The EMBO Journal vol.4 no.5 pp.1293-1300, 1985

# The adenovirus-2 early EIIa transcription unit possesses two overlapping promoters with different sequence requirements for EIa-dependent stimulation

**Deborah A.Zajchowski, Hélène Boeuf and Claude Kédinger**

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Faculté de Médicine, 11 rue Humann, 67085 Strasbourg-Cédex, France

Communicated by C.Kédinger

**The EIa-inducible, EIIa transcription unit of adenovirus-2 is transcribed early in infection from two start sites (+1 or EIIaE1 and −26 or EIIaE2), neither of which is preceded by canonical TATA box elements. Analysis of promoter deletion and linker scanning mutations for *in vivo* transcriptional activity after transfection into HeLa cells has indicated the existence of two overlapping promoters in the EIIaE gene. Two regions, each ~30 nucleotides upstream from start sites EIIaE1 and EIIaE2, function as TATA box substitutes. A sequence centered at position −42 (with respect to the major start site at position +1) is essential for transcription from both sites, while an element further upstream, localized between nucleotides −91 and −62, is also required for efficient EIIaE transcription, with the 3′ border being dispensable for EIIaE2 transcription. Analysis of the entire series of EIIaE mutants, co-transfected with an EIa-containing plasmid, revealed that no unique sequence elements in the EIIaE1 promoter region between −97 and +1 were responsible for the stimulation of EIIaE1 transcription by EIa. In contrast, the EIa-mediated augmentation of EIIaE2 template activity was mainly dependent upon a sequence, the 5′-TTAAATTT-3′ putative TATA box substitute, located around position −59.**

*Key words:* atypical eucaryotic promoter/EIa control of transcription/RNA polymerase B/transient expression

## Introduction

Studies on a number of genes have established that faithful and efficient transcription initiation requires both the TATA box element (located ~30 nucleotides upstream from the start site) and less conserved sequences located at variable distances further upstream (for references, see Grosschedl and Birnstiel, 1980; Corden *et al.*, 1980; Benoist and Chambon, 1981; Breathnach and Chambon, 1981; Hen *et al.*, 1982; McKnight and Kingsbury, 1982; Dierks *et al.*, 1983). In some cases additional *cis*-acting elements have been identified, including enhancer sequences (for references see Yaniv, 1982; Khoury and Gruss, 1983) and control elements which modulate the transcriptional rate in response to exogenous stimuli such as heat shock (Pelham, 1982; Mirault *et al.*, 1982), heavy metals (Karin *et al.*, 1984) or hormones (for references, see Groner *et al.*, 1984).

The usefulness of the study of viral gene expression in elucidating general mechanisms of transcriptional regulation is well established (Flint, 1982). The adenovirus type 2 (Ad2) early EIIa (EIIaE) transcription unit (TU), which encodes a 72 000 dalton DNA-binding protein, provides an intriguing alternative to the classical view of promoter structure-function relationships. Its transcription by RNA polymerase B, early in viral infection, proceeds both from major start sites (EIIaE1), which are not preceded by a consensus TATA sequence, and minor starts (EIIaE2) ~26 nucleotides upstream from the major cap site, which are preceded by a TATA-like element (Mathis *et al.*, 1981; Elkaim *et al.*, 1983; Hashimoto and Green, 1984). An understanding of the role of the peculiar structure of this promoter in the differential transcription of EIIaE1 and EIIaE2 may have general applications, since an increasing number of genes, both viral [adenovirus IVa2 gene (Baker and Ziff, 1981); SV40 late promoter (Brady *et al.*, 1982); hepatitis B surface antigen gene (Cattaneo *et al.*, 1983)] and cellular [HMG-CoA reductase (Reynolds *et al.*, 1984); hypoxanthine phosphoribosyltransferase (Melton *et al.*, 1984)], have been found to lack discernable TATA box elements. In addition, differential regulation of transcription from two start sites in the ϵ-globin gene has recently been described (Allan *et al.*, 1984). The products of the adenovirus immediate early TU, EIa, are required for efficient expression of each of the early viral TU (EIb, EIIa, EIII, EIV and L1) during the infectious cycle (Feldman and Nevins, 1983). Transient expression studies of cloned early viral genes, introduced into cultured cells by transfection (Imperiale *et al.*, 1983; Svensson and Akusjärvi, 1984) or microinjection (Rossini, 1983) have confirmed the involvement of *trans*-acting EIa produced in early gene expression. In the case of the EIIaE gene the EIa-mediated stimulation corresponds to an increased rate of transcription, implying that the EIa products act as positive transcriptional regulators (Leff *et al.*, 1984).

The Ad2 EIIaE TU, therefore, offers the opportunity not only to identify the sequence elements that form an atypical RNA polymerase B promoter, including those which can functionally substitute for the TATA box, but also to study the mechanism of the stimulation of transcription by the EIa products. Elkaim *et al.* (1983) have shown that the EIIaE promoter sequences upstream from position −94 are not required for efficient transcription in the presence and absence of the products of the EIa gene. To define more precisely the sequences required for both accurate transcription and the EIa-mediated stimulation, we have constructed an extensive series of 'linker-scanning' mutants spanning the promoter between −97 and +1 (the major mRNA cap site). Using a transient expression assay in HeLa cells, we show that two overlapping promoters control transcription from the EIIaE start sites. Two independent elements, located around positions −30 and −59, control the efficiency of transcription initiation from the EIIaE1 and EIIaE2 start sites, respectively, while two other elements, one located between the two former ones and the other further upstream, are important for efficient transcription from both promoters. Co-transfection of HeLa cells with the EIIaE promoter mutants and a plasmid carrying the EIa TU revealed that no discrete EIIaE sequences are required for the EIa-mediated stimulation of EIIaE1 transcription. However, stimulation of EIIaE2 clearly requires sequences corresponding to the TATA-like element at position −59.

© IRL Press Limited, Oxford, England.

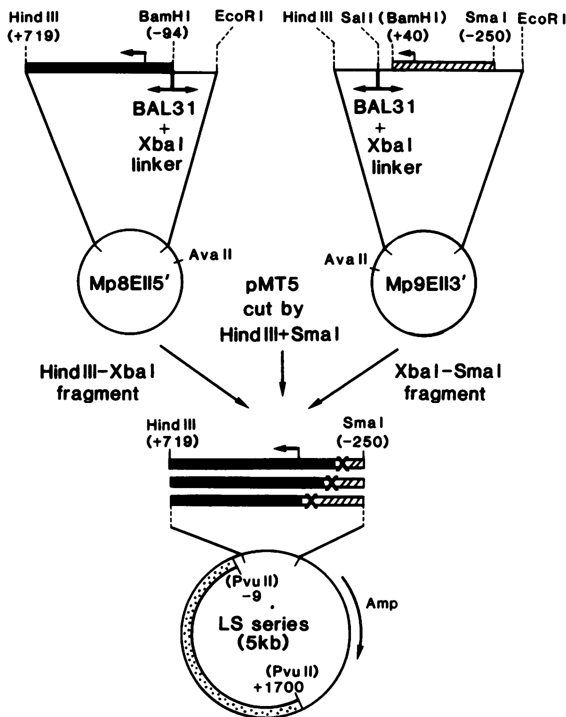


**Fig. 1.** Construction of the LS series of EIIaE promoter mutants. Recombinants Mp8EII5' and Mp9EII3' are families of 5' and 3' deletion mutants produced by *Bal*31 exonuclease digestion from the *Bam*HI and *Sal*I restriction sites, respectively, with an *Xba*I linker sequence inserted at the deletion endpoints (see Materials and methods). Appropriate sized *Hind*III-*Xba*I fragments excised from the 5' deletion family (shaded box) and *Xba*I-*Sma*I fragments (hatched box) from the 3' one were recombined at their *Xba*I ends (X), between the *Hind*III and the *Sma*I sites of the pMT5 vector. This vector, which provides splice and polyadenylation sites to the EIIaE sequences, contains the rabbit β-globin *Pvu*II fragment extending from −9 to about +1700 with respect to the globin cap site (stippled area on the lower diagram), inserted between the *Hind*III and *Pvu*II sites of pBR322. The *Eco*RI to *Hind*III polylinker sequence of M13 mp12 has also been inserted between the *Eco*RI and *Hind*III sites into this vector. The resulting constructs constitute the LS series of EIIaE promoter mutants. The various diagrams are not drawn to scale and the recombinants are represented such as to keep the EIIa gene fragments in the same orientation as in Figure 2. The arrow pointing to the left indicates the direction of transcription from the EIIaE major mRNA start site. The AvaII restriction site and the β-lactamase gene (Amp) are included as landmarks. Restriction sites in parentheses have been lost during construction.

## Results

### *Construction of linker-scanning mutants and in vivo analysis*

Deletion of the sequences between positions −250 and −94 in the EIIaE promoter did not significantly affect the level of transcription from either the major start sites (EIIaE1) or the minor start sites (EIIaE2), whereas deletion to position −63 dramatically reduced transcription from both of these sites (Elkaim *et al.*, 1983). Deletions which decrease constitutive transcription to very low levels by removing essential upstream promoter elements do not allow the identification of sequences located further downstream which could have a role in promoter activity, either alone or in conjunction with elements already deleted and which might be involved in the EIa-mediated induction mechanism. To define more precisely the elements essential for the EIIaE promoter function, we have constructed a series of mutants with clustered base substitutions in the region between position −97 and the cap site of the major EIIaE mRNA.

Employing the linker-scanning strategy of McKnight and Kingsbury (1982), we have produced mutants in which the wild-type EIIaE sequence is substituted by the synthetic *Xba*I linker sequence, 5'-CTCTAGAG-3', at fixed positions spanning the promoter region. This linker sequence was chosen because of its equal AT and GC composition. Mutants were constructed by the recombination of matching pairs of 5' and 3' deletion mutants at the *Xba*I site of the linker sequence inserted at each deletion end-point (see Materials and methods and Figure 1). The recombined EIIaE promoter fragments were ligated into another vector where the EIIa protein-coding sequence has been replaced by the rabbit β-globin coding sequence (Figure 1). This construction was chosen to avoid the possible influence of the production of the 72-K DNA-binding protein on its own gene transcription or stability of its mRNA (Carter and Blanton, 1978; Nevins *et al.*, 1979; Babich and Nevins, 1981), hence simplifying the interpretation of our data.

These linker-scanning mutants (LS series, Figure 2) were tested for their ability to be transcribed after transcription into HeLa cells. The cytoplasmic RNA was isolated and subjected to quantitative S1 nuclease analysis using the $^{32}$P-labeled probes indicated in Figure 2. The results of a typical experiment are shown in Figure 3, where the pattern of S1 nuclease-resistant probe fragments is similar to that obtained with RNA isolated from HeLa cells 4 h after Ad2 infection (Leff *et al.*, 1984). The multiplicity of initiation sites for both the major (EIIaE1) and the minor (EIIaE2) transcripts has been reproducibly observed and confirmed by reverse transcription primer extension assays (data not shown). The amount of RNA initiated at the EIIaE1 and EIIaE2 sites in the various LS recombinants has been quantitated in several separate transcription experiments, each performed with different plasmid DNA preparations. In most experiments a β-globin recombinant was co-transfected with the LS mutants and the level of globin transcription was determined to ascertain that similar transfection efficiencies were observed throughout. The results of such experiments are summarized in Figure 4 (open columns).

### *Distinct TATA box substitute elements control transcription initiation from the EIIaE1 and EIIaE2 start sites*

The absence of a consensus TATA box sequence 30 bp upstream from the major EIIaE initiation site (position +1), and the presence of only a partially homologous element situated ~30 bp upstream from the EIIaE2 start site, raise questions as to the exact nature and location of the promoter elements which functionally substitute for the classical TATA box. Alteration of the TATA-like element at position −59 (5'-TTAAATTT-3') in mutant LS-6052 reduced the level of transcription from the EIIaE2 site without affecting (or even slightly increasing) transcription efficiency from the EIIaE1 site (lane 2, Figure 3A and Figure 4), consistent with its role only in the control of EIIaE2 transcriptional activity.

From the transcription pattern of the LS series (Figure 3A), it is clear that the integrity of the region spanning nucleotides −48 to −19 is crucial for efficient transcription from the EIIaE1 initiation site. Within this region, the mutation carried by the LS-3323 mutant most dramatically affected transcription from EIIaE1 with a reduction to ~2.5% of the wild-type level. The location of the element defined by LS-3323 and LS-2719, 30 nucleotides upstream from the +1 initiation site, suggests that it corresponds to a promoter element which serves a function resembling that provided by the TATA box in other eucaryotic promoters.

Alteration of either of the EIIaE TATA box substitutes did not significantly affect the selection of the initiating nucleotides (see Figure 3A, lanes 1 and 2 for the EIIaE2 starts, lanes 5 and 6

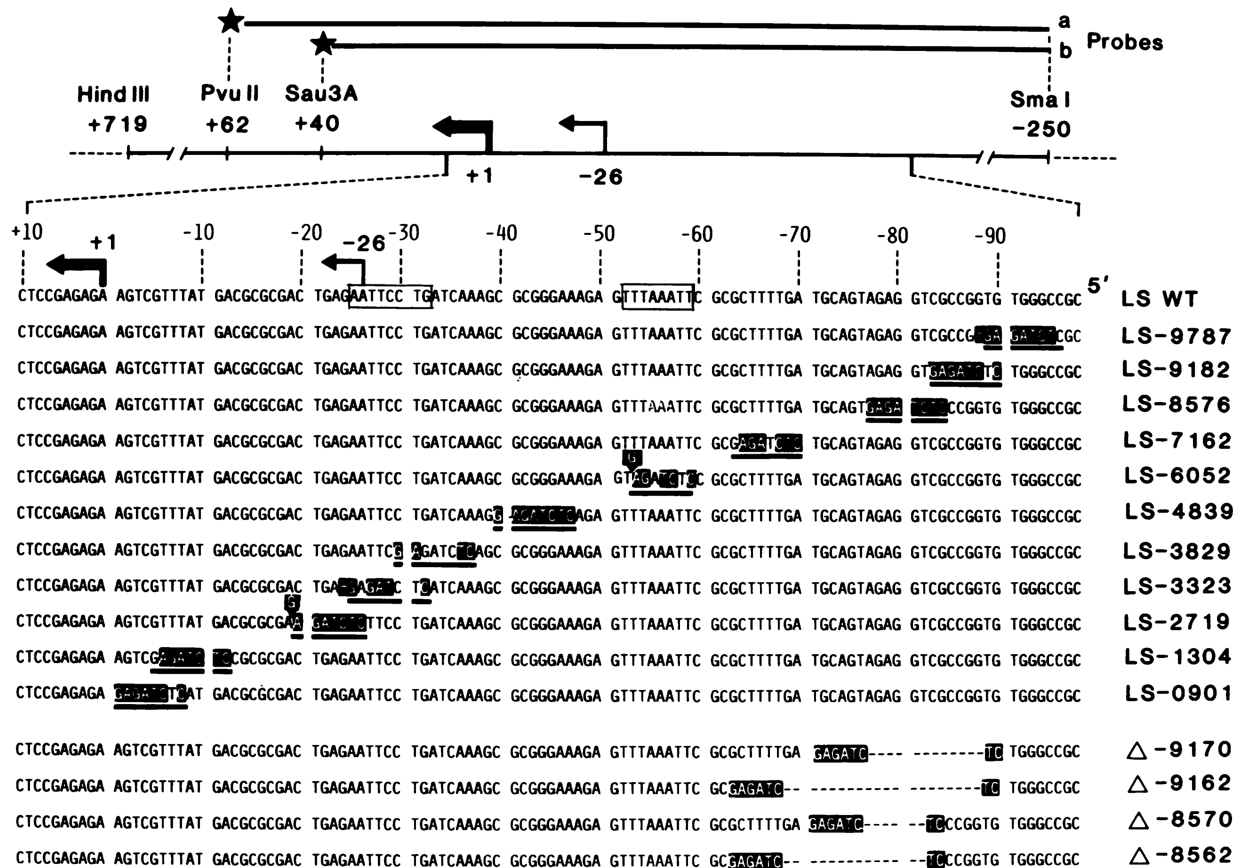


**Fig. 2.** Nucleotide sequences of the EIIaE LS and Δ-mutants. This scheme depicts the *SmaI-HindIII* fragment of the EIIaE transcription unit present in the LS series (see Figure 1). The arrows point to the direction of transcription and mark the position of the major (+1) and minor (−26) EIIaE RNA start sites. Probes, 5′ $^{32}$P-labeled at the *PvuII* (probe a) or *Sau3A* (probe b) sites, used for S1 nuclease analysis, are indicated. The nucleotide sequence between −98 and +10 of the wild-type EIIaE fragment (LS WT) and the corresponding fragment of the LS and Δ series is represented with the two putative TATA box elements boxed in the wild-type sequence. In the LS series, the *XbaI* linker sequence is underlined and those nucleotides which have been altered from the wild-type sequence are stressed by shaded boxes. Single nucleotide insertions (LS-6052 and LS-2719) or deletions (LS-9787 and LS-3323) are arbitrarily indicated on the 3′ border of the linker. The internal deletion mutants (Δ series) show the entire *XbaI* linker boxed between the boundaries of the deletions. The nomenclature of the various mutants is derived from the position of the nucleotides between which the linker sequence has been inserted.

for the EIIaE1 starts) thereby suggesting that additional elements are involved in the process which directs the transcription machinery to the correct initiation sites. Similar observations have been made for TATA box deletion or point mutations of the adenovirus major late (Hen *et al.*, 1982), the herpes thymidine kinase (McKnight and Kingsbury, 1982) and the rabbit β-globin promoters (Dierks *et al.*, 1983).

An increase in transcription from the minor start site at −26 occurs concomitantly with the reduction in transcription from the EIIaE1 site for both LS-3323 and LS-2719, as indicated by the augmented level of probe fragments accumulating at the end-point of homology between the probe and the RNA transcribed from the linker-mutated EII plasmids (arrowheads in Figure 3A). That these transcripts indeed correspond to EIIaE-specific RNA initiated in the −26 region, and not predominantly to random transcripts initiated further upstream, was shown by precise mapping of the corresponding start sites (Figure 5, lanes 5 and 7), using $^{32}$P-labeled probes prepared from the homologous template DNAs, LS-3323 and LS-2719. Similar conclusions were reached by primer extension studies (data not shown and Figure 4). The scatter of EIIaE2 starts in LS-2719 (bracketted area in lane 7 of Figure 5) was probably a consequence of the replacement of the original AAGA initiation sequence by CTCT in the mutant. A quantitation of the transcripts, taking into account the relative specific activities of the probes (see legend to Figure 5), revealed a 2- to 5-fold increase in the level of transcription from the EIIaE2 site upon mutation of the promoter element between −33 and −19. With this increase, however, the total transcription (+1 and −26) from LS-3323 and LS-2719 never exceeded that from the wild-type construct.

Altogether, these results strongly suggest that the sequence elements located ~30 bp upstream from the EIIaE1 and EIIaE2 start sites, respectively, belong to independent promoters, since the major effect of their alteration is to impair transcription only from the corresponding start site. The reason for the increased transcription from the EIIaE2 region upon mutation of the element at position −30 is not clear, and could be due to the incidental creation of more efficient initiation sites around position −26. Alternatively, this result and the fact that a slight increase in transcription from the EIIaE1 site could often be observed after mutation of the TATA-like element at −59 (see above and Figure 4), could indicate reduced competition between the two TATA box substitutes for a limiting transcription factor(s).

Alteration of sequences located downstream from the TATA box substitute at −30 had little effect upon transcription from either the EIIaE1 or EIIaE2 sites as revealed by the template activities of LS-1304 and LS-0901 (data not shown and Figure 4).

### *Overlapping promoter elements are essential for efficient transcription from both the EIIaE1 and EIIaE2 start sites*

Between the TATA box functional counterparts, which separately

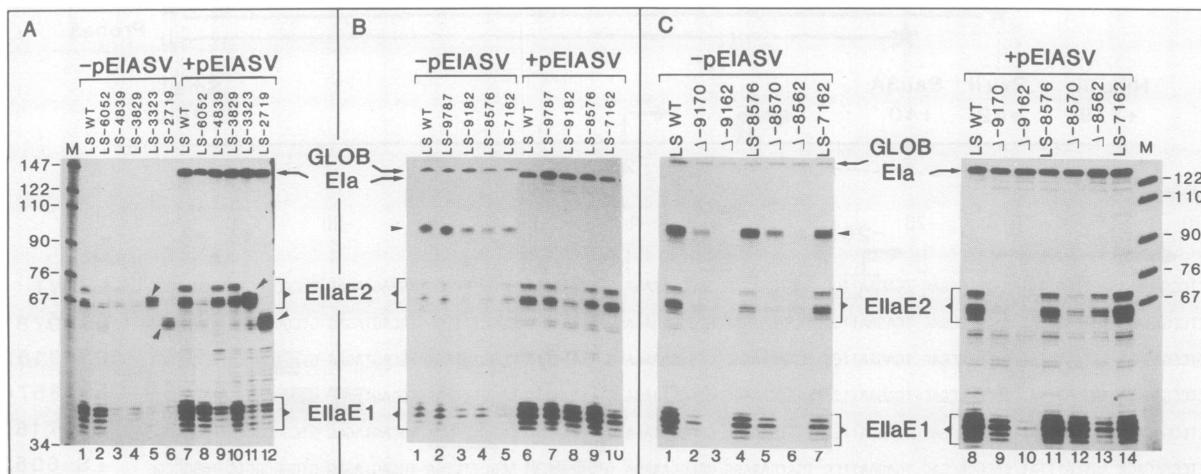


**Fig. 3.** Quantitative S1 nuclease analysis of cytoplasmic RNA from cells transfected with the LS and Δ-series of EIIaE recombinants in the presence and absence of a co-transfected, EIa-containing plasmid. HeLa cells were transfected with 2.5 μg per dish of the LS or Δ-mutants alone (**lanes 1–6, panel A; lanes 1–5, panel B; lanes 1–7, panel C**) or in the presence of 1 μg of pEIASV (**lanes 7–12, panel A; lanes 6–10, panel B; lanes 8–14, panel C**). 2 μg of a polyoma-β-globin recombinant [pβ(244+)β from DeVilliers and Schaffner, 1981], used as an internal control of transfection efficiency, was co-transfected in the experiments shown in **panel B (lanes 1–5)** and **C (lanes 1–7)**. The pEIASV plasmid is a pBR322 recombinant which contains the leftmost *Hpa*I fragment (0–4.5 m.u.) of the Ad2 genome ligated to a 135 bp long SV40 fragment encoding the early and late polyadenylation signals (Elkaim *et al.*, 1983). The probe used to quantitate the EIIaE transcripts was probe b as described in Figure 2. The probe fragments protected by specific transcripts are bracketted and correspond to those bands which have been quantitated by densitometry (see Figure 4). The arrowheads in **panel A** designate transcripts initiated at, or upstream of, the linker-nucleotide substitution (see Results). The probes used to quantitate the specific globin and EIa transcripts were single-stranded restriction fragments corresponding to the coding strand between −84 and +137, and −498 and +129 of the globin and EIa genes, respectively. Arrows point to probe fragments protected by specific transcripts as indicated. Bands marked by the arrowhead in **panel B (lanes 1–5)** and **C (lanes 1–7)** correspond to EIIaE transcripts spliced to a cryptic acceptor site at approximately position +40 in the globin sequence (Wieringa *et al.*, 1983) and revealed by the globin probe. **Lanes M** correspond to $^{32}$P-labeled *Msp*I fragments of pBR322.

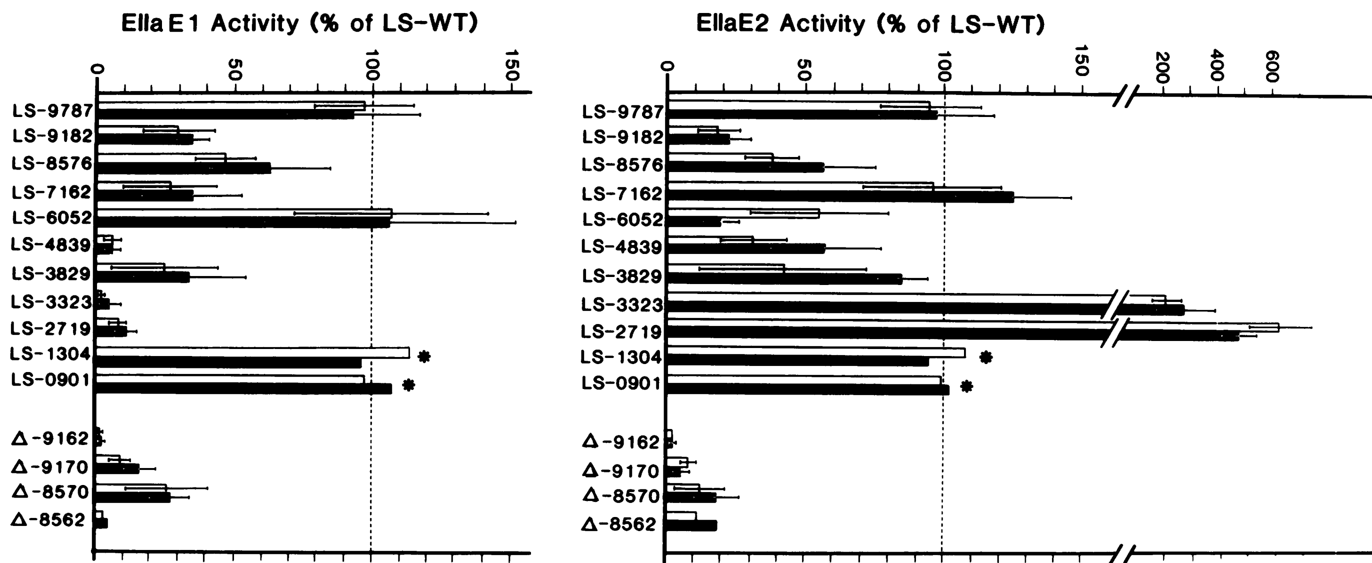


**Fig. 4.** Relative *in vivo* transcriptional activities of the EIIaE1 and EIIaE2 promoters of the LS- and Δ-mutants. Total transcription from the EIIaE1 and EIIaE2 sites (bands bracketted in Figure 3) were quantitated by S1 nuclease mapping in experiments similar to those shown in Figure 3. The intensity of the specific bands was determined by densitometry of autoradiograms exposed for various lengths of time. The results for each mutant, transfected alone (open bars) or with pEIASV (solid bars), are expressed relative to the values found for the LS-WT transfected under the same conditions. The length of each bar corresponds to the mean of five or more experiments, with the standard deviations indicated (those without standard deviations represent the means of 2–3 experiments). The values marked by an asterisk (*) have been obtained by quantitation of a reverse transcriptase primer extension analysis.

control transcription from the EIIaE1 or EIIaE2 sites, lies a sequence element (centered at position −42) consisting of a GC-rich tract flanked by three A residues on each side (see Figure 2). This element is required for efficient transcription from both start sites, since its alteration by the linker substitution in LS-4839 (Figures 3A and 4) strongly reduced EIIaE1 transcription and decreased that of EIIaE2 to ~30% of the wild-type (LS WT) level. The mutation of LS-3829 also decreased transcription from both sites, but the effect might be ascribed to the alteration of either or both the 3′ border of this sequence and the 5′ border of the TATA box substitute element at −30.

The external deletion analyses of Elkaim *et al.* (1983) demonstrated that promoter sequences indispensable for constitutive EIIaE1 template activity were located between positions −94 and −63. Indeed, linker mutations within this region reduced EIIaE1 transcriptional activity to 20–50% of LS-WT (see LS-9182,

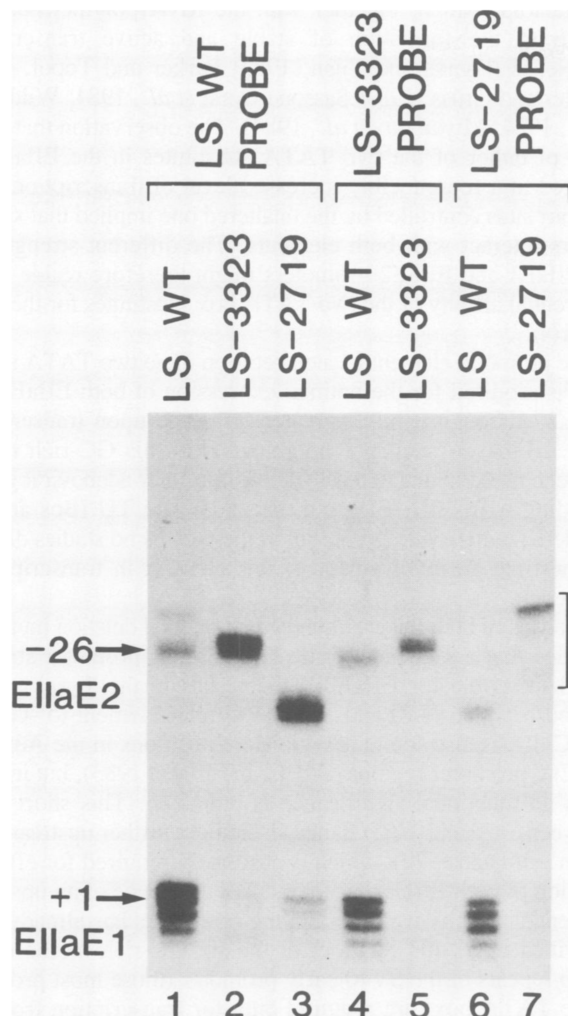


**Fig. 5.** Comparative S1 nuclease analysis of EIIaE-specific trancripts originated from LS-WT, LS-3323 and LS-2719 recombinants. HeLa cells were co-transfected with either LS-WT (**lanes 1, 4** and **6**), LS-3323 (**lanes 2** and **5**) or LS-2719 (**lanes 3** and **7**), as described in the legend to Figure 3. Equivalent amounts of cytoplasmic RNA were hybridized, as indicated, with probes prepared from the *Pvu*II-*Sma*I fragment of recombinants LS-WT (probe a of Figure 2, LS-WT probe), LS-3323 (LS-3323 probe) or LS-2719 (LS-2719 probe), and subjected to S1 nuclease analysis. The relative specific activities of the $^{32}P$-labeled probes homologous to LS-WT, LS-3323 and LS-2719 were determined by comparison of the EIIaE1 (+1) signals obtained for each probe after hybridization with RNA prepared from LS-WT-transfected cells (**lanes 1, 4** and **6**). In this experiment the LS-WT probe is 2- and 2.4-fold more active than the LS-3323 and LS-2719 probes, respectively. Quantification of the EIIaE transcripts initiated around position −26 (bracketted region on the right of the autoradiogram) and at position +1 was performed for **lanes 1, 5** and **7**, and adjusted for the respective probe activities. Arrows point to specific probe fragments as in Figure 3.

-8576, -7162; Figure 3B, lanes 3−5), while mutation of the sequences defined by LS-9787 had no effect (Figure 3B, lane 2). To characterize further this upstream promoter region, we constructed a series of short overlapping deletion mutants (Δ-mutants, see Materials and methods) encompassing the region between −91 and −62. The template activities of these Δ-mutants are depicted in Figures 3C and 4. The smallest internal deletion, Δ-8570, reduced transcriptional activity from the EIIaE1 start sites to ~25% of the wild-type level. Extension of the deletion in either the 5′ (Δ-9170) or the 3′ direction (Δ-8562) further reduced EIIaE1 promoter activity to ~10% and 2% of wild-type levels, respectively. Deletion of the entire region between −91 and −62 reduced transcription to virtually undetectable levels, thereby rigorously defining the borders of the EIIaE1 upstream element required for full promoter activity.

To eliminate the possibility that the higher promoter activity of the Δ-9170 mutants with respect to Δ-9162 was due to the movement of upstream inhibitory sequences closer to the major cap site as a result of the longer internal deletion in the latter mutant, we constructed a Δ-9162 mutant with a double *Xba*I linker. This mutant thereby left the sequences upstream from −91 at the same distance from −62 as in the Δ-9170 mutant. After this spacing correction, no improvement of template efficiency was observed (data not shown), thereby eliminating the possibility that an optimal distance between the start sites and sequences upstream from position −91 was required.

The upstream control element for transcription from the EIIaE2 site is situated between positions −91 and −70, as revealed by the decreased transcription from the LS- and Δ-mutants which span this region (Figure 3B, lanes 3−5, Figure 3C, lanes 2−7 and Figure 4). That the region between −70 and −62 is not essential for transcription from this site is indicated by the differential effect of the LS-7162 mutation which markedly reduced transcription from the EIIaE1 site without affecting EIIaE2 transcription (Figure 3B, lane 5 and 4C, lane 7) as well as by the identical ability of mutants Δ-8570 and Δ-8562 to support EIIaE2 transcription (see Figure 4).

These results raise the possibility that the sequence between −91 and −62 is comprised of two promoter elements, one (between −90 and −70) required for efficient transcription from both the EIIaE1 and EIIaE2 sites, and the other (between −70 and −62) uniquely involved in the transcriptional control from the EIIaE1 site.

### *Stimulation of the EIIaE1 promoter by EIa requires no unique sequence elements in the promoter region encompassing −97 to +1*

The effects of the LS- and Δ-mutations on the level of transcriptional stimulation from the EIIaE promoter by the EIa products are shown in Figure 3A (lanes 7−12), B (lanes 6−10) and C (lanes 8−14), with the quantitated data provided in Figure 4. Due to the variation in the extent of stimulation by EIa that we have observed in different experiments (ranging from 3- to 10-fold), we have represented the template efficiency of each EIIaE mutant relative to that of the 'wild-type' (LS WT) recombinant. It is clear from Figures 3 and 4 (solid bars) that, for any given EIIaE-recombinant co-transfected with the EIa transcription unit (pEIASV plasmid), the amount of transcription initiated from the EIIaE1 site relative to the wild-type EIa-stimulated level was usually similar to that observed when it was transfected alone (compare open and closed bars in Figure 4, corresponding to constitutive and EIa-stimulated transcriptional activity levels, respectively). These observations indicate that the EIa products are capable of equally stimulating transcription from the EIIaE1 start sites of every LS- or Δ-mutant, even those most poorly transcribed and, therefore, that no unique sequence elements are implicated in the EIa-mediated stimulation of EIIaE1 transcription.

### *Stimulation of EIIaE2 transcription by EIa is dependent upon the TATA box substitute sequence around −59*

The conclusion reached for EIIaE1 does not apply to the EIa-inducible, EIIaE2 promoter. On the contrary, modification of the TATA box substitute sequence (LS-6052) resulted in a severe impairment in the inducibility of EIIaE2 transcription by the EIa products (compare lanes 1 and 7 with lanes 2 and 8 in Figure

3A and see Figure 4). In addition, we observed a small but reproducible decrease (2- to 3-fold) in the extent of stimulation of EIIaE2 transcription for the Δ-9170 mutant (compare lanes 1 and 8 with lanes 2 and 9 in Figure 3C and see Figure 4). The fact that LS-mutants in this region are all maximally stimulated implies that the essential element encompassed either a longer nucleotide stretch, or a sequence unaffected by the chosen linker mutations (e.g., between −76 and −70). It is striking that the decrease in stimulation by EIa observed for the EIIaE2 transcription with the LS-6052 and the Δ-mutants does not apply to activity at the EIIaE1 start site, where the stimulation remains identical to that of LS WT. Thus it appears that, in contrast to the regulation by EIa of EIIaE1 promoter activity, the weaker EIIaE2 promoter is comprised of discrete sequence elements specific for the EIa response, with the most critical control associated with the TATA box substitute.

Mutations LS-4839 and LS-3829, which affect constitutive transcription from both EIIaE1 and EIIaE2 start sites, have less effect on EIIaE2 transcription in the presence of pEIASV than in its absence (see Figure 4). The significance of this observation is not clear but it suggests that these sequences may also contribute to the EIa-responsiveness of the EIIaE2 promoter and further stresses the differential requirement of the two promoters for their EIa-mediated response.

## Discussion

### *Requirements for constitutive transcription from the EIIaE promoter*

Three essential elements can be identified in the EIIaE1 and EIIaE2 promoter: one component is promoter-specific and is located ~30 bp upstream from the respective start sites; a second element, common to both promoters, has been mapped between positions −48 and −36; the third element, located between positions −91 and −62, is apparently bipartite, with the sequences delimited by positions −91 to −70 shared by both.

The element centered at position −30 bears no similarity to the TATA consensus sequence found at the equivalent position in most eucaryotic protein-coding genes (Breathnach and Chambon, 1981), yet it is essential for efficient transcription from the EIIaE1 start site. On the other hand, mutation of the corresponding sequence for EIIaE2, which more closely resembles a canonical TATA box, had a less pronounced effect upon EIIaE2 transcription initiation, but clearly reduced its template activity.

The promoter element around position −30 (5′-GTCCTTAA-3′) shares some sequence homology with an element found in the SV40 late promoter around position −30 (5′-GTACCTAA-CC-3′); the requirement of the latter element for promoter function has been contested. Piatak *et al.* (1983) showed that viral mutants in the sequence >19 nucleotides upstream from the major late start site did not alter the specificity or efficiency of transcription initiation, while Brady *et al.* (1982) demonstrated a clear effect of the region around −30 for both *in vivo* and *in vitro* transcription of plasmid constructs bearing this promoter. In addition, Brady *et al.* (1982) have shown that alteration of this element to a sequence (5′-GTATTTAATT-3′), more similar to the consensus TATA box, resulted in improved promoter efficiency. In this respect it is surprising that the EIIaE2 element at position −59 (5′-TTAAATTT-3′), which resembles the modified SV40 sequence, belongs, in fact, to the least effective of the two EIIaE promoters as indicated by the relative transcription from the two start sites.

Accumulating evidence indicates that eucaryotic promoter elements correspond to DNA-binding sites for specific transcription factors, which, together with the RNA polymerase, contribute to the formation of stable and active transcription complexes (Dynan and Tjian, 1983; Parker and Topol, 1984; Schöler and Grüss, 1984; Sassone-Corsi *et al.*, 1984; Wildeman *et al.*, 1984; Miyamoto *et al.*, 1984). The observation that alteration of either of the two TATA substitutes in the EIIaE TU resulted in a reproducibly increased level of transcription from the start sites controlled by the unaltered one implied that similar factors interact with both elements. The different strengths of the EIIaE1 and EIIaE2 promoters might therefore reside in the differential affinity of the two TATA box substitutes for the same factor(s).

The promoter element located between these two TATA substitutes is required for the optimal expression of both EIIaE1 and EIIaE2, although it has a greater influence upon transcription from EIIaE1. A sequence homologous to this GC-rich region has been noted in analogous locations for other adenoviral genes, including EIB, EIII, EIV and the major late TU (Bos and ten Wolde-Kraamwinkel, 1983), but in these cases no studies directly demonstrate an involvement of this element in transcriptional efficiency.

Analysis of both linker-scanning and internal deletion mutations revealed that a sequence, critical for EIIaE1 promoter strength and dispensable for EIIaE2 activity, lies in the T-rich stretch located between −70 and −62. A similar element (5′-G$^{CA}_{TG}$G$^{C}_{T}$-TTTC-3′), is also found at equivalent positions in the Ad5 EIII and EIV promoter regions (Weeks and Jones, 1983), but in these cases its functional significance is unknown. This short motif (between −70 and −62) flanks an element further upstream (between −90 and −70), which is absolutely required for efficient function of both EIIaE promoters. The existence of an upstream element shared by two overlapping promoters has already been described in SV40. Two out of the six GC-rich motifs of the 21-bp repeats of the SV40 early promoter (those most proximal to the TATA box) are required only for transcription from the early-early start sites, while the four others are involved in the transcription from both the early-early and late-early start sites (Baty *et al.*, 1984). It is also noteworthy that the sequence, 5′-GTGGCC-3′, between −90 and −85 in the EIIaE promoter is homologous to an essential upstream promoter element of the rabbit β-globin gene (between −80 and −74; Grosveld *et al.*, 1982; Dierks *et al.*, 1983) and the Ad2 major late promoter (between −66 and −60; Hen, Wintzerith and Chambon, unpublished results). However, *in vitro* transcription competition studies between the EIIaE sequences upstream from −64 and the Ad2 major late promoter have suggested that different factors interact with these two upstream elements (Miyamoto *et al.*, 1984).

Imperiale and Nevins (1984) have ascribed 'enhancer' activity to sequences located between positions −21 and −262 in the EIIa promoter region. Further experiments are required to precisely delimit this enhancer element and to establish its possible identity with one of the upstream elements we describe.

### *Different sequence requirements for EIa-dependent stimulation of EIIaE1 and EIIaE2 transcription*

Analysis of the effect of a variety of clustered point mutations (LS-mutants) and short deletions (Δ-mutants) on EIIaE1 promoter activity revealed that no unique sequence elements between positions −97 and +1 are required for the EIa-dependent stimulation of EIIaE1 transcription. Using a stable transformation assay to analyze the sequence requirements for the EIa-mediated stimulation of an integrated EIIaE-dihydrofolate reductase fusion gene, Kingston *et al.* (1984) inferred that positive regulation of EIIaE1 transcription by EIa does not depend upon the presence

of a set of specific promoter sequences. Similar conclusions have been reached for the Ad5 EIII transcription unit, by Leff *et al.* (1985), who have shown that internal deletions removing significant portions of the promoter, including sequences essential for constitutive expression, do not alter the EIa-dependent stimulation.

Our observation that the mutation in the TATA-like element of the EIIaE2 promoter (LS-6052) always decreased and often abolished the EIa effect upon EIIaE2 transcription indicates that stimulation by EIa can be sequence-specific. That this sequence requirement for the EIa-mediated stimulation applies uniquely to the EIIaE2 promoter is indicated by the unaltered level of stimulation of EIIaE1 transcription in this mutant (LS-6052), compared with the wild-type (LS WT). Minor effects on the EIIaE2 stimulation have also been observed for mutants Δ-9170, LS-4839 and LS-3829. It is of interest that a sequence homologous to a portion of the sequence defined by Δ-9170 (5′-GGAGATGACGTA-3′ located between positions −81 and −68) is responsible for the EIa-mediated stimulation of EIV transcription (5′-GGAAGTGACGTA-3′ localized between −179 and −158 in the EIV upstream region; Gilardi and Perricaudet, 1984). The TATA box, but not the upstream promoter sequences of the human β-globin gene has also been shown to be required for transcription in the presence of EIa (Green *et al.*, 1983), in experiments where deletion and single point mutants were transfected into 293 cells (a cell line transformed with the left end of Ad5).

## Conclusion

Detailed analysis of the adenovirus EIIaE promoter region has revealed the existence, within a 100-bp DNA segment, of two intricate control regions. Constitutive transcription from the EIIaE1 and EIIaE2 start sites requires both common and separate promoter elements, suggesting the involvement of common as well as distinct transcription factors. Although an *in vivo* requirement for EIIaE2 activity has yet to be established, it is not impossible that transcripts originating at the −26 start site belong to a transcription unit different from that encoding the 72-kd DNA-binding protein, and that they correspond to the overlapping EIIb TU encoding the viral pre-terminal protein (Stillman *et al.*, 1981) and DNA polymerase (Stillman *et al.*, 1982).

The sequence requirements for the EIa-caused stimulation of these two transcription units are evidently different. For the EIIaE1 promoter, no unique, discrete sequence element is responsible for assuring the EIa-dependent stimulation of transcription. We cannot exclude, however, that the entire promoter region provides this function if the effect of limited promoter alteration is, in each case, balanced by the surrounding wild-type sequences. On the other hand, positive control of EIIaE2 transcription by the EIa products is dependent upon a specific sequence including the TATA-like element.

Thus from our results and available data, it appears that EIa-responsive genes could be classified according to their sequence requirements for EIa-dependent activation: transcription units for which a unique sequence is implicated (e.g., the adenoviral EIIaE2 and EIV genes and perhaps the human β-globin gene) and those which lack such a sequence (e.g., the adenoviral EIIaE1 and EIII genes). Whether or not the transcription of cellular genes, such as the rat pre-proinsulin (Gaynor *et al.*, 1984), the human heat shock gene (Nevins, 1982), rabbit β-globin (Svensson and Akusjårvi, 1984), and human β-tubulin (Stein and Ziff, 1984), all reported to be increased by EIa products, will have similar requirements, remains to be established.

## Materials and methods

*Construction of linker scanning mutants*

Standard procedures for restriction enzyme digestions, recombinant DNA plasmid preparations, and transformation of *Escherichia coli* (C600 or JM103) were employed.

To generate mutants with deletions extending in a 5′ to 3′ direction from the 5′ side of the Ad2 EIIaE promoter region (5′-deletion mutants), the *Bam*HI-*Hin*dIII fragment from pBX-94 (Elkaim *et al.*, 1983) containing the EII sequences between −94 and +719 (with respect to the EIIaE1 start site) was cloned between the *Bam*HI and *Hin*dIII sites of the polylinker in the mp8 derivative of M13. This recombinant (10 μg) was linearized with *Bam*HI and digested for varying times with 0.1 − 1.0 unit of *Bal*31 exonuclease (BRL, Inc.) essentially as described (Elkaim *et al.*, 1983). A DNA sample from each time-point was sized by restriction enzyme analysis and those which represented deletions spanning the region between −94 and +40 were pooled, treated with bacterial alkaline phosphatase and ligated to a 100-fold molar excess of phosphorylated *Xba*I linkers -5′CTCTAGAG3′- and used to transform JM103. Double-stranded replicative form (RF) DNA was prepared and screened for the presence of the *Xba*I site. The selected recombinants (the Mp8EII5′ series, see Figure 1) were sequenced (Sanger *et al.*, 1980), using an oligonucleotide complementary to the EIIaE sequence between +71 and +86.

A series of mutants that were deleted in a 3′ to 5′ direction (3′-deletion mutants) were generated in a similar manner. The *Sma*I-*Sau*3A fragment of the EIIaE promoter region extending from positions −250 to +40 was cloned between the *Sma*I and *Bam*HI sites in the polylinker of the mp9 derivative of M13. The resulting recombinant was linearized at the *Sal*I site in the polylinker which flanks the lost *Bam*HI site of the insert and ligated to a *Xba*I linker after varying times of *Bal*31 digestion, as described above. Precise positioning of the deletion end-points of selected recombinants (the Mp9EII3′ series, see Figure 1) was achieved by sequencing (Sanger *et al.*, 1980), using as primer either a synthetic oligonucleotide complementary to the EIIaE sequence between −100 and −117 (for short deletions) or the M13 universal pentadecamer primer (for longer deletions).

In order to create recombinants differing from the 'wild-type' sequence only by small clustered nucleotide substitutions, the linker-scanning method first developed by McKnight and Kingsbury (1982) was applied. The linker scanning mutants were constructed by recombining, at the *Xba*I site, purified *Hin*dIII-*Xba*I fragments from the deletions in Mp8EII5′ with the complementary *Xba*I-*Sma*I fragments from the Mp9EII3′ deletion series. Matching pairs of deletion fragments were chosen so as to preserve, within 1 bp, the original length of the *Sma*I-*Hin*dIII fragment in the resulting recombinants. These fragments were mixed with the *Hin*dIII-*Sma*I cut pMT5 vector (a chimeric β-globin recombinant, see legend to Figure 1) and ligated in the same reaction to generate the LS series of mutants carrying the EII sequences between −250 and +719 with the *Xba*I linker sequence replacing the wild-type sequence at varying positions (see Figure 1). The 'wild-type' construct, LS WT, contains the EIIaE *Hin*dIII-*Sma*I fragment (extending from position +719 to position −250; Elkaim *et al.*, 1983) inserted between the unique *Hin*dIII and *Sma*I sites of pMT5.

A set of limited internal deletions was constructed as described for the LS series by recombining appropriate pairs of 5′ and 3′ deletion mutants chosen to generate the desired deletions (Δ-mutants). Each of these LS and Δ-mutants was sequenced (Maxam and Gilbert, 1977) and those used in this study are represented in Figure 2.

*In vivo expression assay*

HeLa cells were transfected, by calcium-phosphate co-precipitation (Banerji *et al.*, 1981), with the indicated amounts of recombinant DNA (see Figure legends), adjusted to 20 μg per 10 cm Petri dish with M13 mp8 RF DNA.

Cytoplasmic RNA was prepared 12 − 16 h after removal of the calcium phosphate DNA precipitate and equivalent RNA samples (5 − 10 μg), quantitated by $A_{260}$ determination, were hybridized with an excess of the appropriate single-stranded $^{32}$P-labeled probes, and digested with S1 nuclease (Boehringer, Mannheim) as previously described (Mathis *et al.*, 1981). RNA-protected probe fragments were sized by 8% polyacrylamide urea gel electrophoresis and quantitation of S1 nuclease analysis was achieved by scanning of the autoradiograms (linearity of film exposition with absorbance was verified).

## Acknowledgements

We are grateful to M.Wintzerith, H.Matthes and A.Staub for synthesizing the sequencing and reverse transcriptase primers, M.Zenke and T.Grundstöm for the pMT5 vector, P.Jalinot for the Mp9EII clone, R.Elkaim for the pBX-94 clone, M.Gilbert and J.L.Weickert for tissue culture and to C.Werlé, B.Boulay, C.Kustchis and C.Aron for preparing the manuscript. A special thanks goes to C.Hauss, whose expert technical assistance was very welcome. For the discussions with Drs P.Chambon and C.Goding, we are very appreciative. This work was supported by the CNRS (ATP 3582), the Ministère de l'Industrie et de la Recherche

(82V1283 and 84V0803), the Association pour le Développement de la Recherche sur le Cancer and the Fondation Simone and Cino del Duca.

## References

Allan,M., Zhu,J., Montague,P. and Paul,J. (1984) *Cell*, **38**, 399-407.
Babich,A. and Nevins,J.R. (1981) *Cell*, **26**, 371-379.
Baker,C.C. and Ziff,E.B. (1981) *J. Mol. Biol.*, **149**, 189-221.
Banerji,J., Rusconi,S. and Schaffner,W. (1981) *Cell*, **27**, 299-308.
Baty,D., Barrera-Saldana,H.A., Everett,R.D., Vigneron,M. and Chambon,P. (1984) *Nucleic Acids Res.*, **12**, 915-931.
Benoist,C. and Chambon,P. (1981) *Nature*, **290**, 304-310.
Bos,J.L. and ten Wolde-Kraamwinkel,H.C. (1983) *EMBO J.*, **2**, 73-76.
Brady,J., Radonovich,M., Vodkin,M., Natarajan,V., Tgoren,M., Das,G., Janik, J. and Salzman,N.P. (1982) *Cell*, **31**, 625-633.
Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **50**, 349-383.
Cattaneo,R., Will,H., Hernandez,N. and Schaller,H. (1983) *Nature*, **305**, 336-338.
Carter,T.H. and Blanton,R.A. (1978) *J. Virol.*, **25**, 664-674.
Corden,J., Wasylyk,B., Buchwalder,A., Sassone-Corsi,P., Kédinger,C. and Chambon,P. (1980) *Science (Wash.)*, **209**, 1406-1414.
DeVilliers,J. and Schaffner,W. (1981) *Nucleic Acids Res.*, **9**, 6251-6264.
Dierks,P., van Ooyen,A., Cochran,M.D., Dobkin,C., Reiser,J. and Weissman,C. (1983) *Cell*, **32**, 695-706.
Dynan,W.S. and Tjian,R. (1983) *Cell*, **32**, 669-680.
Elkaim,R., Goding,C. and Kédinger,C. (1983) *Nucleic Acids Res.*, **11**, 7105-7117.
Feldman,L.F. and Nevins,J.R. (1983) *Mol. Cell. Biol.*, **3**, 829-838.
Flint,S.J. (1982) *Biochem. Biophys. Acta*, **651**, 175-208.
Gaynor,R.B., Hillman,D. and Berk,A.J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1193-1197.
Gilardi,P. and Perricaudet,M. (1984) *Nucleic Acids Res.*, **12**, 7877-7888.
Green,M.R., Treisman,R. and Maniatis,T. (1983) *Cell*, **35**, 137-148.
Groner,H., Kennedy,N., Skroch,P., Hynes,N.E. and Ponta,H. (1984) *Biochim. Biophys. Acta*, **781**, 1-6.
Grosschedl,R. and Birnstiel,M. (1980) *Proc. Natl. Acad. Sci. USA*, **79**, 1431-1436.
Grosveld,G.C., Rosenthal,A. and Flavell,R.A. (1982) *Nucleic Acids Res.*, **10**, 4951-4971.
Hashimoto,S. and Green,M. (1984) *Nucleic Acids Res.*, **12**, 9067-9082.
Hen,R., Sassone-Corsi,P., Corden,J., Gaub,M-P. and Chambon,P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7132-7136.
Imperiale,M.J., Feldman,L.T. and Nevins,J.R. (1983) *Cell*, **35**, 127-136.
Imperiale,M.J. and Nevins,J.R. (1984) *Mol. Cell. Biol.*, **4**, 875-882.
Karin,M., Haslinger,A., Hotlgreve,H., Richards,R.I., Krauter,P., Westphal,H.M. and Beato,M. (1984) *Nature*, **308**, 513-519.
Khoury,G. and Gruss,P. (1983) *Cell*, **33**, 313-314.
Kingston,R.E., Randal,J.K. and Sharp,A.P. (1984) *Mol. Cell. Biol.*, **4**, 1970-1977.
Leff,T., Elkaim,R., Goding,C.R., Jalinot,P., Sassone-Corsi,P., Perricaudet,M., Kédinger,C. and Chambon,P. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 4381-4385.
Leff,T., Corden,J., Elkaim,R. and Sassone-Corsi,P. (1985) *Nucleic Acids Res.*, **13**, 1209-1221.
Mathis,D.J., Elkaim,R., Kédinger,C., Sassone-Corsi,P. and Chambon,P. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 7383-7387.
Maxam,A.M. and Gilbert,W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560-564.
McKnight,S.L. and Kingsbury,R. (1982) *Science (Wash.)*, **217**, 316-324.
Melton,D.W., Konecki,D.S., Brennand,J. and Caskey,C.T. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2147-2157.
Mirault,M.E., Southgate,R. and Delwart,E. (1982) *EMBO J.*, **1**, 1279-1285.
Miyamoto,N.G., Moncollin,V., Wintzerith,M., Hen,R., Egly,J.M. and Chambon,P. (1984) *Nucleic Acids Res.*, **12**, 8779-8799.
Nevins,J.R., Ginsberg,H.S., Blanchard,J-M., Wilson,M.C. and Darnell,J.E. (1979) *J. Virol.*, **32**, 727-735.
Nevins,J.R. (1982) *Cell*, **29**, 913-919.
Parker,C.S. and Topol,S. (1984) *Cell*, **37**, 273-283.
Pelham,H.R.B. (1982) *Cell*, **30**, 517-528.
Piatak,M., Ghosh,P.K., Norkin,L.C., Sherman,M. and Weissman,C. (1983) *J. Virol.*, **48**, 503-520.
Reynolds,G.A., Basu,S.K., Osborne,T.F., Chin,D.J., Gil,G., Brown,M.S., Goldstein,J.L. and Luskey,K.L. (1984) *Cell*, **38**, 279-285.
Rossini,M. (1983) *Virology*, **131**, 49-58.
Sanger,F., Coulson,A.R., Barrell,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161-178.
Sassone-Corsi,P., Wildeman,A. and Chambon,P. (1984) *Nature*, **313**, 458-463.
Schöler,R.H. and Gruss,P. (1984) *Cell*, **36**, 403-411.
Stein,R. and Ziff,E.B. (1984) *Mol. Cell. Biol.*, **4**, 2792-2801.
Stillman,B.W., Lewis,J.B., Chow,L.T., Matthews,M.B. and Smart,J.E. (1981) *Cell*, **31**, 613-623.
Stillman,B.W., Tamanoi,F. and Matthews,M.B. (1982) *Cell*, **31**, 613-623.
Svensson,C. and Akusjärvi,G. (1984) *EMBO J.*, **3**, 789-794.
Weeks,D.L. and Jones,N.C. (1983) *Mol. Cell. Biol.*, **3**, 1222-1234.
Wieringa,B., Meyer,F., Reiser,J. and Weissmann,C. (1983) *Nature*, **301**, 38-43.
Wildeman,A.G., Sassone-Corsi,P., Grundström,T., Zenke,M. and Chambon,P. (1984) *EMBO J.*, **3**, 3129-3133.
Yaniv,M. (1982) *Nature*, **297**, 17-18.

*Received on 14 January 1985*