The EMBO Journal vol.6 no.10 pp.3027–3034, 1987

The bidirectional upstream element of the adenovirus-2 major late promoter binds a single monomeric molecule of the upstream factor

A.C.Lennard[1] and J.M.Egly

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: ICRF, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, UK

Communicated by P.Chambon

The adenovirus-2 major late promoter (Ad2MLP) upstream element (Ad2MLP-UE) contains a sequence of interrupted dyad symmetry. By inverting this element we have found that it functions in a bidirectional manner both *in vivo* and *in vitro*. Footprinting and binding kinetics studies have demonstrated that both orientations of the upstream element bind the sequence-specific upstream factor (UEF) in a similar fashion. These data strongly suggest that the dyad symmetric sequence is sufficient for fully functional binding of the UEF. Binding studies of the UEF to the Ad2MLP-UE indicate that, contrary to prokaryotic palindromic promoter elements which bind multimers of specific factors, the entire Ad2MLP dyad symmetric upstream element binds a single monomeric UEF molecule.

Key words: adenovirus-2/major late promoter/transcription factor/DNA–protein interactions

Introduction

The adenovirus-2 major late promoter (Ad2MLP) contains at least two elements that play important roles in the *in vivo* and *in vitro* transcriptional efficiency of this promoter (Corden *et al.*, 1980; Hen *et al.*, 1982; Miyamoto *et al.*, 1984). A TATA box is located around 30 bp upstream from the site of transcription initiation, and a further upstream element (Ad2MLP-UE) lies between nucleotides −49 and −67 (see Figure 3C). Both of these elements are recognized by sequence-specific factors which are required for efficient transcription from the Ad2MLP. A TATA-box binding factor (Davison *et al.*, 1983; Sawadogo and Roeder, 1985; Tamura *et al.*, in preparation) is indispensable, in addition to at least three other factors and RNA polymerase B (II), for accurate initiation of transcription (Moncollin *et al.*, 1986), whereas the Ad2MLP upstream element factor (UEF) stimulates transcription when bound to the upstream element (Miyamoto *et al.*, 1985; Sawadogo and Roeder, 1985; Chodosh *et al.*, 1986).

The Ad2MLP-UE, like many other promoter elements that bind transcription factors includes a dyad symmetric sequence (see Pabo and Sauer, 1984 for review; Giniger *et al.*, 1985). For the Ad2MLP this sequence of dyad symmetry is recognized by the UEF (Miyamoto *et al.*, 1985). It was therefore of interest to further determine the relevance of this symmetry in transcription from the Ad2MLP. In the present study we show that the Ad2MLP-UE stimulates transcription bidirectionally both *in vivo* and *in vitro*, and that the UEF binds tightly over a similar location in promoters containing wild-type or inverted upstream sequences. We also find that both orientations of the upstream element bind the UEF with the same association and dissociation kinetics.

Although for the few prokaryotic systems studied, dyad symmetric promoter sequences have been shown to bind dimers of sequence-specific factors, e.g. λ-repressor (Pirrotta *et al.*, 1970) and AraC protein (Hendrickson and Schleif, 1985), we present kinetic experiments which indicate that a single monomeric UEF molecule binds stoichiometrically to the Ad2MLP-UE.

Results

An inverted Ad2MLP-upstream element activates transcription in vivo and in vitro as efficiently as in the wild-type orientation

The recombinant pML15 (see Figure 1A) contains an adenovirus-2 major late promoter (Ad2MLP) from *Hin*dIII to *Bam*HI sites (see pM15 in Miyamoto *et al.*, 1985) ligated to a rabbit β-globin gene from −9 to +1650, in pBR322. This construction also includes an internal reference promoter derived from a rabbit β-globin gene to allow reliable comparisons with the constructions described below (Barrera-Saldana *et al.*, 1985). To permit an inversion of the adenovirus-2 major late promoter upstream element (Ad2MLP-UE), a *Xho*I site was created at −42 by site directed mutagenesis (see Materials and methods) to generate pML16 (Figure 1A). A recombinant containing an inverted upstream element (pML16I) was then constructed using synthetic oligonucleotides in such a way as to preserve (i) the sequences flanking the entire region of the Ad2MLP-UE known to interact with the UEF (Figure 3C), and (ii) the relative position of the centre of dyad symmetry (between −57 and −58) in the upstream element with respect to the TATA box, thereby retaining the stereoalignment (Takahashi *et al.*, 1986) of the factors that bind to these promoter elements (Miyamoto *et al.*, 1985; Sawadogo and Roeder, 1985; Carthew *et al.*, 1985). A control recombinant was also constructed which lacks the Ad2MLP-UE (pML16Δ, Figure 1A) by deletion of the upstream element in pML16.

To compare the *in vivo* transcriptional efficiency of the inverted upstream element in pML16I with the wild-type upstream element in pML16, we transfected these plasmids into HeLa cells by calcium phosphate precipitation (see Materials and methods). The quantities of stable RNA transcribed from the recombinants were determined by S1 nuclease mapping using the probe described in Materials and methods (see Figure 1A). We observe that the *in vivo* transcription levels from pML16I (which contains the inverted Ad2MLP-UE; Figure 2A, lane 6) and from pML16 (which contains the wild-type orientation of the Ad2MLP-UE; lane 5) are identical, whereas the transcription from pML16Δ (which does not possess the Ad2MLP-UE; lane 7) results in a 80% decrease. We also show, as a control experiment, that the double point mutation which created the *Xho*I site in pML16 does not significantly impair the stimulatory activity of the Ad2MLP-UE on the transcriptional efficiency from its promoter, since the levels of transcription from pML15 and pML16 (Figure 2A, lanes 4 and 5) are identical. The effect of the upstream element in pML16, compared to pML16Δ, is not apparently as great as previously reported (Hen *et al.*, 1982; Jove and Manley, 1984; Carthew *et al.*, 1985). This is presumably a result of the SV40

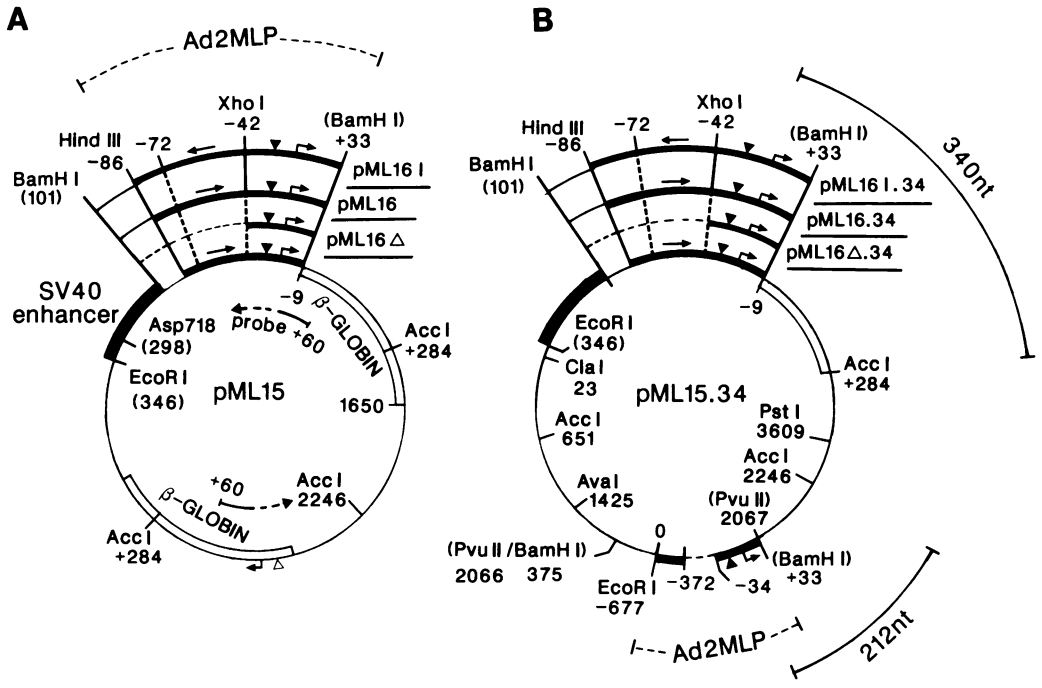

Fig. 1. The pML series of *in vivo* (**A**) and *in vitro* (**B**) recombinants (not drawn to scale). The vectors pML15 and pML15.34 were constructed as described in Materials and methods. Restriction sites in parentheses were destroyed during the construction of these recombinants. The vector pML15 contains the Ad2MLP from *Hin*dIII (−86) to *Bam*HI (+33) linked to a SV40 72 bp enhancer and a segment of the rabbit β-globin gene (open box inside the vector). A *Xho*I site is present at −42 within the Ad2MLP in pML16 and pML16I, this site has been destroyed in pML16Δ. Otherwise, the pML *in vivo* series is identical to pSEGO (Barrera-Saldana *et al.*, 1985). Arrows indicate the orientation of the Ad2MLP upstream element. The limits of the inverted region in pML16I are −43 and −72. *Bam*HI−*Hin*dIII adapters between the Ad2MLP and the SV40 enhancer (pML15, pML16 and pML16I) are represented by thin continuous lines, whereas the *Bam*HI−*Xho*I adapter in pML16Δ is indicated by a thin dotted line. The pML15.34 series is identical to the pML15 series from *Eco*RI (346) to *Acc*I (+284), and also contains a reference Ad2MLP which does not contain the upstream element (Miyamoto *et al.*, 1984).

enhancer element masking to some extent the effect of the upstream element.

For *in vitro* experiments we constructed recombinants containing two Ad2ML promoters (Figure 1B). One, which serves as a reference promoter, has the upstream sequences deleted from −34 to −372 and is inserted into the *Pvu*II site of pBR322; the other has a pML promoter with either a wild-type (pML16.34) or inverted upstream element (pML16I.34), or lacks the upstream element (pML16Δ.34), and is inserted between *Eco*RI and *Acc*I sites of pBR322. The plasmid pML15.34 is identical to pML16.34, but does not contain a *Xho*I site at −42. These double recombinants were cut with *Acc*I to generate templates that result in 340 nt and 212 nt transcripts [from the pML and reference promoters (see Figure 1B), respectively] in reconstituted HeLa cell extract run-off transcription assays (Moncollin *et al.*, 1986). As previously observed by Miyamoto *et al.* (1984), the template containing the wild-type upstream element was consistently 4−6 fold more efficiently transcribed than the template lacking the upstream element (Figure 2B, compare lanes 2 and 4). In agreement with the *in vivo* results described above, the promoter containing an inverted upstream element gave the same level of transcription as the promoter containing a wild-type orientation (compare pML16I.34, lane 3, with pML16.34, lane 2). The *Xho*I site at −42 had no effect on transcription *in vitro* (compare pML15.34, lane 1, with pML16.34, lane 2). Therefore our results demonstrate that the Ad2MLP-UE activates transcription from the Ad2MLP in an orientation-independent manner both *in vivo* and *in vitro*.

The inverted Ad2MLP upstream element and the wild-type orientation similarly bind the UEF

Transcription from the Ad2MLP is mediated by the binding of a DNA sequence-specific factor (UEF) to the upstream element (Miyamoto *et al.*, 1985; Moncollin *et al.*, 1986). Since the transcription levels for the wild-type and inverted upstream elements are identical, it was of interest to define the UEF-upstream sequence interactions for the two orientations of the Ad2MLP-UE by DNase I footprinting (Figure 3A) and dimethylsulfate (DMS) methylation protection experiments (Figure 3B).

The span of DNase I protection for the wild-type Ad2MLP-UE in the non-coding strand of pML15 and pML16 stretches from −49/−50 to −66 (Figure 3A, lanes 7 and 8) and for the inverted upstream element in pML16I (lane 9) protection is from −46/−47 to −66 (see Figure 3C for a summary of the footprints). Similarly for the coding strand, pML15 and pML16 are protected from −49 to −67/−68 and pML16I is protected from −49 to −65/−66 (data not shown). We note that pML16I is protected from DNase I at least three base pairs more than observed for pML16 on the non-coding strand proximal to the TATA sequence. Overall, the DNase I hypersensitive sites for the wild-type and inverted orientations of the upstream element are in equivalent regions. The DNase I hypersensitive sites proximal to the TATA box are on the non-coding strand for both pML16 (at nucleotides C-44, T-45, T-48) and pML16I (at nucleotides A-44, T-45); the hypersensitive sites distal to the TATA box are on the coding strand at nucleotide C-69 for pML16 and at nucleotides T-67, T-70, C-71 for pML16I. These data show that the location of the UEF binding to the Ad2MLP was the same for either orientation of the upstream element (Figure 3C) from approximately nucleotides −49 to −66. The slight changes in the protection from DNase I afforded by the UEF, upon inversion of the upstream element indicate some effect asserted by the flanking sequences on UEF binding. The DNase I hypersensitive sites lie on the same face of the DNA double helix. It is interesting

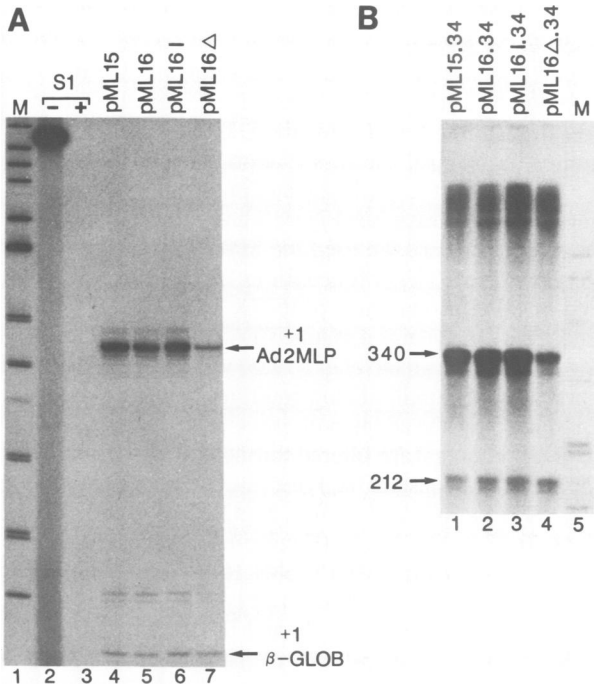

Fig. 2. *In vivo* **(A)** and *in vitro* **(B)** transcription of the pML series of recombinants. **(A)** Transcripts produced *in vivo* after transfection of HeLa cells were analysed by S1 nuclease mapping as described in Materials and methods. The S1 nuclease-resistant bands around 116 nucleotides are from the Ad2MLP. β-GLOB is the S1 nuclease-resistant band from the β-globin internal reference gene. Lanes 2 and 3 are controls of probe incubated with 10 μg of yeast tRNA and further incubated in the presence (+) or absence (−) of S1 nuclease. The duplex of bands above β-GLOB are artefacts occasionally observed when the probe was prepared in the presence of carrier tRNA. **(B)** *In vitro* transcription was carried out using *Acc*I digests of the pML15.34 series as described in Materials and methods. M is a DNA size marker.

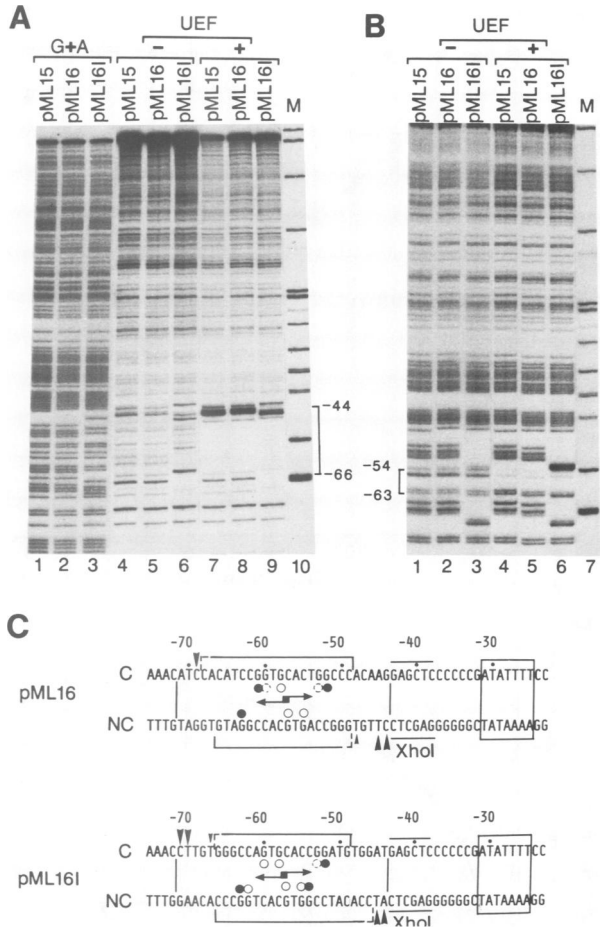

Fig. 3. DNase I footprint **(A)** and DMS methylation protection pattern **(B)** of the non-coding strand of the Ad2MLP in its wild-type (pML15, pML16) and inverted orientation (pML16I). The footprinting reactions were done as described in Materials and methods in the presence (+) or absence (−) of the UEF. M is a DNA size marker. **(C)** The Ad2MLP sequence is given from −23 to −75 in pML16 (wild-type orientation) and pML16I (inverse orientation). Arrowheads indicate DNase I hypersensitive sites, their size reflects the degree of sensitivity. Square brackets denote the extent of protection to DNase I afforded by the UEF (dotted regions are where it was not possible to evaluate protection from DNase I since the naked DNA was not cut at these sites). Open and closed circles represent DMS hypo- and hypermethylation sites respectively. Broken circles designate G residues whose reactivity towards DMS was not resolved, for the wild-type Ad2MLP (pML16) these residues are protected by the UEF (Miyamoto *et al.*, 1985). The TATA region is indicated by the boxed sequence, the dyad symmetry within the upstream element is denoted by the solid arrows, and the limits of the inverted sequence (−43 to −72) are indicated.

to note that with respect to the DNA sequence the three DNase I hypersensitive sites on the non-coding strand of pML16 (at positions C-44, T-45, T-48) are found at exactly their equivalent positions on the coding strand of pML16I (at positions T-67, T-70, C-71) (Figure 3C). The equivalent hypersensitive sites show that their flanking sequences, which change upon inversion of the Ad2MLP-UE, do not affect their sensitivity to DNase I.

We have previously reported that the G residues within the Ad2MLP-UE which are hyper or hypomethylated by DMS, in the presence of the UEF, mostly lie on dyad symmetric residues (Miyamoto *et al.*, 1985). As a consequence of the dyad symmetry these nucleotides are found in the same position on the same strands in our inverted upstream element (see Figure 3C). We therefore determined the DMS methylation patterns for both orientations of the UEF sequence. The pattern of DMS methylation of G residues in the non-coding and coding strands of pML16 can in most cases be mapped directly onto the same strands of pML16I (Figures 3B and C). For example, the protected G residues −55 and −57 in the non-coding strand of pML16´ (Figure 3B, lane 5) are also protected on the non-coding strand in pML16I (lane 6), and the hypermethylated G-63 on the pML16 non-coding strand is also hypermethylated in the non-coding strand of pML16I. When considering the dyad symmetric residues, only one difference in the DMS methylation pattern is generated by inverting the upstream element; G-62, which was not protected on the non-coding strand of pML16 was reproducibly protected in pML16I whereas, as expected, the dyad symmetric equivalent to G-62, G-53 on the coding strand of pML16, is protected from DMS by the UEF for both plasmids. This change indicates an effect of the flanking sequences on UEF binding, as observed above for the DNase I footprint patterns. Since the residues G-61 and T-54 on the coding strand of pML16 are not nucleotides that exhibit dyad symmetry within the Ad2MLP-UE sequence, inversion of the upstream element (pML16I) changes nucleotides −61 and −54 in the coding strand (pML16I) to A and C residues, respectively. Therefore the hypermethylation observed at G-61 in the coding strand of pML16 could not map to −61 in pML16I and the hypermethylation at G-54 in the non-coding strand of pML16I could not map to −54 in pML16. In conclusion we find that the DMS methylation pattern for the dyad symmetric sequence within the inverted Ad2MLP-UE qualitatively resembles the pattern obtained for the wild-type upstream element.

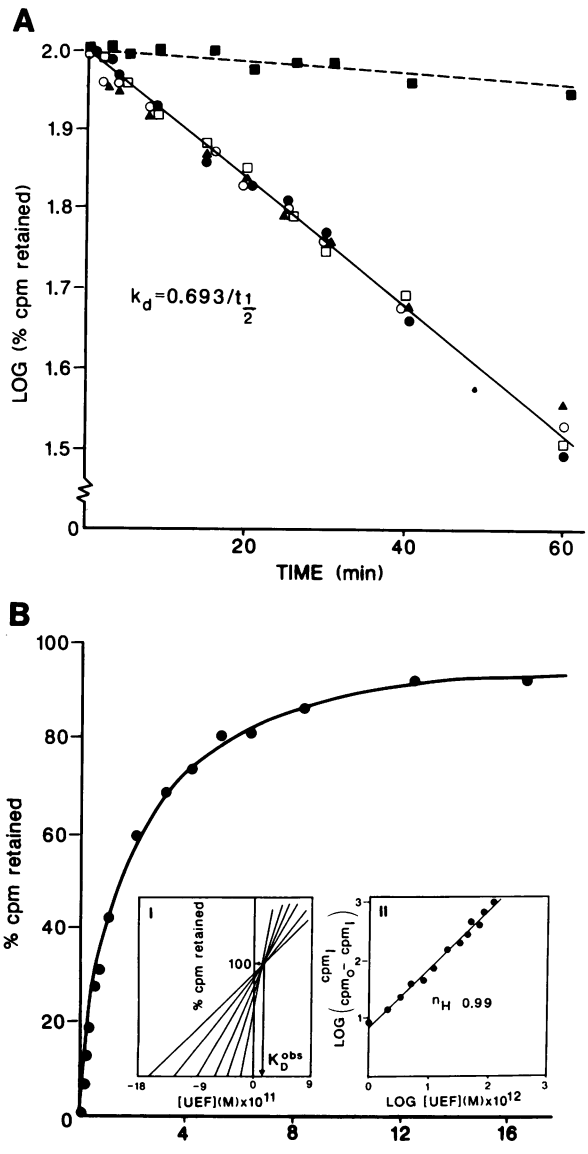

Fig. 4. **(A)** Dissociation kinetics of the Ad2MLP-UEF from the wild-type and inverted upstream element. For the procedure used see Materials and methods. (□) pM677X, (○) pML15, (●) pML16 and (Δ) pML16I were competed with pM677X plasmid, and (■) pM677X was competed with pM1X plasmid. Each point is the average of two independent assays. 180 μl of a 1:100 dilution of UEF in a SP0.34 fraction (Moncollin *et al.*, 1986) was incubated 30 min at 24°C with 0.1 mg/ml BSA, 12 μg poly(dI.dC):(dC.dI) and ~9 ng of [^{32}P]end labelled *Hin*dIII–*Acc*I fragment from pML15, pML16, pML16I or a *Bam*HI–*Sac*II fragment from pM677X, in 1.2 ml of Buffer A. At zero time a 1000 fold excess of pM677X (solid line) or pM1X (dotted line) competitor plasmid DNA was rapidly added, and aliquots of 100 μl were filtered through nitrocellulose at various times. A zero time aliquot was filtered within 10 s after addition of competitor DNA. **(B)** Binding kinetics of the UEF onto the Ad2MLP upstream element. The labelled probe A1 (Materials and methods) at 5×10^{-14} M was incubated with 0–1 ng ($0-17 \times 10^{-11}$ M) of UEF from the SP0.34 fraction in 100 μl at 24°C for 30 min to attain equilibrium and then filtered through nitrocellulose. Insert (I) is a direct linear plot of the binding data to obtain the concentration of UEF required for $[DNA]_{free} = [DNA]_{bound}$ and the maximum amount of DNA bound. K_D is the equilibrium dissociation constant. Insert (II) is a Hill plot of the binding data, where cpm_1 and cpm_0 are the counts retained on the filters and maximum possible counts retained as calculated from the plot in insert (I), respectively.

The UEF binds and dissociates from an inverted upstream element with the same rates as from the wild-type orientation

As a consequence of the similar interactions of the two orientations of the Ad2MLP-UE with the UEF (Figure 3C), one would expect the binding kinetics for the UEF to also be similar. We therefore determined the dissociation (k_d) and association (k_a) rate constants for the wild-type and inverted upstream elements by filter-binding assays (see Figure 4 and Materials and methods). As a source of UEF, we used the SP0.34 fraction (which does not contain the BTF_1; Moncollin *et al.*, 1986).

End-labelled probe from pM677X [*Bam*HI (+33) to *Sac*II (−245); which contains a mutated TATA box and the wild-type Ad2MLP-UE; Miyamoto *et al.*, 1985] was incubated with SP0.34 fraction at 24°C, until binding equilibrium was attained. At different times after the addition of excess cold plasmid 'competitor' pM677X, aliquots were filtered through nitrocellulose. The rate of dissociation constant (k_d) for the UEF from pM677X at 24°C, 50 mM KCl and 3 mM $MgCl_2$ was determined to be $3.08 \pm 0.12 \times 10^{-4}$ s^{-1} (Figure 4A, open squares). Using probes prepared from pML15, pML16 [*Acc*I (+284) to *Hin*dIII (−86), which both contain the wild-type orientations of the upstream elements; see Figure 1] and pML16I [*Acc*I to *Hin*dIII, containing an inverted upstream element], identical dissociation curves were obtained with a rate constant of 3.1×10^{-4} s^{-1} (Figure 4A), identical to that observed for pM677X. In the control experiment using a competitor which is mutated in both the TATA box and the upstream element and can no longer bind the UEF (pM1X; Miyamoto *et al.*, 1985), the upstream factor was not efficiently dissociated from pM677X (Figure 4A; full squares).

Association rate constants (k_a) were determined from the dissociation equilibrium constants (K_D) for the binding reaction (n being the number of UEF molecules):

$$n\text{UEF} + \text{DNA} \underset{k_d}{\overset{k_a}{\rightleftharpoons}} \text{DNA} \cdot (\text{UEF})_n \qquad (1)$$

for which

$$K_D = \frac{k_d}{k_a} \qquad (2)$$

K_D was determined for UEF binding to the wild-type Ad2MLP by two procedures using the filter-binding assay (i) titration with the UEF and (ii) titration with Ad2MLP DNA. Hyperbolic binding curves were obtained by titrating the SP0.34 fraction onto either probe A1 (which contains the wild-type Ad2MLP from +32 to −85 labelled to high specific activity; see Materials and methods, Figure 4B) or the *Bam*HI–*Sac*II probe from pM677X (see above). For either probe, up to 96% of the specific upstream element DNA could be bound. The level of non-specific binding (no more than 5%) was assessed by comparing titrations of probes from pM677X (wild-type upstream element) and pM1X (mutated upstream element) with the SP0.34 fraction (data not shown). Under the conditions used in Figure 4B, where [UEF] > > [DNA], the equilibrium dissociation equation when $[DNA]_{free} = [DNA]_{bound}$ can be reduced to $K_D^{obs} = [UEF]_{total}$. The K_D^{obs} was calculated from binding curves by the procedure of Eisenthal and Cornish-Bowden (1974) (Figure 4B, insert I). This method of analysis transforms the equation for the binding kinetics into the form $x/a + y/b = 1$, where x is the maximum DNA bound, y is the [UEF] when $[DNA]_{free} = [DNA]_{bound}$, a

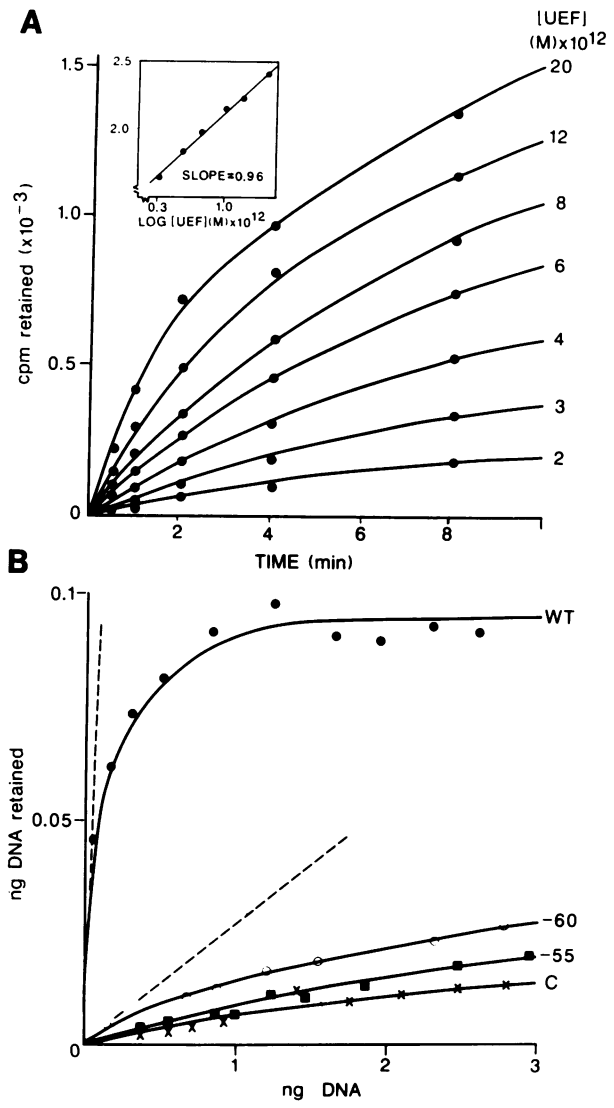

Fig. 5. (A) Kinetics of UEF binding to the Ad2MLP. Aliquots of SP0.34 [(2−20) × 10^{-12} M UEF] in 600 μl were incubated at 24°C for 5 min. At zero time the probe from pM677X (Materials and methods) was added to a final concentration of 2 × 10^{-12} M and aliquots of 100 μl were filtered at the times indicated. (insert) The initial rates from the binding curves were calculated and plotted against the concentrations of UEF. The data shown are an average of duplicate experiments. (B) Binding kinetics of the UEF onto wild-type or mutated Ad2MLP-UE oligonucleotides. 2.3 × 10^{-11} M UEF in the SP0.34 fraction was incubated with 0.1−3 ng upstream element DNA (WT: wild-type sequence from −40 to −72, see Figure 3C for sequence; −60: as for WT except G-60 was changed to T-60; −55: as for WT except C-55 was changed to A-55), in 100 μl at 24°C for 30 min, and then filtered through nitrocellulose. Curve C was obtained using an oligonucleotide of the Ad2MLP from −13 to −35, in place of upstream element DNA.

is the $[DNA]_{bound}$ and b is the [UEF]. A straight line exists for each value of a and b which reach a common intercept to directly give the maximum $[DNA]_{bound}$ and K_D^{obs}. We reproducibly obtained a K_D^{obs} of 1.69 ± 0.68 × 10^{-11} M for the UEF binding to the wild-type Ad2MLP-UE at 24°C in 50 mM KCl and 3 mM $MgCl_2$, corresponding to a free energy change (ΔG) of 61.0 kJ mol^{-1}. The same K_D^{obs} was also obtained by titrating the SP0.34 fraction against excess pM677X probe. Using eqn (2), k_a = 2.21 ± 0.96 × 10^7 $M^{-1}s^{-1}$. It should be noted that the concentration of UEF was determined by titrating the UEF with upstream element DNA and our calculations assumed a stoichiometry of one monomeric UEF molecule per upstream element.

By titrating the SP0.34 fraction with probes prepared from pML15, pML16 (wild-type upstream element) or pML16I (inverted upstream element) labelled at the *Acc*I (+281) and *Hin*dIII (−86) sites, we obtained 2.77 × 10^{-11} M, 1.58 × 10^{-11} M and 2.21 × 10^{-11} M respectively for K_D^{obs} at 24°C, in 50 mM KCl and 3 mM $MgCl_2$ (data not shown). These values are not significantly different from those obtained using probes A1 or pM677X, and allow us to conclude that K_D^{obs}, k_a and k_d are similar for Ad2ML promoters having wild-type or inverted upstream elements.

A single molecule of the Ad2MLP upstream factor binds to the upstream element

Several prokaryotic and eukaryotic proteins that recognize DNA sequences with dyad symmetry have been shown to bind as dimers (Pabo and Sauer, 1984; Giniger *et al.*, 1985). Since the Ad2MLP-UE contains a sequence of dyad symmetry, several possibilities could be envisaged for UEF interaction with the upstream element. These models include a monomeric form of the UEF binding to one defined half of the dyad symmetric DNA sequence which then allows highly cooperative binding of a second UEF molecule to the other half of the upstream element; a low or non-cooperative binding of two UEF molecules to each half of the dyad symmetry; a monomer−dimer equilibrium in solution in which the dimeric form is required for efficient binding to the upstream element; a single monomeric UEF molecule binding to the entire Ad2MLP-UE. We then proceeded to define which of these mechanisms is most likely to occur.

To study the possibility of weakly cooperative or independent binding of two UEF molecules to the Ad2MLP-UE, we determined the kinetic order (m) of the binding reaction for which $v_i = k_a$ [UEF]^m[DNA]. The initial rates (v_i) of binding were obtained from binding curves using increasing concentrations of UEF (from 2 × 10^{-12} to 20 × 10^{-12} M; Figure 5A) by polynomial regression analysis. A plot of log (v_i) against concentration of UEF has a slope that directly gave a reproducible value for m of 0.96 ± 0.08 (Figure 5A, insert) which is independent of the binding stoichiometry used to determine the UEF concentration. We can therefore conclude that non-cooperative binding, or weakly cooperative binding of two (or more) UEF molecules, are highly unlikely since such mechanisms would give values of m greater than one. To further support this conclusion, we investigated the binding of the UEF to upstream element sequences mutated in each half of the dyad symmetry. Assuming that each half of the palindromic sequence can bind one molecule of UEF, mutation of one half should still allow binding of one molecule to the other half of the upstream element. Oligonucleotides of Ad2MLP-UE DNA (see legend to Figure 5) were synthesized in which either C-55 was changed to an A or G-60 was changed to a T. These point mutations were chosen since C-55 and G-60 are both required for optimal stimulation of transcription by the UEF (Yu and Manley, 1984; Miyamoto *et al.*, 1984) and are in equivalent positions within the dyad symmetry. The binding curves for mutants −60 and −55 (Figure 5B) indicate very little, if any, specific binding to the UEF compared to the 'non-specific' binding to oligonucleotide C (Ad2MLP

from −15 to −35). The wild-type sequence bound the UEF with a K_D of 3.2×10^{-11} M. The apparent change in K_D for the UEF compared to that obtained from Figure 4B could be explained by the difference in the lengths of DNA sequence used (Winter and Von Hippel, 1981). The apparent binding to the −60 mutant is not significant, since under the conditions used here where $K_D \times [UEF]_{total} << 1$, the initial tangent (as [DNA] approaches zero) for independent or weakly cooperative binding of two UEF molecules, in which one site is mutated, should theoretically be half that found for the wild-type sequence. This is clearly not the case, since the initial slope for mutant −60 is twenty times weaker than the slope obtained for the wild-type sequence.

The precedent from λ- and cro-repressors suggested that tight dimers of factors in solution bind to DNA sequences of dyad symmetry. Our first order initial rate kinetics of Figure 5A (insert) argue against the formation of such UEF dimers in free solution with a $K_D \geq 10^{-12}$ M. To investigate whether a UEF dimer with a $K_D \leq 10^{-12}$ M can exist, we compared the binding kinetics obtained by adding either concentrated or dilute UEF to upstream element DNA. If a UEF dimer can be generated in concentrated UEF solutions, it will dissociate into monomers upon extensive dilution, which would generate a lag period in the binding kinetics (Pirrotta *et al.*, 1970). Using filter-binding assays with the UEF in the SP0.34 fraction at 1.4×10^{-10} M and 3×10^{-13} M, we observed no evidence for dissociation of a dimeric UEF (data not shown). Furthermore the UEF activity sediments through glycerol gradients as a single peak with the relative molecular mass predicted for a monomeric protein (Moncollin *et al.*, 1986; Chodosh *et al.*, 1986), which is not in agreement with a free stable UEF dimer existing in solution.

We also consider highly cooperative binding of two UEF molecules to be unlikely since the hyperbolic binding curve (Figure 4B) generates a Hill coefficient (n_H) of 0.99 (Figure 4B, insert II). The same value was obtained whether the UEF or DNA was titrated. Cooperative binding of two molecules would give a Hill coefficient ≥ 1. Note that the Hill coefficient is also independent of the binding stoichiometry used to determine the concentration of UEF. We conclude that the most likely mechanism to fit the kinetic data described above is that of a single monomeric UEF molecule binding to the dyad symmetric upstream sequence.

Discussion

The partially dyad symmetric Ad2MLP-UE functions bidirectionally

The stimulation of transcription from the Ad2MLP is mediated by the binding of a specific cellular protein factor, the UEF, onto the upstream element (Ad2MLP-UE) (Miyamoto *et al.*, 1984, 1985; Moncollin *et al.*, 1986). By inverting the entire Ad2MLP-UE (nucleotides −43 to −72) we have constructed a promoter which is as transcriptionally active as the wild-type promoter both *in vivo* and *in vitro*. This orientation-independence is most likely a direct consequence of the presence of the hyphenated palindromic sequence centred between nucleotides −57 and −58. As demonstrated by DNase I footprinting and DMS methylation experiments, the UEF protects and interacts in an essentially identical manner with the wild-type or inverted upstream elements. Furthermore, we obtained the same association and dissociation rate constants for UEF binding to wild-type or inverted upstream elements in the Ad2MLP.

When variations in binding conditions are considered, our values for k_a ($2.21 \times 10\ M^{-1}\ s^{-1}$) and k_d ($3.08 \times 10^{-4}\ s^{-1}$) are in agreement with those previously reported by Chodosh *et al.* (1986) and Sawadogo and Roeder (1986), respectively. In particular we note that our value for k_a, which was obtained by nitrocellulose filter binding assays, agrees well with the value obtained by gel retardation (band shift) assays (Chodosh *et al.*, 1986).

Our values of K_D (1.69×10^{-11} M) for UEF binding to the Ad2MLP indicates strong DNA−protein interactions. It is difficult to compare the published values of K_D for other DNA-binding proteins due to the experimental variation in temperature, ionic strength and the presence or absence of divalent cations. Nevertheless the strength of binding for the UEF to its upstream element appears to be similar to that found for the eukaryotic nuclear factor NF-I (10^{-11} M; Rosenfeld and Kelly, 1986) and cro-repressor ($10^{-10}-10^{-11}$ M; Takeda *et al.*, 1977), but less than λ-repressor ($10^{-12}-10^{-13}$ M; Johnson *et al.*, 1980) and greater than prokaryotic RNA polymerase (10^{-7} M; Bertrand-Burggraf *et al.*, 1984).

The data presented here demonstrate the importance of the dyad symmetry within the upstream element for the binding and function of the UEF. The identical transcription from the Ad2MLP, DNase I footprinting, DMS methylation patterns and UEF binding kinetics for either orientation of the upstream element, suggest that only the sequences that are dyad symmetric are necessary for functional binding of the UEF. This dyad symmetric region defines a minimal binding site of 12 bp for the UEF, which is in agreement with the upstream element domain protected by the UEF from methidiumpropyl-EDTA.Fe[II] (Sawadogo and Roeder, 1986). However, some modulation of the UEF binding by the DNA sequences which flank the dyad symmetric domain may occur since some slight differences are apparent in the DNase I footprints and DMS methylation patterns for the two orientations of the upstream element. As previously discussed in Miyamoto *et al.* (1985), Ad2ML promoters with point mutations at dyad symmetric sites are transcriptionally down *in vivo* and *in vitro* and in certain cases have been shown to no longer bind the UEF. Furthermore, it has been reported that mutations at A-54 or C-61 (Miyamoto *et al.*, 1985, pM7; Shi *et al.*, 1986, CG-61), which are at non-dyad symmetric sites have little effect on the transcriptional activity from the Ad2MLP *in vitro* (70−80 and 100%, respectively). The bidirectional nature of the Ad2MLP-UE and the benign effect of mutating C-61 imply that this residue and its dyad symmetric equivalent A-54, which are the only base pairs to interrupt the palindromic sequence, are not essential for upstream element−UEF interactions.

Three other upstream elements, which bind sequence-specific factors, have been shown to function in a bidirectional manner, the NF-I binding site (de Vries *et al.*, 1987), the hsp upstream sequences (Wei *et al.*, 1986) and the SV40 early promoter 21 bp repeats (Barrera-Saldana *et al.*, 1985; Gidoni *et al.*, 1985). The NF-I binding site and the hsp upstream sequences, like the Ad2MLP-UE, are palindromic DNA sequences, whereas the 21 bp repeats show no apparent symmetry in their sequence. It is highly likely that orientation independence is a general phenomenon for palindromic upstream elements, but it may not be universal for asymmetric elements.

The Ad2MLP-UE binds a single UEF molecule

Many prokaryotic and eukaryotic factors bind to sequences that contain partially palindromic sequences, for example, λ-repressor, cro-repressor, CAP (Pabo and Sauer, 1984), lac repressor (Dunaway and Matthews, 1980), AraC protein (Hendrickson and Schleif, 1985), NF-1 (de Vries *et al.*, 1987), NF-III (Pruijn *et al.*, 1986), Adf-1 (Heberlein *et al.*, 1985), c-fos SRF (Treisman, 1986), GAL4 protein (Giniger *et al.*, 1985) and

the hsp (Parker and Topol, 1984) and GCN4 upstream factors (Hope and Struhl, 1986). For λ- and cro-repressors, CAP, AraC and GCN4 it has been shown that the dyad symmetric sequences bind factors which are dimers in solution, and that the lack repressor binds as a tetramer. It was therefore of interest to determine whether the UEF binds to the Ad2MLP-UE as a monomer or dimer. The kinetic analysis of initial binding velocities which gave the order for the binding reaction as one, and the low level of UEF binding to upstream element DNA mutated in either half of the dyad symmetric sequence, did not support mechanisms that involve independent or weakly cooperative binding of two UEF monomeric molecules. The Hill coefficient ($n_H = 1$) for the binding reaction indicates that highly cooperative binding of two UEF molecules is also unlikely. λ-repressor exists in solution in a monomer−dimer equilibrium, for which only the dimeric form has a high affinity for the operator site (Johnson *et al.*, 1980). The linear relationship between initial binding velocity and UEF concentration obtained in Figure 5A insert, required any possible UEF dimer to be extremely stable with a $K_D \leq 2 \times 10^{-12}$ M. Dilution experiments did not reveal any dissociation down to UEF concentration of 3×10^{-13} M. Using the λ-repressor dimer as a model, with a $K_D \sim 10^{-8}$ M for dimer formation, we consider the formation of a UEF dimer in solution with a $K_D < 3 \times 10^{-13}$ M to be unlikely. Furthermore the UEF sediments through glycerol gradients as a monomer (Moncollin *et al.*, 1986; Chodosh *et al.*, 1986). The best interpretation for our results is that a monomeric UEF molecule binds to the entire dyad symmetric upstream element. A monomeric protein binding to a palindromic DNA sequence implies that the symmetry in the DNA sequence would be reflected in the structure (if not sequence) of the protein. If the monomeric UEF evolved by incomplete duplication of an ancestor gene, then it is possible that the symmetry within the UEF extends further than the DNA binding domain, which would readily account for the bidirectionality of its transcriptional effect.

The start site for the adenovirus-2 IVa2 gene lies 140 bp upstream of the Ad2MLP. The promoter for the IVa2 gene has recently been shown to include the Ad2MLP-UE and presumably the same UEF to potentiate transcription from this gene (Natarajan *et al.*, 1987). This also agrees well with a bidirectional upstream element which binds a symmetrical UEF. Since it has been shown that the UEF interacts with the Ad2MLP-TATA box binding factor (Sawadogo and Roeder, 1986), one may postulate that through the symmetry in the upstream element and the UEF, the UEF could also interact with a potential TATA box factor that may bind to the IVa2 promoter region (Tamura *et al.*, in preparation).

Materials and methods

Construction of recombinants

To construct a double recombinant *in vivo* expression vector for the Ad2MLP, pM15 (Miyamoto *et al.*, 1985) was cut with *Bam*HI at +33, repaired, then cut with *Hin*dIII at −86. The +33 to −86 fragment was then inserted into pSEGO (Barrera-Saldana *et al.*, 1985) between an SV40 72 bp enhancer and a promoterless rabbit β-globin gene using a *Bam*HI−*Hin*dIII adapter (5′ GATCCACGA 3′), resulting in pML15. An *Xho*I site was introduced at −42 of the Ad2MLP by mismatch primer synthesis (Grundström *et al.*, 1985) of an M13mp8 recombinant (M13BA) containing the Ad2MLP from pML15 (*Bam*HI−*Acc*I), thereby creating M13BAX. The *Bam*HI−*Acc*I fragment from M13BAX was then transferred back into a *Bam*HI−*Acc*I vector from pML15 to generate pML16 (Figure 1A). To invert the upstream element in pML16, appropriate oligonucleotides were synthesized (Matthes *et al.*, in press) spanning *Hin*dIII-86 to *Xho*I-43, hybridized and ligated into a pML16 *Hin*dIII−*Xho*I vector thereby generating pML16I. To delete the upstream element, the *Bam*HI−*Xho*I fragment was excised and replaced by a *Bam*HI−*Xho*I adapter (5′ GATCCACGAA 3′) resulting in pML16Δ. *In vitro* expression vectors were constructed using the *Pst*I−*Ava*I fragment from pM34.34 (Miyamoto *et al.*, 1984) which contains an Ad2MLP with the upstream element deleted to act as a reference promoter, an *Ava*I−*Cla*I fragment from pBR322 and *Pst*I−*Cla*I fragments from the pML plasmid series of *in vivo* recombinants (Figure 1B). All cloning was carried out using standard techniques (Maniatis *et al.*, 1982).

Preparation of extracts containing the Ad2MLP-UEF

The steps used to purify the Ad2MLP-UEF were fully described in Moncollin *et al.* (1986). HeLa WCE (Manley *et al.*, 1980) was loaded onto a Heparin-Ultrogel-A4R column. The heparin flow-through peak was applied to a DEAE-cellulose column, and the STF activity eluted at 0.35 M KCl (DEO.35; Egly *et al.*, 1984). The peak eluting at 0.6 M KCl from the heparin column was loaded onto a DEAE-5PW column to generate DEO.15 and DEO.25 fractions. The DEO.15 fraction was applied to a sulphopropyl (SP-5PW; generous gift from Toyosoda, Japan) column and the UEF was eluted at 0.35 M KCl (SPO.35; Moncollin *et al.*, 1986).

Transcription in vivo and in vitro

2×10^6 HeLa cells were transfected with 3 μg of plasmid (Baty *et al.*, 1984); 10 μg of isolated RNA were then analysed by quantitative S1 nuclease mapping on 8% polyacrylamide−8.3 M urea gels (Barrera-Saldana *et al.*, 1985). The probe was prepared by primer extension of M13BA using a ^{32}P-endlabelled oligonucleotide (M6) that hybridized at +39 to +60 in the rabbit β-globin sequence, followed by cleavage with *Eco*RI. The resulting single-stranded probe was isolated from an 8% polyacrylamide−8.3 M urea gel.

The *in vitro* transcription system was essentially as previously described (Miyamoto *et al.*, 1984). The preincubation mix contained 8 μl DEO.15, 3 μl DEO.25 and 1 μl DEO.35 fractions with 0.2 μl (0.002 units) of calf thymus RNA polymerase B (II) in 50 mM KCl, 6 mM $MgCl_2$ and 300 ng of plasmid cut with *Acc*I. After a 15 min preincubation at 24°C, transcription was started by the addition of ATP, GTP and UTP to 1.25 mM, CTP to 6.25 μM and 2−5 μCi [^{32}P]α-CTP. Reactions were stopped after 45 min at 24°C and the products analysed on 5% polyacrylamide−8.3 M urea gels.

DNase I and DMS footprinting

Footprint probes for the non-coding strand of the pML recombinants were prepared from 5 μg of plasmid digested with *Asp718*, dephosphorylated with calf intestinal phosphatase (Boehringer Mannheim), ^{32}P-endlabelled and digested with *Acc*I. The labelled probes were excised from 5% polyacrylamide gels (Miyamoto *et al.*, 1985). The DNase I footprint assay consisted of a 22 μl mix of ~1 ng (10 000 c.p.m.) of labelled DNA template, 2 μl of SP0.34 fraction containing the UEF and 1 μg poly(dI.dC): (dC.dI) (Pharmacia) in Buffer A: 30 mM Tris−HCl (pH 7.9), 3 mM $MgCl_2$, 50 mM KCl, 0.07 mM EDTA, 0.3 mM DTT and 7% (v/v) glycerol. After preincubating for 10 min at 24°C, 45 ng of DNase I (Sigma) in 3 μl of Buffer A was added. The reactions were stopped after incubating at 24°C for 1.5 min, treated and analysed as described by Miyamoto *et al.* (1985). The DMS footprint assay was performed as described above for DNase I except that a 23 μl assay mix containing 4 μl of SP0.34 fraction was chilled 5 min on ice after the preincubation, after which 2 μl of ice-chilled DMS (Aldrich) was added. The reactions were stopped after 1.5 min on ice and analysed as previously described for Maxam and Gilbert DNA sequencing (Maniatis *et al.*, 1982). Maxam and Gilbert DNA sequence markers were prepared as described in Maniatis *et al.* (1982).

Nitrocellulose filter binding assays

Plasmids pM677X and pM1X (which contain a mutated TATA box and a wild-type or mutated Ad2MLP-UE, respectively; Miyamoto *et al.*, 1984; Moncollin *et al.*, 1986) were cut with *Sac*II (−245) and *Bam*HI (+33), dephosphorylated, ^{32}P-endlabelled and the appropriate fragments were isolated from 5% polyacrylamide gels. Alternatively, to obtain a probe of high specific activity, M13BA (2 μg) was hybridized to the oligonucleotide M6 and a partial double-stranded form was obtained using 40 μmol each of [α-^{32}P]ATP, [α-^{32}P]CTP, [α-^{32}P]TTP, GTP (Amersham) and 2 units of the Klenow fragment of DNA polymerase I in 25 μl of nick translation buffer (Maniatis *et al.*, 1982) for 60 min at 24°C. After digestion with *Eco*RI and *Bst*NI the probe (A1, 135 bp) was excised from a 5% polyacrylamide gel and purified as described above. To prepare probes WT, −60, −55 and C, appropriate oligonucleotides were synthesized as described by Matthes *et al.* (in press, see legend to Figure 5B). 100 ng of each non-coding strand oligonucleotide were labelled with [γ-^{32}P]ATP, hybridized to 100 ng of the corresponding coding strand and purified through 5 ml Sephadex G50 columns. 0.04 to 0.7 ng of labelled DNA were then diluted with unlabelled DNA to the quantities shown in Figure 5B.

The standard filter binding assay consisted of incubating the SP0.34 fraction (Moncollin *et al.*, 1986) or labelled DNA with 1 μg poly(dI.dC): (dC.dI) and 0.1 mg/ml BSA in 100 μl of Buffer A for 5 min at 24°C, followed by the addition of ^{32}P-labelled DNA or SP0.34 fraction, respectively. After the appropriate incubation time 100 μl aliquots were filtered through nitrocellulose filters (0.45 μm, Millipore) prewashed in Buffer A which contained 0.1 mg/ml BSA. The filters were washed with 3 × 1 ml BSA-Buffer A, dried and counted in 10 ml

Rotiszint 11 (Roth). The filter binding data were analysed by linear or polynomial regression analysis.

Estimation of DNA, upstream factor and total protein concentration

The DNA concentration of probes prepared from pM677X was determined by ethidium bromide fluorescence using the 'plastic wrap method' of Maniatis *et al.* (1982). The concentration of Ad2MLP-UEF in the SP0.34 fractions was determined by titrating the SP0.34 with ^{32}P-endlabelled fragment from pM677X, in the standard filter binding assay mix. Assuming a 1:1 molar ratio of UEF to Ad2MLP we calculate 0.8–1.2 ng/μl UEF in the SP:0.34 fraction. Total protein concentration was measured using the BCA protein assay kit (Pierce). The SP:0.34 fraction contained 1.0–1.2 mg/ml protein, therefore, on average the SP:0.34 fraction consisted of 0.1% upstream factor.

Acknowledgements

We are very grateful to P.Chambon for his continuous interest in this work and for very helpful advice and discussions. We thank J.F.Lefevre for useful discussions on binding kinetics, N.Miyamoto for pML15 and M13BA, J.M.Chipoulet for technical assistance. A.Staub for oligonucleotide synthesis and especially M.Acker for her expertise in cell culture. We also thank B.Wasylyk, G.Albrecht and V.Moncollin for reading the manuscript and the administrative staff for preparation of the manuscript. A.C.L. was the recipient of a long-term fellowship from EMBO. This work was also supported by the 'Action incitative de Valorisation du CNRS (85–86) and by a grant by MIR-Société IBF (No. 85T0318).

References

Barrera-Saldana,H., Takahashi,K., Vigneron,M., Wildeman,A., Davidson,I. and Chambon,P. (1985) *EMBO J.*, **4**, 3839–3849.

Baty,D., Barrera-Saldana,H.A., Everett,R.D., Vigneron,M. and Chambon,P. (1984) *Nucleic Acids Res.*, **12**, 915–932.

Bertrand-Burggraf,E., Lefèvre,J.F. and Daune,M. (1984) *Nucleic Acids Res.*, **12**, 1697–1706.

Carthew,R.W., Chodosh,L.A. and Sharp,P.A. (1985) *Cell*, **43**, 439–448.

Chodosh,L.A., Carthew,R.W. and Sharp,P.A. (1986) *Mol. Cell. Biol.*, **6**, 4723–4733.

Corden,J., Wasylyk,B., Buchwalder,A., Sassone-Corsi,P., Kedinger,C. and Chambon,P. (1980) *Science*, **209**, 1406–1414.

Davison,B.L., Egly,J.M., Mulvihill,E.R. and Chambon,P. (1983) *Nature*, **301**, 680–686.

de Vries,E., van Driel,W., van den Heuvel,S.J.L. and van der Vliet,P.C. (1987) *EMBO J.*, **6**, 161–168.

Dunaway,M. and Matthews,K.S. (1980) *J. Biol. Chem.*, **255**, 10120–10127.

Egly,J.M., Miyamoto,N.G., Moncollin,V. and Chambon,P. (1984) *EMBO J.*, **3**, 2363–2371.

Eisenthal,R. and Cornish-Bowden,A. (1974) *Biochem. J.*, **139**, 715–720.

Gidoni,D., Kadonaga,J.T., Barrera-Saldana,H., Takahashi,K., Chambon,P. and Tjian,R. (1985) *Science*, **230**, 511–517.

Giniger,E., Vernum,S.M. and Ptashne,M. (1985) *Cell*, **40**, 767–774.

Grundström,T., Zenke,W.M., Wintzerith,M., Matthes,H.W.D., Staub,A. and Chambon,P. (1985) *Nucleic Acids Res.*, **13**, 3305–3316.

Heberlein,U., England,B. and Tjian,R. (1985) *Cell*, **41**, 965–977.

Hen,R., Sassone-Corsi,P., Corden,J., Gaub,M.P. and Chambon,P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7132–7136.

Hendrickson,W. and Schleif,R. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3129–3133.

Hope,I.A. and Struhl,K. (1986) *Cell*, **46**, 885–894.

Johnson,A.D., Pabo,C.O. and Sauer,R.T. (1980) In Grossman and Moldare (eds), *Methods Enzymol.*, Academic Press, New York, Vol. 65, pp. 839–856.

Jove,R. and Manley,J.L. (1984) *J. Biol. Chem.*, **259**, 8513–8521.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Manley,J.L., Fire,A.,Cano,A., Sharp,P.A. and Gefter,M.L. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3855–3859.

Matthes,H.W.D., Staub,A. and Chambon,P. (1987) In Wu,R. (ed.), *Methods Enzymol.* Academic Press Inc. NY, in press.

Miyamoto,N.G., Moncollin,V., Wintzerith,M., Hen,R., Egly,J.M. and Chambon,P. (1984) *Nucleic Acids Res.*, **12**, 8779–8799.

Miyamoto,N.G., Moncollin,V., Egly,J.M. and Chambon,P. (1985) *EMBO J.*, **4**, 3563–3570.

Moncollin,V., Miyamoto,N.G., Zheng,X.M. and Egly,J.M. (1986) *EMBO J.*, **5**, 2577–2584.

Natarajan,V., Madden,M.J. and Salzman,N.P. (1987) *J. Virol.*, **61**, 646–652.

Pabo,C.O. and Sauer,R.T. (1984) *Annu. Rev. Biochem.*, **53**, 293–321.

Parker,C.S. and Topol,J. (1984) *Cell*, **37**, 273–283.

Pirrotta,V., Chadwick,P. and Ptashne,M. (1970) *Nature*, **227**, 41–44.

Pruijn,G.J.M., van Driel,W. and van der Vliet,P.C. (1986) *Nature*, **322**, 656–659.

Rosenfeld,P.J. and Kelly,T.J. (1986) *J. Biol. Chem.*, **261**, 1398–1408.

Sawadogo,M. and Roeder,R.G. (1985) *Cell*, **43**, 165–175.

Shi,X.-P., Lee,R. and Weinmann,R. (1986) *Nucleic Acids Res.*, **14**, 3729–3744.

Takahashi,K., Vigneron,M., Matthes,H., Wildeman,A., Zenke,M. and Chambon,P. (1986) *Nature*, **319**, 121–126.

Takeda,Y., Folkmanis,A. and Echols,H. (1977) *J. Biol. Chem.*, **252**, 6177–6183.

Treisman,R. (1986) *Cell*, **46**, 567–574.

Wei,R., Wilkinson,H., Pfeifer,K., Schneider,C., Young,R. and Guarente,L. (1986) *Nucleic Acids Res.*, **14**, 8183–8188.

Winter,R.B. and Von Hippel,P.H. (1981) *Biochemistry*, **20**, 6948–6960.

Yu,Y.T. and Manley,J.L. (1984) *Nucleic Acids Res.*, **12**, 9309–9321.

Received on April 13, 1987; revised on June 29, 1987