REVIEW

Trans-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses

Nicholas C. Jones,[1] Peter W. J. Rigby,[2] and Edward B. Ziff[3]

[1]Laboratory of Gene Regulation, Imperial Cancer Research Fund, London WC2A 3PX, UK; [2]Laboratory of Eukaryotic Molecular Genetics, National Institute for Medical Research, London NW7 1AA, UK; [3]Department of Biochemistry, New York University Medical Center, New York, New York 10016, USA

Our knowledge of the mechanisms that regulate transcription in higher eukaryotic cells has increased enormously during the past 2 years. Earlier studies, using a combination of in vitro mutagenesis and DNA-mediated gene transfer, identified two distinct types of cis-acting regulatory sequences: promoters, which are located close to the initiation site and act in a position-dependent fashion, and enhancers, which can be located far from the initiation site and act in a position- and orientation-independent fashion. Promoters can be subdivided into proximal elements, including the cap site itself and the TATA box, which is involved in fixing the site of initiation, and distal elements, which can be spread over several hundred base pairs. It is now clear that many promoters, particularly those of 'housekeeping' genes, lack TATA boxes and are instead composed of GC-rich elements that are often located within methylation-free islands (Bird 1986). Transcription controlled by this latter class of promoters often initiates at multiple sites. Many enhancer elements, for example, those of the immunoglobulin, insulin, and elastase genes, impose tissue-specific expression on adjacent promoters.

These cis-acting elements operate by interacting with protein factors, many of which have now been identified by gel retardation and footprinting assays and some of which have been purified to homogeneity. In a few cases the corresponding genes have been cloned. In many of these studies well-characterized cis-acting elements of viruses, particularly of the DNA tumor viruses, have played a major role. At the recent ICRF-sponsored meeting on the Molecular Biology of SV40, Polyoma, and Adenoviruses (held in Cambridge, England, July 27–August 1, 1987) there was a major emphasis on such trans-acting factors. The results reported both identified new factors and considerably clarified the relationships amongst previously described factors.

In this review we will discuss data reported at the meeting and, where appropriate, refer to recently published work. We will concentrate almost exclusively on work with DNA tumor viruses, referring to other systems only when there are direct connections to our major topic. Table 1 lists the factors of relevance, giving the DNA sequences they recognize, their synonyms, and the regulatory systems in which they are known to be involved.

SV40 enhancer

The SV40 enhancer, the first such element to be described, remains the best characterized and is thus often used as the prototype (for reviews and references, see Serfling et al. 1985; Maniatis et al. 1987). It is active in a wide range of cell types and over the last few years has been studied extensively by mutagenesis techniques; the sequences important for full activity have been determined. These studies have defined short DNA segments, or motifs, the integrity of which is required for activity and which have subsequently been shown to bind specific trans-acting nuclear factors. These motifs reside in functional domains or modules which by themselves have no or weak enhancing activity but which act synergistically to give high levels of activity. Zenke et al. (1986) have identified two such functional domains. Another study, involving the characterization of viral revertants of mutants containing mutations within the enhancer sequence, identified three separate domains A, B, and C which are 21, 22, and 15 bp, respectively (Herr and Clarke 1986; Clarke and Herr 1987). The location of these domains is shown in Figure 1.

The enhancer has a modular arrangement, each module containing the binding site for a transcription factor(s). Different modules can replace each other functionally such that the removal of one module can be compensated for by the duplication of another (Herr and Clarke 1986). Individual modules are found associated with other viral and cellular enhancers and promoters. The host range of enhancer activity may depend on the particular modules present, since the corresponding transcription factors may be tissue specific.

Table 1.

Factor	Binding site	Size	Promoter/enhancer
AP1 PEA1	T(T/G)AGTCA	47 kD	SV40/Py enhancers BLE of hMTIIA collagenase stromolysin α_1-anti-trypsin transthyretin MHC-H2[d] AdE3
CREB ATF	(T/G)(T/A)CGTCA	43 kD	Somatostatin (CREB) E1A/E2A/E3/E4 Ad early genes (ATF), c-*fos*, hsp70, tyrosine hydroxylase, α-gonadotropin, VIP, fibronectin, HTLV-II LTR, HTLV-I LTR, BLV LTR
MLTF/USF	GGCCACGTGACC	46 kD	MLP of adenovirus α-fibrinogen mouse Mt1
CTF/NFI	TGGCT(N_3)AGCCAA	52–66 kD	α-, β-globin, hsp70, HSV *tk*, *ras*, Ad2/5 origin AdE3 c-*myc*, albumin
SP1	GGGCGG	105, 95 kD	SV40 early promoter, hMTIIA human ADA, type II procollagen, E1B, HSVIE-3, DHFR, HIV LTR AdITR

Octamer
binding
proteins:
Ubiquitous
  OTF1
  OBP100
  NF111/octB1A
  1gNFA1
  octB3
B-cell specific
  OTF11
  1gNFA2/octB2
  octB1B
	ATTTGCAT	
		
		90 kD
		100 kD
		
		62–58.5 kD
			heavy/light Ig
			histone H2B
			snRNA genes
			SV40 enhancer

E2F	TTTCGCGC	54 kD	adenovirus E2A E1A enhancer
AP3	GGGTGTGGAAAG[a]	?	SV40 enhancer Py enhancer
EBP20	TGTGG(A/T)(A/T)(A/T)G CCAAT		MSV enhancer SV40 enhancer Py enhancer
TFIID	TATA box	?	many genes

Comments	References
AP1 and PEA1 are possibly the same factor (see text). PEA1 activity is low in undifferentiated cells. Binding site is the TPA-responsive element (TRE). AP1 or a very close relative is encoded by the c-*jun* proto-oncogene (see text).	Lee et al. (1987a); Lee et al. (1987b); Angel et al. (1987); Piette and Yaniv (1987); Bohmann et al. (1987)
ATF and CREB are probably the same factor—they are the same size and have identical binding sites; binding site is the cAMP responsive element (CRE).	Montminy et al. (1986); Montminy and Bilezikjian (1987); Lee et al. (1987d); Hurst and Jones (1987)
	Carthew et al. (1985); Sawadago and Roeder (1985); Chodosh et al. (1986); Chodosh et al. (1987); Carthew et al. (1987)
Probably a family of factors (see text). Required for transcriptional stimulation and stimulation of adenovirus DNA replication in vitro. Gene has been cloned and recognizes multiple mRNA species (Tjian, pers. comm.) A half-site is sufficient for binding.	Jones et al. (1987); Lichsteiner et al. (1987)
Human Sp1 cDNA cloned	Briggs et al. (1986); Kadonaga et al. (1987)
The relationship of these factors described by different laboratories is not yet known. Wang et al. (1986) have also described a 74-kD ubiquitous protein that binds to the octamer sequence. Rosales et al. (1987) suggest that NFA1 is a combination of oct B and oct B1B.	Sen and Baltimore (1986); Singh et al. (1986); Staudt et al. (1986); Pruijn et al. (1986); Sive et al. (1987); Margin et al. (1987); Sturm et al. (1987); Rosales et al. (1987); Rosenfield et al. (1987); Fletcher et al. (1987); Scheidereit et al. (1987)
Binding activity detected in infected but not uninfected HeLa cells; also detected in undifferentiated F9 cells.	Kovesdi et al. (1986); Reichel et al. (1987); SivaRaman and Thimmappaya (1987)
Overlaps the core motif of SV40. Induced by TPA.	Tjian et al. (pers. comm.); Chiu et al. (1987)
This protein originally purified by virtue of its ability to bind to the SV40 enhancer, also binds to the CCAAT sequence. It has recently been demonstrated that both EBP20 (enhancer binding activity) and CBP (CCAAT binding activity) reside on one polypeptide encoded by a single gene (S. McKnight, pers. comm.).	Johnson et al. (1987)
Factors binding to the TATA box region not yet purified.	

Table 1—Continued

Factor	Binding site	Size	Promoter/enhancer
AP2	CCCCAGGC	52 kD	SV40 promoter/enhancer Py enhancer/origin pre-proenkephalin collagenase, mouse H2K^b Ad MLP, human hsp70 hMT11A
AP4	CAGCTGTGG	?	SV40 enhancer Py enhancer
GT11-1B			pre-proenkephalin
AP5	CTGTGGAATG	?	SV40 enhancer
EF.E GTII-C			F441 Py enhancer
PEA2	GACCGCA	?	Py enhancer
EF.C	GTTGCN$_2$GGCAAC	?	Py enhancer hepatitis B enhancer
E2aE-Cβ E4EF2	TGGGAATT	?	E2A (E2aE-Cβ) E4 (E4F2)
E4TFI	GGAAGTG	?	E4

[a] This is the SV40 binding site and not a consensus.

GTII motif

This motif lies just upstream of the 72-bp repeat and has been shown by mutagenesis studies to be important (Herr and Clarke 1986; Zenke et al. 1986). Nuclear factors that bind to this motif have been identified and characterized in the laboratories of R. Tjian and M. Karin and were described by T. Williams (University of California, Berkeley) and R. Chiu (University of California, San Diego). The factors bind to similar sequences but appear to be different. The AP5 factor described by Chiu protects the sequence 5′-TGTGGAATGT-3′ from DNase I digestion. Williams reported that some of these same sequences are protected by AP4, but the extent of protection is extended at the 5′ end to include the *Pvu*II restriction site (Fig. 1); consequently, restricting SV40 DNA with *Pvu*II abolishes AP4 binding. Recent evidence from Chambon's laboratory (Xiao et al. 1987) suggests that this region is even more complex; four different factors capable of binding to sequences that contain part or all of the GTII motif have been identified. The factors are distinguished from each other by their cell specificity, their exact binding site as defined by DNase I protection, methylation protection and site-directed mutagenesis, and by their chromatographic behavior. One factor, GT-IIC, is present in HeLa but not lymphoid cells and is likely to be the same as or very similar to AP5. This factor appears to be the most important for enhancer activity in HeLa cells. Mutants pA4, pA5, and pA6 contain alterations that severely affect GT-IIC binding and, as shown in Figure 1, these mutations affect enhancer activity significantly. In contrast, pA2 and pA3, which alter the *Pvu*II restriction site,

have very little effect. There are two other points of interest with respect to the proteins that interact with this motif. The first is that mutations in this region have very little effect on enhancer activity in lymphoid cells, in contrast to the situation in Hela cells (Nomiyama et al. 1987). This would be consistent with the importance of the GT-IIC protein which is not found in lymphoid cells. Second, this factor can bind to a mutant Py enhancer that, unlike its wild-type counterpart, can function efficiently in undifferentiated F9 cells (see below).

GTI motif

This motif is highly homologous to GTII (10/12 bases) and yet the nuclear factors that bind to it are very different. Williams described the purification of a HeLa nuclear factor, AP3, which binds to this motif but not to GTII (Chiu et al. 1987; Mitchell et al. 1987). AP3 appears to be important for enhancer function in HeLa cells since mutations that prevent its binding, e.g., pA12 (Fig. 1), significantly decrease activity.

A factor has been isolated from rat liver that also binds to this motif (Johnson et al. 1987). Whether AP3 is the human homolog of EBP20 is not known at present. However, some relationship between the two is indicated by the finding that antibodies raised against EBP20 show some cross-reactivity with partially pure preparations of AP3 (S. McKnight, pers. comm.; R. Tjian, pers. comm.).

TCI, TCII motifs

Williams discussed a HeLa cell factor, AP2, that binds to

Comments	References
Overlaps the core motif of SV40. AP2 and AP3 together cover the SV40 enhancer core domain. Binds to SV40 TAg. Induced by TPA and cAMP.	Mitchell et al. (1987); Imagawa et al. (1987)
Exact relationship of these two proteins not known.	Tjian et al. (pers. comm.); Xiao et al. (1987)
Exact relationship not known.	Kovesdi et al. (1987); Xiao et al. (1987); Karin et al. (pers. comm.)
Low in undifferentiated F9 cells. Not induced by TPA.	Piette and Yaniv (1987)
	Ostapchuk et al. (1986)
	Jalinot (pers. comm.); Lee (pers. comm.)
	Watanabe (pers. comm.)

the TC motifs (Imagawa et al. 1987; Mitchell et al. 1987). The mutant pA15 has an altered TCI site and, as a result binding of AP2 to TCI, is lost; activity of this mutant enhancer is approximately 2.5-fold lower in HeLa cells. Mutations in TCII have very little effect. Together the AP3 and AP2 binding sites constitute the functional enhancer core domain (C element) described by Herr and Gluzman (1985). The AP2 protein also binds to the 21-bp repeat region of the early promoter but with a lower affinity (Mitchell et al. 1987). Multimers of the core domain, when inserted downstream from a β-globin test gene, can act as an enhancer element with broad cell-type specificity (Ondek et al. 1987; Schirm et al. 1987). Interestingly, it has been shown recently that AP2 can complex with the SV40 large T-antigen protein and, indeed, in the presence of an excess of T antigen, binding of AP2 to the TC motifs is inhibited (Mitchell et al. 1987). In contrast, binding of other factors, such as Sp1, to the 21-bp repeat region, is unaffected. T antigen represses SV40 early gene expression, and therefore it is possible that this repression is at least partially due to its ability to prevent binding of AP2.

Sph/octamer motifs

The enhancer contains a direct repeat of the Sph motif that consists of the 9-bp sequence 5'-AAG(T/C)ATGCA-3'. The motifs appear to be essential for full enhancer activity in HeLa cells (Fig. 1) and bind factors from HeLa cell extracts, the nature of which has not been described fully. The junction of the direct repeat creates a second motif that is very similar to the octamer

motif found in immunoglobulin promoter and enhancer elements as well as the promoters of the snRNA and H2B histone genes. This motif has been shown to bind a variety of factors, some of which are ubiquitous and some specific to lymphoid cells. A HeLa cell nuclear factor of molecular weight 100 kD (OBP100), described by Sturm (Cold Spring Harbor Laboratory, laboratory of W. Herr), binds to the SV40 octanucleotide sequence as well as to a previously unrecognized octanucleotide-related sequence that is immediately adjacent to the classical octanucleotide (Sturm et al. 1987). Interference assays, using modification by both dimethyl sulfate and diethyl pyrocarbonate, imply that the protein makes quite distinct contacts with its two binding sequences. The ability of a single protein to recognize two distinct DNA sequences previously has been documented for the yeast activator protein HAP1 (Pfeifer et al. 1987). These observations suggest that some transcription factors exhibit remarkable flexibility in their DNA sequence recognition. It will not be surprising if OBP100 is identical, or closely related, to the ubiquitous octamer binding protein, which has been designated OTF1 (see Table 1); this, in turn, may be related to the factor NF-III, which is required for adenovirus DNA replication. The octamer sequence in the immunoglobulin control elements contributes to the tissue-specific activity of these genes. Similarly, it appears that the SV40 octamer is important for enhancer activity in lymphoid cells but not in HeLa cells. The mutant p24 (Fig. 1) causes only a slight decrease in enhancer activity in HeLa cells but a 75% decrease in lymphoid cells (Davidson et al. 1986). The complexity of proteins that can bind to this motif is
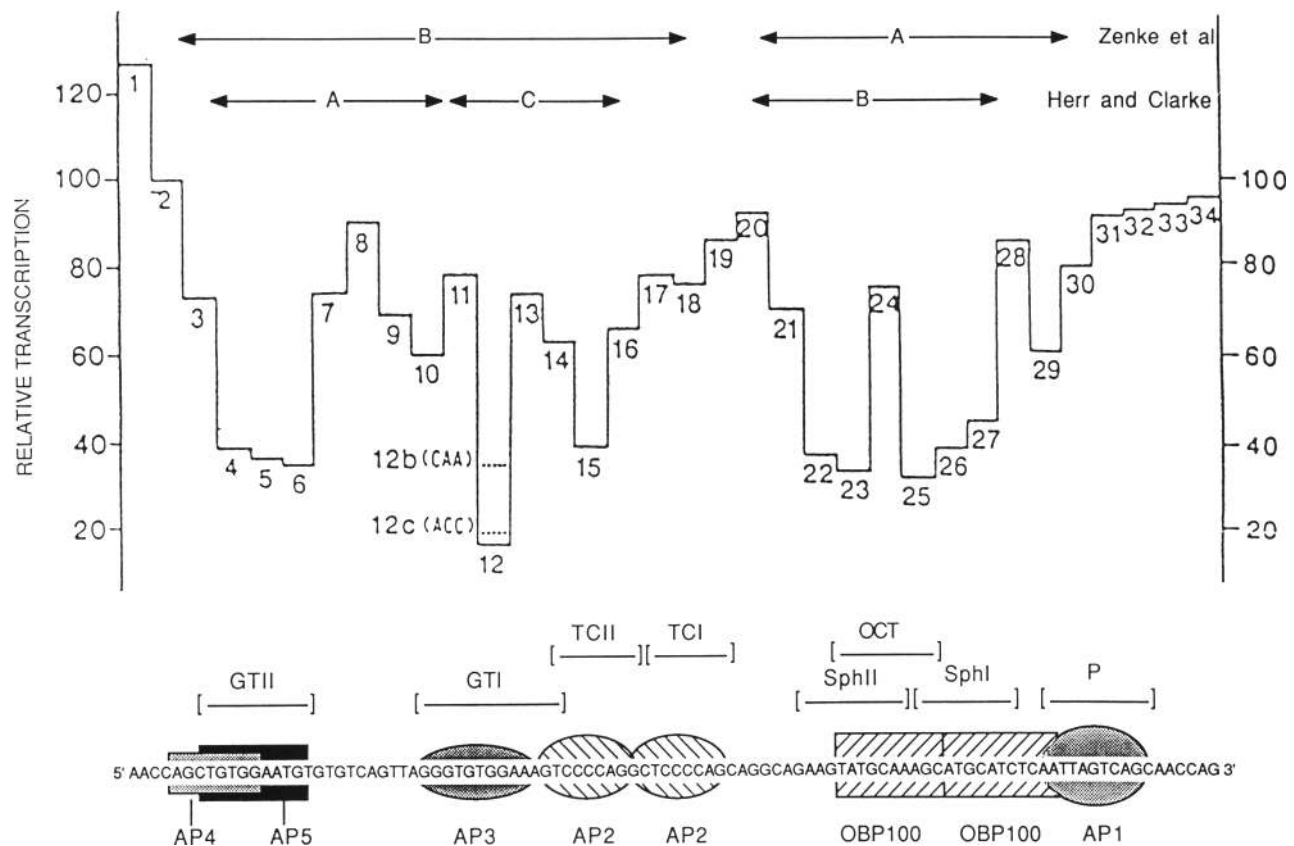
Figure 1. The sequence of the enhancer region of wild-type SV40 containing the distal 72-bp element and its 5'-flanking region is shown. Also shown are the location of sequence motifs that have been identified in the enhancer region and are organized into domains A and B (Zenke et al. 1986) or A, B, and C (Herr and Clarke 1986), the multimerization of which generates enhancer activity. The enhancer–protein interactions of nuclear factors from HeLa cells are indicated diagramatically and explained in detail in the text. The top part of the figure shows the effect of point mutations within the enhancer region on transcription from the EES of the SV40 early promoter; these data are taken from the work of Zenke et al. (1986). The results are expressed relative to the wild-type enhancer region, taken as 100%. The relative transcriptional activity of mutants pA1-pA34 following transient expression in HeLa cells is indicated.

reinforced by the recent work of Rosales et al. (1987) who describe four different factors that bind to this sequence.

P motif

This motif of the enhancer binds the transcription factor AP1, which has been purified by Angel et al. (1987) and Lee et al. (1987b). Mutations that alter this motif have only a modest affect on enhancer activity in both HeLa and lymphoid cells. However, it has recently been shown that the AP1 binding site can confer inducibility by phorbol esters (Angel et al. 1987; Lee et al. 1987b). A number of genes that are induced by TPA contain AP1 sites within their promoter regions and an oligonucleotide containing this site can confer inducibility to a heterologous promoter. It appears, therefore, that AP1 activity can be modulated by TPA-induced cellular changes, possibly related to TPA's role in the activation of protein kinase C. Not all TPA-inducible promoters contain AP1 binding sites, which suggests that the activity of other factors can be affected similarly. Indeed,

Chiu reported that the binding sites for the factors AP2, AP3, and AP5 can also confer TPA responsiveness (Chiu et al. 1987; Imagawa et al. 1987).

Polyoma enhancer

This enhancer was first localized to a 246-bp fragment extending from the *BclI* site at nucleotide position 5021 to the *PvuII* site at position 5262 (Fig. 2). Subsequently, it was shown to consist of a number of domains that can replace each other functionally and interact with nuclear proteins from a variety of cell types. The protein binding sites found in some of the domains resemble those found in other enhancer and promoter elements.

A number of reports at the meeting described the factors that interact with the various domain sequences. Much of this work has centered on the A domain, a 22-bp sequence that, when oligomerized, can function as an independent enhancer element (Veldman et al. 1985). At least three different nuclear proteins have been found to bind to this element (Fig. 2), two of which were first described and have since been characterized extensively

in the laboratory of M. Yaniv (Institut Pasteur, Paris). These factors, which bind to immediately adjacent sites, have been termed PEA1 and PEA2. From the nature of the binding sequence and the characteristics of the factor, PEA1 is almost certainly closely related to AP1 and may be the murine equivalent. Both Hirai (Paris) and Satake (Kyoto University), from the laboratory of Y. Ito, reported that the activity of PEA1 was induced by TPA but that it was undetectable in undifferentiated F9 cells (Kryszke et al. 1987). Upon differentiation of these cells into parietal endoderm, PEA1 activity appeared, suggesting the interesting possibility that this factor is important in the control of gene expression during early embryonic development. PEA1 is certainly important for in vivo enhancer activity, since a mutant, B1, described by Martin (University of Texas, Austin, W. Folk's laboratory), which is defective for viral replication and transcription, contains several point mutations, one of which, at position 5115, destroys binding of PEA1. Reversion of this mutation not only restores PEA1 binding but also enhancer function. The PEA1 factor, which binds to a very similar sequence but which has been shown by competition experiments to be distinct, was also reported by Hirai and Satake to be virtually undetectable in undifferentiated F9 cells. Williams reported that AP4 also binds in this region over the *PvuII* restriction site.

A factor that binds to the adjacent C element, called EF-C, has been characterized in P. Hearing's laboratory and was described by Ostapchuk (State University of New York, Stony Brook). This factor is present in a wide range of cells (Ostapchuk et al. 1986) and the region of DNase I protection conferred by binding corresponds almost exactly to the C element defined genetically by Veldman et al. (1985). Integrity of this region is important for enhancer function in vivo. A sequence homologous to the EF-C binding site is found in a region of the hepatitis B virus genome shown to behave as an enhancer element in liver cells and to bind a specific nuclear factor. Binding of EF-C to the polyoma enhancer was efficiently competed by the hepatitis virus enhancer sequences.

The B enhancer element contains a sequence similar to the SV40 core sequence described earlier. In SV40 this sequence was found to bind AP3 and AP2, and Williams reported that these same two factors specifically bound over the core homology present in the B element.

The polyoma enhancer functions very poorly in undifferentiated cells. This can be explained, at least partially, by the absence of the factors that interact with the A element (PEA1[AP1], PEA2). Mutant viruses have been isolated that replicate in such undifferentiated cells, and these viruses were found to contain altered enhancer elements whose activity did not depend upon the differentiated state of the cell. The mutant enhancers were of two types: those that have duplicated sequences (e.g., EC PCC4 mutants that have duplicated A domains) and those that have single point mutations (e.g., EC F9 mutants).

The polyoma virus mutant F441 contains an A→G mutation at nucleotide 5230 which enables the enhancer to function efficiently in a variety of undifferentiated EC cells. It has now been shown that this single change creates binding sites for two different transcription factors: Williams reported that pure NFI/CTF can bind to this region of the mutant much more efficiently than to the wild-type enhancer and Ostapchuk reported the acquisition of a binding site for the factor they term
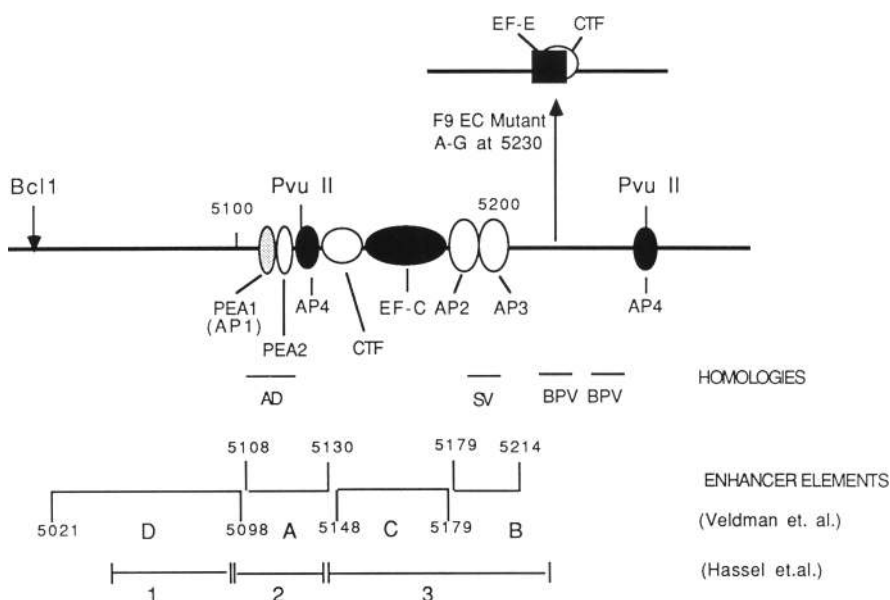

Figure 2. Schematic of the polyoma enhancer region. The DNA–protein interactions within this region are summarized and described in detail in the text. Homologies to other viral enhancer regions are indicated: (Ad) adenovirus E1A enhancer homology (Hearing and Shenk 1983); (SV) SV40 core enhancer element homology (Weiher et al. 1983); (BPV) bovine papilloma virus enhancer homology (Lusky et al. 1983). Also shown are the positions of the functional elements of the enhancer as defined by Veldman et al. (1985) and Hassel et al. (1986). At the top of the figure, the protein–DNA interactions specific to the F9.EC mutant, F441, capable of efficient replication in undifferentiated F9 cells and containing a nucleotide substitution at nucleotide 5230, are indicated.

EF-E. Other recent reports also describe the binding of this latter factor to the mutant F9 enhancer (Kovesdi et al. 1987; Xiao et al. 1987). The binding site for EF-E is identical to the GTII motif of the SV40 enhancer; therefore, this factor is likely to be identical or related to AP5.

Adenovirus early genes

Adenovirus E1A gene products stimulate transcription from a number of viral and cellular genes, including the early adenovirus genes E1B, E2A, E3, and E4 (for review, see Berk 1986). The mechanism by which E1A protein transactivates remains unclear. It is very unlikely to act directly since purified E1A protein does not bind specifically to DNA. In addition, mutagenesis studies on the early promoters have failed to detect sequences specific for E1A stimulation. Rather, it appears that the same promoter sequences are necessary both for uninduced and E1A-induced expression. These functionally important sequences bind nuclear factors that must then mediate the E1A effect. How this is accomplished is unknown, but it could involve either an increase in the amount of a particular factor or modification of a preexisting factor. Such modification could alter the DNA binding properties of the factor or some other activity. Clearly, there is a need to characterize in detail the factors that interact with the early promoters. A great deal of effort in a number of laboratories has been devoted to this goal.

The early gene promoters have been found to interact with multiple nuclear proteins (Fig. 3). However, the pattern of interactions is very different from one promoter to the next and only one or possibly two factors have been found to interact with more than one of the early genes. A brief description of the protein–DNA interactions follows:

E1B

The E1B promoter directs rightward transcription from base pair 1699 in the Ad2 sequence. A series of deletion and linker-scanning mutations have defined two major elements critical for efficient transcription (Wu et al. 1987). One element, extending from -49 to -38, is a high-affinity binding site for the transcription factor Sp1 that binds to a number of other promoters, including the SV40 early gene promoter. The second element contains the TATA-box sequence and presumably interacts with a TATA-box binding protein(s). Interestingly, mutations in this sequence decrease uninduced expression and obliterate stimulation by E1A. Parks (Pennsylvania State University College of Medicine, laboratory of D. Spector) also reported binding of four other factors, upstream of the Sp1 site, although the removal of sequences upstream of -49 has little affect on E1B transcription.

E2A

A number of laboratories have concentrated their efforts

on the E2A gene which is transcribed from an initiation site at roughly 75 map units on the adenovirus genome. The promoter has been investigated extensively by deletion and linker-scanning mutagenesis (for references, see Berk 1986). Analysis of these mutants demonstrated that 79 bp of 5' sequence were sufficient for full activity; however, deletion to -70 severely impaired activity. Binding of a factor between positions -70 and -79 has been demonstrated. This factor, found in uninfected and infected cells, is almost certainly identical to a factor that also interacts with the E3 and E4 promoters. The factor, originally named E2A.EF, has been renamed activating transcription factor or ATF (Lee et al. 1987d). The ATF binding site (T/G)(A/T)CGTCA, is identical to a sequence found in promoters that are induced by raising cAMP levels, such as that of the somatostatin gene. Additionally, this sequence has been shown to mediate cAMP inducibility and to bind a 43-kD phosphoprotein (CREB) (Montminy and Bilezikjian 1987). The ATF protein is also a 43-kD phosphoprotein, and ATF binding can be competed efficiently by sequences from the somatostatin promoter (Hurst and Jones 1987). It seems very likely, therefore, that CREB and ATF are the same factor. Indeed, a report by Leza (State University of New York, Stony Brook, laboratory of P. Hearing) demonstrated that expression of the E4 promoter was induced by raising cAMP levels dependent, at least partly, on the ATF binding site at -50 (see below).

The mutagenesis studies also indicated that important elements were present in the region between -69 and -33. Initially identified and characterized in the laboratory of J. Nevins, this region contains two binding sites for the factor E2F, whose activity appears to increase dramatically upon adenovirus infection. Undifferentiated F9 cells were also found to contain this factor, a result consistent with the ability of these cells to transcribe from the E2A promoter in the absence of E1A. Upon differentiation this ability is lost and so too is E2F activity. An update on E2F characterization was presented by Yee (Rockefeller University) who reported that the protein has now been purified from infected cells and has a molecular weight of 54 kD. There is an additional E2F binding site further upstream in the E2A promoter and two binding sites within the E1A enhancer region; these sites were described by Jalinot (Strasbourg) and Hardy (Princeton University) from the Kedinger and Shenk laboratories, respectively. However, these workers could find no evidence of E2F inducibility and readily detected the factor in extracts from uninfected cells. The reason for this major discrepancy is unknown.

Finally, the E2A promoter contains a binding site for a third factor described by Jalinot and termed E2aE-Cβ. This factor may be related or identical to E4F2, which binds to the E4 promoter (see below).

E3

Three regions of the E3 promoter, centered approximately over the TATA box, position -55 and position -91, have been shown by mutagenesis studies to be

critical both for uninduced and E1A-induced transcription. Factors binding to each of these regions were described by Hurst (ICRF) from the laboratory of N. Jones, Weinmann (Wistar), and Garcia (University of California, Los Angeles, laboratory of R. Gaynor). The factor that binds to the sequences between -44 and -68 is ATF. The region between -81 and -103 contains a sequence that is very similar to the ATF consensus and is bound specifically by a nuclear protein. However, competition experiments described by Hurst showed that
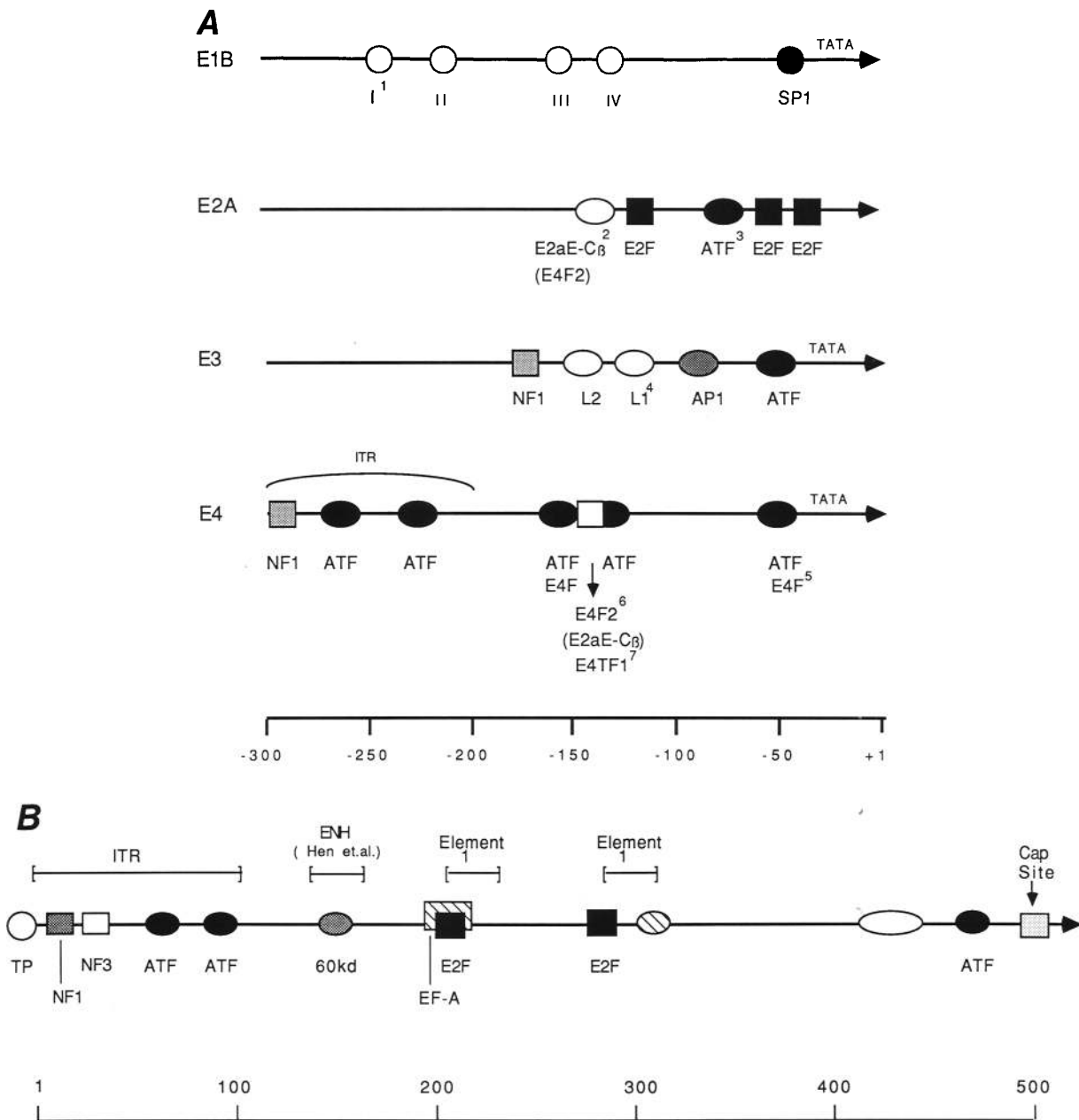

Figure 3. (A) Summary of the DNA–protein interactions within the upstream promoter region of the adenovirus early genes. The position of the binding sites and names of the factors are indicated. The arrowhead denotes the direction of transcription. *Footnotes*: [1]Footprints protected by factors I–IV are as follows: I is TGGGCGTCGC, which may bind Sp1; II is GCTGTGGAATGT, which is identical to the GTII motif of the SV40 enhancer; III is CCCCAGGC, which may bind AP2; IV is TGTGGTTA. [2]This factor contains the same binding site as the E4F2 factor that binds at about -150 on the E4 promoter. [3]The factors E2A-EF (SivaRaman et al. 1986), E3.F2 (Hurst and Jones 1987), and E4F1 (Lee et al. 1987c) have been renamed ATF (Lee et al. 1987d). The conclusion that these factors are probably identical is based on identical binding sites, size, and chromatographic behavior. [4]L1 and L2 are lymphoid cell specific. [5]E4F binds the same sequences as ATF, but its activity is induced upon infection and it is only seen in whole-cell extracts. [6]See note 2. [7]Binds in the same regions as E4F2 but the binding site is different. (B) Summary of the DNA–protein interactions within the upstream region of the E1A gene. The factors that bind to sequences centered at around positions 310 and 440 are referred to as γ and delta in the text.

this protein is not ATF but more likely AP1 (Hurst and Jones 1987). The consensus for AP1 is T(G/T)AGTCA(G/C), the difference between this and the ATF consensus being an A residue instead of C at the third position. The findings illustrate how such a minor difference can change the pattern of factor binding. Further upstream, between positions −154 and −183, there is a binding site for NF1/CTF, a factor that acts not only on transcription but is also required for efficient initiation of adenovirus DNA replication. Removal of this site, however, has only a minor detrimental effect on E3 transcription, whereas removal of the AP1 or ATF sites has significant effect. Garcia also reported two other factors that bound to specific sequences in the E3 promoter (L1 and L2) but which are lymphoid cell specific. These factors also interact with the LTR sequences of the HIV genome.

E4

A number of talks were presented on the factors that interact with this promoter, and the picture emerging is very complex. Within the promoter region there are three binding sites for the ATF factor, one between positions −45 and −50 and the other two in a region between −140 and −170. This latter region can act as an E1A-inducible enhancer element. Two additional factors were reported to bind to the −50 region. Reinberg (University of Medicine and Dentistry of New Jersey) described a factor active in the in vitro transcription of the E4 promoter that has been purified to near homogeneity and has a molecular weight of 94–107 kD. The relationship of this factor to ATF is not clear. Raychaudhuri (Rockefeller University, laboratory of J. Nevins) described a factor that appears to be distinct from ATF by virtue of its methylation interference pattern, its chromatographic behavior, and competition studies. Interestingly, this factor, E4F, is only readily detected in whole-cell extracts and its activity is increased 10- to 20-fold by infection with wild-type virus. E4F also appears to bind to the E4 enhancer region which is very complex since two other factors were reported to bind here, in addition to ATF and E4F. E4F2 binding was reported by Lee (Harvard University, M. Green's laboratory); the binding site is identical to that of the E2AE-Cβ factor found to interact with the E2A promoter (see above). In addition binding of yet another distinct protein, E4TFI, was reported by Watanabe (University of Tokyo, laboratory of H. Handa). So at least four different factors can bind within this small DNA sequence. Very careful mutagenesis studies will be required to unravel the importance of each individual factor to the transcriptional activity of this promoter.

E1A

The special regulatory role of the E1A products in the viral life cycle, in particular the requirement for the 289-amino-acid protein to trans-activate other early promoters, is reflected by E1A's expression from the earliest stages of infection. Indeed, the E1A promoter is active by virtue of a cis-acting, enhancer-dependent mechanism and functions in the absence of viral protein synthesis. Unlike the other adenoviral early promoters, it does not require E1A for its activity. Nevertheless, E1A promoter activity is induced to yet higher levels by the E1A trans-activation function and therefore E1A is autostimulatory.

The 5'-flanking region of the E1A gene (see Fig. 3B) extends from the left terminus of the genome (nucleotide 1) to the E1A cap site (nucleotide 499). This region contains, in addition to E1A's transcription regulatory elements, one of two copies of the adenovirus inverted terminal repeat (ITR). The ITR (nucleotides 1–102) is the viral DNA replication origin and is exactly repeated at the right end of the genome in inverted orientation.

The location of the ITR upstream from the E1A promoter juxtaposes viral replication and transcription elements. Indeed, the ITR itself contains binding sites for both bona fide DNA replication factors and proteins that serve as transcription factors for other viral and cellular genes. The DNA 5' terminal residue at the extreme left end of the genome is covalently linked to the adenovirus 'terminal protein' (TP), which serves as a primer for DNA synthesis. Just adjacent to this linkage site are a site for binding the virus-encoded DNA polymerase and a binding site for NF1/CTF. Adjacent lies a site for a third protein, nuclear factor III (NFIII), which is a member of the octamer binding family of transcription factors.

The major E1A enhancer elements lie outside the ITR. As we proceed from the ITR toward the E1A cap site, we encounter a region (nucleotides 155–178) that contains a TGTGG motif and was defined previously by deletion mutagenesis as an enhancer (Hen et al. 1983). L. Clark (St. Andrews, laboratory of R. Hay) reported purification of a 60-kD protein from uninfected HeLa cells, the binding site of which overlaps this motif. The protein also binds with high affinity to enhancer elements in SV40, polyoma, and Rous sarcoma virus. Binding to the SV40 enhancer is over the GTI motif; however, this protein is distinct from AP3 and EBP20 (R. Hay, pers. comm.). Additional E1A enhancer elements lie downstream. Hearing and Shenk (1986) have described a sequence, AGGAAAGTGA(C/A)A(A/T), called element I, which is present twice within the E1A flanking region downstream from the element identified by Hen et al. Although dispensable for replication, retention of at least one copy of element I is essential for E1A expression. Nested between the two copies of element I lie copies of a second sequence motif, element II. Both element I repeats are adjacent to binding sites for the factor E2F, which also binds to the E2A promoter region. Interestingly, as with the E2A promoter, the E2F sites that precede the E1A gene are phased helically and for E1A lie 63 bp apart. Whether this helical phasing is important for activity is not yet known. At least one factor appears to interact both with element I and with a neighboring E2F site. Bruder (State University of New York, Stony Brook, laboratory of P. Hearing) described the

factor EFA, which binds tightly to the upstream copy of element I, at a position that overlaps the adjacent (upstream) E2A binding site. A factor described by Hardy (Princeton, laboratory of T. Shenk) designated γ and detected through its bandshift activity and footprint pattern, binds to a region that overlaps (but does not coincide with) the downstream copy of element I.

Yet closer to the E1A cap site, Hardy has also described a binding site for an additional factor, δ, which thus far is also known only through its footprint pattern and ability to induce a bandshift. No phenotype has yet been described for mutations in the δ binding site. Still closer to the cap site is an additional ATF binding site. Finally, I. Kruczek (Munich) used mobility shift assays and transfection competition to demonstrate a protein that binds to an inverted terminal repeat sequence of Ad12 at residue 443 (corresponding to residues 488–506 of Ad5), which directly overlaps the Ad12 E1A cap site, and which may stimulate E1A promoter activity.

How does E1A trans-activate?

Most current theories focus on the possibility that E1A acts to increase the concentration or activity of a set of transcription factors. The best example of a factor being modulated during adenovirus infection is the factor E2F. The increase in activity of this factor could be due to an increase in the amount of the factor or the modification of a preexisting form that cannot bind efficiently to DNA. Preliminary evidence presented by Yee suggests the latter possibility. In the cases of other factors that bind to the inducible promoters but can be readily detected in uninfected cells, E1A-induced modification is also suspected; but, as a result of the modification some activity of the factor other than DNA binding may be affected. This other activity could involve interactions with other transcription factors (e.g., the TATA-box factor) or even the polymerase itself. What kind of modification could be induced? A distinct possibility is phosphorylation, suggested by observations made with the AP1 and CREB factors. AP1 is induced by TPA, which is well known to activate protein kinase C. This suggests that AP1 activation could involve direct phosphorylation by PKC or, more likely, a PKC-initiated phosphorylation cascade. The CREB protein appears to mediate the transcriptional response of some promoters to increasing cAMP levels. Recently, it was shown that in response to rising cAMP levels the phosphorylation state of the CREB factor was increased (Montminy and Bilezikjian 1987). Within the next year the materials necessary to answer such questions (cDNA sequences that encode the factors, antibodies) should become available; a better understanding of E1A action may ensue.

Another possible mode of E1A action would be a direct interaction between the E1A protein and some or all of the factors that bind to the early promoters. The bound E1A may then facilitate interactions with other proteins that may constitute the transcriptional complex. It was reported by Weinmann (Wistar) and Hurst (ICRF) that indeed some interaction between E1A and the factors AP1 and ATF could be demonstrated; the factors are removed from a nuclear extract following passage over an E1A affinity column. The significance of these interactions is very unclear at present, but future studies will be awaited with interest.

Another advance that will undoubtedly facilitate the study of E1A action is the development of an in vitro transcription system that faithfully mimicks E1A activation. Such in vitro activation was reported recently from Harter's laboratory and was described at the meeting by Datta (New Jersey Medical School, Newark). They have shown that E1A protein synthesized and purified from E. coli can increase transcription of both the E2A and major late promoters when added to HeLa cell transcription extracts (Spangler et al. 1987). This stimulation was blocked by adding antibodies specific for the E1A protein. These results suggest that E1A action does not involve the synthesis of new protein but rather the modification of preexisting proteins.

Although the details of how E1A trans-activates remains largely unknown, a great deal more progress has been made regarding the critical regions or domains of the E1A protein necessary for this activation. Extensive mutagenesis studies, in a number of different laboratories, have identified only one essential region that is predominantly encoded by the unique portion of the 13S mRNA. This region, which is 46 amino acids in size, is highly conserved amongst different adenovirus serotypes and contains an arrangement of four Cys residues similar to the repetitive sequences found in TFIIIA. These residues are thought to form a metal binding core, and Culp (Smith Kline and French, M. Rosenberg's laboratory) showed that the product of the 13S E1A mRNA but not the 12S mRNA was capable of binding zinc. Mutations that affect any one of the four Cys residues resulted in loss of binding as well as loss of trans-activation. The critical involvement of this 46-amino-acid region in trans-activation was demonstrated most clearly by a series of experiments described by Green (St. Louis University Medical Center). He showed that a synthetic 49-amino-acid peptide containing these residues can activate the E2A early gene in microinjected HeLa cells. Peptides comprised of other E1A residues failed to activate (Lillie et al. 1987).

Discussion

The results reported at the meeting illuminated a number of general principles and defined technical problems that will have to be faced in future studies. We shall now comment on these wider issues.

In early work, factors were defined generally by the DNA sequence to which they bound, because even the size of the protein was not known. However, it is now clear that in several cases multiple factors can recognize the same consensus sequence, and so this feature alone is not sufficient for factor classification. One of the clearest examples of such a factor family involves the proteins that bind to the octamer sequence described earlier; multiple factors that interact with this con-

sensus can be found, some of which are cell-type specific. Another example involves the consensus CCAAT sequence, which is critical for the transcription of numerous cellular and viral genes, e.g., the β-globin gene, herpes simplex *tk* gene, and human heat shock gene. A multiplicity of factors that recognize this consensus have been described. Nuclear factor I (NFI) was originally defined by its activity in the adenovirus in vitro DNA replication system but was subsequently shown also bind to the CCAAT consensus (Jones et al. 1987). This NFI/CTF factor was purified to apparent homogeneity by sequence-specific DNA affinity chromatography and was found to consist of a set of proteins with apparent molecular weights of 52–66 kD. The exact relationship of these proteins to each other is not entirely clear, but recently several cDNAs encoding NFI/CTF have been cloned and appear to recognize a single gene but multiple transcripts (R. Tjian, pers. comm.). The heterogeneity of polypeptides could therefore be explained by alternative splicing of a primary transcript. Gander (University of Munich, laboratory of E. Winnacker) described studies that identified several proteins possessing NFI activity, not all of which, however, could bind CCAAT. Moreover, Graves et al. (1986) described the purification of a liver protein, CBP, that binds to the CCAAT sequence of the HSV *tk* promoter; this protein is identical to EBP20 (see Table 1). Recent work has identified proteins that bind to the CCAAT boxes of major histocompatibility complex genes and the adenovirus E2L promoters (Dorn et al. 1987; Goding et al. 1987). These proteins are by many criteria distinct from NFI/CTF. It remains to be seen whether each of these polypeptides recognizes a particular variant of the consensus sequence or whether binding is controlled by sequences flanking the CCAAT motif. Moreover, the possibility that the various CCAAT binding factors are derived from heterotypic interactions between multiple subunits must also be entertained. It will also be of great interest to ascertain whether the various members of the family are subject to tissue-specific, developmental, or cell-cycle regulation. Not all proteins that bind to sequences overlapping the CCAAT motif need to be activation proteins. A protein found in sea urchin embryos binds over a CCAAT motif present in the sperm histone H2B-1 gene promoter and in so doing prevents interaction with a positive-acting CCAAT-binding factor, resulting in a block to transcription (Barberis et al. 1987). This CCAAT-displacement factor is not present in the testis, where the gene is expressed. This interesting result suggests the possibility that the CCAAT region of other vertebrate genes may also mediate negative as well as positive regulation. These examples of multiple proteins that can bind to what we presently perceive as the same sequence means that more discriminatory assays will have to be applied to distinguish between factors that may themselves be closely related. Methylation interference footprinting may be useful in this regard, particularly when combined with the UV cross-linking procedure which allows determination of the size of the binding polypeptide.

Yet another possible factor family are the proteins that interact with the TATA box. Subfractionation of in vitro transcription extracts has resulted in the description of a fraction, TFIID, that is absolutely required for transcription and contains a protein(s) that binds to the TATA-box region of the adenovirus major late promoter. So far, however, this factor has resisted all attempts at purification. Data presented at the meeting showed that two viral transactivators, E1A (Simon, Rockefeller, laboratories of J. Nevins and N. Heintz; Wu, University of California, Los Angeles, laboratory of A. Berk) and the immediate-early gene product of pseudorabies virus (Abmayr, Rockefeller, laboratory of R. Roeder) may act via the TATA-box region. Moreover, Leong (University of California, Los Angeles, A. Berk's laboratory) showed that there are higher levels of TFIID activity in extracts prepared from adenovirus-infected cells than in those prepared from uninfected cells. However, not all TATA boxes respond to *trans*-activation, and Simon defined a consensus sequence TATAA for *trans*-activator responsive TATA boxes. In contrast, the TATA box of the early control region of SV40, sequence TATTTA, does not respond. A clear implication of these data is that there are, at the least, functional subsets of TFIID and that there may be a family, or families, of proteins interacting with the TATA-like sequence motifs. The TATAA motif is found in the adenovirus early E3 promoter. However, mutagenesis studies on this promoter have demonstrated that removal of this motif decreases the efficiency of transcription but that activation by E1A can still take place. Additionally, sequences upstream of the TATAA can confer E1A inducibility to other promoters that are not activated efficiently. These results reinforce the possibility that E1A can activate via a variety of different transcription factors and, in the case of a promoter such as the E3 promoter, full activation may involve modulation of the TATA-binding factor as well as factors that bind to the upstream elements.

It is also clear that the precise assay conditions adopted can significantly affect the results obtained. The laboratories of Thimmappaya and Nevins have previously reported two distinct factors that bind to the E2A promoter. Data reported at this meeting make it clear that the apparent discrepancy is due to the exact conditions employed in the assays used. The use of synthetic polymers such as poly(dI : dC) as a nonspecific competitor favors identification of the ATF factor, whereas the use of salmon sperm DNA favors the identification of E2F (SivaRaman and Thimmappaya 1987; Yee et al. 1987). In addition, two factors binding to a similar sequence of the E4 promoter have been described (see earlier), the different results apparently arising from the use of nuclear extracts by one group and whole-cell extracts by the other.

Most assays depend on the interaction of naked DNA with either cell extracts or purified factors. However, there is now clear evidence that such assays need not reflect the in vivo situation. A recently published study (Becker et al. 1987) serves to highlight this type of problem. These workers have studied factors that in-

teract with the upstream region of the tyrosine amino-transferase gene in expressing and nonexpressing cells. Extracts of both cell types produce exactly the same footprinting pattern on naked DNA, yet an elegant in vivo footprinting procedure using whole cells reveals the expected differences in interactions between cells in which the gene is expressed and those in which it is not. So in this case the factors that presumably mediate regulation are present in equal quantities in both cell types and one is led to the conclusion that their ability to bind is controlled either by accessory proteins not detectable by footprinting assays or by long-range chromatin structure. Jalinot described studies in which in vivo footprinting was used to analyze proteins interacting with the adenovirus E2 promoter. Both in vivo footprints and in vitro gel retardation assays show the same interactions with this promoter in cells infected with wild-type adenovirus and the E1A deletion mutant *dl*312, despite the fact that some interaction between E1A products and the E2 promoter is required for efficient transcription. Gralla (University of California, Los Angeles) reported primer-extension footprinting experiments designed to probe in vivo interactions with SV40 DNA. He has been able to detect changes in binding, presumably of Sp1, to the GC-rich sequences as the infection proceeds. However, his results highlight one problem that will necessarily bedevil many studies in viral systems, namely, that the cells contain large numbers of copies of the viral genome, only a small proportion of which are transcriptionally active, and that proportion may not be detectable against the large excess of inactive templates.

AP1, originally defined as a factor binding to the SV40 enhancer (see earlier), is turning out to be a molecule of considerable interest. It binds to a consensus sequence identified in genes that respond to TPA (Angel et al. 1987; Lee et al. 1987b), and the binding activity of this factor is increased in TPA-treated cells. The sequence recognized by AP1 is identical to that recognized by the *Saccharomyces cerevisiae* factor GCN4, and Struhl (1987) has shown that the DNA-binding domain of GCN4 is functionally interchangeable with that of the avian oncogene v-*jun*, recently isolated from a new sarcoma virus. Anti-peptide antisera against v-*jun* recognize purified human AP1, and peptide sequences from AP1 match perfectly with the nucleotide sequence of human c-*jun*. Moreover, a β-galactosidase–human c-*jun* fusion protein is recognized by the anti-peptide sera and has a DNA binding activity indistinguishable from that of AP1 (Bohmann et al. 1987). More recently, Angel et al. (1988) have reported similar data. Thus, it is extremely likely that AP1 is encoded by the c-*jun* proto-oncogene. However, for the reasons discussed above this identity is not yet absolutely certain. It is possible that c-*jun* encodes a family of proteins, one of which is AP1, and in this context it is important to note that the most purified preparations of AP1 currently available contain several polypeptides, only one of which has been shown to have the DNA binding properties which define AP1. Moreover, there may be other genes related to c-*jun* which encode proteins with the same DNA binding

properties as AP1 but different functions and, indeed, Ryder et al. (1988) have identified a c-*jun*-related gene whose transcription is induced by serum stimulation of quiescent fibroblasts. It has long been clear on theoretical grounds that one class of oncogene should encode transcription factors, and there is circumstantial evidence that the c-*myc* and c-*fos* gene products have such activities. However, the work with AP1/GCN4/v-*jun* is the first clear demonstration that a gene encoding a defined transcription factor can act as an oncogene.

As described earlier, the PEA1 factor which binds to the polyoma enhancer recognizes a sequence indistinguishable from the AP1 consensus. Hirai reported that SV40-transformed mouse cells, grown in low serum, contain fivefold higher levels of this factor than their untransformed parents, grown under the same conditions, raising the possibility that increased activity of the murine c-*jun* gene product plays a role in SV40 transformation. However, the identity of AP1 and PEA1 is not yet proven. Hirai showed that the TPA or serum induction of PEA1 activity in NIH-3T3 cells requires de novo protein synthesis, as judged by experiments using the inhibitors anisomycin and cycloheximide, whereas the TPA induction of AP1 in HeLa cells is insensitive to such protein synthesis inhibitors (Angel et al. 1987). Hirai also observed an increase in PEA1 activity in NIH-3T3 cells treated with forskolin, and other cAMP-phosphodiesterase inhibitors, suggesting that the factor can also be activated by a pathway involving a cAMP-dependent protein kinase. However, it is not yet clear that TPA and cAMP activate exactly the same polypeptide (M. Yaniv, pers. comm.). It is not known if de novo synthesis is required in this case. Thus, while it is clear that PEA1 and AP1 recognize the same DNA sequence, the data discussed above clearly admit the possibility that PEA1 is either another product of c-*jun* or that it is encoded by a related gene such as that described by Nathans.

The issues discussed in the context of AP1 will also apply to studies of other factors. The clear demonstrations that multiple proteins can bind to the same DNA sequence and the strong suggestions that the genes encoding transcription factors can give rise to multiple polypeptides mean that precise definitions of the regulatory roles and biochemical activities of these factors can only be achieved when the corresponding cDNAs have been cloned and expressed, thereby allowing access to large quantities of defined polypeptide. Such developments will also allow the generation of high-quality antibodies with which to probe the functional interactions between factors. We can look forward confidently to a period of rapid progress in this field and expect that the first complete description of the regulation of a eukaryotic transcriptional control element will come from studies of DNA tumor viruses.

References

Angel, P., M. Imagawa, R. Chiu, B. Stein, R.J. Imbra, H.J. Rahmsdorf, C. Jonat, P. Herrlich, and M. Karin. 1987.

Phorbol ester-inducible genes contain a common *cis* element recognized by a TPA-modulated *trans*-acting factor. *Cell* **49**: 729–739.

Angel, P., E. Allegretto, S. Okino, K. Hattorri, W.J. Boyle, T. Hunter, and M. Karin. 1988. The *jun* oncogene encodes a sequence specific *trans*-activator similar to AP-1. *Nature* (in press).

Barberis, A., G. Superti-Furga, and M. Busslinger. 1987. Mutually exclusive interaction of the CCAAT-binding factor and of a displacement protein with overlapping sequences of a histone gene promoter. *Cell* **50**: 347–359.

Becker, P.B., S. Rippert, and G. Schutz. 1987. Genomic footprinting reveals cell type-specific DNA binding of ubiquitous factors. *Cell* **51**: 435–443.

Berk, A.J. 1986. Adenovirus promoters and E1A transactivation. *Annu. Rev. Genet.* **20**: 45–79.

Bird, A.P. 1986. CpG-rich islands and the function of DNA methylation. *Nature* **321**: 209–213.

Bohmann, D., T.J. Bos, A. Admon, T. Nishimura, P.K. Vogt, and R. Tjian. 1987. Human proto-oncogene c-*jun* encodes a DNA binding protein with structural and functional properties of transcription factor AP1. *Science* **238**: 1386–1392.

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

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

Carthew, R.W., L.A. Chodosh, and P.A. Sharp. 1987. The major late transcription factor binds to the mouse metallothionein 1 promoter. *Genes Dev.* **1**: 973–980.

Chodosh, L.A., R.W. Carthew, and P.A. Sharp. 1986. A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. *Mol. Cell. Biol.* **6**: 4723–4733.

Chodosh, L.A., R.W. Carthew, J.G. Morgan, G.R. Crabtree, and P.A. Sharp. 1987. The adenovirus major late transcription factor activates the rat *X*-fibrinogen promoter. *Science* **238**: 684–688.

Chiu, R., M. Imagawa, R.J. Imbra, J.R. Bockoven, and M. Karin. 1987. Multiple *cis*- and *trans*-acting elements mediate the transcriptional response to phorbol esters. *Nature* **329**: 648–651.

Clarke, J. and W. Herr. 1987. Activation of mutated simian virus 40 enhancers by amplification of wild-type enhancer elements. *J. Virol.* **61**: 3536.

Davidson, I., E. Fromental, P. Augereau, A. Wildeman, M. Zenke, and P. Chambon. 1986. Cell-type specific protein binding to the enhancer of simian virus 40 in nuclear extracts. *Nature* **323**: 544–548.

Dorn, A., J. Bollekens, A. Stauls, C. Benoist, and D. Mathis. 1987. A multiplicity of CCAAT box-binding proteins. *Cell* **50**: 863–872.

Fletcher, C., N. Heintz, and R.G. Roeder. 1987. Purification and characterization of OTF1, a transcription factor regulating cell cycle expression of a human histone H2b gene. *Cell* **51**: 773–781.

Garcia, J., F. Wu, and R. Gaynor. 1987. Upstream regulatory regions required to stabilise binding to the TATA sequence in an adenovirus early promoter. *Nucleic Acids Res.* **15**: 8367–8385.

Goding, C.R., S.M. Temperley, and F. Fisher. 1987. Multiple transcription factors interact with the adenovirus-2 EII-late promoter: Evidence for a novel CCAAT recognition factor. *Nucleic Acids Res.* **15**: 7761–7780.

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

Hassell, J.A., W.J. Muller, and C.R. Mueller. 1986. The dual role of the polyoma enhancer in transcription and DNA replication. *Cancer Cells* **4**: 561–569.

Hearing, P. and T. Shenk. 1983. The adenovirus type 5 E1A transcriptional control region contains a duplicated enhancer element. *Cell* **33**: 695–703.

———. 1986. The adenovirus type 5 enhancer contains two functionally distinct domains: One is specific for E1A and the other modulates transcription of all early units in *cis*. *Cell* **45**: 229–236.

Herr, W. and J. Clarke. 1986. The SV40 enhancer is composed of multiple functional elements that can be compensated for one another. *Cell* **45**: 461–470.

Herr, W. and Y. Gluzman. 1985. Duplications of a mutated simian virus 40 enhancer restore its activity. *Nature* **313**: 711–714.

Hen, R., E. Borrelli, P. Sassoni-Corsi, and P. Chambon. 1983. An enhancer is located 340 bp upstream from the adenovirus-2 E1a cap site. *Nucleic Acids Res.* **11**: 8747–8761.

Hurst, H.C. and N.C. Jones. 1987. The identification of factors that interact with the E1A-inducible adenovirus E3 promoter. *Genes Dev.* **1**: 1132–1146.

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

Johnson, P.F., W.H. Landschulz, B.J. Graves, and S.L. McKnight. 1987. Identification of a rat liver protein that binds to the enhancer core element of three animal viruses. *Genes Dev.* **1**: 133–146.

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

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

Kovesdi, I., R. Reichel, and J.R. Nevins. 1986. Identification of a cellular transcription factor involved in E1A *trans*-activation. *Cell* **45**: 219–228.

Kovesdi, I., M. Satake, K. Furukawa, R. Reichel, Y. Ito, and J.R. Nevins. 1987. A factor discriminating between the wild-type and a mutant polyoma enhancer. *Nature* **328**: 87–89.

Kryszke, M.H., J. Piette, and M. Yaniv. 1987. Induction of a factor that binds to the polyoma virus A enhancer on differentiation of embryonal carcinoma cells. *Nature* **328**: 254–256.

Lee, W., A. Haslinger, M. Karin, and R. Tjian. 1987a. Activation of transcription by two factors that bind promoter and enhancer sequences of the human metallothionein gene and SV40. *Nature* **325**: 368–372.

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

Lee, K.A.W. and M.R. Green. 1987c. A cellular transcription factor E4F1 interacts with an E1a-inducible enhancer and mediates constitutive enhancer function *in vitro*. *EMBO J.* **6**: 1345–1353.

Lee, K.A.W., T.Y. Hai, L. SivaRaman, B. Thimmappaya, H.C. Hurst, N.C. Jones, and M.R. Green. 1987d. A cellular transcription factor ATF activates transcription of multiple

E1A-inducible adenovirus early promoters. *Proc. Natl. Acad. Sci.* **84:** 8355–8359.

Lichtsteiner, S., J. Wuarin, and U. Schibler. 1987. The interplay of DNA-binding proteins on the promoter of the mouse albumin gene. *Cell* **51:** 963–973.

Lillie, J.W., P.M. Loewenstein, M.R. Green, and M. Green. 1987. Functional domains of adenovirus type 5 E1a proteins. *Cell* **50:** 1091–1100.

Lusky, M., L. Berg, H. Weiher, and M. Botchan. 1983. Bovine papilloma virus contains an activator of gene expression at the distal end of the early transcription unit. *Mol. Cell. Biol.* **3:** 1108–1122.

Maniatis, T., S. Goodbourn, and J.A. Fischer. 1987. Regulation of inducible and tissue-specific gene expression. *Science* **236:** 1237–1244.

Mitchell, P.J., C. Wang, and R. Tjian. 1987. Positive and negative regulation of transcription in vitro: Enhancer binding protein AP2 is inhibited by SV40 T antigen. *Cell* **50:** 847–861.

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

Montminy, M.R., K.A. Sevarino, J.A. Wagner, G. Mondel, and R.H. Goodman. 1986. Identification of a cyclic AMP responsive element within the rat somatostatin gene. *Proc. Natl. Acad. Sci.* **86:** 6682–6686.

Nomiyama, H., C. Fromental, J.H. Xiao, and P. Chambon. 1987. Cell-specific activity of the constituent elements of the simian virus 40 enhancer. *Proc. Natl. Acad. Sci.* **84:** 7881–7885.

Ondek, B., A. Shepard, and W. Herr. 1987. Discrete elements within the SV40 enhancer region display different cell-specific enhancer activities. *EMBO J.* **6:** 1017–1025.

Ostapchuk, P., J.F.X. Diffley, J.T. Bruder, B. Stillman, A.J. Levine, and P. Hearing. 1986. Interaction of a nuclear factor with the polyomavirus enhancer region. *Proc. Natl. Acad. Sci.* **83:** 8550–8554.

Pfeifer, K., T. Prezant, and L. Guarente. 1987. Yeast Hap1 activator binds to two upstream activation sites of different sequence. *Cell* **49:** 19–27.

Piette, J. and M. Yaniv. 1987. Two different factors bind to the α-domain of the polyoma virus enhancer, one of which also interacts with the SV40 and c-*fos* enhancers. *EMBO J.* **6:** 1331–1337.

Pruijn, G.J.M., W. van Driel, and P.C. van der Vliet. 1986. Nuclear factor III, a novel sequence-specific DNA-binding protein from HeLa cells stimulating adenovirus DNA replication. *Nature* **322:** 656–659.

Reichel, R., T. Kovesdi, and J.R. Nevins. 1987. Developmental control of a promoter-specific factor that is also regulated by the E1A gene product. *Cell* **48:** 501–506.

Rosales, R., M. Vigneron, M. Macchi, I. Davidson, J.H. Xiao, and P. Chambon. 1987. In vitro binding of cell-specific and ubiquitous nuclear proteins to the octamer motif of the SV40 enhancer and related motifs present in other promoters. *EMBO J.* **6:** 3015–3025.

Rosenfeld, P.J., E.A. O'Neill, R.J. Wides, and T.J. Kelly. 1987. Sequence specific interactions between cellular DNA binding proteins and the adenovirus origin of replication. *Mol. Cell. Biol.* **7:** 875–886.

Ryder, K., L. Lau, and D. Nathans. 1988. A gene activated by growth factors is related to the oncogene v-*jun*. *Proc. Natl. Acad. Sci.* (in press).

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

Schneidereit, C., A. Heguy, and R.G. Roeder. 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.

Schirm, S., J. Jiricny, and W. Schaffner. 1987. The SV40 enhancer can be dissected into multiple segments, each with a different cell type specificity. *Genes Dev.* **1:** 65–74.

Sen, R. and D. Baltimore. 1986. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell* **46:** 705–716.

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

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

SivaRaman, L. and B. Thimmappaya. 1987. Two promoter-specific host factors interact with adjacent sequences in an E1A-inducible adenovirus promoter. *Proc. Natl. Acad. Sci.* **84:** 6112–6116.

Sive, H.L., and R.G. Roeder. 1986. Interaction of a common factor with conserved promoter and enhancer sequences in histone H2B, immunoglobulin, and U2 small nuclear RNA (snRNA) genes. *Proc. Natl. Acad. Sci.* **83:** 6382–6386.

Spangler, R., M. Bruner, B. Dalie, and M.L. Harter. 1987. Activation of adenovirus promoters by the adenovirus E1A protein in cell-free extracts. *Science* **237:** 1044–1046.

Staudt, L.M., H. Singh, R. Sen, T. Wirth, P.A. Sharp, and D. Baltimore. 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. *Nature* **323:** 640–643.

Struhl, K. 1987. The DNA-binding domains of the *jun* oncoprotein and the yeast GCN4 transcriptional activator proteins are functionally homologous. *Cell* **50:** 841–846.

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

Veldman, G.M., S. Lupton, and R. Kamen. 1985. Polyoma enhancer contains multiple redundant sequence elements that activate both DNA replication and gene expression. *Mol. Cell. Biol.* **5:** 649–658.

Wang, J., K. Nishiyama, K. Araki, D. Kitamuta, and T. Wantanabe. 1987. Purification of an octamer sequence (ATG-CAAAT)–binding protein from human B cells. *Nucleic Acids Res.* **15:** 10105–10116.

Weiher, H., M. Konig, and P. Gruss. 1983. Multiple point mutations affecting the simian virus 40 enhancer. *Science* **219:** 626–631.

Wu, L., D.S.E. Rosser, M.G. Schmidt, and A.J. Berk. 1987. A TATA box implicated in E1a transcriptional activation of a simple adenovirus 2 promoter. *Nature* **326:** 512–515.

Xiao, J.H., I. Davidson, D. Ferrandon, R. Rosales, M. Vigneron, M. Macchi, F. Ruffernach, and P. Chambon. 1987. One cell-specific and three ubiquitous nuclear proteins bind in vitro to overlapping motifs in the domain B1 of the SV40 enhancer. *EMBO J.* **6:** 3005–3013.

Yee, A.S., R. Reichel, I. Kovesdi, and J. Nevins. 1987. Promoter interaction of the E1A-inducible factor E2F and its potential role in the formation of a multi-component complex. *EMBO J.* **6:** 2061–2068.

Zenke, M., T. Grundtstrom, H. Matthes, M. Wintzereith, C. Schatz, A. Wildeman, and P. Chambon. 1986. Multiple sequence motifs are involved in SV40 enhancer function. *EMBO J.* **5:** 378–397.

Trans-acting protein factors and the regulation of eukaryotic transcription: lessons from studies on DNA tumor viruses.

N C Jones, P W Rigby and E B Ziff

References	This article cites 72 articles, 21 of which can be accessed free at: **http://genesdev.cshlp.org/content/2/3/267.full.html#ref-list-1**
License	
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**