* **64,** 573–584
85 Williams, S.C., Cantwell, C.A., and Johnson, P.F. (1991). A family of C/EBP-related proteins capable of forming covalently linked leucine zipper dimers in vitro. *Genes Dev.* **5,** 1553–1567
86 Cao, Z., Umek, R.M., and McKnight, S.L. (1991). Regulated expression of three C/EBP isoforms during adipose conversion of 3T3-L1 cells. *Genes Dev.* **5,** 1538–1552
87 Descombes, P., Chojkier, M., Lichsteiner, S., Falvey, E., and Schibler, U. (1990). LAP, a novel member of the C/EBP gene family, encodes a liver-enriched transcriptional activator protein. *Genes Dev.* **4,** 1541–1551
88 Friedman, A.D., Landschulz, W.H., and McKnight, S.L. (1989). CCAAT/enhancer binding protein activates the promoter of the serum albumin gene in cultured hepatoma cells. *Genes Dev.* **3,** 1314–1322
89 Staudt, L.M. and Lenardo, M.J. (1991). Immunoglobulin gene transcription. *Annu. Rev. Immunol.* **9,** 373–398.
90 Wirth, T., Staudt, L., and Baltimore, D. (1987). An octamer

oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. *Nature* **329,** 174–178

91 Dreyfus, M., Doyen, N., and Rougeon, F. (1987). The conserved decanucleotide from the immunoglobulin heavy chain promoter induces a very high transcriptional activity in B-cells when introduced into an heterologous promoter. *EMBO J* **6,** 1683–1690

92 Schaffner, W. (1989). How do different transcription factors binding the same DNA sequence sort out their jobs? *Trends Genet.* **5,** 37–39

93 Clerc, R.B., Corcoran, L.M., LeBowitz, J.H., Baltimore, D., and Sharp, P.A. (1988). The B-cell-specific Oct-2 protein contains POU box and homeo box-type domains. *Genes Dev.* **2,** 1570–1581

94 Sturm, R.A., Das, G., and Herr, W. (1988). The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a homeo box subdomain. *Genes Dev.* **2,** 1582–1599

95 Herr, W., Sturm, R.A., Clerc, R.G., Corcoran, L.M., Baltimore, D., Sharp, P.A., Ingraham, H.A., Rosenfeld, M.G., Finney, M., Ruvkan, G., and Horvitz, H.R. (1988). The POU domain: a large conserved region in the mammalian pit-1, oct-1, oct-2 and *Caenorhabditis elegans* unc-86 gene products. *Genes Dev.* **2,** 1513–1516

96 Yoza, B.K., Roeder, R.G. (1990). Identification of a novel factor that interacts with an immunoglobulin heavy-chain promoter and stimulates transcription in conjunction with the lymphoid cell-specific factor OTF2. *Mol. Cell. Biol.* **10,** 2145–2153

97 Pierani, A., Heguy, A., Fujii, H., and Roeder, R.G. (1990). Activation of octamer-binding transcription factor I (OTF-1) or OTF-2 and requirement of an additional B-cell-specific component for optimal transcription of immunoglobulin promoters. *Mol. Cell. Biol.* **10,** 6204–6215

98 Muller-Immergluck, M.M., Schaffner, W., and Matthias, P. (1990). Transcription factor Oct-2A contains functionally redundant activating domains and works selectively from a promoter but not from a remote enhancer position in non-lymphoid (HeLa) cells. *EMBO J.* **9,** 1625–1634

99 Kemler, I., Bucher, E., Seipel, K., Muller-Immergluck, M.M., and Schaffner, W. (1991). Promoters with the octamer DNA motif (ATGCAAAT) can be ubiquitous or cell type-specific depending on binding affinity of the octamer site and Oct-factor concentration. *Nucl. Acids Res.* **19,** 237–242

100 Tanaka, M., Lai, J.-S., and Herr, W. (1992). Promoter-selective activation domains in Oct-1 and Oct-2 direct differential activation of an snRNA and mRNA promoter. *Cell* **68,** 755–767

101 Luo, Y., Fujii, H., Gerster, T., and Roeder, R.G. (1992). A novel B cell-derived coactivator potentiates the activation of immunoglobulin promoters by octamer-binding transcription factors. *Cell* **71,** 231–241

102 Gregory, P.D., Sawadogo, M., and Roeder, R.G. (1990). The adenovirus major late transcription factor USF is a member of the helix-loop-helix group of regulatory proteins and binds to DNA as a dimer. *Genes Dev.* **4,** 1730–1740

103 Annweiler, A., Muller-Immergluck, M., and Wirth, T. (1992). Oct2 transactivation from a remote enhancer position requires a B-cell-restricted activity. *Mol. Cell. Biol.* **12,** 3107–3116

104 Donaldson, L. and Capone, J.P. (1992). Purification and characterization of the carboxyl-terminal transactivation domain of Vmw65 from herpes simplex virus type 1. *J. Biol. Chem.* **267,** 1411–1414

105 O'Hare, P. and Williams, G. (1992). Structural studies of the acidic transactivation domain of the Vmw65 protein of herpes simplex virus using ^{1}H NMR. *Biochemistry.* **31,** 4150–4156