Cell, Vol. 116, 247–257, January 23, 2004, Copyright ©2004 by Cell Press

Regulation of RNA Polymerase II Transcription by Sequence-Specific DNA Binding Factors

Review

James T. Kadonaga*
Section of Molecular Biology, 0347
University of California, San Diego
La Jolla, California 92093

In eukaryotes, transcription of the diverse array of tens of thousands of protein-coding genes is carried out by RNA polymerase II. The control of this process is predominantly mediated by a network of thousands of sequence-specific DNA binding transcription factors that interpret the genetic regulatory information, such as in transcriptional enhancers and promoters, and transmit the appropriate response to the RNA polymerase II transcriptional machinery. This review will describe some early advances in the discovery and characterization of the sequence-specific DNA binding transcription factors as well as some of the properties of these regulatory proteins.

Timing is everything. With regard to the subject of this essay, my own timing was perhaps a bit late, but not too late. I will therefore attempt to provide, based on my experiences over the past 20 years, a brief overview of some of the key advances in our understanding of the sequence-specific DNA binding factors that regulate transcription of protein-coding genes by RNA polymerase II. It is my aim to present an accurate and balanced perspective, but this endeavor will inevitably result in unintentional omissions and errors for which I apologize.

In eukaryotes, there are tens of thousands of protein-coding genes, each of which has its own specific program of transcriptional control. Much of the specificity of these programs is affected by sequence-specific DNA binding proteins that bind to the proximal promoter and distal transcriptional regulatory regions (such as enhancers and silencers). That is, sequence-specific DNA binding transcription factors (henceforth termed "sequence-specific factors") interpret and transmit the information that is encoded in the primary DNA sequence to the factors and cofactors that mediate the synthesis of RNA transcripts from the DNA template. Thus, the sequence-specific factors collectively function as the key interface between genetic regulatory information and the transcription system (Figure 1).

Transcription is a complex process that relies on the collective action of the sequence-specific factors along with the core RNA polymerase II transcriptional machinery, an assortment of coregulators that bridge the DNA binding factors to the transcriptional machinery, a number of chromatin-remodeling factors that mobilize nucleosomes, and a variety of enzymes that catalyze the covalent modification (e.g., acetylation, deacetylation, phosphorylation, dephosphorylation, methylation, ubiquitylation, deubiquitylation, and ADP-ribosylation) of histones and other proteins. Hence, the modern researcher in transcriptional control has much to think about. It is an unfortunate consequence of this complexity that the critical role of the sequence-specific factors is sometimes overlooked.

In an earlier time, our view of transcription was not so complicated. I will therefore go back in time to the early 1980s and describe some of our ideas and knowledge (or what I remember of what we thought we knew) regarding sequence-specific factors. I will then briskly travel forward in time through the present and into the future. This account reflects some selected thoughts and impressions, and is not meant to be comprehensive. To capture the flavor of this earlier era in transcriptional regulation, I have cited "vintage" papers and reviews. I hope that this perspective will be of some interest and utility to a broad range of scientists, particularly those individuals who have more recently become involved in the transcription field.

Identification of Sequence-Specific DNA Binding Transcription Factors

By the early 1980s, a few fundamental concepts in eukaryotic transcription had been established. The three RNA polymerases (I, II, and III) had been isolated, and RNA polymerase II was found to be responsible for the transcription of protein-coding genes. The C-terminal domain (CTD), a repeated heptapeptide motif at the C terminus of the largest subunit of RNA polymerase II, would soon be discovered. The "general" transcription factors (TFIID, TFIIB, etc.) were being fractionated, but the complete set of proteins that constitute the general/basal factors was not yet known. An early deletion analysis of a eukaryotic promoter was carried out with the *his3* gene in yeast (Struhl, 1981), and a particularly comprehensive analysis of mammalian *cis*-acting promoter elements was carried out with the herpes simplex virus thymidine kinase gene (McKnight and Kingsbury, 1982). These and other studies revealed a few well-defined *cis*-control elements such as the TATA box (Goldberg, 1979), CCAAT box, GC box, heat shock element (Pelham, 1982), and glucocorticoid response element (Chandler et al., 1983). Transcriptional enhancers were discovered as remarkable *cis*-acting sequences that act at long distances (>1 kbp) from either upstream or downstream of the RNA start site (for vintage reviews, see Khoury and Gruss, 1983; Serfling et al., 1985). In prokaryotes, it was well established that sigma factors and sequence-specific DNA binding proteins play an important role in gene-specific regulation of transcription. However, it was still a matter of speculation whether or not eukaryotic regulatory factors would have related properties.

One of the earliest reports of sequence-specific DNA binding by an RNA polymerase II transcription factor (Tjian, 1978) is that of a hybrid adenovirus-SV40 large T antigen protein (D2 protein) that had been shown to repress transcription from the SV40 early promoter, to activate transcription from the SV40 late promoter, and to function in SV40 DNA replication. Studies of the intact SV40 large T antigen confirmed that it binds to a specific region in SV40 that overlaps the early promoter and

*Correspondence: jkadonaga@ucsd.edu

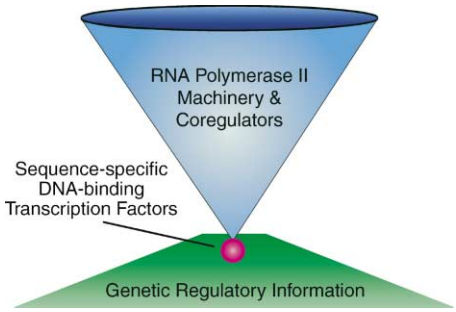

Figure 1. Sequence-Specific DNA Binding Transcription Factors Interpret and Transmit Genetic Regulatory Information

In this diagram, sequence-specific factors are depicted as the apex at the interface of the vast array of genetic regulatory information and the inverted cone of the RNA polymerase II transcriptional machinery and coregulators.

lies adjacent to the origin of DNA replication (Myers et al., 1981).

Another notable account of sequence-specific DNA binding by a eukaryotic transcription factor is that of TFIIIA, which is involved in the transcription of 5S RNA genes by RNA polymerase III. In these studies, purified TFIIIA was found to bind to a specific region in 5S RNA promoters (Engelke et al., 1980; Sakonju et al., 1981) as well as to 5S RNA itself (Honda and Roeder, 1980).

By the early 1980s, several RNA polymerase II factors were found to be sequence-specific DNA binding proteins. For instance, purified glucocorticoid receptor was observed to bind specifically to sequences in the long terminal repeat of the murine mammary tumor virus (Payvar et al., 1981, 1983; Scheidereit et al., 1983). Specificity protein 1 (Sp1), which was initially identified as a factor that selectively activates transcription from the SV40 early promoter (Dynan and Tjian, 1983a), was isolated and found to bind to the GC box motifs in the 21 bp repeat elements of SV40 (Dynan and Tjian, 1983b). Studies in *Drosophila* revealed binding of a partially purified factor (which is now known to be TFIID) to the TATA box motif (Parker and Topol, 1984a). In addition, heat shock transcription factor (also known as heat shock activator protein or HSF), which is involved in heat-induced transcription of heat shock genes, was partially purified and found to bind to the heat shock response element (Parker and Topol, 1984b). Furthermore, several yeast regulatory factors, which include GAL4 protein (Giniger et al., 1985; Bram and Kornberg, 1985), MAT α2 protein (Johnson and Herskowitz, 1985), and GCN4 (Hope and Struhl, 1985), were found to be sequence-specific DNA binding proteins.

A few other DNA binding transcription factors that were identified in the mid 1980s are as follows. USF (also known as MLTF) binds upstream of the TATA box in the adenovirus major late promoter (Sawadogo and Roeder, 1985; Carthew et al., 1985; Miyamoto et al., 1985). CAT binding protein, which is now known as C/EBP, was found to interact with the CCAAT box motif (Graves et al., 1986). Also, IgNF-A (now called Oct-1) was found in nuclear extracts by a novel application of the gel mobility shift assay (Singh et al., 1986).

The IgNF-A paper (Singh et al. 1986) not only reported a new factor, but it additionally inspired many researchers to use the gel shift assay (Garner and Revzin, 1981; Fried and Crothers, 1981) in their studies of promoter and enhancer binding factors. Similarly, the studies of sequence-specific DNA binding by TFIIIA (Engelke et al., 1980), T antigen (Myers et al., 1981), Sp1 (Dynan and Tjian, 1983b), and the glucocorticoid receptor (Payvar et al., 1983) are excellent early descriptions of the use of the DNase I footprinting assay (Galas and Schmitz, 1978) in the analysis of eukaryotic DNA binding factors. The analysis of the yeast GCN4 protein led to the development of new techniques, such as the detection of specific DNA-protein complexes with radiolabeled in vitro-translated factors (Hope and Struhl, 1985), the use of differently sized proteins for the deduction of factor binding stoichiometry (Hope and Struhl, 1987), and a general method (now typically referred to as SELEX) for isolating DNA recognition elements of sequence-specific DNA binding proteins (Oliphant et al., 1989).

In mammalian cells, the transcriptional activity of eukaryotic promoters was commonly analyzed by using the cell-based transient transfection assay (Mulligan and Berg, 1980), such as with the CAT reporter gene (Gorman et al., 1982). The cotransfection assay allowed the analysis of the properties of a transcription factor (introduced from one transfected template) in conjunction with a reporter gene (from the cotransfected template) (see, for example Imperiale et al., 1983; Green et al., 1983; Giguère et al., 1986). As a powerful complement to the cell-based assays, in vitro transcription assays (Wu, 1978; Weil et al., 1979; Manley et al., 1980; Rio et al., 1980; Wasylyk et al., 1980; Handa et al., 1981; Dignam et al., 1983) allowed the fractionation of the general/basal transcriptional machinery. These in vitro transcription systems were then further developed to enable the biochemical analysis of *cis*-acting DNA elements and *trans*-acting factors. These latter assays played an important role in the discovery and characterization of the sequence-specific factors (see, for example, Dynan and Tjian, 1983a).

Purification and Cloning of Sequence-Specific DNA Binding Transcription Factors

By the mid to late 1980s, the importance and generality of sequence-specific activators had become well established, and hence, there was vigorous activity directed toward the identification and characterization of DNA binding proteins that interact with promoter and enhancer elements. Many of these experiments were carried out with crude cell extracts or with partially purified protein fractions. Therefore, the next major challenge was to purify these low-abundance regulatory factors. This aim was achieved by the development of sequence-specific DNA-affinity chromatography (Rosenfeld and Kelly, 1986; Kadonaga and Tjian, 1986; Wu et al., 1987) (Figures 2 and 3). This widely applicable technique enabled the purification of many sequence-specific factors, which include Sp1 (Briggs et al., 1986; Kadonaga and Tjian, 1986); CTF/NF-I (Rosenfeld and Kelly, 1986; Jones et al., 1987); HSF (Wu et al., 1987); AP-1 (Lee et al., 1987; Angel et al., 1987); AP-2 (Mitchell et al., 1987; Imagawa et al., 1987); Oct-1 (OTF-1, OBP100; Fletcher et al., 1987; Sturm et al., 1987); Oct-2 (OTF-2; Scheidereit

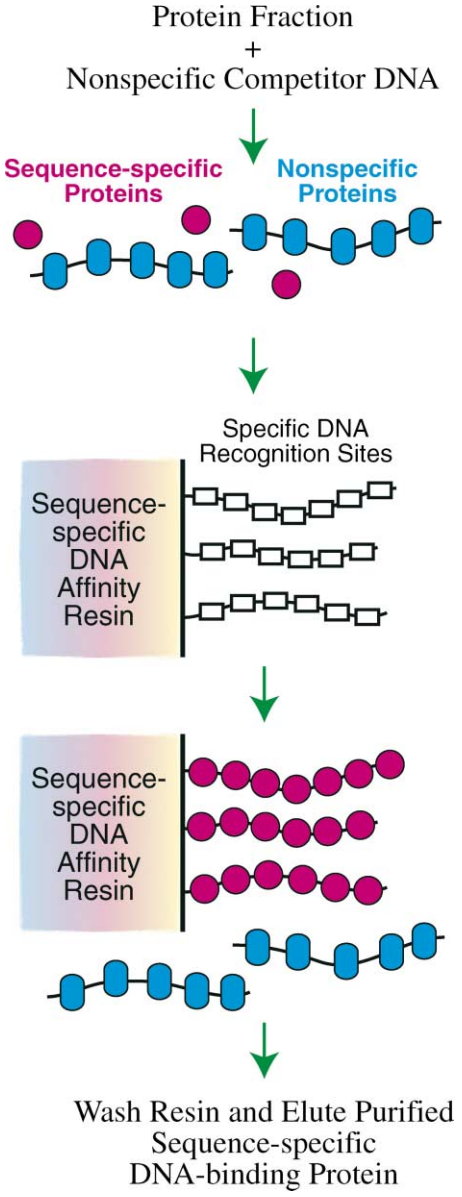

Figure 2. Purification of Sequence-Specific DNA Binding Proteins by Affinity Chromatography

The sequence-specific DNA affinity resin is prepared by CNBr-mediated coupling of multimerized synthetic oligonucleotides that contain the recognition site of the desired factor. To separate the sequence-specific DNA binding proteins from nonspecific DNA binding proteins, a nonspecific competitor DNA is added to the crude protein fraction (or extract) before application to the DNA affinity resin. The nonspecific DNA binding proteins bind to the nonspecific competitor DNA and flow through the DNA affinity resin, whereas the sequence-specific proteins bind to their recognition sites on the resin. The method depicted in this figure was described by Kadonaga and Tjian (1986), and has been used for the purification of many sequence-specific DNA binding factors, some of which are described in this review.

et al., 1987); SRF (Treisman, 1987); CREB (Montminy and Bilezikjian, 1987); ATF (Hai et al., 1988); and HNF1 (Courtois et al., 1988). The purification of these regulatory proteins was soon followed by the cloning of genes encoding these sequence-specific factors, which includes the cloning of one of my favorite factors, Sp1 (Kadonaga et al., 1987). In addition, a southwestern hybridization method (based on sequence-specific DNA binding) for the screening of expression libraries was developed (Singh et al., 1988) and further facilitated the cloning of many other transcription factors.

Figure 3. Transcription Factor Factory

A photograph (taken by the author) of the Tjian laboratory cold room in 1986. Sequence-specific DNA affinity columns for Sp1, CTF/NF-I, AP-1, AP-2, and others are shown. The DNA affinity columns were typically stacked in tandem to allow the simultaneous purification of multiple sequence-specific factors from a single protein fraction.

In a separate but related line of experiments, genes encoding steroid hormone receptors, such as the glucocorticoid receptor (Miesfeld et al., 1984; Weinberger et al., 1985; Govindan et al., 1985) and estrogen receptor α (Walter et al., 1985), were cloned by using immunological methods. These receptors were purified prior to the development of sequence-specific DNA-affinity chromatography on the basis of their unique ability to bind to nonspecific DNA resins in a ligand-dependent manner. The availability of purified steroid hormone receptors allowed the generation of antibodies, which were then used to clone the corresponding genes. Within a few years, a superfamily of genes encoding a variety of nuclear receptors had been isolated (for vintage reviews, see Yamamoto, 1985; Evans, 1988).

Hence, by the late 1980s, many sequence-specific factors had been identified, purified, and cloned (for a vintage review, see Mitchell and Tjian, 1989). This work established the paradigm that the binding of sequence-specific factors to their cognate regulatory elements in promoters and enhancers is responsible for the communication of the genetic information that is encoded in the primary DNA sequence to the RNA polymerase II transcriptional machinery.

Properties of Sequence-Specific DNA Binding Transcription Factors

The field was then faced with several new challenges—notably, to understand how these factors work (for vintage reviews, see Struhl, 1987; Ptashne, 1988). At the time, several new directions were apparent. First, by using insights gained from the study of prokaryotic tran-

scription factors, it was possible to design experiments that revealed functional subregions of the transcription factors (in the Supplement to this issue of *Cell*). This goal was pursued, in part, by the generation and characterization of mutant and chimeric proteins. A second aim was to complete the fractionation and purification of the general transcription machinery, because it was apparent that we could not understand the function of the sequence-specific factors without knowledge of the components of the basic RNA polymerase II transcription apparatus. (It is also interesting to note that several components of the basal transcription machinery have since been found to possess properties that are reminiscent of bacterial sigma factors.) A third aim was to search for potential coregulatory factors that are distinct from the sequence-specific factors and the basal RNA polymerase II machinery. The existence of such cellular coregulators was suggested, for instance, by the analysis of viral transactivators that had a strong influence on transcription but did not appear to bind directly to DNA. A fourth and somewhat controversial aim was to investigate the role of chromatin in the regulation of transcriptional activity. These studies were inspired by the fact that chromatin is the natural state of the DNA template in vivo. These and other research pursuits led to an explosion of discoveries, some of which are highlighted below.

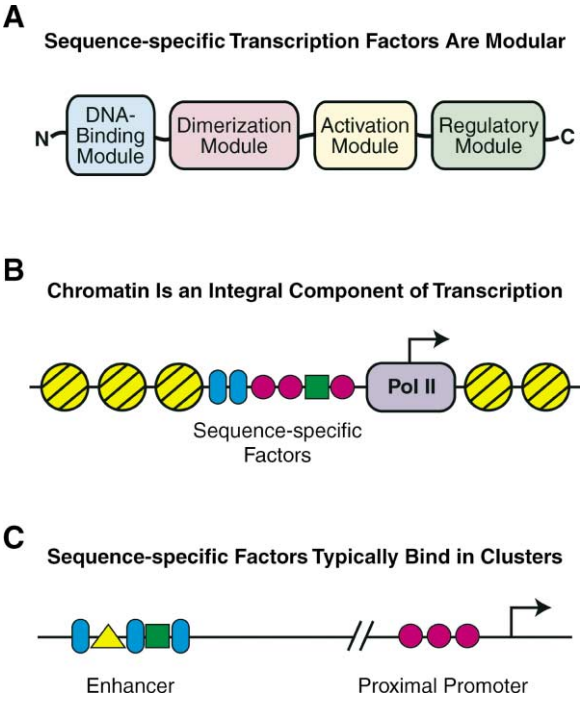

Figure 4. Some Properties of Sequence-Specific DNA Binding Transcription Factors

(A) Sequence-specific factors are composed of functional modules.
(B) Chromatin is an integral component in the function of sequence-specific factors.
(C) Recognition sites for sequence-specific factors tend to be located in clusters. These and other properties of sequence-specific factors are described in the text.

Sequence-Specific Transcription Factors Are Modular

A typical sequence-specific factor has a DNA binding module linked to one or more activation or repression modules as well as perhaps a multimerization module and a regulatory module (Figure 4A). This remarkable and unexpected modular nature of eukaryotic transcription factors was first revealed in the analysis of the yeast GAL4 transcription factor (Brent and Ptashne, 1985; see Supplement to this issue of *Cell*). In this work, the DNA binding region of LexA (a bacterial sequence-specific DNA binding protein) was fused to the entire GAL4 protein to yield a hybrid LexA-GAL4 protein that activates transcription in a LexA binding-site-dependent manner. These results indicated that the GAL4 transcriptional activation module can function autonomously when fused to the heterologous LexA DNA binding module.

The first well-defined DNA binding module was the helix-turn-helix motif, which was originally discovered in prokaryotic DNA binding proteins. Since some of the early eukaryotic transcription factors were found to be helix-turn-helix proteins, it seemed, for a time, that this one motif would be used universally for the binding of proteins to DNA. It later became apparent, however, that nature has generated many distinct DNA binding modules, which include the homeodomain (a variant of the helix-turn-helix), zinc finger (of which there are different types), leucine zipper, helix-loop-helix, HMG1 domain, and others (for a vintage review, see Pabo and Sauer, 1992). Both the leucine zipper (Landschultz et al., 1988) and helix-loop-helix (Murre et al., 1989) motifs with their associated basic regions are somewhat unusual because they perform dual DNA binding and dimerization functions. The discovery of the leucine zipper was particularly enlightening not only because it revealed a novel DNA binding and dimerization motif, but also because its structure was deduced from the primary amino acid sequence. This rare feat of deductive reasoning was repeated in the description of the helix-loop-helix motif.

Unlike the well-defined DNA binding motifs, transcriptional activation modules have generally been structurally more elusive. For example, the fusion of random *Escherichia coli* genomic DNA fragments with the coding sequence of the DNA binding region of the yeast GAL4 protein resulted in the generation of many functional fusion proteins that activate transcription. These hybrid activators appear to have no properties in common except for a net negative charge in the activating peptide derived from the *E. coli* DNA sequences (Ma and Ptashne, 1987; see Supplement to this issue of *Cell*). Analysis of GAL4-peptide fusions also revealed that a synthetic amphipathic α helix with one acidic face and one hydrophobic face can function as a transcriptional activation region (Giniger and Ptashne, 1987). Parallel studies of the yeast GCN4 protein revealed that its activation region is a distinct module enriched in acidic amino acid residues (Hope and Struhl, 1986) that appear to contribute to the activation function (Hope et al., 1988). These results collectively led to the "acid blob" theory of transcriptional activation, in which an amphipathic α helix with a negatively charged face can function as a transcriptional activation module (for vintage reviews, see Sigler, 1988; Ptashne, 1988).

At that point in time, it seemed possible that acid blobs might constitute the predominant and perhaps only transcriptional activation module. Nature did not turn out to be so simple, however. In addition to acid blobs, transcriptional activation regions have been found to contain other distinct motifs. The first of the nonacidic activation regions was found in transcription factor Sp1, which contains multiple glutamine-rich activation motifs (Courey and Tjian, 1988). Other transcriptional activation motifs include proline-rich regions (see, for example Mermod et al., 1989) and hydrophobic β sheets (see, for example Leuther et al., 1993; Van Hoy et al., 1993). More recently, transcriptional activation regions, such as those in the HPV E2 protein and in AF-2 regions of nuclear receptors have been defined more precisely by X-ray structural analyses as interaction sites for specific coregulators. It seems likely that more specific structure-function insights into transcriptional activation and repression regions will emerge in the future.

Transcriptional activation modules have also been found to associate with DNA binding modules in *trans* via protein-protein interactions. This concept was introduced in the analysis of the herpes virus VP16 transactivator protein (Triezenberg et al., 1988a, 1988b; Preston et al., 1988; O'Hare and Goding, 1988). In these studies, VP16 was found to function as a transcriptional coactivator via protein-protein interactions with one or more sequence-specific factors in the host cell. In addition, this work led to the creation of the well-known GAL4-VP16 fusion protein (Sadowski et al., 1988), which binds to GAL4 recognition sites via the GAL4 DNA binding module and activates transcription via the potent VP16 activation module.

Some transcription factors are controlled by regulatory modules. For example, nuclear receptors are transcription factors that contain a regulatory module that is located on the same polypeptide as the DNA binding and transcriptional activation modules. The existence of a distinct regulatory module in nuclear receptors was revealed in an analysis of the glucocorticoid receptor (Picard et al., 1988). In these experiments, the steroid binding domain was found to be sufficient to confer hormone regulation when it is moved from the C terminus of the glucocorticoid receptor (its normal location) to the N terminus of the protein, or even when it is fused to a different unrelated protein (E1A). IκB is an example of a regulatory module that is not covalently attached to the transcription factor (NF-κB proteins) that it regulates (Baeuerle and Baltimore, 1988). Instead, IκB functions as a detachable regulatory subunit that modulates the activity and cellular location of NF-κB. Thus, as seen with transcriptional activation modules, regulatory modules can act both in *cis* or in *trans* relative to their specific target proteins.

Sequence-Specific Factors Regulate Transcription Via Recruitment of Coactivators and Corepressors

How do sequence-specific factors work? Current evidence indicates that the sequence-specific factors function mainly by recruitment of transcriptional coactivators and corepressors to the DNA template via protein-protein interactions (for a review, see Ptashne and Gann, 1997). These cofactors then act both directly and indirectly to regulate the activity of the RNA polymerase II transcriptional machinery at the core promoter.

Over the past decade, a large number of transcriptional coregulators have been identified and characterized (for some recent reviews, see Glass and Rosenfeld, 2000; Goodman and Smolik, 2000; Lemon and Tjian, 2000; Strahl and Allis, 2000; Zhang and Reinberg, 2001; McKenna and O'Malley, 2002; Narlikar et al., 2002; Freiman and Tjian, 2003; Hampsey and Reinberg, 2003). Many but not all of these coactivators and corepressors are recruited to the DNA template via interactions with the sequence-specific factors. Some coregulators are direct intermediaries between the sequence-specific factors and the general/basal transcriptional machinery. For instance, there are specific interactions between sequence-specific factors and TAF (TBP-associated factor) subunits of the TFIID component of the basal machinery. (For a notable early analysis of the interaction between TFIID and a sequence-specific factor, see Horikoshi et al., 1988.) In addition to TAFs, many other coactivator complexes (which include TRAP, SMCC, Mediator, SRB complex, CRSP, DRIP, NAT, p300/CBP, and others) can serve as a bridge between the sequence-specific factors and the general/basal transcriptional machinery. Another distinct class of cofactors are chromatin-related coregulators (described briefly below), which are also thought to be recruited by the sequence-specific factors. The chromatin-related coregulators affect transcription indirectly by remodeling nucleosomes or by covalent modification of histones (e.g., by acetylation, methylation, phosphorylation, ubiquitylation, and ADP-ribosylation) or the DNA template. Other coregulators are not recruited by the sequence-specific factors, but instead interact directly with RNA polymerase II and modulate the efficiency of transcriptional elongation.

These phenomena reveal the diversity and complexity of transcriptional regulation. It is notable that the theme of recruitment via protein-protein interactions is a concept that originally arose in the analysis of transcriptional regulation in prokaryotes. Yet, in eukaryotes, there are also other modes of regulation such as the covalent modification of histones, nucleosome remodeling, and the formation of higher order chromatin structures. In the future, it will be important to understand the contexts in which each of the coregulators is required for transcriptional control and how these factors work in concert to potentiate the transcriptional signals that emanate from the sequence-specific factors.

Sequence-Specific Factors Can Be Regulated by Posttranslational Modifications

Some sequence-specific factors have been found to be regulated by posttranslational modifications. An excellent example of this phenomenon is the phosphorylation of CREB (cyclic AMP response element binding protein) (Gonzalez and Montminy, 1989). CREB activates transcription of cyclic AMP-inducible genes via binding to CRE (cyclic AMP response element) motifs. First, cyclic AMP stimulates protein kinase A, which phosphorylates CREB at serine residue 133. Then, upon phosphoryla-

tion, CREB becomes activated and stimulates transcription from cyclic AMP-responsive promoters. Another example of posttranslational modification of transcription factors is the *O*-linkage of *N*-acetylglucosamine in many sequence-specific factors, which include Sp1, AP-1, AP-2, CTF/NF-I, Zeste, GAGA factor, and Adf-1 (Jackson and Tjian, 1988). The function of the *O*-linked *N*-acetylglucosamine monosaccharide residues remains to be clarified. Some results suggest a role of this modification in transcriptional repression (see for example Yang et al., 2002a). Sequence-specific factors can also be acetylated. For example, the acetylation of p53 increases the affinity of its binding to DNA (Gu and Roeder, 1997). Many transcription factors, which include the sequence-specific proteins, are ubiquitylated (for reviews, see Muratani and Tansey, 2003; Freiman and Tjian, 2003). Moreover, ubiquitylation of a LexA-VP16 fusion protein was observed to be required for its ability to activate transcription in vivo (Salghetti et al., 2001). There are many interesting and important avenues to pursue in the area of posttranslational modification of transcription factors.

Sequence-Specific Factors Are Often Members of Multiprotein Families

It is amusing to remember that some of us had naively thought (back in the early 1980s) that there might be perhaps only 10 to 50 sequence-specific factors that controlled the transcription of tens of thousands of genes by some sort of combinatorial mechanism. Today, it is apparent that combinatorial mechanisms are used in transcriptional regulation (for a vintage paper on combinatorial mechanisms, see Britten and Davidson, 1969). To our surprise, however, there are thousands of sequence-specific factors in addition to the nearly 100 polypeptides that constitute the core RNA polymerase II machinery.

Many sequence-specific factors are members of multiprotein families. For instance, as noted above, nuclear receptors are members of a superfamily of related proteins (for a vintage review, see Evans, 1988). AP-1 consists of Fos, Jun, and many other Fos- and Jun-related proteins as well as ATF and CREB-like proteins (for a vintage review, see Curran and Franza, 1988). CTF/NF-I is a family of proteins that appear to result from alternative RNA splicing (Santoro et al., 1988). NF-κB is yet another family of proteins (see, for example Baldwin, 1996). p53, p63, and p73 are also a protein family (see, for example Yang et al., 2002b). Even the original Sp1 has turned out to be a member of the Sp family of proteins (Kingsley and Winoto, 1992). Thus, there are many families of transcription factors. Within each family, the members often display closely related or essentially identical DNA binding properties but distinct activation functions. It remains a significant challenge to elucidate the molecular bases for the unique functional specificities of individual members of each transcription factor family.

Chromatin Is an Integral Component in the Function of Sequence-Specific Factors

Chromatin is the natural state of the DNA template. Yet, for many years, chromatin had been commonly viewed as an insubstantial subject area in which eccentric individuals could do little harm to our general understanding of gene regulation. By the early 1990s, however, new experimental methods and approaches facilitated chromatin research and led to several different lines of evidence that chromatin is an integral participant in the regulation of transcription (Figure 4B). One early concept was that transcription factors function primarily to counteract chromatin-mediated global repression of basal transcription in the absence of activators (for some contemporary reviews, see Grunstein, 1990; Felsenfeld, 1992; Kornberg and Lorch, 1992; Wolffe, 1992; Adams and Workman, 1993; Paranjape et al., 1994). This "antirepression" model is distinct but not mutually exclusive from a "true activation" model (the prevailing thought at the time) in which sequence-specific factors function to increase the rate of the intrinsic transcription process. Current data support both of these hypotheses, but also reveal an unforeseen complexity in the involvement of chromatin with transcription that includes chromatin-remodeling factors (ATP-dependent enzymes that mobilize nucleosomes) and a variety of histone-modifying enzymes. It appears that sequence-specific factors recruit chromatin-remodeling factors and histone-modifying enzymes, which in turn function to rearrange chromatin structure (for instance, to relieve chromatin-mediated repression) as well as to modify histones in a specific fashion that promotes the desired gene activation or repression.

From a broader perspective, eukaryotic transcription factors have coexisted with chromatin for hundreds of millions of years. Over this time, the process of transcriptional regulation has evolved to function optimally in chromatin. Moreover, it appears that not only sequence-specific factors, but also components of the core transcriptional machinery have evolved to interact specifically with distinct features of chromatin (e.g., binding to acetylated histones via bromodomains of coregulators, TAFs, and so on). Hence, we humans may have been surprised to see the integral role of chromatin in transcription, but the transcription factors themselves have been partners with chromatin for quite a long time.

Recognition Sites for Sequence-Specific Factors Tend to Be Located in Clusters

Individual eukaryotic sequence-specific factors generally bind to DNA with relatively low specificity. This phenomenon has been observed not only in vitro, but also in vivo (see, for example Walter et al., 1994). Thus, the precise control of gene transcription requires a higher degree of specificity than that typically afforded by the binding of a single sequence-specific factor to DNA. Instead, the high degree of specificity and potency of promoter and enhancer binding factors appears to be accomplished by the utilization of multiple factor recognition sites in composite *cis*-regulatory arrays (Figure 4C). Hence, a cluster of several short (about 6 to 8 bp) recognition sites, such as in an enhancer region, would be rarely encountered in the genome, even though a single recognition site might be common. It is also relevant to note that multiple sequence-specific factors in a cluster typically function synergistically and activate transcription more strongly than a single factor alone

(see, for instance Carey et al., 1990; Laybourn and Kadonaga, 1992). In this manner, the specificity of gene activation by sequence-specific factors derives from the use of multiple, clustered *cis* binding sites in conjunction with the synergistic enhancement of transcription that is achieved with multiple *trans*-acting factors.

Some Other Properties of Sequence-Specific DNA Binding Transcription Factors

I have neither the knowledge nor the desire to provide a comprehensive description of all of the known functions of all of the sequence-specific factors. I will, however, briefly mention some additional properties of these factors.

First, sequence-specific factors have been found to interact with transcriptional insulator (also known as "boundary") elements, which function to block the spreading of the influence of either positive DNA elements (such as enhancers) or negative DNA elements (such as a silencers, or heterochromatin-like repressive effects) (for some recent reviews, see Bell et al., 2001; Labrador and Corces, 2002; Kuhn and Geyer, 2003). Sequence-specific factors that have been found to act at insulator elements include BEAF-32, suppressor of Hairy-wing, CTCF, and Zw5. It will be important to determine the specific features of these factors that enable them to function in transcriptional insulation.

Second, sequence-specific activators can stimulate transcription elongation as well as initiation (see, for example Rougvie and Lis, 1988; Yankulov et al., 1994; Blau et al., 1996). In fact, some factors primarily stimulate initiation, whereas other factors stimulate predominantly elongation or both initiation and elongation. It will be interesting to study further the relation between promoter and enhancer binding proteins and transcription elongation factors.

Third, it is useful to mention that a subset of sequence-specific activators, which include Sp1 and CBF/NF-Y, are commonly found in the proximal promoter region of genes, such as from -250 to -30 relative to the $+1$ transcription start site. Some of these proximal promoter factors might function most effectively near the start site because they interact with the core transcriptional machinery. Moreover, some of the promoter proximal factors might also act as a conduit between distal enhancers and the basal/general transcriptional machinery, possibly as tethering factors that link or loop enhancer complexes to the core transcription complex (see, for example Calhoun et al., 2002).

Lastly, a study of the SV40 enhancer (in which individual cells were analyzed rather than extracts prepared from a pool of cells) revealed that the enhancer appears to increase the probability of transcription in any particular cell rather than the general level of transcription in all cells (Weintraub, 1988; for a review, see Blackwood and Kadonaga, 1998). Thus, it may be informative to examine the effects of enhancers or transcriptional activators upon individual cells rather than the average of a population of cells.

What Lies Ahead in the Future?

I don't have a mystical crystal ball, but I do have an old cobalt blue aspirin bottle (dated "1899") that was pulled out of an old garage dump in Colorado. This blue bottle is not exactly Aladdin's lamp, but it does reveal visions of the future...

What should we do? First, there is every reason to conclude that we should continue doing what we have been doing. That is, we should continue to develop in vitro systems to decipher the molecular mechanisms by which sequence-specific factors and cofactors regulate transcription. We should devise novel assays for the discovery and isolation of new activities. We should determine low as well as high resolution structures of transcription factors and cofactor complexes. We should study the regulation of transcription complex assembly and disassembly. We should also further investigate the role of intranuclear localization in transcriptional regulation. Ultimately, we need to resolve the underlying logic that governs the networks of gene expression. In other words, we should try to understand the nature of nature.

Transcriptional regulation is undoubtedly a complicated process, but it is important not to forget that the binding of sequence-specific factors to their cognate recognition sequences is a key step in the transmission of genetic information from the primary DNA sequence to the downstream transcriptional machinery. I am optimistic that we may someday begin to understand a gene expression code that reveals much of the transcriptional program of each gene during development and differentiation based on its primary DNA sequence. To this end, it remains essential to identify the *cis*-elements and *trans*-acting factors (i.e., the sequence-specific DNA binding proteins) that regulate the transcription of each gene. Thus, I believe that it is important for us to continue to "bash" (i.e., perform detailed and systematic analyses of) promoters and enhancers, even though such studies are often thought to be mundane.

At present, a huge amount of information is being generated from genomic analyses, which includes gene expression profiles as well as chromatin immunoprecipitation (ChIP) results with specific transcription factors, coregulators, histone modifications, and so on. To interpret and to integrate these data, it will be necessary to have comprehensive and accurate knowledge of the transcription process. To this end, we will need to devote more effort toward the biochemical analysis of transcriptional regulation. In this area, it will be particularly important to develop further the use of chromatin templates with highly purified transcription factors. These in vitro reconstitution experiments are challenging, yet they are essential for the mechanistic knowledge of transcription that we need.

We also need to dissolve artificial barriers that sometimes separate "fields of research." We often segregate biochemical phenomena, such as transcription and chromatin dynamics or transcription and RNA processing, when such phenomena are intertwined in the cell. In the future, we will likely experience a coalescence of many traditionally defined fields of research.

What else could we do? I'd like to suggest that we devote more effort toward studying the unnatural—that is, phenomena that do not (at least as far as we know) occur in nature. There are many more things that *can* happen than *do* happen, and many of these novel phenomena will be not only interesting, but also occasion-

ally of practical application. In this regard, it might be relevant to make a comparison to the field of organic chemistry, where few chemists today study naturally occurring chemical reactions. Instead, they use their knowledge of chemistry to make new molecules by using new chemical reactions. Perhaps by the next century, there may be many more biologists who create new and unnatural phenomena and biological systems.

In conclusion, the discovery, purification, cloning, and characterization of sequence-specific DNA binding transcription factors has been a fascinating and wonderful scientific journey. These proteins are of nearly universal biological importance—it is hard to imagine biological phenomena that do not involve these regulatory factors. For instance, the classic screen for segmentation defects in *Drosophila* (Nüsslein-Volhard and Wieschaus, 1980) resulted in the identification of genes that mostly encode transcription factors. I can also remember the disappointment of some developmental biologists when homeodomain-containing proteins were found to be sequence-specific transcription factors.

Notably, the pathway that has led to our current knowledge of these factors has been neither linear nor progressive, but rather has involved many revisions, or "revolutions" (Kuhn, 1970), of hypotheses and paradigms. For instance, an example of a model that has changed is the role of chromatin in the regulation of transcription. Of course, I mean not to criticize paradigms that were later corrected, because the revision of hypotheses is what science is all about. It is natural and appropriate for lucubrations in the study to be followed by experiments in the laboratory. I look forward to seeing how the concepts described in this essay are changed by new revolutions in the field. There remain many exciting new frontiers in transcription for scientific revolutionists.

Acknowledgments

Special thanks to Kathy Jones and Robert Tjian for their invaluable advice and reminiscences. I express my gratitude to Steve McKnight, Keith Yamamoto, Kevin Struhl, Robert Tjian, Kathy Jones, Alexandra Lusser, Jer-Yuan Hsu, and Tammy Juven-Gershon for critical reading of this manuscript and very useful suggestions for its improvement. Research in my laboratory is supported by grants from the National Institutes of Health (GM41249, GM46995, GM58272) and the Volkswagen Stiftung (I/79 725).

References

Adams, C.C., and Workman, J.L. (1993). Nucleosome displacement in transcription. Cell *72*, 305–308.

Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R.J., Rahmsdorf, H.J., Jonat, C., Herrlich, P., and Karin, M. (1987). Phorbol ester-inducible genes contain a common cis element recognized by a TPA-modulated trans-acting factor. Cell *49*, 729–739.

Baeuerle, P.A., and Baltimore, D. (1988). IκB: a specific inhibitor of the NF-κB transcription factor. Science *242*, 540–546.

Baldwin, A.S., Jr. (1996). The NF-κB and IκB proteins: new discoveries and insights. Annu. Rev. Immunol. *14*, 649–683.

Bell, A.C., West, A.G., and Felsenfeld, G. (2001). Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. Science *291*, 447–450.

Blackwood, E.M., and Kadonaga, J.T. (1998). Going the distance: a current view of enhancer action. Science *281*, 60–63.

Blau, J., Xiao, H., McCracken, S., O'Hare, P., Greenblatt, J., and Bentley, D. (1996). Three functional classes of transcriptional activation domains. Mol. Cell. Biol. *16*, 2044–2055.

Bram, R.J., and Kornberg, R.D. (1985). Specific protein binding to far upstream activating sequences in polymerase II promoters. Proc. Natl. Acad. Sci. USA *82*, 43–47.

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

Briggs, M.R., Kadonaga, J.T., Bell, S.P., and Tjian, R. (1986). Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science *234*, 47–52.

Britten, R.J., and Davidson, E.H. (1969). Gene regulation for higher cells: a theory. Science *165*, 349–357.

Calhoun, V.C., Stathopoulos, A., and Levine, M. (2002). Promoter-proximal tethering elements regulate enhancer-promoter specificity in the *Drosophila Antennapedia* complex. Proc. Natl. Acad. Sci. USA *99*, 9243–9247.

Carey, M., Lin, Y.-S., Green, M.R., and Ptashne, M. (1990). A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. Nature *345*, 361–364.

Carthew, R.W., Chodosh, L.A., and Sharp, P.A. (1985). An RNA polymerase II transcription factor binds to an upstream element in the adenovirus major late promoter. Cell *43*, 439–448.

Chandler, V.L., Maler, B.A., and Yamamoto, K.R. (1983). DNA sequences bound specifically by glucocorticoid receptor in vitro render a heterologous promoter hormone responsive in vivo. Cell *33*, 489–499.

Courey, A.J., and Tjian, R. (1988). Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. Cell *55*, 887–898.

Courtois, G., Baumhueter, S., and Crabtree, G.R. (1988). Purified hepatocyte nuclear factor 1 interacts with a family of hepatocyte-specific promoters. Proc. Natl. Acad. Sci. USA *85*, 7937–7941.

Curran, T., and Franza, B.R., Jr. (1988). Fos and Jun: the AP-1 connection. Cell *55*, 395–397.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. *11*, 1475–1489.

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

Dynan, W.S., and Tjian, R. (1983b). The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell *35*, 79–87.

Engelke, D.R., Ng, S.Y., Shastry, B.S., and Roeder, R.G. (1980). Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. Cell *19*, 717–728.

Evans, R.M. (1988). The steroid and thyroid hormone receptor superfamily. Science *240*, 889–895.

Felsenfeld, G. (1992). Chromatin as an essential part of the transcriptional mechanism. Nature *355*, 219–223.

Fletcher, C., Heintz, N., and Roeder, R.G. (1987). Purification and characterization of OTF-1, a transcription factor regulating cell cycle expression of a human histone H2b gene. Cell *51*, 773–781.

Freiman, R.N., and Tjian, R. (2003). Regulating the regulators: lysine modifications make their mark. Cell *112*, 11–17.

Fried, M., and Crothers, D.M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res. *9*, 6505–6525.

Galas, D., and Schmitz, A. (1978). DNAase footprinting: a simple method for the detection of protein-DNA binding specificity. Nucleic Acids Res. *5*, 3157–3170.

Garner, M.M., and Revzin, A. (1981). A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. Nucleic Acids Res. *9*, 3047–3060.

Giguère, V., Hollenberg, S.M., Rosenfeld, M.G., and Evans, R.M.

(1986). Functional domains of the human glucocorticoid receptor. Cell *46*, 645–652.

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.

Giniger, E., Varnum, S.M., and Ptashne, M. (1985). Specific DNA binding of GAL4, a positive regulatory protein of yeast. Cell *40*, 767–774.

Glass, C.K., and Rosenfeld, M.G. (2000). The coregulator exchange in transcriptional functions of nuclear receptors. Genes Dev. *14*, 121–141.

Goldberg, M.L. (1979). Sequence analysis of Drosophila histone genes. PhD thesis, Stanford University, Stanford, California.

Gonzalez, G.A., and Montminy, M.R. (1989). Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. Cell *59*, 675–680.

Goodman, R.H., and Smolik, S. (2000). CBP/p300 in cell growth, transformation, and development. Genes Dev. *14*, 1553–1577.

Gorman, C.M., Moffat, L.F., and Howard, B.H. (1982). Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. *2*, 1044–1051.

Govindan, M.V., Devic, M., Green, S., Gronemeyer, H., and Chambon, P. (1985). Cloning of the human glucocorticoid receptor cDNA. Nucleic Acids Res. *13*, 8293–8303.

Graves, B.J., Johnson, P.F., and McKnight, S.L. (1986). Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. Cell *44*, 565–576.

Green, M.R., Treisman, R., and Maniatis, T. (1983). Transcriptional activation of cloned human beta-globin genes by viral immediate-early gene products. Cell *35*, 137–148.

Grunstein, M. (1990). Histone function in transcription. Annu. Rev. Cell Biol. *6*, 643–678.

Gu, W., and Roeder, R.G. (1997). Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell *90*, 595–606.

Hai, T.W., Liu, F., Allegretto, E.A., Karin, M., and Green, M.R. (1988). A family of immunologically related transcription factors that includes multiple forms of ATF and AP-1. Genes Dev. *2*, 1216–1226.

Hampsey, M., and Reinberg, D. (2003). Tails of intrigue: phosphorylation of RNA polymerase II mediates histone methylation. Cell *113*, 429–432.

Handa, H., Kaufman, R.J., Manley, J., Gefter, M., and Sharp, P.A. (1981). Transcription of Simian virus 40 DNA in a HeLa whole cell extract. J. Biol. Chem. *256*, 478–482.

Honda, B.M., and Roeder, R.G. (1980). Association of a 5S gene transcription factor with 5S RNA and altered levels of the factor during cell differentiation. Cell *22*, 119–126.

Hope, I.A., and Struhl, K. (1985). GCN4 protein, synthesized in vitro, binds *HIS3* regulatory sequences: Implications for general control of amino acid biosynthetic genes in yeast. Cell *43*, 177–188.

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

Hope, I.A., and Struhl, K. (1987). GCN4, a eukaryotic transcriptional activator protein, binds as a dimer to target DNA. EMBO J. *6*, 2781–2784.

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*, 635–640.

Horikoshi, M., Hai, T., Lin, Y.-S., Green, M.R., and Roeder, R.G. (1988). Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. Cell *54*, 1033–1042.

Imagawa, M., Chiu, R., and Karin, M. (1987). Transcription factor AP-2 mediates induction by two different signal-transduction pathways: protein kinase C and cAMP. Cell *51*, 251–260.

Imperiale, M.J., Feldman, L.T., and Nevins, J.R. (1983). Activation of gene expression by adenovirus and herpesvirus regulatory genes acting in trans and by a cis-acting adenovirus enhancer element. Cell *35*, 127–136.

Jackson, S.P., and Tjian, R. (1988). O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. Cell *55*, 125–133.

Johnson, A.D., and Herskowitz, I. (1985). A repressor (MAT α2 product) and its operator control expression of a set of cell type specific genes in yeast. Cell *42*, 237–247.

Jones, K.A., Kadonaga, J.T., Rosenfeld, P.J., Kelly, T.J., and Tjian, R. (1987). A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. Cell *48*, 79–89.

Kadonaga, J.T., and Tjian, R. (1986). Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. USA *83*, 5889–5893.

Kadonaga, J.T., Carner, K.R., Masiarz, F.R., and Tjian, R. (1987). Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell *51*, 1079–1090.

Khoury, G., and Gruss, P. (1983). Enhancer elements. Cell *33*, 313–314.

Kingsley, C., and Winoto, A. (1992). Cloning of GT box-binding proteins: a novel Sp1 multigene family regulating T-cell receptor gene expression. Mol. Cell. Biol. *12*, 4251–4261.

Kornberg, R.D., and Lorch, Y. (1992). Chromatin structure and transcription. Annu. Rev. Cell Biol. *8*, 563–587.

Kuhn, T.S. (1970). The Structure of Scientific Revolutions, 2nd ed. (Chicago: The University of Chicago Press).

Kuhn, E.J., and Geyer, P.K. (2003). Genomic insulators: connecting properties to mechanism. Curr. Opin. Cell Biol. *15*, 259–265.

Labrador, M., and Corces, V.G. (2002). Setting the boundaries of chromatin domains and nuclear organization. Cell *111*, 151–154.

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.

Laybourn, P.J., and Kadonaga, J.T. (1992). Threshold phenomena and long-distance activation of transcription by RNA polymerase II. Science *257*, 1682–1685.

Lee, W., Mitchell, P., and Tjian, R. (1987). Purified transcription factor AP-1 interacts with TPA-inducible enhancer elements. Cell *49*, 741–752.

Lemon, B., and Tjian, R. (2000). Orchestrated response: a symphony of transcription factors for gene control. Genes Dev. *14*, 2551–2569.

Leuther, K.K., Salmeron, J.M., and Johnston, S.A. (1993). Genetic evidence that an activation domain of GAL4 does not require acidity and may form a β sheet. Cell *72*, 575–585.

Ma, J., and Ptashne, M. (1987). A new class of yeast transcriptional activators. Cell *51*, 113–119.

Manley, J.L., Fire, A., Cano, A., Sharp, P.A., and Gefter, M.L. (1980). DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. Proc. Natl. Acad. Sci. USA *77*, 3855–3859.

McKenna, N.J., and O'Malley, B.W. (2002). Combinatorial control of gene expression by nuclear receptors and coregulators. Cell *108*, 465–474.

McKnight, S.L., and Kingsbury, R. (1982). Transcriptional control signals of a eukaryotic protein-coding gene. Science *217*, 316–324.

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

Miesfeld, R., Okret, S., Wikström, A.-C., Wrange, Ö., Gustafsson, J.-Å., and Yamamoto, K.R. (1984). Characterization of a steroid hormone receptor gene and mRNA in wild-type and mutant cells. Nature *312*, 779–781.

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

Mitchell, P.J., Wang, C., and Tjian, R. (1987). Positive and negative regulation of transcription in vitro: enhancer-binding protein AP-2 is inhibited by SV40 T antigen. Cell *50*, 847–861.

Miyamoto, N.G., Moncollin, V., Egly, J.-M., and Chambon, P. (1985).

Specific interaction between a transcription factor and the upstream element of the adenovirus-2 major late promoter. EMBO J. *4*, 3563–3570.

Montminy, M.R., and Bilezikjian, L.M. (1987). Binding of a nuclear protein to the cyclic-AMP response element of the somatostatin gene. Nature *328*, 175–178.

Mulligan, R.C., and Berg, P. (1980). Expression of a bacterial gene in mammalian cells. Science *209*, 1422–1427.

Muratani, M., and Tansey, W.P. (2003). How the ubiquitin-proteasome system controls transcription. Nat. Rev. Mol. Cell Biol. *4*, 192–201.

Murre, C., McCaw, P.S., and Baltimore, D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. Cell *56*, 777–783.

Myers, R.M., Rio, D.C., Robbins, A.K., and Tjian, R. (1981). SV40 gene expression is modulated by the cooperative binding of T antigen to DNA. Cell *25*, 373–384.

Narlikar, G.J., Fan, H.-Y., and Kingston, R.E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. Cell *108*, 475–487.

Nüsslein-Volhard, C., and Wieschaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. Nature *287*, 795–801.

O'Hare, P., and Goding, C.R. (1988). Herpes simplex virus regulatory elements and the immunoglobulin octamer domain bind a common factor and are both targets for virion transactivation. Cell *52*, 435–445.

Oliphant, A.R., Brandl, C.J., and Struhl, K. (1989). Defining the sequence specificity of DNA-binding proteins by selecting binding sites from random-sequence oligonucleotides: Analysis of yeast GCN4 protein. Mol. Cell. Biol. *9*, 2944–2949.

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

Paranjape, S.M., Kamakaka, R.T., and Kadonaga, J.T. (1994). Role of chromatin structure in the regulation of transcription by RNA polymerase II. Annu. Rev. Biochem. *63*, 265–297.

Parker, C.S., and Topol, J. (1984a). A *Drosophila* RNA polymerase II transcription factor contains a promoter-region-specific DNA-binding activity. Cell *36*, 357–369.

Parker, C.S., and Topol, J. (1984b). A *Drosophila* RNA polymerase II transcription factor binds to the regulatory site of an *hsp70* gene. Cell *37*, 273–283.

Payvar, F., Wrange, Ö., Carlstedt-Duke, J., Okret, S., Gustafsson, J.-Å., and Yamamoto, K.R. (1981). Purified glucocorticoid receptors bind selectively in vitro to a cloned DNA fragment whose transcription is regulated by glucocorticoids in vivo. Proc. Natl. Acad. Sci. USA *78*, 6628–6632.

Payvar, F., DeFranco, D., Firestone, G.L., Edgar, B., Wrange, Ö., Okret, S., Gustafsson, J.-Å., and Yamamoto, K.R. (1983). Sequence-specific binding of glucocorticoid receptor to MTV DNA at sites within and upstream of the transcribed region. Cell *35*, 381–392.

Pelham, H.R.B. (1982). A regulatory upstream promoter element in the Drosophila hsp 70 heat-shock gene. Cell *30*, 517–528.

Picard, D., Salser, S.J., and Yamamoto, K.R. (1988). A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor. Cell *54*, 1073–1080.

Preston, C.M., Frame, M.C., and Campbell, M.E. (1988). A complex formed between cell components and an HSV structural polypeptide binds to a viral immediate early gene regulatory DNA sequence. Cell *52*, 425–434.

Ptashne, M. (1988). How eukaryotic transcriptional activators work. Nature *335*, 683–689.

Ptashne, M., and Gann, A. (1997). Transcriptional activation by recruitment. Nature *386*, 569–577.

Rio, D., Robbins, A., Myers, R., and Tjian, R. (1980). Regulation of simian virus 40 early transcription in vitro by a purified tumor antigen. Proc. Natl. Acad. Sci. USA *77*, 5706–5710.

Rosenfeld, P.J., and Kelly, T.J. (1986). Purification of nuclear factor I by DNA recognition site affinity chromatography. J. Biol. Chem. *261*, 1398–1408.

Rougvie, A.E., and Lis, J.T. (1988). The RNA polymerase II molecule at the 5′ end of the uninduced *hsp70* gene of *D. melanogaster* is transcriptionally engaged. Cell *54*, 795–804.

Sadowski, I., Ma, J., Triezenberg, S., and Ptashne, M. (1988). GAL4–VP16 is an unusually potent transcriptional activator. Nature *335*, 563–564.

Sakonju, S., Brown, D.D., Engelke, D., Ng, S.Y., Shastry, B.S., and Roeder, R.G. (1981). The binding of a transcription factor to deletion mutants of a 5S ribosomal RNA gene. Cell *23*, 665–669.

Salghetti, S.E., Caudy, A.A., Chenoweth, J.G., and Tansey, W.P. (2001). Regulation of transcriptional activation domain function by ubiquitin. Science *293*, 1651–1653.

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.

Sawadogo, M., and Roeder, R.G. (1985). Interaction of a gene-specific transcription factor with the adenovirus major late promoter upstream of the TATA box region. Cell *43*, 165–175.

Scheidereit, C., Geisse, S., Westphal, H.M., and Beato, M. (1983). The glucocorticoid receptor binds to defined nucleotide sequences near the promoter of mouse mammary tumour virus. Nature *304*, 749–752.

Scheidereit, C., Heguy, A., and Roeder, R.G. (1987). Identification and purification of a human lymphoid-specific octamer-binding protein (OTF-2) that activates transcription of an immunoglobulin promoter in vitro. Cell *51*, 783–793.

Serfling, E., Jasin, M., and Schaffner, W. (1985). Enhancers and eukaryotic gene transcription. Trends Genet. *1*, 224–230.

Sigler, P.B. (1988). Transcriptional activation: acid blobs and negative noodles. Nature *333*, 210–212.

Singh, H., Sen, R., Baltimore, D., and Sharp, P.A. (1986). A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. Nature *319*, 154–158.

Singh, H., LeBowitz, J.H., Baldwin, A.S., Jr., and Sharp, P.A. (1988). Molecular cloning of an enhancer binding protein: isolation by screening of an expression library with a recognition site DNA. Cell *52*, 415–423.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Struhl, K. (1981). Deletion mapping a eukaryotic promoter. Proc. Natl. Acad. Sci. USA *78*, 4461–4465.

Struhl, K. (1987). Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. Cell *49*, 295–297.

Sturm, R., Baumruker, T., Franza, B.R., Jr., and Herr, W. (1987). A 100-kD HeLa cell octamer binding protein (OBP100) interacts differently with two separate octamer-related sequences within the SV40 enhancer. Genes Dev. *1*, 1147–1160.

Tjian, R. (1978). The binding site on SV40 DNA for a T antigen-related protein. Cell *13*, 165–179.

Treisman, R. (1987). Identification and purification of a polypeptide that binds to the c-fos serum response element. EMBO J. *6*, 2711–2717.

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

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

Van Hoy, M., Leuther, K.K., Kodadek, T., and Johnston, S.A. (1993). The acidic activation domains of the GCN4 and GAL4 proteins are not α helical but form β sheets. Cell *72*, 587–594.

Walter, P., Green, S., Greene, G., Krust, A., Bornert, J.-M., Jeltsch, J.-M., Staub, A., Jensen, E., Scrace, G., Waterfield, M., and Chambon, P. (1985). Cloning of the human estrogen receptor cDNA. Proc. Natl. Acad. Sci. USA *82*, 7889–7893.

Walter, J., Dever, C.A., and Biggin, M.D. (1994). Two homeo domain proteins bind with similar specificity to a wide range of DNA sites in *Drosophila* embryos. Genes Dev. *8*, 1678–1692.

Wasylyk, B., Kedinger, C., Corden, J., Brison, O., and Chambon, P. (1980). Specific in vitro initiation of transcription on conalbumin and ovalbumin genes and comparison with adenovirus-2 early and late genes. Nature *285*, 367–373.

Weil, P.A., Luse, D.S., Segall, J., and Roeder, R.G. (1979). Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. Cell *18*, 469–484.

Weinberger, C., Hollenberg, S.M., Ong, E.S., Harmon, J.M., Brower, S.T., Cidlowski, J., Thompson, E.B., Rosenfeld, M.G., and Evans, R.M. (1985). Identification of human glucocorticoid receptor complementary DNA clones by epitope selection. Science *228*, 740–742.

Weintraub, H. (1988). Formation of stable transcription complexes as assayed by analysis of individual templates. Proc. Natl. Acad. Sci. USA *85*, 5819–5823.

Wolffe, A.P. (1992). Chromatin: Structure and Function (San Diego, CA: Academic Press).

Wu, C., Wilson, S., Walker, B., Dawid, I., Paisley, T., Zimarino, V., and Ueda, H. (1987). Purification and properties of *Drosophila* heat shock activator protein. Science *238*, 1247–1253.

Wu, G.J. (1978). Adenovirus DNA-directed transcription of 5.5S RNA in vitro. Proc. Natl. Acad. Sci. USA *75*, 2175–2179.

Yamamoto, K.R. (1985). Steroid receptor regulated transcription of specific genes and gene networks. Annu. Rev. Genet. *19*, 209–252.

Yang, X., Zhang, F., and Kudlow, J.E. (2002a). Recruitment of *O*-GlcNAc transferase to promoters by corepressor mSin3A: coupling protein *O*-GlcNAcylation to transcriptional repression. Cell *110*, 69–80.

Yang, A., Kaghad, M., Caput, D., and McKeon, F. (2002b). On the shoulders of giants: p63, p73 and the rise of p53. Trends Genet. *18*, 90–95.

Yankulov, K., Blau, J., Purton, T., Roberts, S., and Bentley, D.L. (1994). Transcriptional elongation by RNA polymerase II is stimulated by transactivators. Cell *77*, 749–759.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev. *15*, 2343–2360.