

Specific interaction of cellular factors with the B enhancer of polyoma virus

Jacques Piette, Marie-Hélène Kryszke and Moshe Yaniv

Unité des Virus Oncogènes, Département de Biologie Moléculaire, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cedex 15, France

Communicated by M. Yaniv

Specific interactions between proteins from mouse 3T6 cells and the enhancer sequence of polyoma virus were detected using the method of band shifting on polyacrylamide gels. Proteins eluted from 3T6 nuclei using a buffer containing 0.55 M NaCl, formed a stable complex with the B enhancer of polyoma virus. At least two different factors are involved in this interaction. The contact sites which were mapped on the DNA sequence using DNase I footprinting correspond to a GC-rich palindrome surrounded by two sequences homologous respectively to the immunoglobulin and to the immunoglobulin and SV40 enhancers. Moreover *Bal31* deletion analysis confirmed that similar sequences are required for the formation of the complex. In spite of a common function and partial sequence homology among some enhancers, neither the polyoma A enhancer, the mouse immunoglobulin heavy chain gene enhancer, nor the origin-promoter-enhancer region of SV40 efficiently competed with the polyoma B enhancer for the binding of these molecules.

Key words: DNA-protein interactions/enhancers/polyoma virus

Introduction

Enhancers are DNA sequences that stimulate transcription *in cis* in a position- and orientation-independent manner even when they are located >1000 nucleotides from a gene (for review see Yaniv, 1982; Gruss, 1984). Initially discovered in SV40 and other viruses (Banerji *et al.*, 1981; Moreau *et al.*, 1981), enhancers have also been found in the immunoglobulin genes of mammalian cells (Banerji *et al.*, 1983; Gillies *et al.*, 1983; Neuberger, 1983; Picard and Schaffner, 1984). Many hypotheses can be envisaged to explain their action: for instance as bi-directional entry sites for transcriptional proteins; as organizers of chromatin structure; as regulators of the superhelical density of DNA domains; or as attachment sites to transcriptionally active membrane or matrix sites. However, almost nothing is known yet about the exact mechanism of enhancer function and the molecules involved in this process. Therefore, we began by looking for the existence of proteins which interact with enhancer sequences. For these studies, we chose the polyoma enhancer element. This structure was first defined as a 244-bp sequence located to the late side of the origin of replication (de Villiers and Schaffner, 1981; Figure 1) and has been subdivided into two independently functioning regions. The A element corresponding to the *BclI-PvuII* fragment, and the B element corresponding to the *PvuII-PvuII* fragment (Herbomel *et al.*, 1984; Figure 1). These elements were further analysed and the essential sequences determined using deletion analysis (Herbomel *et al.*, 1984; Veldman *et al.*, 1985), which has revealed a mosaic organization of the activating

sequences. The A (α) or B (β) enhancer elements were found not only to be necessary for the expression from the early promoter but also for the replication of the virus (Tyndall *et al.*, 1981; Luthman *et al.*, 1982; Muller *et al.*, 1983). The replication function of the A or B element could be replaced by heterologous enhancers suggesting that it is the enhancer function *per se* that is necessary to activate replication (de Villiers *et al.*, 1984). Thus, the same sequences are probably involved in both transcription and replication.

A variety of host range mutations which permit growth of the virus in non-permissive embryonal carcinoma cells were located in the enhancer sequences of polyoma (Katinka *et al.*, 1980). This result suggests that enhancer sequences may serve as modulators for cell type-specific gene expression (Yaniv, 1982).

The polyoma enhancer is functionally analogous to the 72-bp repeat of SV40. This region of SV40 is included in a segment of viral DNA which is free of nucleosomes and exhibits a marked DNase I sensitivity (Saragosti *et al.*, 1980). The altered chromatin structure is in part associated with the enhancer sequence itself (Fromm and Berg, 1983; Jongstra *et al.*, 1984) and could be a general feature of active regulatory regions in eukaryotes (Elgin, 1984). DNase I hypersensitivity is also displayed by the polyoma virus enhancer (Herbomel *et al.*, 1981). The existence of specific cellular DNA-binding proteins was proposed to explain the formation and maintenance of an altered chromatin conformation in the region of promoters or enhancers (Herbomel *et al.*, 1981; Cereghini and Yaniv, 1984). A search for such proteins has recently revealed a protein that binds to the upstream control sequences of the chicken β -globin gene (Emerson and Felsenfeld, 1984). Chromatin assembly *in vitro* in the presence of this protein creates a nuclease-sensitive site similar to that observed *in vivo*. Also, studies using the *Drosophila* heat shock promoter have revealed two specific proteins, one which binds the TATA box and another which binds the heat shock regulatory element sequences that are exposed to DNase I in chromatin (Wu, 1984a, 1984b; Parker and Topol, 1984). Indirect evidence for cell-specific factors which interact with enhancers include the observation of species and tissue specificity in their activity (Laimins *et al.*, 1982; de Villiers *et al.*, 1982). The tissue specificity is most pronounced in the case of the Ig heavy chain gene enhancer, active only in lymphocytes (Gillies *et al.*, 1983; Banerji *et al.*, 1983). Also *in vivo* or *in vitro* transcription competition experiments using the SV40 enhancer strongly suggest that cellular proteins interact with such elements (Schöller and Gruss, 1984; Wildeman *et al.*, 1984; Sassone-Corsi *et al.*, 1985). Finally, dimethyl sulphate protection experiments performed *in vivo* with the mouse immunoglobulin heavy chain gene enhancer were consistent with the tissue-specific binding of molecules to the enhancer (Ephrussi *et al.*, 1985).

In this paper we show directly that there are specific interactions of cellular proteins, present in mouse 3T6 cells with the B enhancer of polyoma virus. The site of interaction is precisely mapped on the viral DNA sequence using DNase I footprinting and *Bal31* deletion analysis. The behaviour of the specific DNA-

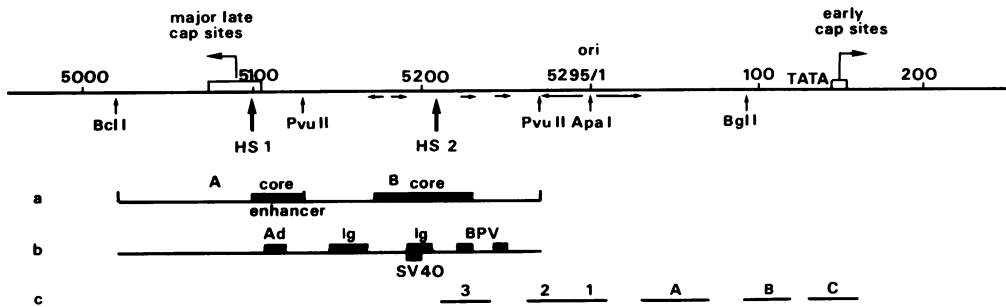


Fig. 1. The regulatory region of polyoma virus. The numbering of nucleotides is according to Tyndall *et al.* (1981). The restriction sites for *BclI*, *PvuII*, *ApaI* and *BglII* are indicated by small vertical arrows; major late cap sites (Cowie *et al.*, 1981) and early cap sites (Cowie *et al.*, 1982) are indicated by boxes with arrows, palindromes or direct repeats by small horizontal arrows; the two DNase I hypersensitive sites mapped in the chromatin (Herbomel *et al.*, 1981) by large vertical arrows termed HS1 and HS2. Also indicated are the origin region (ori) and the TATA box sequence of the early promoter. (a) The 244-bp polyoma *BclII-PvuII* enhancer region is subdivided into the A (*BclII-PvuII*) and B (*PvuII-PvuII*) enhancer elements with the enhancer core sequences identified by Herbomel *et al.* (1984) represented by black bars. (b) The homologies between the polyoma enhancer and the enhancers of adenovirus (Hearing and Shenk, 1983), Ig heavy chain gene enhancer (Banerji *et al.*, 1983), BPV (Weiher and Botchan, 1984) and SV40 (Weiher *et al.*, 1983) are represented by black bars. (c) Binding sites of large T antigen as deduced from footprinting experiments are shown (Cowie and Kamen, 1984; Dilworth *et al.*, 1984).

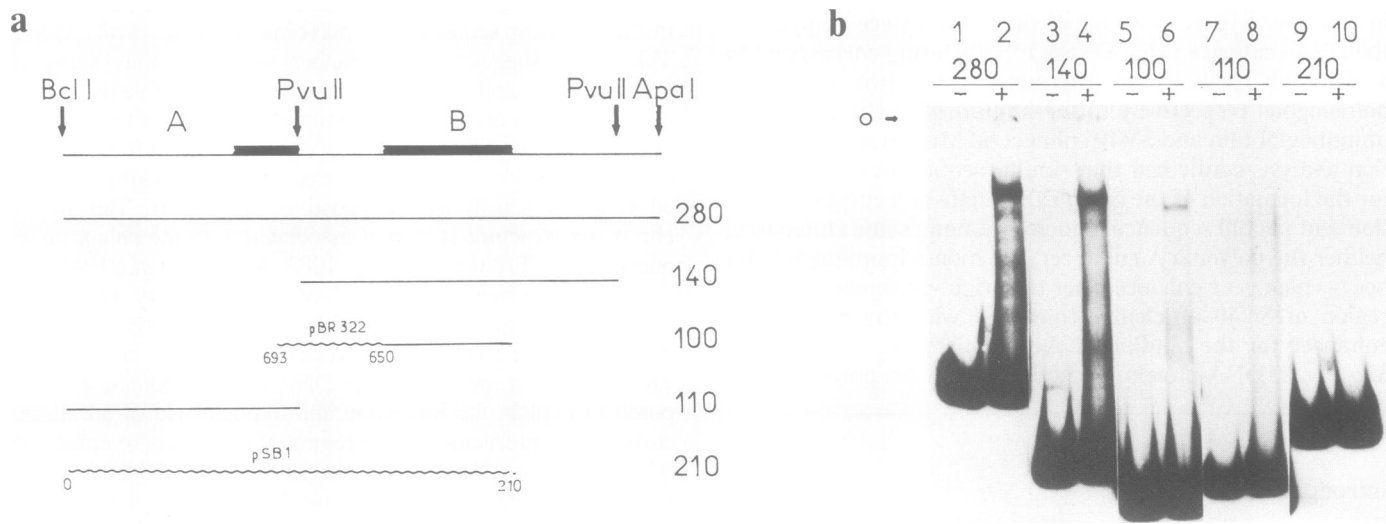


Fig. 2. Retardation of protein-DNA complexes on polyacrylamide gel (a) Description of DNA fragments used in the experiments is given in Figure 2a. Straight lines represent DNA of the polyoma enhancer region, wavy lines DNA of prokaryotic origin. The 280-bp fragment contains the entire enhancer region, the 140-bp fragment, the B enhancer, the 100-bp fragment, the core of the B enhancer plus plasmid vector DNA, i.e., nucleotides 650–693 of pBR322 (Sutcliffe, 1979), the 110-bp fragment, the A enhancer and the 210-bp fragment exclusively plasmid vector DNA, i.e., nucleotides 0–210 of pSB1 (Herbomel *et al.*, 1983). (b) DNA-protein incubations were made as explained in Materials and methods and loaded on a 5% polyacrylamide gel. About 10^{-2} pmol of the fragments shown in a were used as probes in the presence of 1 μ g salmon sperm DNA and incubated without (–) or with (+) 12 μ g protein of a 0.4 M nuclear extract. (1) and (2) = fragment 280; (3) and (4) = fragment 140; (5) and (6) = fragment 100; (7) and (8) = fragment 110; (9) and (10) = fragment 210. After electrophoresis the gel was fixed, dried and autoradiographed.

protein complex in polyacrylamide gels suggests that several protein molecules are interacting with the enhancer sequences.

Results

Proteins from a nuclear extract of mouse 3T6 cells interact specifically with the B enhancer of polyoma virus

To study the interaction of cellular components with the enhancer sequence of polyoma virus, we have chosen to use a method which detects the slower migration of DNA-protein complexes with respect to free DNA on polyacrylamide gels (Garner and Revzin, 1981). This method is quite sensitive, it can detect a complex with a binding constant of $10^5/M$ (Fried and Cortners, 1981). In practice, we mixed a radioactively labelled DNA fragment containing the enhancer sequence with a nuclear extract. The sample was then electrophoresed through a polyacrylamide gel, which separates the naked fragment from the DNA-protein complexes.

To eliminate the non-specific binding of nuclear proteins to test DNA, high amounts of cold sonicated salmon sperm DNA or restricted plasmid DNA were added as carrier. Polyoma or prokaryotic labelled DNA fragments of similar size to the enhancer fragment but lacking enhancer activity were analysed under the same conditions to make sure that all non-specific binding was eliminated (Strauss and Varshavsky, 1984).

As detailed in Materials and methods, nuclei were prepared, washed with 0.1 M NaCl and then resuspended in a higher NaCl concentration to elute nuclear proteins that may interact with DNA. As is shown in Figure 2b, the migration of the polyoma 244-bp *BclII-PvuII* complete enhancer fragment was specifically retarded by nuclear factor(s) present in the 0.4 M NaCl wash of 3T6 cell nuclei (lane 2). In fact, most of the binding activity was eluted with a stepwise wash between 0.3 and 0.55 M NaCl; it was defined as the 0.55 M wash fraction. The proteins of this fraction represent <2% of the total cellular proteins. Most of

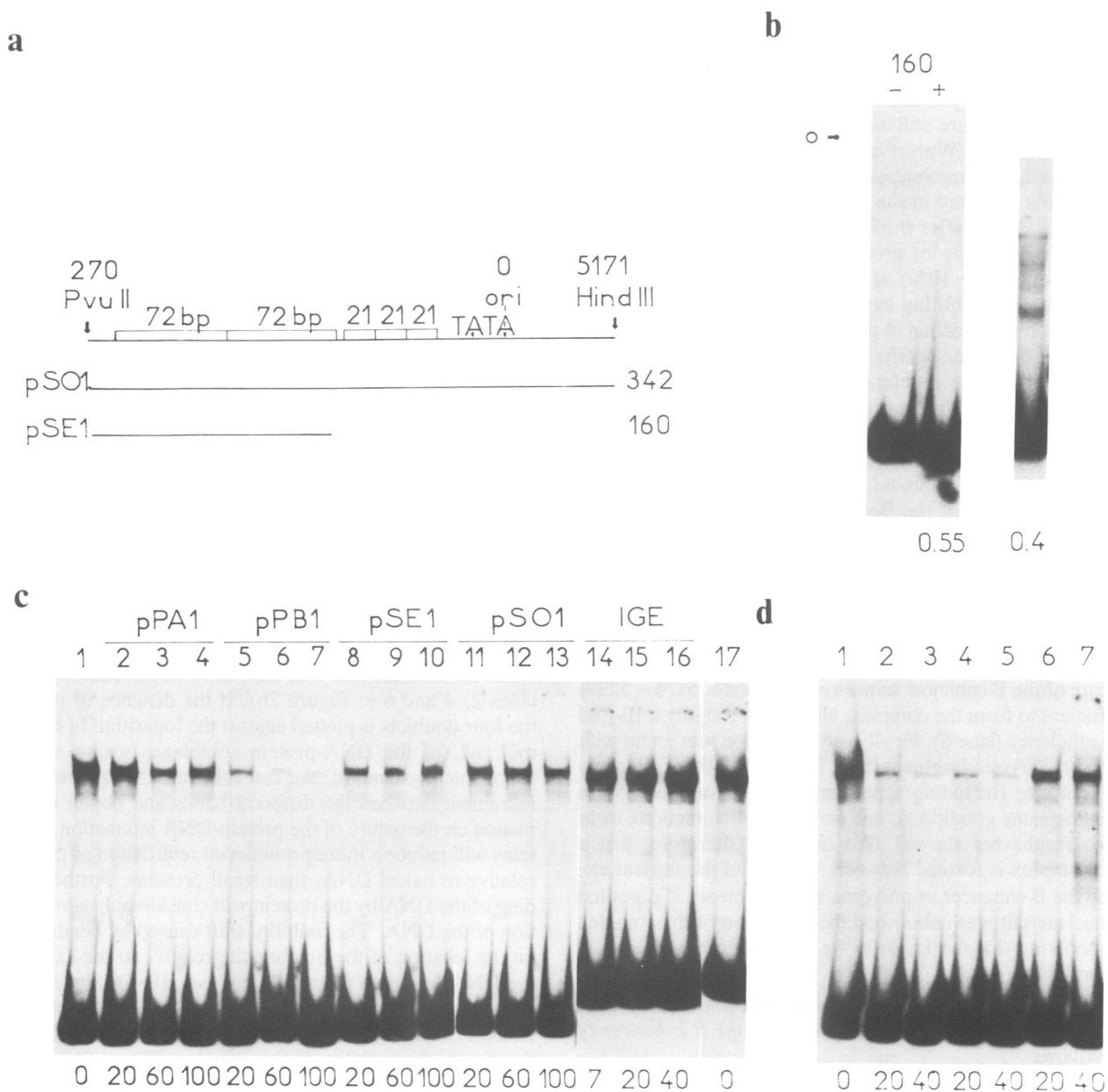


Fig. 3. Competition experiments. (a) SV40 fragments used as competitors. The regulatory region of SV40 is shown. The 72-bp repeats and 21- (22-) bp repeats are boxed. The TATA box of the early promoter and the origin region are indicated. The 342-bp fragment contains the whole origin-promoter-enhancer region, whereas the 160-bp fragment contains only the 72-bp repeats. The competitor fragments were cloned in pML2 and the plasmids obtained were called pPA1 for the A enhancer of polyoma (*BclI-PvuII* fragment), pPB1 for the B enhancer of polyoma (*PvuII-PvuII* fragment), pSE1 for the 160-bp SV40 enhancer fragment and pSO1 for the 342-bp SV40 origin fragment. IGE is the mouse heavy chain gene *EcoRI-PstI* enhancer fragment inserted in the pBR322 *BamHI* site (Sassone-Corsi *et al.*, 1985). (b) Binding of the 0.4 M and 0.55 M extracts to the SV40 enhancer. DNA-protein incubations were made as explained in Materials and methods and loaded on a 5% polyacrylamide gel. About 10^{-2} pmol of the 160-bp SV40 enhancer fragment as shown in Figure 3a, was used as probe in the presence of 1 μ g salmon sperm DNA and incubated without (-) or with (+) 1.5 μ g protein of a 0.55 M nuclear extract or with 0.4 M nuclear extract (right gel). (c) The incubations were made as explained in Materials and methods. The 140-bp B enhancer fragment was used as probe in the presence of different amounts of *EcoRI*-linearized pML2 and pML2 derivatives containing the competitor fragments and incubated with 12 μ g protein of a 0.4 M-nuclear extract except for IGE where 1.5 μ g of a 0.55 M nuclear extract was used. The competitor DNA was mixed with the DNA probe before the addition of the protein fraction. No competitor was added in lane 1; pPA1 was added as competitor in lanes 2-4, pPB1 in lanes 5-7, pSE1 in lanes 8-10, pSO1 in lanes 11-13 and IGE in lanes 14-16. Respectively 1 μ g (20-fold molar excess over probe), 3 μ g (60-fold molar excess) and 5 μ g (100-fold molar excess) of competitor DNA were added, except for IGE where 0.5 μ g (7-fold molar excess), 1.5 μ g (20-fold molar excess) and 3 μ g (40-fold molar excess) were added. The same molar excess of pPB1 displayed a strong competition with the extract used for the competition with IGE. Lane 17 = no competitor added with the extract used for IGE competition. The total amount of cold DNA was brought to 5 μ g with pML2. (d) The DNA-protein incubation was made as described in Materials and methods with the modifications explained below. The 140-bp B enhancer fragment was used as probe in the presence of pML2 and pPB1 and incubated with 4 μ g protein of a 0.4 M nuclear extract. (1) 2 μ g pML2 were used. (2) 1 μ g pML2 was used and 1 μ g pPB1 was added at the same time as the labelled fragment. (3) 2 μ g pPB1 were added at the same time as the labelled fragment. (4) 1 μ g pML2 was used and 1 μ g of pPB1 was added and incubated with proteins without addition of labelled fragment for 10 min, only then was the labelled fragment added and incubation continued for another 10 min. (5) Same as 4 but 2 μ g pPB1 were used without pML2. (6) 1 μ g pML2 was used and the labelled fragment was added and incubated with proteins without addition of pPB1 for 10 min, only then was 1 μ g pPB1 added and incubation continued for another 10 min. (7) Same as 6 but 2 μ g pPB1 were used without pML2.

the non-histone proteins are already eluted from nuclei at 0.3 M NaCl, and only about one fifth of this amount is present in the subsequent 0.55 M wash. No prominent histone bands are detectable on SDS-polyacrylamide gels of this extract: the core nucleosome histones are still associated with the chromatin at this salt concentration (Worcel *et al.*, 1983), and histone H1 was eliminated during an ammonium sulphate precipitation step. The matrix proteins also stay in the nucleus under these conditions (Lewis *et al.*, 1984). The 0.55 M extract contained traces of nuclear DNA (20 ng/ μ g of protein).

Only a fraction (5–10%) of the labelled DNA fragment is retarded in the band-shifting experiments. The specific factors are certainly a minor fraction of the proteins in the extract (among which are many non-specific DNA-binding proteins), and therefore increasing the amount of protein in a reaction leads inevitably to non-specific interactions with the probe (results not shown).

The *PvuII-PvuII* fragment, or the B enhancer displayed the same pattern of retarded bands, with the same mobility relative to free DNA, as that of the *BclI-PvuII* 244-bp fragment (lane 4). This suggested that the complex was formed with the B enhancer; indeed we found no specific interaction between the A enhancer and the 0.4 M nuclear extract (lane 8). A faint band was nevertheless seen at longer exposures: this binding activity was more prominent in 1 M washes of nuclei but was not specific for the A enhancer as revealed by competition experiments. The 55-bp core of the B enhancer sequence (nucleotides 5175–5229) was sufficient to form the complex, albeit with roughly a 10-fold weaker efficiency (lane 6). Finally, no interaction was found with a 210-bp prokaryotic fragment (lane 10). Three other fragments of different size (including a polyoma coding fragment) were tested in the same conditions, and no specific interactions were detected (results not shown). We conclude, therefore, that a specific complex is formed between factor(s) of the nuclear extract and the B enhancer of polyoma virus. Protease K digestion of the nuclear salt wash abolished the formation of the complex clearly showing that the factor(s) we are studying is at least in part a protein.

Other enhancer or promoter elements tested compete very inefficiently for the binding of cellular factors to the B enhancer of polyoma virus

Using the same approach, we detected a strong interaction of proteins from the 0.4 M nuclear extract with the 342-bp origin-promoter-enhancer fragment or the 160-bp enhancer fragment of SV40 virus. The 0.55 M nuclear extract was much less reactive however (Figure 3B). The existence of a homology with the SV40 and Ig enhancers (Weiher *et al.*, 1983; Banerji *et al.*, 1983) prompted us to perform a series of competition experiments to test if factors binding to these enhancer elements were identical to those binding to the B enhancer of polyoma (see Figure 3a and c). To do so, the different fragments were cloned in either pBR322 or in its derivative, pML2 (Lusky and Botchan, 1981), and the *EcoRI* linearized plasmids were used as competitors. Only the pML2 derivative containing the B enhancer (pPB1) competed strongly for binding to the labelled B enhancer fragment (Figure 3C, lanes 5–7). The other derivatives containing the polyoma A enhancer (pPA1, lane 2–4), the SV40 72-bp repeat (pSE1, lanes 8–10), the SV40 origin-promoter-enhancer region (pSO1, lanes 11–13) or mouse heavy chain gene *EcoRI-PstI* enhancer fragment (IGE, lanes 14–16) competed weakly or not at all with the B enhancer (we used up to a 100-fold molar excess of competitor). It is clear from Figure 3 that only for the homologous fragment the decrease in the amount of complex is roughly pro-

portional to the molar excess of the competing DNA.

The order in which the different components were added to the reaction is crucial: no competition was observed if the labelled fragment was added before the cold competitor DNA (see Figure 4d, lanes 6–7), whereas efficient competition was observed when the cold competitor DNA was added first (lanes 4–5) or mixed with the radioactive fragment before the addition of the protein (lanes 2–3). As only a 20- or 40-fold molar excess of competitor was used in these experiments, a residual binding to the probe is still observed. These results suggest that the half-life time of the specific DNA-protein complex is relatively long since no significant competition was observed when excess non-radioactive homologous fragment was added after the formation of the complex. All together these studies demonstrate that the complex formed between the B enhancer and the nuclear factors is relatively stable and quite specific for this enhancer sequence.

Several proteins are interacting with the polyoma B enhancer

The distribution of DNA on the polyacrylamide gels had a quite reproducible pattern (Figure 4). The fastest migrating species corresponded to the naked DNA fragment. The slower migrating bands, made up of four doublets, corresponded to the DNA-protein complexes. The most abundant doublet dd' (the separation between the two sub-bands is not visible in the figures) was the slowest to migrate. The mobility of this band increased with the decrease in the size of the DNA fragment containing the B enhancer element that was used to form the complex (compare lanes 2, 4 and 6 in Figure 2b). If the distance of migration of the four doublets is plotted against the logarithm of the apparent mol. wt. of the DNA-protein complexes, using naked DNA fragments as standard, we can calculate the apparent mol. wt. increment between the different bands and obtain some information on the nature of the protein-DNA interaction. Large proteins will induce a more pronounced retardation of the complex, relative to naked DNA, than small proteins. Furthermore bending of the DNA by the protein will considerably slow the migration of the DNA. The mobility shift caused by bending depends on the location of the binding site relative to the extremities of the DNA fragment (Wu and Crothers, 1984). Since, in our experiments, the location of the core element of enhancer B relative to the extremities of the DNA fragment did not seem to affect drastically the mobility of the complex (Figure 2b), we believe that its shift is mainly due to the increase in mass brought about by the binding of proteins. The mobility shift of free DNA to position aa' (Figure 4) is probably due to interaction with a relatively high mol. wt. protein. The fact that the mol. wt. increment between the first three doublets (a to b and b to c) is constant suggests that the same protein may be added in the second and third steps of the band ladder (see Figure 2 of Fried and Crothers, 1981, and discussion of this figure). The last doublet is slightly shifted: this could be due to the addition of a different subunit or to some conformational alteration in the final complex. Altogether these results strongly suggest that the quantal change in mobility of the bands is caused by the binding of several polypeptide chains. Finally, with some preparations and in some experiments other faint bands were visible: this could be explained by the presence of another binding activity in some preparations or more probably by the partial degradation of the proteins of the complex during the preparation of the extract or during incubation.

The fact that even when lower amounts of protein are used the upper band is always the major one suggests that the binding of the different factors is cooperative, or alternatively that a pre-existing multimeric factor is partly dissociated in the gel.

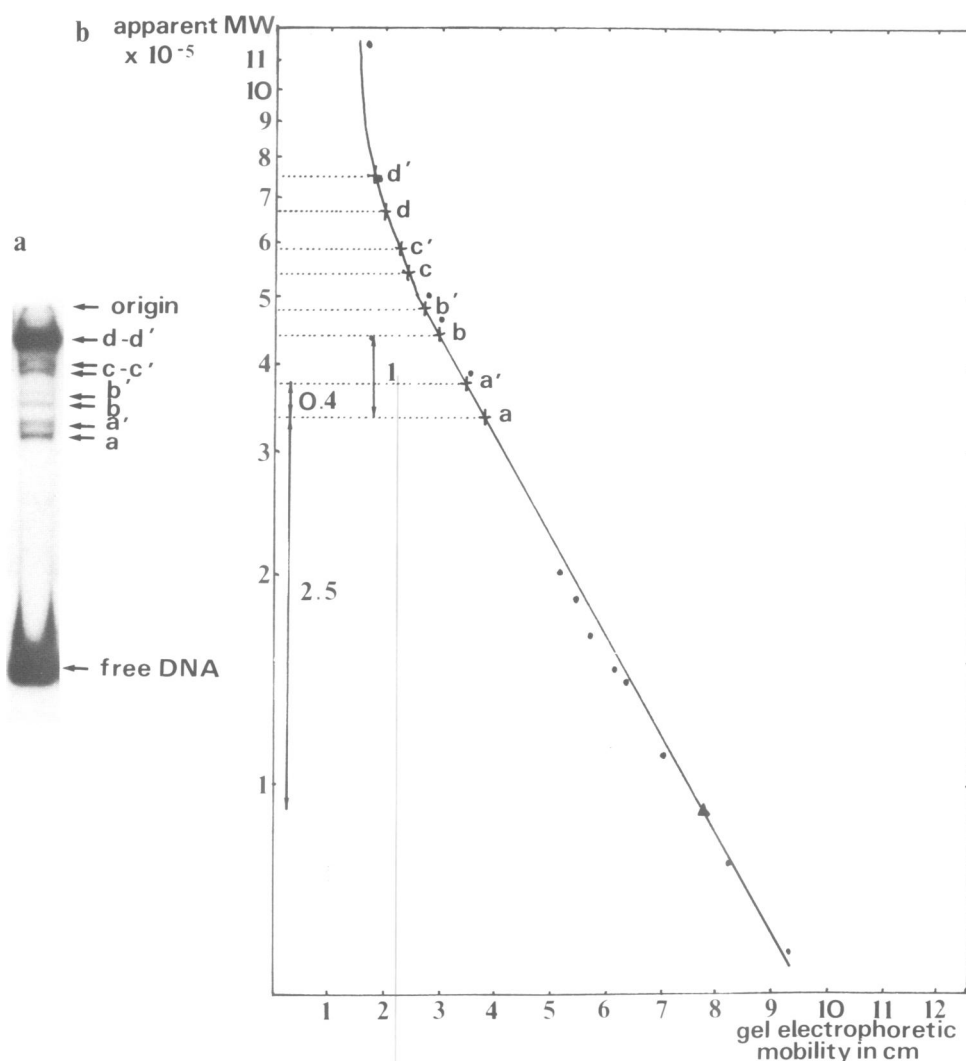


Fig. 4. Interpretation of the band pattern observed on polyacrylamide gels. (a) Pattern of intermediate bands as observed on 7.5% polyacrylamide gels. The 140-bp fragment was used as a probe in the presence of 5 μ g salmon sperm DNA and incubated with 4.8 μ g protein of a 0.4 M nuclear extract. The different bands are labelled by letters. On less exposed autoradiographs, the upper band can be seen as a doublet. (b) The apparent mol. wt. of the intermediates is plotted against the electrophoretic mobilities of the corresponding bands in a 5% polyacrylamide gel. DNA fragments used as markers are indicated by dots. The different intermediates are represented by crosses. The apparent mol. wt. increment for three presumed subunits is given in the ordinate.

The DNA sequences protected by binding of the cellular factor(s) correspond to essential sequences as defined by in vivo functional tests

To try to localize more precisely the DNA sequences interacting with the cellular factors, we performed DNase I footprinting experiments (Galas and Schmitz, 1978). As only a small fraction of DNA molecules were specifically complexed in our binding assay, we decided to separate the complexed fragment from the free DNA on a polyacrylamide gel after the DNase I digestion and to compare the patterns obtained from both fragments on a DNA sequencing gel. As both control DNA and complexed DNA are incubated together with DNase I and subsequently separated on the basis of the existence of a stable complex, the differences in the DNase I pattern between both DNAs is most probably due to the stable interaction of some proteins with the shifted DNA fragment. The results of such an experiment are shown in Figure 5. Even after separation of the complex, the protection observed is not complete for several possible

reasons. The upper band may still be contaminated with the non-specific complexes that smeared on the preparative gel. Moreover, the protection against DNase I cleavage by enhancer binding proteins may be only partial. To obtain a quantitative estimate of the protection pattern we scanned the autoradiograph shown in Figure 5 with a microdensitometer and measured the height of the corresponding peaks. To visualize protection or increased sensitivity we used a graphic representation similar to that of Ogata and Gilbert (1978). The logarithm of ratios between the peak heights of control and complex channels is displayed relative to the nucleotide sequence. If we consider as significant for protection values >0.15 and for enhancement of cleavage values <-0.15 , then a protected region can be discerned including the GC-rich palindrome and the SV40 and the Ig enhancer homology. The pattern of protection is schematized in Figure 6 by brackets: the upper strand is partially protected from nucleotide 5140 to nucleotide 5200. In addition to protection, we observe an increase in the frequency of several DNase I

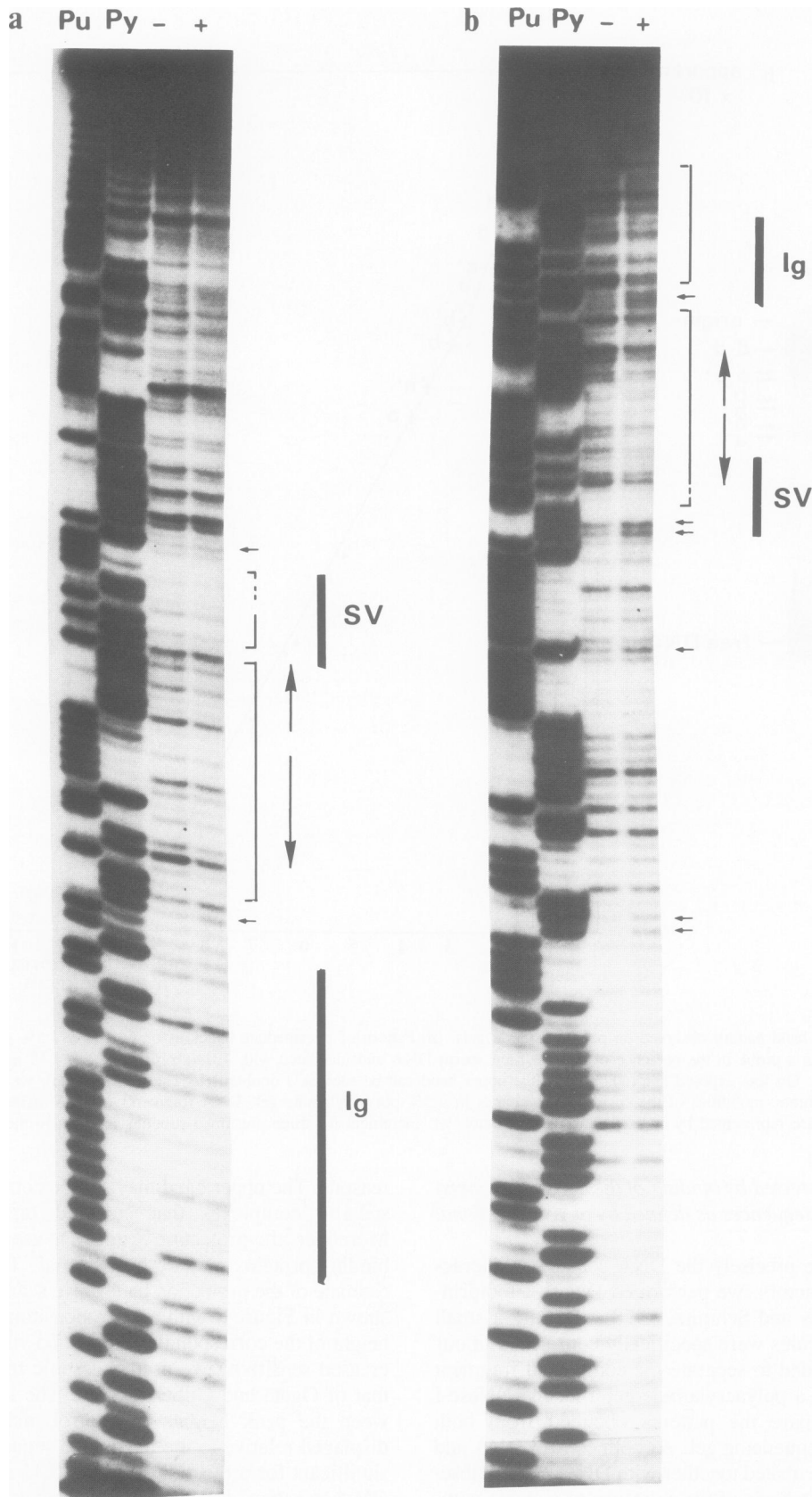


Fig. 5. DNase I footprinting of the B enhancer of polyoma virus. The experiment was performed as explained in Materials and methods. **(a)** Footprint of the 3'-labelled lower strand as depicted in Figure 6. Pu column = G+A chemical degradation products, Py column = C>T chemical degradation products, - column stands for protein-free DNA corresponding to the lower band of a retardation gel, + column stands for complexed DNA corresponding to the uppermost band of a retardation gel. **(b)** Footprint of the 3'-labelled upper strand as given in Figure 6. Symbols are as in (a). Protected regions are indicated by brackets along the gel and enhanced digestion by arrows. Also indicated are the GC-rich palindrome and the SV40 and Ig enhancer sequence homologies.

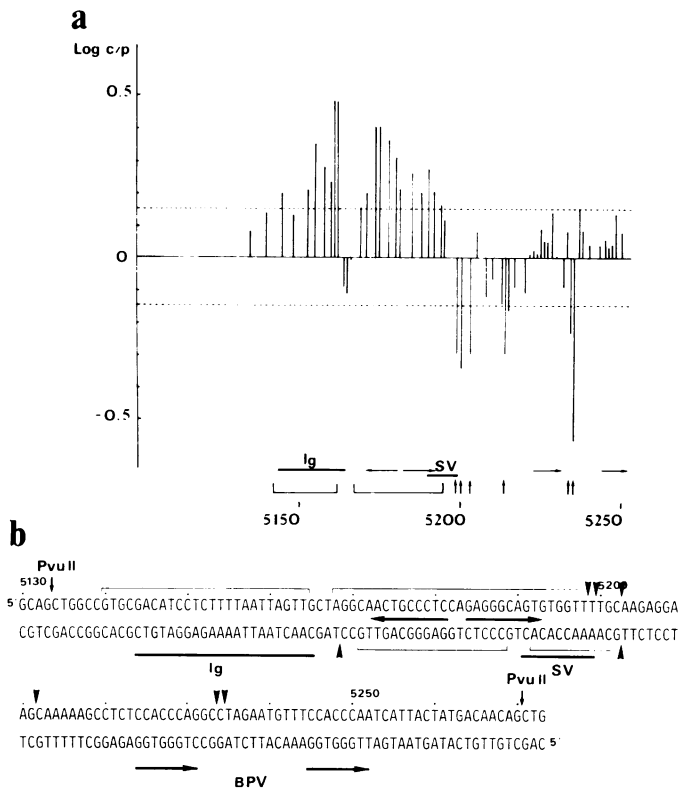


Fig. 6. (a) Densitometer scanning of the footprinting experiment. The autoradiograph shown in Figure 5 was scanned with a densitometer and analysed according to Ogata and Gilbert (1978). The logarithm of the ratio of the peak height of the control DNA band (-) to that of the protein complexed DNA band (+) (log c/p) was displayed in the ordinate against the position of the bands in the DNA sequence. A positive value corresponds to protection and a negative one to enhancement of the DNase I digestion. Only values >0.15 and <-0.15 are considered as significant. Also indicated are the GC-rich palindrome, the sequences homologous to Ig and SV40 enhancers and the short repeats homologous to the BPV enhancer. Protected regions are indicated by brackets, enhancements of digestion by vertical arrows. (b) Pattern of the DNase I footprint on the DNA sequence. The sequence is numbered as in Tyndall *et al.* (1981). Protected regions are indicated by brackets, enhanced digestion by arrows. Also indicated are: the GC-rich palindrome by two horizontal divergent arrows; the homology with the SV40, Ig and BPV enhancer sequences by black bars.

cleavage sites as indicated in Figures 5 and 6 by arrows. The pattern of protection of the other strand is less obvious, a partial protection of the GC-rich palindrome is visible (from nucleotide 5158 to nucleotide 5203). Minor contacts or structural alterations in the DNA outside this region cannot be completely excluded since we observe an increase in the cleavage of certain sites in the origin-proximal sequences (Figure 6). The protected region corresponds to the minimal essential sequences for the activation of transcription and replication as defined in functional tests by different groups (Tyndall *et al.*, 1981; Muller *et al.*, 1983; Herbomel *et al.*, 1984; Veldman *et al.*, 1985; see Discussion).

Identification of the minimal sequences required for the formation of the complex

Are the sequences defined by the footprinting experiment sufficient and necessary for the formation of a specific complex? To answer this question we prepared a set of fragments of different length by *Bal31* digestion: all these fragments have the same labelled end, only the other end is variable (see Materials and methods). The position of this variable end on the DNA sequence

can be determined on a sequencing gel. The set of unidirectional deletions was incubated with a 0.55 M nuclear extract, run on a polyacrylamide gel and the retarded fragments were eluted and analyzed on a sequencing gel (Figure 7).

The ability of the fragments deleted from the late side of the B enhancer to form a stable complex with the 0.55 M nuclear extract decreased drastically upon removal of the Ig homology. Then, the nuclease entry into the GC-rich palindrome totally abolished the formation of the complex (compare lanes 11 and 12 in Figure 7). With the fragments deleted from the early side of the B enhancer, a gradual decrease of the binding (to $\sim 10\%$) is first observed after the deletion of the origin-proximal bovine papilloma virus (BPV) homology. The deletion of the beginning of the GC-rich palindrome results in a complete loss of protein-fragment interaction (Figure 7, lanes 5 and 6).

These results suggest that the Ig and BPV homology regions may play a role in the interaction of the B enhancer with the 0.55 M nuclear extract, nevertheless their removal does not abolish totally the formation of this complex. A non-specific size effect cannot be excluded at this stage: smaller fragments being less easily bound than larger ones. However the observation that fragments with equal size can differ in their binding capability depending on the direction of the exonuclease digestion (compare lanes 6 and 12) makes this interpretation less probable. In conclusion, only fragments containing at least the GC-rich palindrome are able to form a stable complex with the 0.55 M nuclear extract. Sequences on each side of this region increase the stability of the complex.

Discussion

The methodology used in our work is a powerful technique for detecting sequence-specific DNA-protein interactions despite some technical imperfections. In the first step, components from nuclear extracts are analysed for specific electrophoretic retardation of the DNA fragment of interest, each particular complex giving rise to a discrete electrophoretic species (Fried and Crothers, 1981). The precise site of interaction with the DNA fragment can then be mapped by using partial deletions or competing fragments or more accurately by DNase I footprinting (or other methods to analyse DNA-protein contacts) followed by separation of the protein-DNA complex from the rest of the DNA on a polyacrylamide gel and comparison of the patterns obtained from both samples on a sequencing gel.

We found that proteins from a nuclear extract from 3T6 cells bind specifically to the polyoma virus B enhancer. The factors interacting with the B enhancer are washed from nuclei of 3T6 cells with salt between 0.3 and 0.55 M. The pattern of bands obtained on polyacrylamide gels suggests that the complex is composed of at least two subunits, one larger and one or possibly several smaller ones. The complex could be formed by the cooperative interaction of the different molecules, the larger protein interacting first with the DNA. Alternatively, an enhancer-binding particle could pre-exist in the cell. The faster migrating bands would then represent dissociation products of this particle. It must be stressed that the footprint reflects the DNA-protein interaction in the complex isolated from the uppermost band, thus containing the whole set of factors or subunits of the particle. This band disappears when the assay is performed in 0.3 M NaCl, although lower bands are still visible. One or more components of the complex are probably dissociating at this ionic strength.

A few potential candidates for the binding factors we detect can be ruled out by the absence of strong competition with defined DNA sequences: the TATA box binding factor (Davison *et al.*,

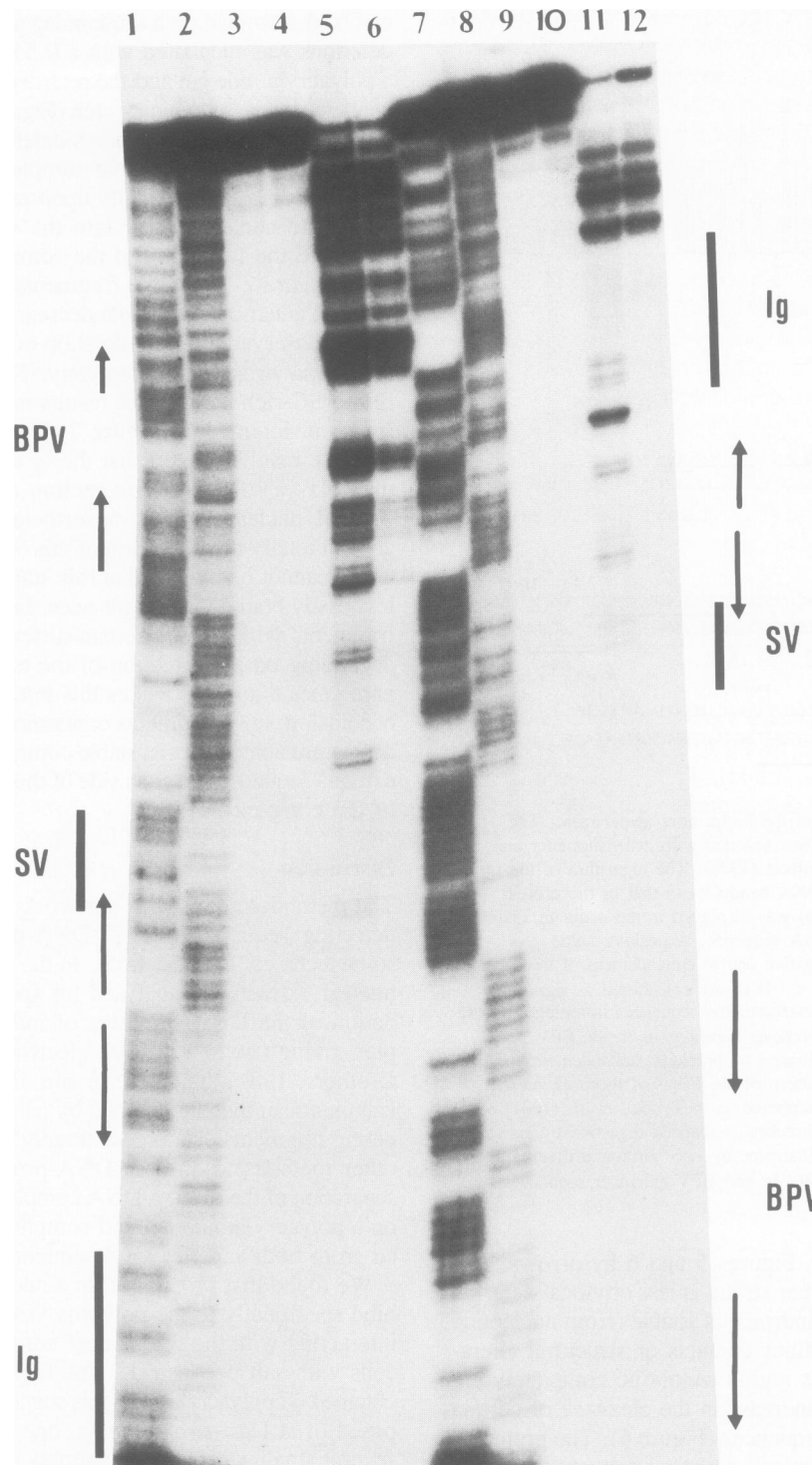


Fig. 7. *Bal31* deletion analysis. A mixture of fragments deleted unidirectionally from either the origin proximal or the late region proximal *PvuII* sites were labelled at their invariable 3' end and incubated with 2 μ g of a 0.55 M nuclear extract for 10 min. Complexed fragments were separated from free DNA on a 5% polyacrylamide gel, the DNA eluted and loaded on a 6% sequencing gel. Deleted but not complexed fragments were taken as controls together with purine or pyrimidine sequence ladders. **Lanes 1–6**, DNA fragments labelled at the late side (*PvuII* site at nucleotide 5133); **lanes 7–12**, DNA fragments labelled at the early side (*PvuII* site at nucleotide 5270); **lanes 1 and 7**, purine sequence reactions; and **lanes 2 and 8**, pyrimidine sequence reactions; **lanes 3 and 6**, control fragments; **lanes 4 and 10**, retarded fragments of a short *Bal31* digestion; **lanes 5 and 11**, control fragments; and **lanes 6 and 12**, retarded fragments of a long *Bal31* digestion.

1983), the Sp1 factor binding to the GGGCGG motifs of the 21-bp repeats of SV40 (Dyan and Tjian, 1983) that was proposed by Mueller *et al.* (1984) to bind to partially homologous polyoma sequences or a putative factor binding to the Weiher-Gruss enhancer consensus sequence (Weiher *et al.*, 1983).

The enhancer-binding proteins that we detected are specific for the B element of polyoma. No strong competition was observed when a large excess of the polyoma A, immunoglobulin heavy chain or SV40 enhancer sequences were added. In agreement with these results, Sassone-Corsi *et al.* (1985) did not find strong

competition between the entire polyoma enhancer and the SV40 enhancer in their enhancer-dependent *in vitro* transcription assay. The factors seem, thus, to be quite specific for the B enhancer, although it is not excluded that, in addition to a specific factor, other proteins can interact with several enhancers and are generally involved in the process of enhancement. It is perhaps significant that the GC-rich palindrome, that appears to be crucial for the recognition of the polyoma B enhancer, does not show any extensive homology to the Ig heavy chain enhancer (Banerji *et al.*, 1983; Gillies *et al.*, 1983). In contrast, it is surrounded by two protected sequences showing an almost perfect homology with Ig enhancer sequences. In cells where the Ig enhancer is active, another specificity factor will be required to promote binding to the homologous sequences. Comparison with the σ -like factors conferring promoter specificity to RNA polymerase in *Bacillus subtilis* is tempting here (Losick and Pero, 1981).

The polyoma DNA sequences interacting with the proteins are similar to those required *in vivo* for viral transcription and replication. The 55-bp core enhancer sequence (nucleotides 5175–5229) is still capable of complex formation, although with a largely lower efficiency than the whole B fragment (nucleotides 5130–5625). Similarly *in vivo* this element is less active than the entire B fragment. The origin-proximal (right hand) limit of this region was mapped by Veldman *et al.* (1985) between nucleotides 5179 and 5214. The left hand border was functionally defined (by DNA replication assay) to be between nucleotides 5131 and 5182 (Muller *et al.*, 1983). The GC-rich palindrome is protected in DNase I digestion experiments of chromatin and a hypersensitive site was found at the right hand side of the Weiher-Gruss consensus sequence homology (Herbomel *et al.*, 1981), again in concordance with our results. The GC-rich palindrome is also conserved in all polyoma PCC4 host range mutants, stressing its importance for the enhancer function (Melin *et al.*, 1985).

Using *Bal31* deletions and DNase I footprinting we tried to define more precisely the nucleotides within the B enhancer involved in DNA-protein contacts. The deletion analysis has shown that the GC-rich palindrome is absolutely required for complex formation. Sequences on its late side (the IgH homology) and on its early side (the SV40 and part of the BPV1 homology) drastically increase the capability to form stable complexes. The DNase I protection experiments show that altogether ~30 bases are weakly protected along the late strand (bottom strand in Figure 5) and ~60 bases are more strongly protected along the early strand (upper strand). The protected region includes the GC-rich palindrome and the IgH and SV40 homologies. In addition, certain DNase I hypersensitive sites are observed next to the protected sequences and the BPV1 homology. These results are in agreement with our hypothesis that several proteins interact with the enhancer sequences. The entire complex is more stable than partially formed entities. Although no clear protection against DNase I was detected for all the sequences involved in the formation of the most stable complex, it can be envisaged that some of the DNA sequences are interacting with proteins without being strongly protected against the nuclease. Furthermore, the fact that the size of the residual complex formed on the core sequence is comparable with that of the complex formed with a whole B fragment (Figure 2) suggests that protein-protein contacts are perhaps as important as protein-DNA interactions. After recognition of the GC-rich palindrome by the large subunit, the other factors might interact with the bound protein and the complex might subsequently be stabilized by the interaction of these proteins with neighbouring DNA sequences.

The asymmetry in the interaction we observe with both DNA strands was also noted by others for the interaction of Sp1 with the GGGCGG upstream sequences in the 21-bp repeat of the SV40 early promoter (Gidoni *et al.*, 1984), for that of the positive transcription factor (TF III B) with the 5S gene of *Xenopus* (Sakonju and Brown, 1982) or the factor interacting with the histone H3 gene promoter of *Drosophila* (Parker and Topol, 1984). This may be a general characteristic of factors interacting with eukaryotic control regions, reflecting the (still unknown) mechanism of action or the mechanism of transmission of the differentiated state to the cellular progeny (see Brown, 1984).

As previously mentioned, two enhancers (A and B) were defined in the non-coding region of polyoma by cloning viral DNA fragments upstream or downstream of the p-collagen CAT expression vector (Herbomel *et al.*, 1984). Several other groups found that the B element (the *PvuII*-*PvuII* fragment) can enhance the expression of different genes like BPV early region, tk or tk CAT (Lusky *et al.*, 1983; Linney and Donerly, 1983). Furthermore, the same polyoma fragment with a single point mutation increases the strength of this enhancer in EC cells (Linney and Donnerly, 1983) or in both EC cells and fibroblasts (Herbomel *et al.*, 1984). The fragment isolated from mutant strains can enhance tk expression in the four possible orientations. Only Mueller *et al.* (1984) failed to detect enhancement of tk stable transformants with a polyoma fragment missing the A enhancer. They proposed that the B element is functioning only as an upstream sequence of the early promoter. However, in their plasmid constructions the B element was included in a context very different from that of Linney and Donnerly or of Herbomel *et al.* (distance from promoter, presence of a polyoma early transcription unit, etc.). The observation that both enhancers are able to activate replication also contradicts the upstream promoter sequence hypothesis for the B element (Muller *et al.*, 1983).

At this point we can consider why we have not detected any specific interaction with the A enhancer. It is clear that the B element binding factors are not interacting with the A enhancer. Both enhancers may act by a very dissimilar mechanism or the factor interacting with A may not be present in our extracts. The A binding factor(s) may have been degraded or disassembled during the preparation of the extract. The integrity of the factor is probably most important for its interaction with DNA; in this respect it is worth noting that the specific binding activity with the B enhancer was lost after passage of the extract on a phosphocellulose column (J. Piette, unpublished observations).

Of course, nothing is learnt here about the functional role of the characterized complex or the mechanism by which it could exert this role. This will require the development of an *in vitro* system in which the polyoma enhancer is active in stimulating transcription (see Sassone-Corsi *et al.*, 1984 for SV40) or replication (see Li and Kelly, 1984 for SV40). Another approach will be the use of antibodies raised against the factors involved in the complex formation to analyse the location and role of the factors in the cell. The purification of these factors is presently under way.

Materials and methods

Preparation of nuclear extracts

Nuclear extracts were prepared according to Siebenlist *et al.* (1984) with modifications. Mouse 3T6 cells were grown in Eagle's modified medium supplemented with 7% neonatal calf serum to a density of 10^7 cells/10-cm dish. Cells were rinsed twice with phosphate buffered saline and scraped from the dishes; the cell pellet was conserved in four volumes of Tris Dulbecco containing 30% glycerol at -70°C .

The content of 10 dishes was centrifuged at 1500 r.p.m. and resuspended in 10 ml A buffer (10 mM Hepes pH 8, 50 mM NaCl, 0.5 M sucrose, 1 mM EDTA, 0.5 mM spermidine, 0.15 mM spermine, 0.5% Triton X-100, 1 mM PMSF, 7 mM EtSH). The cells were disrupted in a Dounce homogenizer by 20 strokes with an A pestle, and centrifuged for 10 min at 1000 g. The nuclear pellet was rinsed twice with 10 ml A buffer and resuspended in 3 ml B buffer with 100 mM NaCl (B buffer = 10 mM Hepes pH 8, 25% glycerol, 0.1 mM EDTA, 0.5 mM spermidine, 0.15 mM spermine, 1 mM PMSF, 7 mM EtSH). After 15 min of gentle agitation the nuclei were centrifuged at 1000 g and resuspended in 3 ml B buffer with 400 mM NaCl. After 30 min of gentle agitation, the nuclei were centrifuged at 2000 g and the supernatant was brought to 45% saturation with $(\text{NH}_4)_2\text{SO}_4$. Proteins were precipitated by centrifugation for 15 min at 12 000 r.p.m. The pellet was resuspended in 200 μl B buffer with 100 mM NaCl. The nuclear extract (= 0.4 M nuclear extract) was aliquoted and conserved at -70°C .

Alternatively the first rinse with B buffer was performed at 0.3 M NaCl and the nuclei washed at 0.55 M NaCl, this is called the 0.55 M nuclear extract.

The activity of the newly prepared extracts was tested each time and the optimal amount of proteins was used in the experiments explaining the variable quantities of extracts appearing in the legends to figures. Protein concentrations were determined by the method of Bradford (1976). DNA concentrations by DAPI-fluorescence.

Analysis of DNA-protein complexes on polyacrylamide gels

The DNA fragments used as probes were 3' labelled with the large fragment of DNA polymerase I and purified on a polyacrylamide gel after separation of the labelled extremities by restriction (Maxam and Gilbert, 1980).

A few nanograms of 3'-labelled DNA were incubated at 30°C with the indicated amount of sonicated salmon sperm DNA and nuclear extract in 10 mM Hepes pH 8, 17.5% glycerol, 0.1 mM EDTA, 100 mM NaCl, 10 mM MgCl_2 , 2 mM DTT and 100 $\mu\text{g}/\text{ml}$ BSA in a total volume of 50 μl . After 15 min 5 μl dye were added (1 x TBE, 25% glycerol, 1 mg/ml xylene cyanol) and after gentle mixing 30 μl samples were loaded immediately on a 5% or 7.5% polyacrylamide gel in 0.5 x TBE buffer. The sample was allowed to enter the gel at 250 V and migration was at 185 V (Garner and Revzin, 1981). The gel was fixed in 10% acetic acid and 10% methanol, dried and exposed overnight.

Competition experiments

The fragments used for competition were cloned between the *EcoRI* and *SalI* sites of pML2 (Lusky and Botchan, 1981) with the exception of plasmid IGE which was a gift from P.Sassone-Corsi (Sassone-Corsi *et al.*, 1985). The incubation was performed as explained above except that *EcoRI*-linearized and phenol-extracted pML2 were used instead of salmon sperm DNA as carrier DNA. The *EcoRI*-linearized and phenol-extracted pML2 derivatives were used as competitors. The final amount of cold DNA was brought to a constant amount with pML2. The order of addition of competitor and labelled DNA was as indicated in Figure 3.

DNase I footprinting

The nuclear protein fractions were mixed with a DNA fragment labelled at a single 3' end in the presence of carrier DNA as described above. After incubation for 15 min, 5 μl of a 25 mM CaCl_2 and 50 mM MgCl_2 solution were added together with 5 μl of a 10 $\mu\text{g}/\text{ml}$ DNase I solution. The digestion was stopped after 1 min by adding 5 μl of 0.1 M EDTA and the mixture was loaded on a 5% polyacrylamide gel as indicated above. After migration, the gel was exposed for 2 h at 4°C and the bands corresponding to the DNA-protein complex (+) and free DNA (-) were cut out and eluted (Maxam and Gilbert, 1980). The DNase I-treated fragments were loaded on an 8% sequencing gel altogether with the G + A and C > T chemical degradation products obtained with the same labelled fragment to localize the DNase I pattern on the DNA sequence (Maxam and Gilbert, 1980).

Bal31 deletion analysis

Plasmid pPB1 containing the polyoma *PvuII*-4 fragment between *EcoRI* and *SalI* sites was cut with either *EcoRI* or with *SalI* and treated with *Bal31* nuclease for different times. The protruding ends were then filled in with Klenow DNA polymerase I fragment and the resulting blunt end fragments were cut with *SalI* or *EcoRI*, respectively, and 3' labelled with Klenow fragment. The mixture of unidirectionally deleted fragments was then eluted from a preparative polyacrylamide gel and incubated with a protein extract as described above. The mixture was loaded on a 5% polyacrylamide gel. After migration, the gel was exposed for 12 h at 4°C and the band corresponding to the DNA-protein complex was cut out and eluted (Maxam and Gilbert, 1980). The retarded fragments were loaded on a 6% sequencing gel together with control DNA and the G + A and C > T chemical degradation products obtained with an undeleted fragment labelled at the same end to localize the end point of the deletions on the DNA sequence.

Acknowledgements

We thank J.Rubinstein and P.Herbomel for helpful comments on the manuscript. F.Arnos for technical help and J.Ars for the preparation of the manuscript. J.Piette is a recipient of a long-term EMBO fellowship. This work was supported by grants from the CNRS (LA 270 and ATP Biologie Moléculaire du Gène), the ARC and the Fondation pour la Recherche Médicale Française.

References

- Banerji, J., Rusconi, S. and Schaffner, W. (1981) *Cell*, **27**, 299-308.
 Banerji, J., Olson, L. and Schaffner, W. (1983) *Cell*, **33**, 729-740.
 Bradford, M.M. (1976) *Ann. Biochem.*, **72**, 248-254.
 Brown, D.D. (1984) *Cell*, **37**, 359-365.
 Cereghini, S. and Yaniv, M. (1984) *EMBO J.*, **3**, 1243-1253.
 Cowie, A. and Kamen, R. (1984) *J. Virol.*, **52**, 750-760.
 Cowie, A., Tyndall, C. and Kamen, R. (1981) *Nucleic Acids Res.*, **9**, 6305-6322.
 Cowie, A., Jat, P. and Kamen, R. (1982) *J. Mol. Biol.*, **152**, 225-255.
 Davison, B.L., Egly, J.M., Mulvihill, E.R. and Chambon, P. (1983) *Nature*, **301**, 680-686.
 de Villiers, J. and Schaffner, W. (1981) *Nucleic Acids Res.*, **9**, 6251-6264.
 de Villiers, J., Olson, L., Banerji, J. and Schaffner, W. (1982) *Cold Spring Harbor Symp. Quant. Biol.*, **47**, 911-919.
 de Villiers, J., Schaffner, W., Tyndall, L., Lupton, S. and Kamen, R. (1984) *Nature*, **312**, 242-246.
 Dilworth, S.M., Cowie, A., Kamen, R. and Griffin, B.E. (1984) *Proc. Natl. Acad. Sci. USA*, **79**, 1959-1963.
 Dynan, W.S. and Tjian, R. (1983) *Cell*, **32**, 669-680.
 Elgin, S.R. (1984) *Nature*, **309**, 213-214.
 Emerson, B.M. and Felsenfeld, G. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 95-99.
 Ephrussi, A., Church, G.M., Tonegawa, S. and Gilbert, W. (1985) *Science (Wash.)*, **227**, 134-140.
 Fried, M. and Crothers, D.M. (1981) *Nucleic Acids Res.*, **9**, 6505-6525.
 Fromm, M. and Berg, P. (1983) *Mol. Cell. Biol.*, **3**, 991-1006.
 Galas, B.J. and Schmitz, A. (1978) *Nucleic Acids Res.*, **5**, 3157-3170.
 Garner, M.M. and Revzin, A. (1981) *Nucleic Acids Res.*, **9**, 3047-3060.
 Gidoni, D., Dynan, W.S. and Tjian, R. (1984) *Nature*, **312**, 409-413.
 Gillies, S.D., Morrison, J.L., Oi, V.T. and Tonegawa, S. (1983) *Cell*, **33**, 717-728.
 Gruss, P. (1984) *DNA*, **3**, 1-5.
 Hearing, P. and Shenk, T. (1983) *Cell*, **33**, 695-703.
 Herbomel, P., Saragosti, S., Blangy, D. and Yaniv, M. (1981) *Cell*, **25**, 651-658.
 Herbomel, P., de Crombrughe, B. and Yaniv, M. (1983) in Silver, L.M., Martin, G.K. and Strickland, S. (eds.), *Teratocarcinoma Stem Cells, Cold Spring Harbor Conf. Cell Prolif.*, **10**, 285-294.
 Herbomel, P., Bourachot, B. and Yaniv, M. (1984) *Cell*, **39**, 653-662.
 Jongstra, J., Reudelhuber, T.L., Oudet, P., Benoist, C., Chae, C.B., Jeltsch, J.-M., Mathis, B. and Chambon, P. (1984) *Nature*, **307**, 708-714.
 Katinka, M., Yaniv, M., Vasseur, M. and Blangy, D. (1980) *Cell*, **20**, 393-399.
 Laimins, L.A., Khoury, G., Gorman, C., Howard, B. and Gruss, P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6453-6457.
 Lewis, C.D., Lebkowski, J.S., Daby, A.K. and Laemmli, U.K. (1984) *J. Cell Sci.*, **Suppl. 1**, 103-122.
 Li, J.L. and Kelly, T.J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6973-6977.
 Linney, E. and Donerly, S. (1983) *Cell*, **35**, 693-699.
 Losick, R. and Pero, J. (1981) *Cell*, **25**, 582-584.
 Lusky, M. and Botchan, M. (1981) *Nature*, **293**, 79-82.
 Lusky, M., Berg, L., Weiher, H. and Botchan, M. (1983) *Mol. Cell. Biol.*, **3**, 1108-1122.
 Luthman, H., Nilsson, M.G. and Magnusson, G. (1982) *J. Mol. Biol.*, **161**, 553-550.
 Maxam, A. and Gilbert, W. (1980) *Methods Enzymol.*, **65**, 499-560.
 Melin, F., Pinon, H., Reiss, C., Kress, C., Montreau, N. and Blangy, D. (1985) *EMBO J.*, **4**, 1799-1803.
 Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M.P. and Chambon, P. (1981) *Nucleic Acids Res.*, **9**, 6047-6068.
 Mueller, M.R., Mes-Masson, A.M., Bouvier, M. and Hassel, J.A. (1984) *Mol. Cell Biol.*, **4**, 2594-2609.
 Muller, W.J., Mueller, C.R., Mes, A.M. and Hassel, J.A. (1983) *J. Virol.*, **47**, 586-599.
 Neuberger, M.S. (1983) *EMBO J.*, **2**, 1373-1378.
 Ogata, R. and Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA*, **72**, 5851-5854.
 Parker, C.J. and Topol, J. (1984) *Cell*, **36**, 357-369.
 Picard, D. and Schaffner, W. (1984) *Nature*, **307**, 80-82.
 Sakonju, S. and Brown, D.D. (1982) *Cell*, **31**, 395-405.
 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. and Chambon,P. (1985) *Nature*, **313**, 458-463.
- Saragosti,S., Moyne,G. and Yaniv,M. (1980) *Cell*, **20**, 65-73.
- Schöller,H.R. and Gruss,P. (1984) *Cell*, **36**, 403-411.
- Siebenlist,U., Hennighausen,L., Battey,J. and Leder,P. (1984) *Cell*, **37**, 381-391.
- Strauss,F. and Varshavsky,A. (1984) *Cell*, **37**, 889-901.
- Sutcliffe,J.G. (1979) *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 77.
- Tyndall,C. La Mantia,G., Thacker,C.M., Favaloro,J. and Kamen,R. (1981) *Nucleic Acids Res.*, **9**, 6231-6250.
- Veldman,G.M., Lupton,J. and Kamen,R. (1985) *Mol. Cell. Biol.*, **5**, 649-658.
- Weiher,H. and Botchan,M.R. (1984) *Nucleic Acids Res.*, **12**, 2901-2916.
- Weiher,H., Konig,M. and Gruss,P. (1983) *Science (Wash.)*, **219**, 626-631.
- Worcel,A., Gargiulo,G., Udwardy,A., Louis,C. and Schedl,P. (1983) *Nucleic Acids Res.*, **11**, 421-439.
- Wildeman,A.G., Sassone-Corsi,P., Grundström,T., Zenke,M. and Chambon,P. (1984) *EMBO J.*, **3**, 3129-3133.
- Wu,C. (1984a) *Nature*, **309**, 229-234.
- Wu,C. (1984b) *Nature*, **311**, 81-84.
- Wu,H.M. and Crothers,D.M. (1984) *Nature*, **308**, 509-513.
- Yaniv,M. (1982) *Nature*, **297**, 17-18.

Received on 11 April 1985; revised on 22 July 1985