Transcriptional Regulation in Plants: The Importance of Combinatorial Control

Karam B. Singh
Department of Molecular, Cell and Developmental Biology, University of California, Los Angeles, California 90095–1606

GLOSSARY OF TERMS

Combinatorial control: use of a discrete number of transcription factors in different combinations to give rise to a wide spectrum of expression patterns.

Enhanceosome: a higher-order nucleoprotein complex that is formed by the binding of a specific combination of transcription factors to the transcriptional regulatory sequences of a particular gene.

General transcription factors: components of the Pol II transcription-initiation complex that are thought to be common to all Pol II promoters.

Holoenzyme: a large protein complex that is preformed off of the DNA and contains Pol II and many of the other components of the Pol II transcription-initiation complex.

Pol II transcription-initiation complex: a large protein machine that contains dozens of polypeptides and assembles near the transcription start site.

Transcriptional activators: class of regulatory proteins that are gene specific and function to increase transcription from target promoters.

Transcriptional synergy: when specific combinations of transcription factors give rise to significantly higher levels of transcription than the sum of the additive effects obtained when each factor is assayed individually.

During their development and differentiation, plants need to integrate a wide range of tissue, developmental, and environmental signals to regulate complex patterns of gene expression. The regulation of seed-storage protein gene expression, resulting in expression during specific stages of seed development but not in other parts of the plant, is a striking example of tissue and developmental control and is of considerable agricultural importance. Plants also have unique needs and strategies for responding to changes in their environment. When light strikes an etiolated leaf, numerous genes encoding chloroplastic, mitochondrial, peroxisomal, and cytosolic proteins are activated. Similarly, a number of biotic and abiotic stresses cause a battery of genes to be activated as part of the plant-defense/stress response. A major level at which gene expression is regulated is the initiation of transcription, and this is reflected in the percentage of the genome dedicated to transcription factors in plants and other eukaryotes. For example, an analysis of 1.9 Mb of Arabidopsis genomic sequence from chromosome 4 revealed that about 15% of the genes with predicted or known functions were involved in transcription, a percentage similar to what has been found in other eukaryotes (Bevan et al., 1998).

In eukaryotic cells, genomic DNA is complexed with proteins to form chromatin. One of chromatin’s major roles is to facilitate the packaging of DNA in the nucleus, but the structure of chromatin also leads to a general suppression of gene activity. For gene activation and transcription to occur, the chromatin in the vicinity of the gene must be remodeled to allow access for transcription factors and the recruitment of the RNA polymerase II (Pol II) transcription-initiation complex. Transcription factors play important and diverse roles in gene expression, including chromatin remodeling and recruitment/stabilization of the Pol II transcription-initiation complex. Transcription factors, which come in many shapes and sizes, can be divided into a number of functional classes, with some proteins belonging to more than one class. A major class of transcription factors is activators and repressors. These proteins bind to specific DNA sequences found only in certain promoters and are instrumental in giving rise to genespecific regulation. A second class of transcription factors are coactivators or corepressors. These proteins mediate the transcriptional effects of specific activators/repressors, in some cases by remodeling chromatin. Whereas this group of transcription factors are typically not able to bind to DNA on their own, they can still be promoter-specific as a result of protein-protein interactions with specific activators and repressors. A third class comprises the general transcription factors, which are important components of the Pol II transcription-initiation complex. A fourth class is architectural transcription factors that are also involved in remodeling DNA, e.g. by inducing bends that facilitate the binding of other proteins to the promoter.

In this Update I will address how transcription factors regulate gene transcription in plants, as well as relying on

1 This work was supported by grants from the National Institutes of Health and the U.S. Department of Agriculture.

* Corresponding author; e-mail kramSingh@ucla.edu; fax 1–310–206–3987.

Abbreviations: ABRC, ABA-responsive complex; ABRE, ABA-responsive element.
advances in other systems. The focus will be on activators and the importance of combinatorial control. First, I will comment on chromatin, chromatin remodeling, and the Pol II transcription-initiation complex, since it is the recruitment and/or activity of the transcription-initiation complex that is regulated by the gene-specific transcription factors, and this regulation occurs in the context of chromatin.

**CHROMATIN REMODELING AND TRANSCRIPTIONAL REGULATION**

Chromatin has several levels of structural organization, with the basic unit being the nucleosome core, which consists of 146 bp of DNA wrapped around a histone octamer. The presence of nucleosomes affects the accessibility of DNA to other proteins, including transcription factors and the Pol II transcription-initiation complex. Higher orders of chromatin structure are also likely to affect transcription, e.g. by leading to the organization of chromatin into active and silent regions. Exciting progress has been made recently on chromatin remodeling, including the involvement of histone acetylation/deacetylation on nucleosome conformation/stability, and the identification of novel protein complexes that cause nucleosome disruption (for recent reviews, see Cairns, 1998; Struhl, 1998).

Histone acetylation was first reported in 1964 and at the time was proposed to play a role in the regulation of transcription (Allfrey et al., 1964). Studies during the next three decades supported this proposal, although until recently the molecular mechanisms involved were not known. Acetylation of histones occurs on Lys residues in the amino-terminal tails that protrude from the surface of the nucleosome. Acetylation neutralizes the positive charge of the histone tails and consequently causes a reduction in their affinity for DNA. This leads to changes in nucleosome conformation and may lead to unfolding of the nucleosome. Thus, acetylation of histones is normally correlated with transcriptional activity by facilitating the access of transcription factors to the DNA, whereas deacetylation of histones is correlated with transcriptional repression.

A growing number of histone acetylases and histone deacetylases have been identified recently. Significantly, many of these proteins had already been associated with transcription, with some having been identified previously as components of the Pol II transcription-initiation complex, and others initially identified as coactivators or corepressors. For example, the yeast Gcn5 and mammalian p300/CBP proteins were initially identified as coactivators and were subsequently found to be histone acetylases (Bannister and Kouzarides, 1996; Brownell et al., 1996; Ogryzko et al., 1996). In the case of Gcn5, the histone acetylase activity has been shown to be required for coactivator function in vivo (Candau et al., 1997), and overexpression of Gcn5 leads to increased histone acetylation at promoter regions of genes regulated by Gcn5 (Kuo et al., 1998). Histone acetylation has also been observed in plants (Belyaev et al., 1997) and a maize histone deacetylase has been identified (Lusser et al., 1997).

Other forms of chromatin remodeling involving multi-protein complexes with ATP-dependent chromatin remodeling activities have also been observed. A good example is the SWI2/SNF2 complex, which was initially discovered through genetic studies as a transcriptional regulator of specific genes in yeast (Stern et al., 1984; Neigeborn and Carlson, 1984). A number of lines of evidence link the yeast SWI2/SNF2 complex, which has a size of about 2,000,000 D, to chromatin remodeling (for review, see Cairns, 1998). Chromatin remodeling does not occur at promoters normally regulated by the SWI2/SNF2 complex in strains with a defective SWI2/SNF2 complex. In addition, the purified SWI2/SNF2 complex was able to cause chromatin remodeling to occur in vitro. SWI2/SNF2-related complexes have also been identified in Drosophila melanogaster and humans. In D. melanogaster three additional chromatin-remodeling complexes have been identified, including NURF, which is able to stimulate the binding of a number of transcription factors to chromatin templates (Tsukiyama and Wu, 1996). The identification of distinct multiprotein complexes involved in chromatin remodeling raises important questions regarding the role each plays in transcriptional regulation, how they are targeted to specific promoters, and whether they interact with other types of chromatin-remodeling activities, such as histone acetylases.

**THE POL II TRANSCRIPTION-INITIATION COMPLEX IS A LARGE PROTEIN MACHINE**

The Pol II transcription-initiation complex has a size in excess of 2,500,000 D (Fig. 1). Pol II itself is a complex enzyme that in yeast consists of 14 subunits. Whereas Pol II catalyzes RNA synthesis, numerous other proteins are required for promoter recognition and accurate transcription initiation. These include the six general transcription factors (Fig. 1), as well as a growing number of accessory proteins. Some of these accessory proteins may serve as coactivators or may be involved in chromatin remodeling, whereas others probably provide regulatory functions that remain to be elucidated. Some of the general transcription factors are complex, e.g. TFIIID is composed of TATA box binding protein (TBP) and a number of TBP-associated factors (TAFs).

The general transcription factors and some of the accessory proteins were first identified biochemically using in vitro transcription systems. Based on the in vitro studies a stepwise model was proposed for the assembly of the transcription-initiation complex starting with the binding of TFIIID to the TATA box (for review, see Roeder, 1996). In this model a major function of transcriptional activators was to facilitate the stepwise assembly of the transcription-initiation complex. However, some of the Pol II in cells has been shown to exist in a large protein complex(es) called the holoenzyme (for review, see Greenblatt, 1997), suggesting that to a large extent the transcription-initiation complex may already be formed in the absence of DNA. A major role of transcriptional activators may be to recruit the holoenzyme to specific promoters and/or to stabilize the transcription-initiation complex once bound at the core promoter.
Until recently, the composition of the transcription-initiation complex was not considered to be a major area of regulation. However, this view is changing with reports of multiple forms of the holoenzyme and the discovery of accessory proteins such as the TAFs. TAFs, which were initially identified as coactivators required by certain activators to function in vitro, also function in core promoter recognition in vivo (Shen and Green, 1997), and TAF250 has histone acetylase activity (Mizzen et al., 1996). Cell-type-specific TAFs have also been identified, e.g. the B-cell-type-specific TAF105 (Verrijzer and Tjian, 1996). The potential for TFIID to serve as a major point of regulation has also been suggested by the discovery of multiple TBP s in some species, including Arabidopsis. In D. melanogaster, in addition to a ubiquitous TBP, a second tissue-specific TBP, TRF (TBP-related factor), has been identified that may be important for neural-specific patterns of gene regulation (Hansen et al., 1997). Recently, a TFIID complex that does not contain TBP has been isolated from human cells (Wiertz et al., 1998). Surprisingly, this complex, called TFTC (TBP-free TAF-containing complex), was shown to be able to substitute for normal TFIID and support transcription (TBP-free TAF-containing complex), was shown to be able to substitute for normal TFIID and support transcription initiation, and it serves as a coactivator required by certain activators to function in vitro, also function in core promoter recognition in vivo (Shen and Green, 1997), and TAF250 has histone acetylase activity (Mizzen et al., 1996). Cell-type-specific TAFs have also been identified, e.g. the B-cell-type-specific TAF105 (Verrijzer and Tjian, 1996). The potential for TFIID to serve as a major point of regulation has also been suggested by the discovery of multiple TBP s in some species, including Arabidopsis. In D. melanogaster, in addition to a ubiquitous TBP, a second tissue-specific TBP, TRF (TBP-related factor), has been identified that may be important for neural-specific patterns of gene regulation (Hansen et al., 1997). Recently, a TFIID complex that does not contain TBP has been isolated from human cells (Wiertz et al., 1998). Surprisingly, this complex, called TFTC (TBP-free TAF-containing complex), was shown to be able to substitute for normal TFIID and support transcription from both TATA-box-containing promoters and TATA-less promoters, using an in vitro transcription system. Although these exciting results demonstrate that multiple transcription-initiation complexes exist and offer additional opportunities for the regulation of transcription, it remains likely that the major level of transcriptional control is mediated by transcriptional activators and repressors.

**PLANTS CONTAIN FAMILIES OF DNA-BINDING PROTEINS THAT BEHAVE AS TRANSCRIPTIONAL ACTIVATORS**

Whereas there has been relatively little work on the general transcriptional machinery in plants in comparison with animals and yeast, an increasing number of transcriptional regulatory proteins have been identified in plants. Although transcriptional regulators can be activators or repressors, and, in some cases, the same protein can serve both functions, the focus of this Update is on activators. Typically, activators have a modular structure consisting of discrete domains responsible for specific DNA binding, transcriptional activation, and in some cases dimerization and/or other forms of protein-protein interactions. Whereas domain-swap experiments have shown that DNA-binding domains and activation domains can operate independently when fused to a heterologous protein, there is growing evidence that in their native protein context they may sometimes functionally communicate (for review, see Lefstin and Yamamoto, 1998).

Plant transcription factors contain a variety of structural motifs that allow for binding to specific DNA sequences. Most of these DNA-binding domains were first identified in transcription factors isolated in animal/yeast systems. In many cases the structure of the DNA-binding domain bound to DNA has been determined, allowing a good understanding of how these DNA-binding domains function. In bZIP transcription factors, DNA binding and dimerization are mediated by the bZIP motif, which consists of a region rich in basic amino acids and an adjacent bZIP that consists of a 4-3-heptad repeat of hydrophobic and nonpolar residues (for review, see Hurst, 1995). The bZIP is required for dimerization of the protein prior to binding to DNA, whereas the basic region contacts the DNA-recognition site. Crystal structure of the bZIP motif of the yeast GCN4 protein complexed with DNA demonstrated that the bZIP motifs resembled a helical forceps as they gripped the major groove of DNA (Ellenberger et al., 1992; Fig. 2). The bZIP region of each monomer is packed together in a coiled coil, whereas the basic region passes through the major groove and makes a number of contacts with specific DNA bases and the phosphate backbone.

Many transcription factors have been broadly classified into a number of families on the basis of the type of DNA-binding domain, such as bZIP, homeodomain, helix-turn-helix, helix-loop-helix, and zinc-finger proteins. In some cases these can be large protein families. A good example are MYB family members, which contain a specific type of the helix-turn-helix motif of about 50 amino acids, which serves as the DNA-binding domain. In Arabidopsis more than 80 MYB family members have already been identified (Romero et al., 1998). Some plant transcription factors have DNA-binding domains that appear to be
unique to plants. For example, the AP2/EREBP family of plant transcription factors, found in a range of higher plants, contain a conserved, approximately 60- to 70-amino acid region required for DNA binding, a part of which has been predicted to form an amphipathic α-helix (Okamuro et al., 1997). The large family of viviparous 1 (VP1)-related proteins, which includes the maize VP1 protein (McCarty et al., 1991) and the Arabidopsis ARF1 and Monopteros proteins (Ulmasov et al., 1997; Hardtke and Berleth, 1998), represents a second class of plant-specific DNA-binding proteins that play important roles in plants. Although plant transcription factors typically have only a single DNA-binding domain, there are examples, such as the GT-2 and APETALA2 proteins, in which there are two related DNA-binding domains (Ni et al., 1996; Okamuro et al., 1997). In the case of GT-2, a rice DNA-binding protein that interacts with GT-box promoter elements in the rice phytochrome A gene, the different DNA-binding domains have been shown to discriminate between closely related GT-box sequences (see Ni et al., 1996, and refs. therein). For some transcription factors the ability to form either homodimers and/or heterodimers with related family members is a prerequisite for DNA binding. The ability to form specific heterodimers is a form of combinatorial control, and this can expand the number of DNA target sequences that can be recognized, as well as allow different combinations of activation domains to be recruited to a promoter element.

Several classes of transcriptional activation domains have been identified and some can be classified on the basis of their amino acid composition/properties. For example, there are acidic, Gln-, Pro-, and Ser/Thr-rich activation domains. Other strong activation domains have been identified that are not particularly rich in any specific amino acid. Many activators contain more than one activation domain. Relatively little is known about the structures of activation domains and how they function, although there have been many reports of interactions of specific types of activation domains with one or more components of the transcription-initiation complex.

The notion that a major function of activation domains is recruitment of the transcriptional machinery to a given promoter has gained support from studies of yeast involving activator bypass experiments (for review, see Ptashne and Gann, 1997). In these experiments fusion of a DNA-binding domain to different components of the holoenzyme was sufficient to confer transcription on genes containing the corresponding DNA-binding site, and activators were no longer required. However, recruitment of the transcriptional machinery may not be the only way that activation domains function. For example, some activation domains may function by altering the conformation of the transcription-initiation complex assembled at the core promoter, facilitating promoter escape at some step during transcription elongation. Controlling both the type and arrangement of activation domains brought to a specific promoter is a major way that transcriptional control is manifested, which is discussed below.

**COMBINATORIAL CONTROL IS IMPORTANT FOR TRANSCRIPTIONAL REGULATION AND CAN LEAD TO FORMATION OF HIGHER-ORDER NUCLEOPROTEIN COMPLEXES**

It is becoming clear that a major mechanism underlying eukaryotic transcriptional regulation is combinatorial control. Many genes are regulated by multiple transcriptional activators by virtue of having a specific set of protein-binding sites in their promoters. At any given time, a distinct set of transcription factors bind to these different sites to give rise to higher-order nucleoprotein complexes that have been called enhanceosomes (for a recent review, see Carey, 1998). Specific interactions between proteins that form an enhanceosome, as well as interactions with components of the general transcriptional machinery, can then lead to cooperativity in DNA binding and transcriptional synergy. The composition of the enhanceosome assembled at a given promoter may change in response to environmental, developmental, or other signals. It is possible to form a much larger number of distinct enhanceosomes by using a discrete number of transcription factors in different combinations. A corollary to this is that, through combinatorial interactions, a given transcription factor can play multiple roles and help regulate different genes whose expression is induced by distinct signals. There is a growing body of evidence from animal systems, using both in vivo and in vitro studies, for the existence and importance of enhanceosomes in mediating transcrip-
The regulation of pigment production in maize is one of the best-characterized examples in plants for the importance of combinatorial interactions in gene regulation (for recent reviews, see Mol et al., 1996, 1998). The biosynthesis of one class of maize pigments, the anthocyanins, is regulated by both developmental and environmental signals. UV light induces anthocyanin expression in the epidermis of leaves and petioles, whereas developmental cues induce their expression in a number of organs such as kernels, seedlings, and leaves. The developmental regulation of anthocyanin expression is the result of combinatorial interactions between two distinct families of plant transcription factors. The first is the C1 family, which are Myb-related regulatory proteins and consist of the closely related C1 and P1 proteins (Paz-Ares et al., 1987; Cone et al., 1993). The second is the R family, which are encoded by the B and R loci (Ludwig and Wessler, 1990). R family members contain the b/HLH (basic helix-loop-helix) motif and also share a high degree of amino acid homology with each other. The members of each family differ in their tissue-specific expression patterns, which in turn reflect the patterns of pigmentation observed in maize. However, individual family members are not sufficient to induce the anthocyanin biosynthetic genes, but, rather, a member from each family must be coexpressed in a particular tissue for anthocyanin biosynthesis to occur. Thus, genetic studies have demonstrated that the combination of R and C1 is responsible for pigmentation in the kernel, whereas the combination of B and P1 is responsible for pigmentation in mature tissues of the plant, such as the husk leaves (for review, see Mol et al., 1996). The importance of combinatorial control was also illustrated by transgenic studies showing that the expression of just C in tobacco or Arabidopsis was not sufficient to increase anthocyanin expression but required the coexpression of R (Lloyd et al., 1992).

The precise mechanism by which combinatorial interactions between the Myb and b/HLH proteins leads to transcriptional activation is not known. Promoter sequences required by C1 and B/R for activation have been mapped in the promoters of anthocyanin biosynthetic genes (Sainz et al., 1997; Lesnick and Chandler, 1998, and refs. therein). Several possibilities for why C1 requires B or R for transcriptional activation were proposed by Sainz et al. (1997). One possibility is that B/R acts to increase C1 DNA-binding specificity to promoters in anthocyanin biosynthetic genes, although this seems less likely, since C1 is able to bind to a specific site in the a1 gene promoter in the absence of B or R. Another possibility is that the C1 Myb domain inhibits the C1 activation domain and that this inhibition is relieved through the interaction with B or R. It is interesting that, whereas C1 has been shown to have a strong activation domain, careful analysis of B did not reveal any activation domains. The B/R proteins may also assist with the nuclear localization of C1. Whereas C1 absolutely requires either R or B to activate the anthocyanin biosynthetic gene promoters including the a1 gene, another Myb member, P1, is able to activate a subset of the anthocyanin biosynthetic genes, including the a1 gene, without a requirement for either the B or R proteins (Grote-Wold et al., 1994). Whether P1 is truly acting independently or is interacting with another b/HLH protein is not known.

The plant hormone ABA regulates a number of processes in plants, including helping to mediate the response to a number of environmental stresses, as well as the generation of specific expression patterns during seed development. ABA-induced gene expression is an important part of ABA action (for recent reviews, see Shen and Ho, 1997; Busk and Pages, 1998). A detailed analysis of the cis-acting sequences required for ABA-induced gene expression has been performed on a number of ABA-regulated genes, such as the wheat EM gene and two barley genes, HVA1 and HVA22. The first ABRE to be identified was an ACGT-containing sequence in the EM promoter (Guiltnan et al., 1990). ABREs have now been found in the promoters of many ABA-responsive genes, including the rice rab and barley HVA genes. There are ABA-responsive genes that do not contain ABREs, and other cis-acting sequences have been shown to function in ABA-responsive gene expression in some of these promoters.

The ABRE is similar to a family of sequences called the G-box, which also contain an ACGT core and are present in a number of gene promoters that respond to different environmental conditions, such as UV light, anoxaerobiosis, and wounding (Menkens et al., 1995). A bZIP protein called EmBP1 that specifically binds to the ABRE has been identified (Guiltnan et al., 1990). EmBP1 may be part of a larger protein-DNA complex, which includes VP1 and GF14 proteins (Shultz et al., 1998). In maize VP1 has been shown by genetic analysis to be important for mediating certain ABA responses during seed maturation (Carson et al., 1997, and refs. therein). Although VP1 is unable to directly bind the ABRE, it is able to transactivate the Em promoter through the ABRE sequences, presumably via protein-protein interaction with other proteins, such as EmBP1, that are recruited to the ABRE (McCarty et al., 1991; Vasil et al., 1995).

Analyses of truncated forms of VP1 show that the DNA-binding domain (B3) is not required for gene activation mediated through the ABRE (Carson et al., 1997). It is interesting that the activation of ABRE-coupled genes requires a second conserved basic domain of VP1, which in vitro can stimulate the DNA-binding activity of a broad spectrum of transcription factors, including EmBP1 (Hill et al., 1996). However, the actual order of events remains unclear because in vivo footprinting studies did not detect any major differences in ABRE binding when wild-type and vp1 mutant embryos were compared (Busk and Pages, 1997). GF14s are plant 14-3-3 proteins that were initially identified as part of a G-box protein-DNA complex (for review, see Ferl, 1996). Although their precise functions...
remain to be elucidated, they may play a role in promoting protein-protein interactions. GF14, which is unable to directly bind DNA, was found to interact with both EmBP1 and VP1 and could therefore provide a structural link in the ABA-responsive protein-DNA complex proposed by Shultz et al. (1998).

The promoter context can be important for how ABRE/G-box sequences give rise to specific expression patterns. A detailed analysis of the barley HVA1 and HVA22 promoters (Shen and Ho, 1995; Shen et al., 1996) demonstrated that the promoter unit necessary and sufficient to mediate the ABA response consisted of an ABRE and a closely linked sequence, called a coupling element (CE), that collectively constitute an ABRC. As shown in Figure 3, the HVA1 and HVA22 promoters each had a different ABRC, called ABRC3 and ABRC1, and distinct CEs, called CE1 and CE3, respectively. CE1 and CE3 differed both in their nucleotide sequence and their position relative to the ABRE. Whereas the ABRE sequences were interchangeable, the CEs were only partially exchangeable, with CE3 exhibiting flexibility in terms of its position relative to the ABRE. In contrast, CE1 could function only when placed distal to the ABRE. Further underscoring the differences between ABRC1 and ABRC3 was the finding that VP1 was able to enhance transcription only through ABRC3 but not ABRC1. It is interesting that synthetic promoters containing an ABRE and both CE1 and CE3 were substantially more ABA-responsive. Whereas distinct ABRCs may allow for the regulation of ABA-responsive genes in response to different environmental and/or physiological cues, a number of important questions still need to be addressed. For example, the protein(s) that acts through CE1 and CE3 needs to be identified and the mechanisms by which ABA leads to activation of specific bZIP proteins need to be understood. In rice, de novo protein synthesis may play a role, since the expression of a bZIP protein called OSBZ8 is rapidly induced by ABA (Nakagawa et al., 1996).

The maize VP1 protein illustrates an important aspect of combinatorial control whereby a given activator can have different roles depending on the promoter context. As mentioned earlier, VP1, acting as an activator, is involved as part of an ABA-responsive protein-DNA complex to mediate the ABA-induced expression of the wheat Em gene. In a different promoter context, VP1 acts as a repressor of α-amylase expression (Hoecker et al., 1995). VP1 also plays a role in anthocyanin gene expression by transactivating the C1 promoter through a DNA sequence called the Sph element (Kao et al., 1996, and refs. therein). The Sph element is quite distinct from the ABRE, which is important for VP1 transactivation of the Em promoter. While full-length VP1 is unable to bind to the Sph element, the 120-amino acid B3 domain, located at the C terminus of VP1, is able to bind to the Sph element with high specificity (Suzuki et al., 1997). An exciting possibility is that other regions of VP1 inhibit the DNA-binding activity of the B3 domain and that this inhibition may be relieved through protein-protein interactions between VP1 and other proteins to enable VP1 binding to the C1 promoter.

**Dof/bZIP INTERACTIONS MAY REGULATE SPECIFIC PATTERNS OF PLANT GENE EXPRESSION**

There is evidence that in addition to ABA responses, regulation of some other patterns of plant gene expression are mediated in part by combinatorial interactions between bZIP proteins and other types of transcription factors binding to closely linked sites. A good example is the interaction of bZIP and Dof transcription factors. Dof proteins are a new class of plant transcription factors that contain a single zinc-finger DNA-binding domain that is highly conserved in plants (for review, see Yanagisawa, 1996). Dof proteins have been shown to interact specifically with bZIP proteins, and this interaction results in stimulation of bZIP binding to DNA target sequences in plant promoters (Chen et al., 1996). The Arabidopsis glutathione S-transferase-6 gene (GST6) promoter contains a number of Dof-binding sites closely linked to another promoter sequence called the ocs element. The ocs elements are a family of 20-bp DNA promoter sequences that are important for the expression of a number of pathogen and plant genes and are the binding sites for bZIP proteins.

The regulation of seed-storage protein gene expression in maize is another case in which Dof/bZIP interactions may play an important role. A major class of maize seed-storage proteins are the 22- and 19-kD zein proteins (for review, see Aukerman and Schmidt, 1994). As shown in Figure 4, two cis-acting sequences, the O2 (Opaque 2) box and the prolamin box, have been found to be important for 22-kD zein expression in endosperm tissue. The O2 box is

---

**Figure 3.** A schematic representation of two ABRCs assembled at the promoters of ABA-responsive genes. A, Promoter sequences from the barley HVA22 gene that form ABRC1. Sequences of the ABRE (A3) and the CE (CE1) are shown. bZIP proteins bind to A3 as dimers, whereas the protein(s) that bind to CE1 remains to be identified. B, Promoter sequences from the barley HVA1 gene that form ABRC3. The sequences of the ABRE (A2) and the CE (CE3) are shown. bZIP proteins bind to A2 as dimers, whereas the protein(s) that bind to CE3 remains to be identified. The maize VP1 protein, which can transactivate from ABRC3 but not ABRC1, is also shown, although the precise position of VP1 in the complex remains to be established.
the binding site for a bZIP protein called O2, and molecular genetic studies have demonstrated that O2 is required for the expression of the 22-kD zein promoter in maize endosperm. The prolamin box is found in the promoters of all zein genes, as well as in storage protein genes in other cereals. In maize the prolamin box is the binding site for a Dof protein called PBF (Vicente-Carbajosa et al., 1997). PBF and O2 are expressed in an identical fashion in maize endosperm tissue, where they accumulate just before activation of zein gene expression. PBF interacts specifically with O2 but not another bZIP protein tested (Vicente-Carbajosa et al., 1997). The PBF/O2 interaction is likely to be important, since a functional prolamin box is required for O2-dependent activation in maize and for activation by related bZIP proteins in other cereals (Albani et al., 1997; Vicente-Carbajosa et al., 1997, 1998). It is interesting that other classes of zein genes lack an O2-binding site but contain the PBF site. In these cases it is not known whether PBF interacts with other transcription factors or is able to act alone.

ENHANCEOSOME FORMATION ON THE IFNβ PROMOTER CAN SERVE AS A USEFUL PARADIGM FOR PLANT STUDIES

The above examples provide evidence for the importance of combinatorial control in transcriptional regulation of plant gene expression. However, the situation in a number of cases is likely to be more complex with additional proteins playing a role. In animal systems a number of good examples of combinatorial control have been identified and proteins playing a role. In animal systems a number of cases is likely to be more complex with additional pro-

![Figure 4](image)

**Figure 4.** A schematic representation of DNA-protein interactions involved in expression of the maize 22-kD zein gene. The core sequences of the prolamin (P) box and the O2 box are shown. The bZIP protein O2 binds to the O2 box as a dimer. The Dof protein PBF binds to the prolamin box, although at this stage it is not known if PBF binds as a monomer or as some form of multimer. The dashed arrows represent potential interactions between PBF and O2.

![Figure 5](image)

**Figure 5.** A schematic representation of the virus-inducible IFNβ enhanceosome. The positions of the four positive regulatory domains are shown. The cooperative assembly of the enhanceosome leads to an optimal arrangement of activation domains for interactions with various components of the transcription-initiation complex. See the text for further details.
not itself activate transcription, was lost. HMGI(Y) appears to function by binding to DNA and inducing conformational changes that facilitate protein-protein interactions and cooperative DNA binding of the other proteins in the enhanceosome. HMGI(Y) is also able to directly interact with ATF2 and p50. A competition assay was used to assay the stability of different complexes assembled at the IFN\(\beta\) enhancer. These studies demonstrated that there was a good correlation between stability of the enhanceosome and transcriptional activity. The stability of the enhanceosome was further increased in the presence of components of the general transcriptional machinery, suggesting that the enhanceosome and transcription-initiation complexes reciprocally stabilize assembly. The IFN\(\beta\) enhanceosome should serve as a useful paradigm for studies of transcriptional control in plants.

**CONCLUSIONS**

I have reviewed some basic features of eukaryotic transcription with an emphasis on plant activators and the importance of combinatorial control. The IFN\(\beta\) promoter illustrates nicely how a high degree of specificity and transcriptional activity can be generated by the cooperative assembly and stability of a higher-order protein-DNA complex called an enhanceosome. It is the precise arrangement of the regulatory sequences and their corresponding transcription factors on the DNA that facilitate the many protein-protein interactions both within the enhanceosome and between the enhanceosome and components of the transcription-initiation complex. It is these multiple protein-protein and protein-DNA interactions that in large part dictate the amount of transcription that will occur at a given promoter under a specific set of conditions. The IFN\(\beta\) enhanceosome also demonstrates the importance of in vivo studies to validate/determine those combinatorial interactions identified initially at the molecular/biochemical level that are actually being used to regulate gene expression in vivo. This is certainly the case with the regulation of anthocyanin biosynthesis in maize, in which there is substantial genetic evidence that the interactions between C and R proteins are critical for transcriptional control.

The analysis of transcriptional control in plants will continue to be an exciting field of research. The rapid progress being made on the isolation of important regulatory proteins, the development of in vitro transcription systems, and the use of powerful genetic screening approaches for additional mutants using promoter/reporter gene fusions will facilitate further studies of transcriptional control in plants, which should provide valuable insight into the mechanisms underlying various aspects of plant growth and development and lead to agricultural benefits. The potential for agriculture is illustrated by studies of the molecular basis of plant tolerance to freezing, an important research area, since freezing temperatures have placed major limitations on agricultural productivity. Through the identification of cold-regulated genes and the careful analysis of their expression, an Arabidopsis transcriptional activator called CBF1 has been identified (Stockinger et al., 1997), overexpression of which increased the freezing tolerance of Arabidopsis plants (Jaglo-Ottosen et al., 1998). If similar mechanisms are used for freezing tolerance in important crop plants, the use of CBF1 and/or its orthologs may help make other plants more cold-tolerant.

**ACKNOWLEDGMENTS**

I would like to thank Drs. Mike Carey, Maarten Chrispeels, Don McCarty, and Bob Schmidt for their helpful comments concerning the manuscript, and Dr. Stephen Harrison and Cell Press for permission to use Figure 2. I apologize to the many colleagues whose work or original publications I was unable to cite because of space constraints.

Received July 13, 1998; accepted August 11, 1998. Copyright Clearance Center: 0032-0889/98/118/1111/10.

**LITERATURE CITED**


Carson CB, T Hattori, L Rosenkranz, V Vasil, IK Vasil, PA Peterson, DR McCarty (1997) The quietest/colorless alleles of viviparous1 show that the conserved B3 domain of VP1 is not essential for ABA regulated gene expression in the seed. Plant J 12: 1231–1240
transcriptional control


Vicente-Carbajosa J, Moose SP Parsons RL, Schmidt RJ (1997) A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. Proc Natl Acad Sci USA 94: 7685–7690

