The EMBO Journal vol.4 no.13A pp.3563-3570, 1985

Specific interaction between a transcription factor and the upstream element of the adenovirus-2 major late promoter

N.G.Miyamoto[1], V.Moncollin, J.M.Egly and P.Chambon

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

[1]Present address: Division of Biological Research, The Ontario Cancer Institute, and Department of Medical Biophysics, University of Toronto, 500 Sherbourne Street, Toronto, Ontario M4X 1K9, Canada

Communicated by P.Chambon

Stimulation of *in vitro* transcription mediated by the upstream element of the adenovirus-2 major late promoter (Ad2MLP) involves its recognition by a specific *trans*-acting factor present in a HeLa whole-cell extract. DNase I footprinting and dimethylsulfate methylation protection experiments were used to determine, at the nucleotide level, upstream sequences which interact with this transcription factor. The ability of upstream element mutants to bind the transcription factor correlates directly with the efficiency of transcription from the corresponding Ad2ML promoters *in vivo* and *in vitro*. Competition footprinting experiments show that the transcription factor, which binds to the upstream element of the Ad2MLP, can also interact, but with a lower affinity, with the upstream elements of the Ad2E2a and rabbit β-globin promoters, both of which display some sequence homology to the 'interacting' region of the Ad2MLP upstream element. The transcription factor does not, however, interact with the upstream elements of either the Ad2E2L, Ad5E3, SV40 early, herpes virus thymidine kinase or chicken conalbumin promoters.

Key words: adenovirus-2/major late promoter/transcription factor/transcription efficiency

Introduction

Several sequence elements located upstream from the mRNA start site are required *in vivo* for accurate and efficient transcription of eucaryotic protein-coding genes by RNA polymerase B (II) (for references and reviews, see Benoist and Chambon, 1981; Breathnach and Chambon, 1981; Hen *et al.*, 1982; Dierks *et al.*, 1983; Baty *et al.*, 1984; Chambon *et al.*, 1984; McKnight *et al.*, 1984; Wasylyk, 1985). The *in vivo* effect of these elements has been reproduced *in vitro* for a number of RNA polymerase B promoters, using either S-100 (Weil *et al.*, 1979), whole cell (Manley *et al.*, 1980) or nuclear (Dignam *et al.*, 1983) extracts of HeLa cells. The TATA box located 25-30 bp upstream from the start site is required for both accurate and efficient *in vitro* transcription (Corden *et al.*, 1980; Wasylyk *et al.*, 1980; Breathnach and Chambon, 1981; Hu and Manley, 1981; Grosveld *et al.*, 1981; Lee *et al.*, 1982; Wasylyk *et al.*, 1983). Elements located further upstream (40-110 bp from the start site) are required for maximally efficient *in vitro* transcription from the promoters of a variety of viral and cellular genes (Tsuda and Suzuki, 1981; Grosschedl and Birnstiel, 1982; Hen *et al.*, 1982; Hansen and Sharp, 1983; Jove and Manley, 1984; Miyamoto *et al.*, 1984; Vigneron *et al.*, 1984; Barrera-Saldana *et al.*, 1985). Stimulation of *in vitro* transcription by the SV40 enhancer element has also been documented (Sassone-Corsi *et al.*, 1984, 1985; Wildeman *et al.*, 1984). Recent competition and footprinting experiments, using either crude extracts of HeLa cells or partially purified transcription factors, showed that these various promoter elements may be specifically recognized by transcription factors *in vitro*: (i) the TATA box element appears to bind a general transcription factor (Davison *et al.*, 1983; Parker and Topol, 1984a) in the presence of another general stimulatory transcription factor, 43-kd STF (Egly *et al.*, 1984); (ii) the upstream elements of the SV40 early promoter (21-bp repeat region) (Dynan and Tjian, 1983; Gidoni *et al.*, 1984; Barrera-Saldana *et al.*, 1985), the *Drosophila* heat shock *hsp70* gene (Parker and Topol, 1984b) and the adenovirus-2 major late promoter (Ad2MLP) (Miyamoto *et al.*, 1984) bind specific factors required for efficient *in vitro* transcription; (iii) the SV40 enhancer element appears to bind specific *trans*-acting factors (Wildeman *et al.*, 1984, 1985; Sassone-Corsi *et al.*, 1985).

We have reported that sequences located upstream from the TATA box of the Ad2MLP, between -34 and -97, are required for efficient transcription *in vivo* in a transient expression assay and *in vitro* using a HeLa whole cell extract (WCE) (Hen *et al.*, 1982). Results obtained in our laboratory (Hen *et al.*, 1982; Miyamoto *et al.*, 1984) and that of Manley and co-workers (Jove and Manley, 1984; Yu and Manley, 1984), have defined an upstream element for the Ad2MLP located between -51 and -66. An *in vitro* competition assay was used to demonstrate that the stimulation of transcription mediated by the upstream element of the Ad2MLP involves a specific factor, since preincubation of a WCE with a DNA fragment containing the wild-type, but not the mutated upstream sequence resulted in decreased transcription from an Ad2MLP template containing the upstream element (Miyamoto *et al.*, 1984).

We have purified a transcription factor, free of any other known RNA polymerase B transcription factor, which is required for the stimulation of transcription mediated by the upstream element of the Ad2MLP (Moncollin *et al.*, 1985). Using both DNase I footprinting and dimethyl sulfate (DMS) methylation protection experiments, we show here that this transcription factor strongly interacts with specific sequences in the upstream element of the Ad2MLP. Competition footprinting experiments also reveal that the upstream elements of the promoters of the Ad2E2a and rabbit β-globin, but not the Ad2E2L, Ad5E3, SV40 early, herpes virus thymidine kinase (TK) or chicken conalbumin, transcription units, interact weakly with this transcription factor.

Results

The upstream element of the Ad2MLP is required for efficient transcription in vitro using a reconstituted system

We have previously used an *in vitro* competition assay to prove that efficient transcription from the Ad2MLP involves the stable binding of a specific *trans*-acting factor(s) present in a HeLa WCE to sequences upstream from the TATA box (Miyamoto *et al.*,

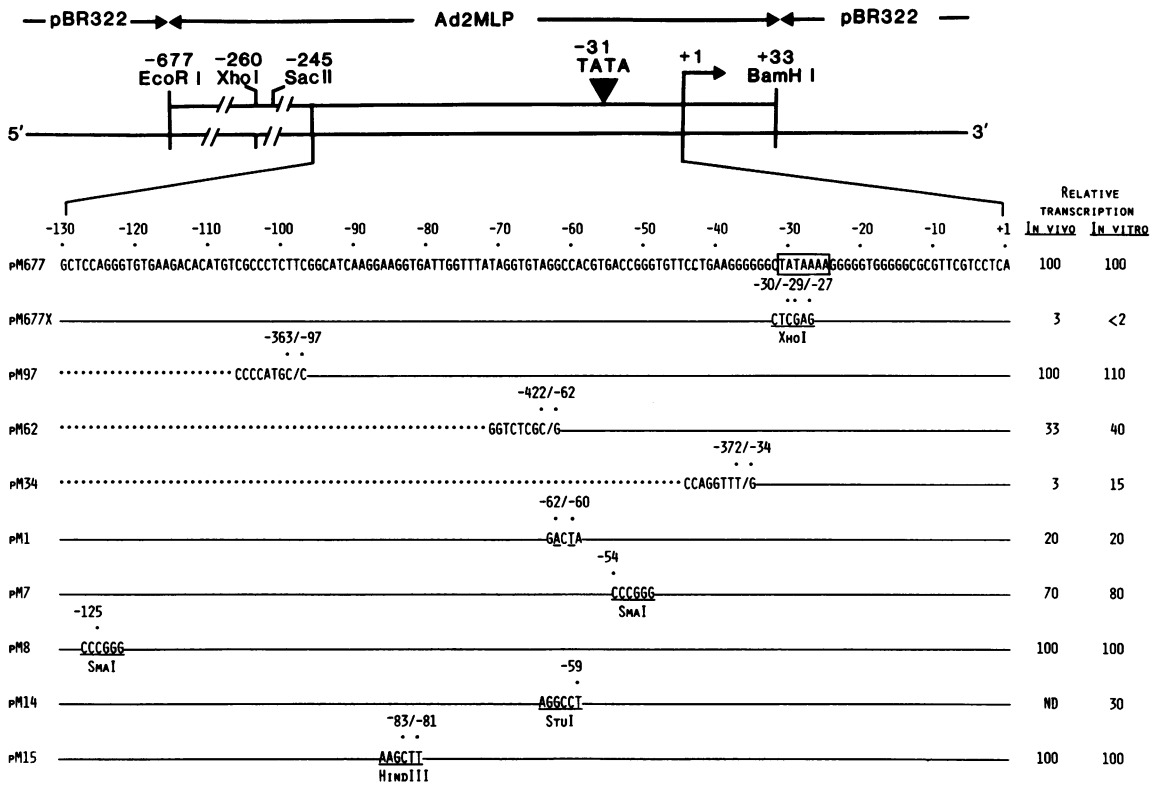

Fig. 1. Sequence and transcription efficiency of the pM series of Ad2MLP-containing recombinants. The nucleotide sequence of the non-coding strand of the Ad2MLP between −130 and +1 is shown (Akusjarvi and Pettersson, 1979) for recombinant pM677 and a series of upstream sequence-mutants (Materials and methods). The transcription efficiency of these recombinants was determined *in vivo* using a transient expression assay (Hen *et al.*, 1982; Hen, 1985) and *in vitro* by run-off transcription using a HeLa WCE (Hen *et al.*, 1982; Miyamoto *et al.*, 1984; and data not shown). Transcription efficiency is expressed as a percentage of the level of wild-type Ad2MLP recombinant pM677 transcription, taken as 100%. N.D.: not determined.

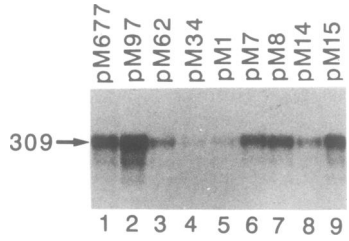

Fig. 2. Requirement for upstream sequences of the Ad2MLP for efficient transcription *in vitro*. The wild-type Ad2MLP recombinant pM677 and the series of upstream sequence mutants (Figure 1) were transcribed using *in vitro* reconstituted system as described in Materials and methods. The specific templates (100 ng plasmid), indicated at the top of each lane, were linearized at the *Sal*I site, resulting in 309 nucleotide run-off transcripts. Only the relevant portion of the gel autoradiogram is shown.

1984). Transcription of the wild-type Ad2MLP and a series of recombinants containing deletions or point mutations in the Ad2MLP upstream element (Figure 1) was analysed here using an *in vitro* reconstituted transcription system. This system consists of partially purified transcription factors obtained by chromatographic fractionation of a HeLa WCE on Heparin-Ultrogel and DEAE-5PW (h.p.l.c.) (see Materials and methods and Moncollin *et al.*, 1985). In addition to RNA polymerase B and 43-kd STF (Egly *et al.*, 1984), two fractions from the DEAE-5PW (h.p.l.c.) column the 150 mM (DE150) and 250 mM (DE250) KCl salt-eluted fractions are absolutely required to reconstitute transcription *in vitro* from the Ad2MLP. The DE150 fraction contains at least two general transcription factors and the factor which specifically stimulates transcription from the Ad2MLP by interacting with the upstream element (Moncollin *et al.*, 1985 and see below). The DE250 fraction contains the TATA box factor (J.M.Egly and T.Tamura, unpublished data).

That sequences located upstream from the TATA box are required for efficient *in vitro* transcription using the reconstituted system is shown by the correlation between the position of the various upstream sequence deletion and point mutations and the relative transcription efficiencies of the corresponding templates (Figures 1 and 2). In particular, recombinants containing point mutations located in the upstream element of the Ad2MLP (pM1 and pM14) are transcribed less efficiently than the wild-type Ad2MLP, pM677 (Figure 2, compare lanes 5 and 8 with lane 1; see also the deletion mutants pM62 and pM34, lanes 3 and 4). Transcription from mutant pM7 was also decreased *in vivo* and *in vitro*, but to a lower extent (the values given in Figure 1 correspond to several experiments which gave very similar results). In contrast, recombinants with point mutations located further upstream (pM8 and pM15) are transcribed at the wild-type level (Figure 2, compare lanes 7 and 9 with lane 1 and lane 2). Thus, the reconstituted *in vitro* system faithfully reproduces the upstream element requirement for efficient transcription from the Ad2MLP previously observed *in vivo* and *in vitro* using a HeLa WCE (Figure 1).

Specific protein interactions on the Ad2MLP upstream element

The possibility that the stimulation of *in vitro* transcription mediated by the upstream element of the Ad2MLP involves the direct interaction of a protein factor(s) present in the DE150 fraction with specific sequences in the upstream element was in-

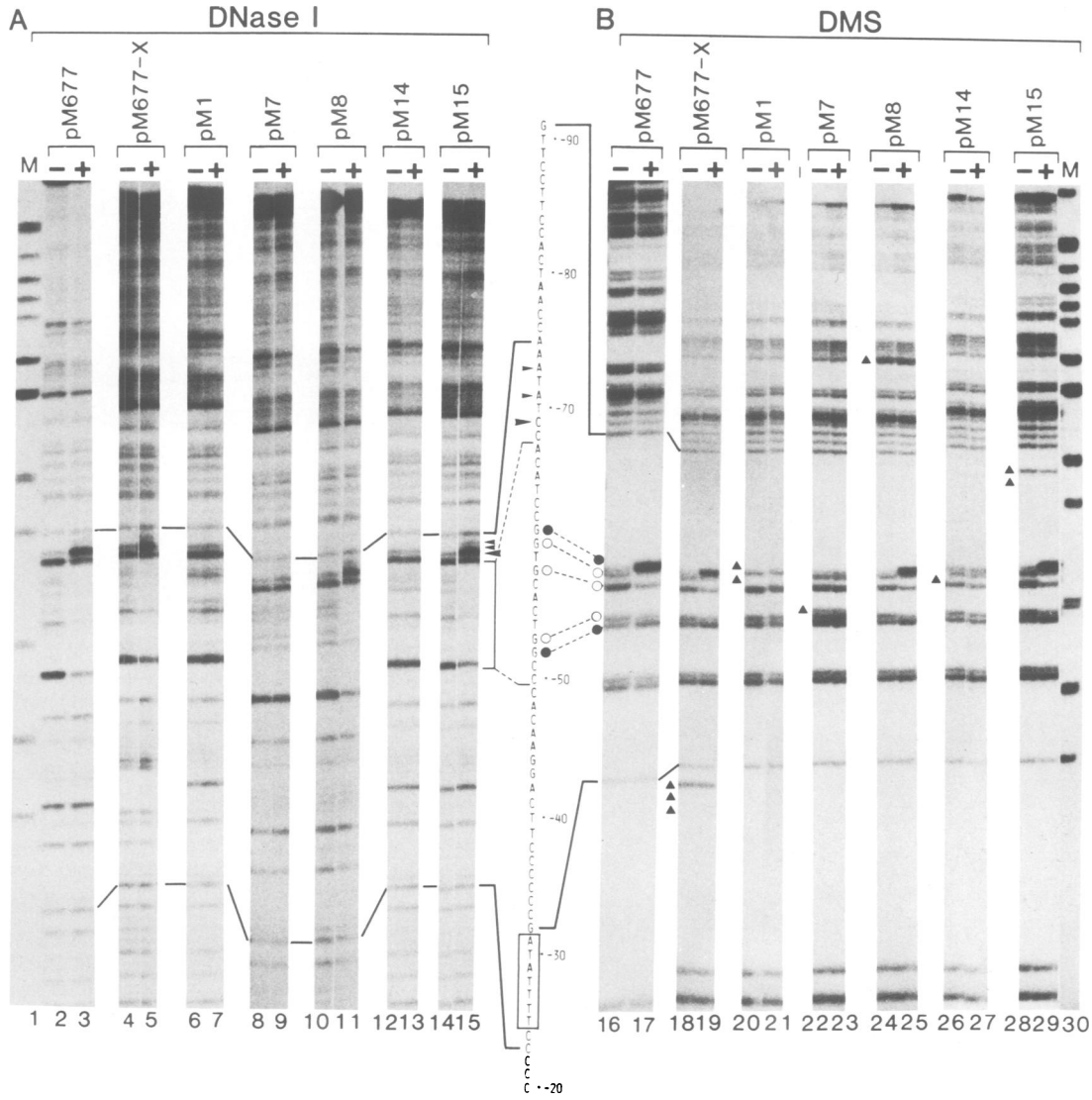

Fig. 3. Protein interaction with the coding strand of the Ad2MLP. DNase I footprinting (**A**) and DMS methylation protection (**B**) experiments using the wild-type Ad2MLP recombinant pM677 and the upstream sequence point mutants, ^{32}P 5′ end-labelled on the coding strand, were as described in Materials and methods. Footprinting reactions for each Ad2MLP recombinant, as indicated, were carried out in the absence (−) or presence (+) of either 2 μl or 4 μl of the DE150 fraction for DNase I or DMS footprinting, respectively. Arrows indicate an increase in DNase I digestion (hypersensitive sites) and the solid line indicates the region protected from DNase I digestion in the presence of the protein fraction. Closed or open circles indicate G residues that are either more or less accessible to DMS methylation, respectively, in the presence of the protein fraction. Triangles indicate the position of point mutations for the various Ad2MLP upstream mutants. The DNA sequence of the coding strand of the Ad2MLP between −91 and −20 is shown. The TATA box sequence is denoted by the boxed region. M: size markers, ^{32}P 5′ end-labelled *Msp*I fragments of pBR322.

vestigated using both DNase I footprinting (Galas and Schmitz, 1978) and DMS methylation protection (Gilbert *et al.*, 1976) experiments.

DNase I footprinting and DMS methylation protection on the coding strand. 'Footprinting' templates, ^{32}P 5′ end-labelled on the coding strand (see Materials and methods), were subjected to DNase I footprinting (Figure 3A) in the absence or presence of the DE150 protein fraction. The upstream regions of pM677, pM8 and pM15 between −67 and −50 were selectively protected in the presence of the DE150 fraction, whereas nucleotides at −73, −71 and −69 became hypersensitive to DNase I digestion (Figure 3A, compare lanes 3, 11 and 15 with lanes 2, 10 and 14, respectively); no other changes were reproducibly observed. In contrast, no footprints were observed on the upstream elements of pM1 and pM14 (Figure 3A, compare lanes 7 and 13 with lanes 6 and 12, respectively). The extent of DNase I protection (P) between −67 and −50 (in particular at residues −50) and the DNase I hypersensitivity (HS) at residue −69, constitutes a useful indicator of a specific protein footprint on the Ad2MLP upstream element (see Figure 6). Thus, wild-type DNase I footprints are observed on the upstream element point mutations which display a wild-type phenotype, but not on those recombinants which exhibit a marked mutant phenotype. Although not apparent in the experiment shown (Figure 3A, lanes 8 and 9), we have often observed a weak footprint on pM7 (data not shown), which is in agreement with the observation that pM7 is transcribed with greater efficiency than either pM1 or pM14 *in vitro* (Figure 2).

DMS methylation protection experiments were performed with the same series of Ad2MLP DNA templates (Figure 3B). As in the case of DNase I footprinting (Figure 3A), similar changes

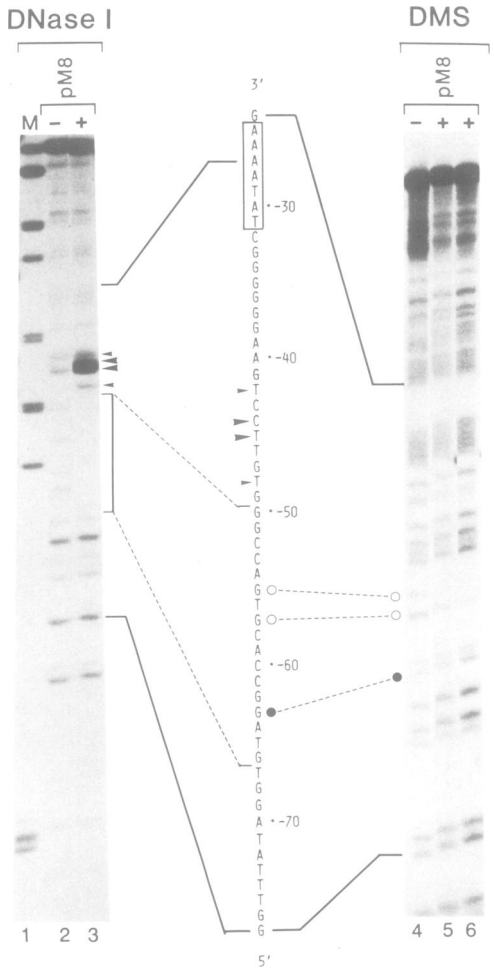

Fig. 4. Protein interaction with the non-coding strand of the Ad2MLP. DNase I footprinting (**left panel**) and DMS methylation protection (**right panel**) experiments, using the Ad2MLP recombinant pM8 ^{32}P 5′ end-labelled on the non-coding strand, were as described in Materials and methods. Reactions were in the absence (−) or presence (+) of 2 μl (**lanes 3** and **5**) or 4 μl (**lane 6**) of the DE150 fraction. Symbols are as in Figure 3. The DNA sequence of the non-coding strand of the Ad2MLP between −77 and −24 is shown.

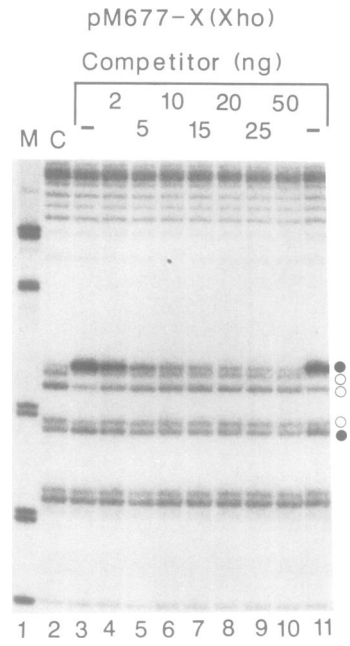

Fig. 5. Competition of DMS methylation protection on the coding strand of the Ad2MLP by the homologous promoter upstream element DNA fragment. DMS methylation protection experiments using 1 ng of the wild-type Ad2MLP recombinant pM677, ^{32}P 5′ end-labelled on the coding strand, were performed (see Materials and methods) after a 10-min pre-incubation of the DE150 fraction with various amounts of the wild-type upstream element-containing DNA fragment pM677-X(*Xho*), as indicated in ng. The total amount of DNA fragment in each reaction was adjusted to 100 ng, by supplementing with pBR322 (RI/RV/*Bam*) fragments. pM677-X(*Xho*) is the *Xho*I fragment (coordinates −260 to −31) of pM677-X (see Figure 1 and Miyamoto *et al.*, 1984) and pBR322(RI/RV/*Bam*) denotes the *EcoRI-Eco*RV (coordinates 0−187) and *Eco*RV-*Bam*HI (coordinates 188−375) fragments of pBR322 (Sutcliffe, 1978). Symbols are as in Figure 3. C: the reaction shown in lane 2 contains no protein.

in the DMS methylation patterns were observed in pM677, pM677-X, pM8 and pM15 in the presence of the DE150 fraction (Figure 3B). The G residues at −61 and (to a lesser extent) −52 were found to be more accessible to, and those at positions −60, −58 and −53 were found to be protected from, methylation by DMS (Figure 3B, compare lanes 17, 19, 25 and 29 with lanes 16, 18, 24 and 28, respectively, for pM677-X see also Figure 5, compare lanes 2 and 3). Also in agreement with the above DNase I footprinting results (Figure 3A), no changes could be discerned in the DMS methylation pattern in the presence of the DE150 fraction for either pM1 (lanes 20 and 21) or pM14 (lanes 26 and 27), whereas very little modification could be seen for pM7 (lanes 22 and 23).

DNase I footprinting and DMS methylation protection on the non-coding strand. To analyze protein-DNA interaction on the non-coding strand of the Ad2MLP, a ^{32}P 5′ end-labelled DNA fragment was prepared from pM8 (see Figure 1 and Materials and methods). pM8, which displays a wild-type phenotype (Figures 1−3), was used as a source of footprint probe because the point mutation present in pM8 creates a *Sma*I (*Xma*I) restriction site at a convenient distance from the upstream element region. The left panel of Figure 4 shows the DNase I digestion pattern on naked DNA (lane 2) and on DNA incubated with the DE150 fraction (lane 3). The upstream region between −66 and −50 is consistently and significantly protected from DNase I digestion, whereas residues at positions −48, −45, −44 and −42 become more sensitive. As in the case of the coding strand footprints (Figure 3A), little or no changes in the DNase I footprints have been observed in the presence of the DE150 fraction on non-coding strand prepared from mutants pM1 and pM14 (data not shown). The DMS methylation patterns of pM8 incubated in the absence (lane 4) or presence (lanes 5 and 6) of the DE150 fraction is shown in the right panel of Figure 4. The G residues at position −55 and, to a lesser extent, at position −57 are significantly protected from methylation by DMS in the presence of the DE150 fraction. In addition, scanning of the autoradiograms from this and similar DMS methylation protection experiments indicate that the G residue at −63 is relatively more exposed to DMS when pM8 is incubated with the DE150 fraction (data not shown; compare also the relative intensities of the Gs at positions −62 and −63 in lanes 4 and 6 of Figure 5).

DNA binding specificity of the Ad2MLP upstream element factor

Pre-incubation of the DE150 fraction and the other required transcription factors with a DNA fragment [pM677-X(*Xho*) fragment, described in the legend to Figure 5] containing the upstream sequence of the Ad2MLP between −260 and −31 resulted in a decrease of the transcription efficiency of a subsequently add-

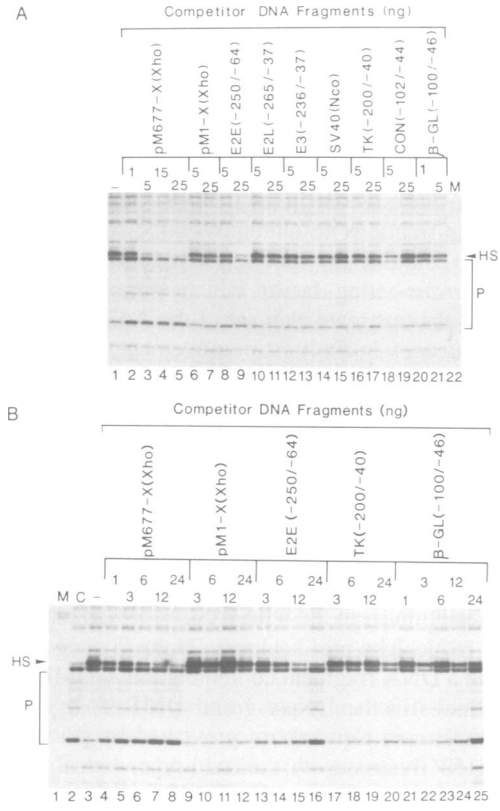

Fig. 6. Competition of DNase I footprinting on the coding strand of the Ad2MLP by homologous and heterologous promoter upstream element DNA fragments. DNase I footprinting using 1 ng of the wild-type Ad2MLP recombinant pM677, ^{32}P 5′ end-labelled on the coding strand, was as described in Materials and methods after a 10-min pre-incubation of either (**A**) 2 μl of the DE150 fraction or (**B**) 4 μl of the HAP120 (see text and Moncollin *et al.*, 1985) fraction with various amounts of competitor 'upstream' DNA fragment, as indicated. The total amount of DNA fragment in each reaction was adjusted to 50 ng, by supplementation with pBR322(RI/RV/*Bam*). Fragments pM677-X(*Xho*) and pBR322(RI/RV/*Bam*) are as in Figure 5. pM1-X(*Xho*) is the *Xho*I fragment of pM1-X, which differs from pM677-X by the introduction of two point mutations at −62 and −60 (see Figure 1 and Miyamoto *et al.*, 1984); E2E(−250/−64) is the *Sma*I-*Xba*I fragment from −250 to −64 of the Ad2 E2a promoter present in LN-64 (Zajchowski *et al.*, 1985); E2L(−265/−37) is the *Sma*I-*Dde*I fragment, which contains the *Fnu*DII-*Dde*I fragment from −265 to −37 of the Ad2 E2L promoter present in pL (Leff and Chambon, 1986); E3(−236/−37) is the *Eco*RI-*Bam*HI fragment, which contains the *Eco*RI-*Xma*I fragment from −236 to −37 of the Ad5E3 promoter present in E3XB (Miyamoto *et al.*, 1984); SV40(*Nco*) is the *Nco*I fragment (SV40 coordinates 333−38), with an *Sph*I deletion from −200 to −128 of the SV40 early promoter present in pHB3 (Vigneron *et al.*, 1984); TK(−200/−40) is the *Cla*I-*Hind*III fragment, which contains the *Pvu*II-*Bst*NI fragment from −200 to −40 of the herpesvirus TK promoter present in pTK$_1$ (Miyamoto *et al.*, 1984); CON(−102/−44) is the *EcoRI-Eco*RV fragment, which contains the *Alu*I-*Bst*NI fragment from −102 to −44 of the chicken conalbumin promoter present in Con60C (Miyamoto *et al.*, 1984); β-GL(−100/−46) is the *Pst*I-*Hind*III fragment from −100 to −46 of the rabbit β-globin promoter present in pDPVU1070 (Dierks *et al.*, 1983) [with the exception of β-GL(−100/−46), which is 55 bp in length, all the DNA fragments used here are of similar size (180−240 bp)]. P and HS indicate regions protected from or hypersensitive to, respectively, DNase I digestion in the presence of the protein fraction. C: the reaction shown in **lane 2** in (**B**) contains no protein.

ed Ad2MLP template containing a wild-type upstream element. In contrast, pre-incubation with a similar DNA fragment [pM1-X(*Xho*) fragment, described in the legend to Figure 5] containing the two point mutations at −62 and −60 in pM1, did not reduce the transcription efficiency of the subsequently added Ad2MLP template (data not shown). These results, and the excellent correlation which exists between the efficiency of *in vivo* and *in vitro* transcription (Figures 1 and 2) from the various Ad2MLP mutant templates and DNase I and DMS 'footprints' on their respective upstream elements (Figure 3), provides strong evidence that the footprints are caused by the transcription factor responsible for stimulation of transcription mediated by the Ad2MLP upstream element. To add further support that the 'footprints' are directly related to the specific binding of the *trans*-acting factor, we studied the ability of the pM677-X(*Xho*) DNA fragment which contains the upstream sequence of the Ad2MLP, but no other functional promoter element such as the TATA box, to compete the DNase I and DMS 'footprints' on the wild-type Ad2MLP recombinant pM677 (Figures 5 and 6). Figure 5 shows that the DMS methylation pattern produced on the upstream element of pM677 in the presence of the DE150 fraction (compare lanes 3 and 11 with lane 2) can be competed by prior incubation of the protein fraction with the pM677-X(*Xho*) fragment in a DNA concentration-dependent fashion (lanes 4−10). The pM1-X(*Xho*) DNA fragment, which contains the same sequence as pM677-X(*Xho*) with the double point mutation present in pM1 and does not exhibit any 'footprint' on its upstream sequence (Figure 3), did not compete at all under identical conditions (data not shown). Similarly, the wild-type pM677-X(*Xho*) fragment, but not the mutant pM1-X(*Xho*) fragment, prevented the DNAse I footprint caused by the factor present in the DE150 fraction (Figure 6A, compare lane 1 with lanes 2−5, and lanes 6 and 7; and see also below, Figure 6B, lanes 2−12).

The *in vitro* competition assay described previously (Miyamoto *et al.*, 1984) indicated that the *trans*-acting factor responsible for the upstream element-mediated stimulation of transcription from the Ad2MLP is promoter-specific. Pre-incubation of the WCE with DNA fragments containing the upstream element sequences from a number of heterologous promoters did not adversely affect transcription from a subsequently added Ad2MLP template containing a wild-type upstream element. To examine further the DNA binding specificity of the Ad2MLP upstream element transcription factor, we performed DNase I competition footprinting experiments with DNA fragments containing the upstream element promoter sequences of Ad2 E2a [E2E(−250/−64)], Ad2 E2L [E2L(−265/−37)], Ad5 E3 [E3(−236/−37)], SV40 early [SV40(*Nco*I), containing the 21- and 72-bp repeat regions], herpes virus thymidine kinase [TK(−200/−40)], chicken conalbumin [CON(−102/−44)] and rabbit β-globin [β-GL(−100/−46)] (Figure 6). Neither the E2L, E3, SV40, TK, CON nor β-GL DNA competitor fragments had any significant effect on the DNase I footprint in the presence of the DE150 fraction (Figure 6A, compare lanes 10−22 with lane 1). However, the E2E(−250/−64) fragment slightly inhibited the footprinting of the transcription factor present in the DE150 fraction on the Ad2MLP upstream element (Figure 6A, compare lanes 8 and 9 with lane 1 and lanes 2−5). Since a DNA sequence homology exists between the upstream elements of the Ad2E2a, rabbit β-globin and Ad2ML promoters (see Elkaim *et al.*, 1983 and Figure 7) in the regions known to be involved in transcription stimulation by these upstream elements *in vivo* and/or *in vitro* (Grosveld *et al.*, 1982; Hen *et al.*, 1982; Dierks *et al.*, 1983; Elkaim *et al.*, 1983; Miyamoto *et al.*, 1984; Zajchowski *et al.*, 1985) we repeated the DNase I competition footprinting using a more highly purified preparation of the Ad2MLP upstream transcription factor.

The HAP120 fraction contains the Ad2MLP upstream element transcription factor activity, but does not appear to contain any

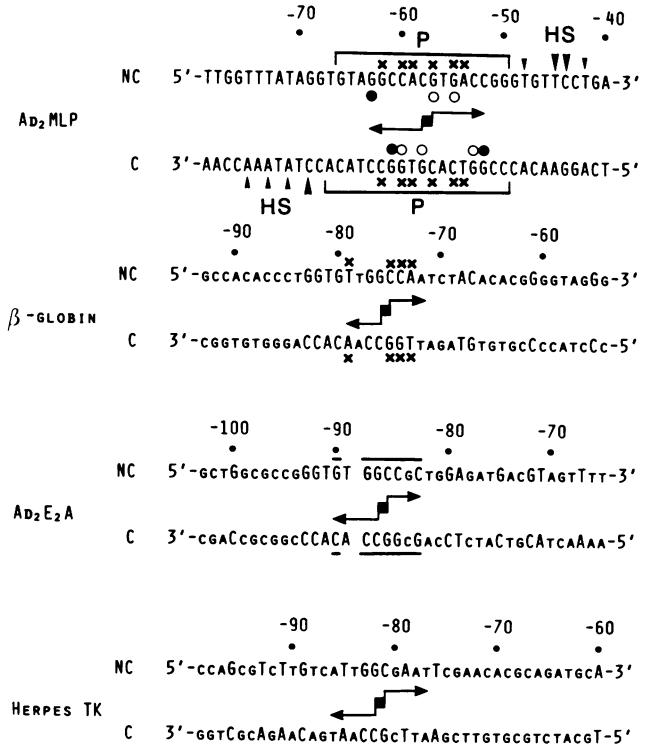

Fig. 7. Summary of DNase I and DMS 'footprinting' on the Ad2MLP and sequence comparisons. The nucleotide sequence of the RNA non-coding (NC) and coding (C) strands of the upstream regions of the Ad2ML, rabbit β-globin, Ad2E2a and TK genes are shown. Sequence homology between the four promoter regions are indicated by capital letters. Symbols used to indicate regions of the Ad2MLP upstream element either protected (P) from or hypersensitive (HS) to DNase I, and G residues that are more (filled circles) or less (open circles) accessible to DMS are as in Figures 3 and 4. The location of mutations known to be detrimental to Ad2MLP activity are indicated by crosses (see text). The bases pointed out by crosses in the β-globin sequences denote the positions of point mutants shown to be important for efficient *in vivo* expression from the β-globin promoter (Grosveld *et al.*, 1982; Dierks *et al.*, 1983). Nucleotide changes in a linker-scanner mutation of the Ad2E2a promoter that dramatically reduces *in vivo* transcription efficiency (Zajchowski *et al.*, 1985) are indicated by lines. The regions exhibiting a dyad symmetry are indicated by arrows pointing in opposite orientations.

other activity required for reconstitution of *in vitro* transcription (Moncollin *et al.*, 1985). Figure 6B shows that the upstream element DNase I footprint obtained with this HAP120 fraction could be competed by the wild-type pM677-X(*Xho*) (lanes 4–8), but not the mutant pM1-X(*Xho*) (lanes 9–12), competitor DNA fragment. The E2E(−250/−64) DNA fragment also inhibited footprinting of the factor (Figure 6B, lanes 13–16). Although the β-GL(−100/−46) DNA fragment did not appear to affect the DNAse I footprint using the DE150 fraction (Figure 6A), a significant competition by the same fragment was observed using the HAP120 fraction (Figure 6B, lanes 21–25). By comparison of the sizes of the competitor DNA fragments (see legend to Figure 6) and the amounts required to observe similar reductions in the DNase I-hypersensitive site at residue −69 or increases in DNase I protection at residue −50, we estimate that the AdE2a and rabbit β-globin upstream elements have binding affinities for the Ad2MLP upstream transcription factor present in the HAP120 fraction which are at least 5- and 20-fold lower, respectively, than the affinity of the Ad2MLP upstream element.

Discussion

Sequence and transcription factor requirements for binding of the Ad2MLP upstream element factor

The requirement for specific sequences upstream from the TATA box of the Ad2MLP for efficient transcription *in vivo* and/or *in vitro* using a HeLa WCE (see Figure 1), has been faithfully reproduced using a reconstituted *in vitro* system (Figure 2). One of the protein fractions absolutely required for efficient *in vitro* transcription, the DE150 fraction (see Moncollin *et al.*, 1985), contains a *trans*-acting factor which stimulates transcription mediated by the upstream element of the Ad2MLP and binds to specific sequences in the upstream element.

Three lines of evidence support the conclusion that the observed DNase I and DMS 'footprints' on the Ad2MLP upstream element are due to the binding of the specific *trans*-acting factor responsible for the upstream element-mediated stimulation of transcription: (i) a strong correlation is observed between the ability of an Ad2MLP mutant template to bind the upstream element factor and to be efficiently transcribed *in vitro* in the reconstituted system (Figures 2 and 3); (ii) the efficiency of *in vitro* transcription from the wild-type Ad2MLP template can be effectively reduced by pre-incubation of the *in vitro* reconstituted system with a DNA fragment containing the Ad2MLP upstream sequence; and (iii) the DNase I and DMS 'footprints' on the Ad2MLP upstream element are prevented by prior incubation of the DE150 fraction with the wild-type, but not a mutated, Ad2MLP upstream element (Figures 5 and 6).

The results of the analysis of interaction between the upstream element factor and sequences of the Ad2MLP, as revealed by DNase I footprinting and DMS methylation protection experiments, are summarized in Figure 7. The region of interaction (from −66/−67 to −50) corresponds to the Ad2MLP upstream element defined by previous *in vivo* and/or *in vitro* transcription analyses using upstream sequence deletion and point mutants of the Ad2MLP. The Ad2MLP point mutants pM1 (mutations at −60 and −62) and pM14 (mutation at −59), which are much less efficiently transcribed *in vitro* than the wild-type pM677, do not interact with the upstream element factor, whereas point mutant pM7 (mutation at −54), whose transcription is also decreased, but to a lesser extent, interacts weakly with the factor (see Figure 1). The results of Yu and Manley (1984), who showed that G to A transitions at either −55 or −57 resulted in a marked decrease in the efficiency of *in vitro* transcription using a HeLa WCE, are also consistent with our footprinting data, which clearly demonstrate that binding of the upstream element factor to the Ad2MLP significantly protects these two G residues from methylation by DMS (Figure 4).

All of these mutations which are detrimental to the activity of the Ad2MLP upstream element are located within a region of dyad symmetry (Figure 7). Prokaryotic regulatory DNA sites that contain dyad sequence signals interact with protein dimers exhibiting 2-fold rotational symmetry (Takada *et al.*, 1983 and references therein). That a protein(s) with such a symmetry does in fact recognize the sequence of dyad symmetry shown in Figure 7 is strongly suggested by the pattern of DMS methylation which allows the identification of specific G residues that form intimate contacts with sequence-specific binding proteins (Gilbert *et al.*, 1976). Most of the G residues which exhibit decreased or enhanced methylation in the presence of the upstream element factor are indeed symmetrically situated in the two halves of the region of dyad symmetry. It is also noteworthy that the Ad2MLP

upstream element is centered at position −57/−58, whereas the TATA box is centered at position −28, i.e., three helical turns downstream. It is thus possible that both the upstream element and TATA box factors interact with their respective recognition sites on the same side of the DNA helix. A very similar spatial arrangement exists in the SV40 early promoter between the TATA box and the upstream GC-motif I to which another specific transcription factor (Sp1) binds (Gidoni *et al.*, 1984; Barrera-Saldana *et al.*, 1985).

Whereas protection of the upstream sequences from cleavage by DNase I indicates the physical blocking of the enzyme by the upstream element factor, the generation of increased susceptibility to DNase I probably involves DNA structural alterations (Fox and Waring, 1984) induced in the vicinity of the binding region by the upstream element factor. The DNase I-hypersensitive sites are located both upstream and downstream of the DNase I-protected region, on the coding and non-coding strands, respectively (Figure 7). Sites of increased susceptibility to DNase I are also induced both upstream and downstream from the SV40 early promoter upstream region (the 21-bp repeat region) upon binding of its cognate factor (Barrera-Saldana *et al.*, 1985; Wildeman *et al.*, 1985).

For most of the footprinting experiments described here, a DE150 fraction was used as source of upstream element factor. This protein fraction also contains two general RNA polymerase B transcription factors (Moncollin *et al.*, 1985). Nevertheless, the Ad2MLP upstream element factor must interact independently with its specific recognition sequence, since identical DNase I and DMS footprints (Figure 6 and data not shown) are obtained using the HAP120 fraction which contains the upstream element factor but no known general transcription factor activity (Moncollin *et al.*, 1985). The same DNase I and DMS footprints were also observed using the HO.6 protein fraction containing RNA polymerase B and all of the transcription factors required for efficient *in vitro* RNA synthesis from the Ad2MLP (see Davison *et al.*, 1983) (data not shown). Furthermore, a functional TATA box region is not absolutely required for binding of the upstream element factor, since wild-type DNase I and DMS footprints are seen on the TATA box mutant pM677-X template (Figure 3) and on a TATA box-depleted template (unpublished observation). Thus the presence of other transcriptional factors does not detectably alter the interaction of the upstream element factor with its specific recognition site. Experiments are in progress to determine whether the stability of the binding of the upstream element factor may be affected by the presence of other transcriptional factors, in particular the TATA box factor and 43-kd STF, which are involved in the formation of a stable preinitiation complex at the Ad2MLP TATA box region (Davison *et al.*, 1983; Egly *et al.*, 1984).

Promoter specificity of the Ad2MLP upstream element factor

As previously noted (Elkaim *et al.* 1983), there is a marked DNA sequence homology [the CAAT box (Benoist *et al.*, 1980)] between the upstream sequences of the Ad2ML, Ad2E2a and rabbit β-globin promoters (Figure 7), in regions important for the efficient expression of these genes *in vivo* and/or *in vitro* (Grosveld *et al.*, 1982; Hen *et al.*, 1982; Dierks *et al.*, 1983; Elkaim *et al.*, 1983; Miyamoto *et al.*, 1984; Zajchowski *et al.*, 1985). There is also a weaker homology with a sequence present in the upstream region of the herpesvirus TK gene (Figure 7) (McKnight *et al.*, 1981). Our present competition footprinting experiments show that the Ad2MLP upstream factor can interact, albeit with a low affinity, with the Ad2E2a and rabbit β-globin promoter upstream sequences; in contrast, using this assay no interaction could be detected with either the Ad2E2L, Ad5E3, SV40 early, herpesvirus TK or conalbumin promoters (Figure 6). These observations are consistent with our 'footprinting' results which show that at least part of the sequence common to the Ad2MLP, Ad2E2a and rabbit β-globin promoters belongs to the binding region for the upstream element factor. However, the homology sequence, GGTGTNGGCCPu, is not sufficient for efficient binding of the upstream element factor, since the upstream region of the Ad2E2a and rabbit β-globin promoters containing this sequence does not stably bind the upstream element factor, as judged by the lack of a DNase I footprint (data not shown). Sequences located downstream from the homology sequence are clearly required in the Ad2MLP for efficient binding and activity of the upstream element factor. In this respect it is noteworthy that the regions of dyad symmetry which are possibly present in the β-globin Ad2E2a and TK upstream regions do not correspond to that identified here in the Ad2MLP upstream element (Figure 7). In addition, a mutation at position −69 in the β-globin upstream element has no significant effect on the globin promoter activity (Dierks *et al.*, 1983). Thus, the Ad2MLP upstream element factor may belong to a family of upstream factors derived from an ancestor whose specificity resided in the homology sequence, whereas the present specificity would be more related to the adjacent sequences.

Materials and methods

Construction of recombinants

The construction of the pM series of Ad2MLP recombinants has been described previously (Miyamoto *et al.*, 1984). Additional Ad2MLP mutants used here (see Figure 1) were similarly constructed. pM62 contains a deletion upstream from −62 (Hen *et al.*, 1982). pM7 and pM8 contain point mutations at positions −54 and −125, generating *Sma*I sites at −52 and −125, respectively; pM14 contains a point mutation at −59, generating a *Stu*I site at −62; pM15 contains two point mutations at −83 and −81, generating a *Hind*III site at −85 (Hen, 1985).

In vitro transcription

The standard *in vitro* transcription run-off assay using a HeLa WCE was as described previously (Miyamoto *et al.*, 1984), except that the amount of specific truncated DNA template was reduced to 100 ng per reaction. *In vitro* transcription using partially purified protein fractions derived from a HeLa WCE is described in detail elsewhere (Moncollin *et al.*, 1985). Briefly protein fractions were pre-incubated for 15 min at 24°C in a 20 μl reaction mixture containing 50 mM Tris-HCl (pH 7.9), 6 mM $MgCl_2$, 0.1 mM EDTA, 0.5 mM dithiothreitol (DTT), 50 mM KCl, 8% glycerol and 100 ng of the specific truncated DNA template. After the pre-incubation period, 4 μl of a mixture containing 6 mM $MgCl_2$, 50 mM KCl, 1.25 mM each of ATP, GTP and UTP and 6.25 μM CTP containing 2−5 μCi [α-^{32}P]CTP was added to start RNA synthesis. The protein fractions utilized in the reconstituted *in vitro* transcription system were: 12 μl and 4 μl of DEAE 150 mM (DE150) and 250 mM (DE250) KCl salt-eluted fractions, respectively, as source of upstream and general transcription factors (Moncollin *et al.*, 1985), 2 μl of a DE0.35 fraction (Egly *et al.*, 1984) as source of 43-kd STF; and 0.2 μl (0.002 units) of partially purified calf thymus polymerase B (Davison *et al.*, 1983). The standard competition assay was as described above, except that 120 ng of competitor DNA fragment was added in the pre-incubation mixture and 100 ng of the specific DNA template was added with the nucleotides to start RNA synthesis. Reactions were stopped after 45 min at 24°C, and the RNA run-off transcripts were purified and analyzed on 5% acrylamide-8.3 M urea gels as described in Davison *et al.* (1983).

DNase I footprinting

The standard DNase I footprinting assay (Galas and Schmitz, 1978) consisted of a 10-min pre-incubation at 24°C in a 18 μl reaction mixture containing variable amounts of the indicated protein fractions, ~1 ng (10 000 c.p.m.) of the labelled DNA template, 50 ng of unlabelled carrier pBR322 DNA fragments, 30 mM Tris-HCl (pH 7.9), 6 mM $MgCl_2$, 50 mM KCl, 0.07 mM EDTA, 0.3 mM DTT and 7% glycerol. After the pre-incubation period, 2 μl of a mixture containing 20 ng of DNase I (Worthington) in the same buffer was added and the reaction mixtures incubated at 24°C for 3 min. The standard competition footprinting assay was as described above except that the 5′ end-labelled specific DNA template was added after the initial pre-incubation period in the presence of variable amounts

of competitor DNA fragment (as indicated in the figure legends) and the reaction mixture was incubated for an additonal 10 min at 24°C prior to DNase I addition. Reactions were stopped at variable times, empirically determined to give 'balanced' DNase I digestion patterns. DNA digestion products were purified by phenol [0.05 M Tris-HCl (pH 7.9)-saturated] extraction followed by ethanol precipitation, and analyzed on 8% acrylamide-8.3 M urea gels followed by autoradiography.

DMS methylation protection

The standard DMS methylation protection assay (Gilbert *et al.*, 1976) was as described above for DNase I footprinting, except that the reaction mixture (50 μl) contained 100 ng of unlabelled carrier pBR322 DNA and was cooled on ice after the pre-incubation period prior to addition of 2 μl DMS. After 1 min at 0−4°C, reactions were stopped and prepared for analysis on 8% acrylamide − 8.3 M urea gels as described in the Maxam and Gilbert DNA sequencing protocol (Maniatis *et al.*, 1982).

Preparation of footprinting templates

DNA templates were unique ^{32}P 5′ end-labelled on one strand, either upstream (RNA non-coding strand) or downstream (RNA coding strand) from the Ad2MLP (Maniatis *et al.*, 1982). For labelling the non-coding strand, pM8 was linearized at the *Xma*I (*Sma*I) site at −125 (Figure 1). After dephosphorylation of 5′ ends with calf intestinal phosphatase (Boehringer Mannheim), 5′ end-labelling with [γ-^{32}P]ATP and T4 polynucleotide kinase, and digestion with *Bam*HI at +33, the resulting *Sma*I-*Bam*HI DNA fragment was purified on a 6% acrylamide gel and eluted. For labelling the coding strand, the wild-type pM677 and the series of upstream element Ad2MLP point mutants (Figure 1) were linearized at the *Bam*HI site at +33 and their 5′ ends were dephosphorylated and subsequently ^{32}P-phosphorylated as described above. DNA footprinting templates were then prepared by digestion either at the *Xho*I site at −260 or at the *Sac*II site at −245 (for pM677-X) and purified as above.

Acknowledgements

We thank M.P.Brocard, A.Kempf and J.L.Plassat for skilful technical assistance; M.Gilbert for HeLa cell culture; B.Boulay and C.Werlé for photography and illustration respectively; C.Aron for typing the manuscript. H.Boeuf, R.Hen, T.Leff, C.Weissman and M.Wintzerith for gifts of recombinants. This work was supported by the CNRS (ATP 3582), the INSERM (PRC 124026), the Fondation pour la Recherche Médicale, The Ministère de la Recherche et de la Technologie (82V1283) and the Association pour le Développement de la Recherche sur le Cancer). N.G.Miyamoto was a recipient of a Medical Research Council of Canada Post-Doctoral Fellowship.

References

Akusjarvi,G. and Pettersson,U. (1979) *J. Mol. Biol.*, **134**, 143-158.
Barrera-Saldana,H.A. Takahashi,K., Vigneron,M., Wildeman,A.G., Davison,I. and Chambon,P. (1985) *EMBO J.*, in press.
Baty,D., Barrera-Saldana,H., Everett,R., Vigneron,M. and Chambon,P. (1984) *Nucleic Acids Res.*, **12**, 915-932.
Benoist,C. and Chambon,P. (1981) *Nature*, **290**, 304-310.
Benoist,C., O'Hare,K., Breathnach,R. and Chambon,P. (1980) *Nucleic Acids Res.*, **8**, 127-142.
Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **50**, 349-383.
Chambon,P., Dierich,A., Gaub,M.P., Jakowlev,S., Jongstra,J., Krust,A., LePennec,J.P., Oudet,P. and Reudelhuber,T. (1984) *Recent Prog. Horm. Res.*, **40**, 1-42.
Corden,J., Wasylyk,B., Buchwalder,A., Sassone-Corsi,P., Kédinger,C. and Chambon,P. (1980) *Science (Wash.)*, **209**, 1406-1414.
Davison,B.L., Egly,J.M., Mulvihill,E.R. and Chambon,P. (1983) *Nature*, **301**, 680-686.
Dierks,P., Van Ooyen,A., Cochran,M.D., Dobkin,C., Reiser,J. and Weissman,C. (1983) *Cell*, **32**, 695-706.
Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475-1489.
Dynan,W.S. and Tjian,R. (1983) *Cell*, **35**, 79-87.
Egly,J.M., Miyamoto,N.G., Moncollin,V. and Chambon,P. (1984) *EMBO J.*, **3**, 2363-2371.
Elkaim,R., Goding,C. and Kédinger,C. (1983) *Nucleic Acids Res.*, **11**, 7105-7117.
Fox,K.R. and Waring,M.J. (1984) *Nucleic Acids Res.*, **12**, 9271-9285.
Galas,D. and Schmitz,A. (1978) *Nucleic Acids Res.*, **5**, 3157-3170.
Gidoni,D., Dynan,W.S. and Tjian,R. (1984) *Nature*, **312**, 409-413.
Gilbert,W., Maxam,A.M. and Mirzabekov,A.D. (1976) in Kjeldgaard,N.C. and Maaloe,O. (eds.), *Control of Ribosome Synthesis*, Munksgaard, Copenhagen, pp. 139-143.
Grosschedl,R. and Birnstiel,M.L. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 297-301.
Grosveld,G.C., Shewmaker,C.K., Jat,P. and Flavell,R.A. (1981) *Cell*, **25**, 215-226.
Grosveld,G.C., Rosenthal,A. and Flavell,R.A. (1982) *Nucleic Acids Res.*, **10**, 4951-4971.
Hansen,U. and Sharp,P.A. (1983) *EMBO J.*, **2**, 2293-2303.
Hen,R. (1985) Thèse d'Etat, Unviersité Louis Pasteur, Strasbourg, France.
Hen,R., Sassone-Corsi,P., Corden,J., Gaub,M.P. and Chambon,P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7132-7136.
Hu,S.L. and Manley,J.L. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 820-824.
Jove,R. and Manley,J.L. (1984) *J. Biol. Chem.*, **259**, 8513-8521.
Lee,D.C., Roeder,R.G. and Wold,W.S.M. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 41-45.
Leff,T. and Chambon,P. (1986) *J. Mol. Cell. Biol.*, in press.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.
Manley,J.L., Fire,A., Cano,A., Sharp,P.A. and Gefter,M.L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3855-3859.
McKnight,S.L., Gavis,E.R. and Kingsbury,R. (1981) *Cell*, **25**, 385-398.
McKnight,S.L., Kingsbury,R.C., Spence,A. and Smith,M. (1984) *Cell*, **37**, 253-262.
Miyamoto,N.G., Moncollin,V., Egly,J.M. and Chambon,P. (1984) *Nucleic Acids Res.*, **12**, 8779-8799.
Moncollin,V., Miyamoto,N.G., Zheng,X.M., Egly,J.M. and Chambon,P. (1985) in press.
Parker,C.S. and Topol,J. (1984a) *Cell*, **36**, 357-369.
Parker,C.S. and Topol,J. (1984b) *Cell*, **37**, 273-283.
Sassone-Corsi,P., Dougherty,J.P., Wasylyk,B. and Chambon,P. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 308-312.
Sassone-Corsi,P., Wildeman,A.G. and Chambon,P. (1985) *Nature*, **313**, 463.
Shenk,T. (1981) *Cur. Top. Microbiol. Immunol.*, **93**, 25-46.
Sutcliffe,J.G. (1978) *Nucleic Acids Res.*, **5**, 2721-2728.
Takada,Y., Ohlendorf,D.H., Anderson,W.F. and Matthews,B.W. (1983) *Science (Wash.)*, **221**, 1020-1026.
Tsuda,M. and Suzuki,Y. (1981) *Cell*, **27**, 175-182.
Vigneron,M., Barrera-Saldana,H.A. Baty,D., Everett,R.E. and Chambon,P. (1984) *EMBO J.*, **3**, 2373-2382.
Wasylyk,B. (1985) *Genetic Engineering*, in press.
Wasylyk,B., Derbyshire,R., Guy,A., Molko,D., Roget,A., Teoule,R. and Chambon,P. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 7024-7028.
Wasylyk,B., Wasylyk,C., MAtthes,H., Wintzerith,M. and Chambon,P. (1983) *EMBO J.*, **2**, 1605-1611.
Weil,P.A., Luse,D.S., Segall,J. and Roeder,R.G. (1979) *Cell*, **18**, 469-484.
Wildeman,A.G., Sassone-Corsi,P., Grundström,T., Zenke,M. and Chambon,P. (1984) *EMBO J.*, **3**, 3129-3133.
Wildeman,A.G., Zenke,M., Schatz,C., Wintzerith,M., Grunström,T., Matthes,H., Takahashi,T. and Chambon,P. (1985) *Nature*, in press.
Yu,Y.T. and Manley,J.L. (1984) *Nucleic Acids Res.*, **12**, 9309-9312.
Zajchowski,D.A., Boeuf,H. and Kédinger,C. (1985) *EMBO J.*, **4**, 1293-1300.

Received on 25 August 1985; revised on 27 September 1985