The EMBO Journal vol.5 no.2 pp.387-397, 1986

Multiple sequence motifs are involved in SV40 enhancer function

Martin Zenke[1], Thomas Grundström[2], Hans Matthes, Marguerite Wintzerith, Christian Schatz, Alan Wildeman[3] and Pierre Chambon

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

Present addresses: [1]Differentiation Programme, European Molecular Biology Laboratory, Postfach 102209, D-6900 Heidelberg, FRG
[2]Unit for Applied Cell and Molecular Biology, University of Umeå, S-901 87 Umeå, Sweden
[3]Department of Molecular Biology and Genetics, University of Guelph, Guelph, Ontario N1G 2W1, Canada

Communicated by P.Chambon

A systematic mutagenesis of the SV40 enhancer indicates that it spans ~100 bp and is composed of at least two distinct DNA domains which exhibit very little enhancing activity on their own. Their association results in a 400-fold enhancement of transcription, virtually irrespective of their relative orientation and, to some extent, of the distance between them. Enhancer activity can also be generated by duplication of either domain. We show also that the activity of each domain is due to the presence of several specific sequence motifs. These motifs are found assorted in different combinations in other viral and cellular enhancers.

Key words: transcription/site-directed mutagenesis/promoter/RNA polymerase B(II)/simian virus 40

Introduction

Enhancers, originally identified in SV40 (Moreau *et al.*, 1981; Banerji *et al.*, 1981) are *cis*-acting DNA segments which dramatically stimulate transcription from RNA polymerase class B (II) homologous and heterologous promoters, in an orientation-independent manner and over long distances. Many other viral and several cellular enhancers have since been described (for reviews and references, see Chambon *et al.*, 1984; Yaniv, 1984; Gruss, 1984; Wasylyk, 1985; Serfling *et al.*, 1985; Khoury and Gruss, 1983). In contrast to the SV40 enhancer, which is active in a variety of different cell lines, some enhancers exhibit a pronounced host cell preference (De Villiers and Schaffner, 1981; Laimins *et al.*, 1982; Kriegler and Botchan, 1983), and strict cell lineage specificity is associated with cellular enhancers, most notably immunoglobulin gene enhancers (Banerji *et al.*, 1983; Gilles *et al.*, 1983; Neuberger, 1983; Picard and Schaffner, 1984; Queen and Baltimore, 1983 and references therein).

Recent *in vivo* (Schöler and Gruss, 1984; Mercola *et al.*, 1985) and *in vitro* (Sassone-Corsi *et al.*, 1984, 1985; Wildeman *et al.*, 1984; Sergeant *et al.*, 1984) studies have indicated that transcriptional enhancement involves the interaction of enhancers with specific *trans*-acting factor(s). Although different enhancers do not share extensive sequence homologies, short and degenerate consensus sequences have been pointed out (Chambon *et al.*, 1984; Weiher *et al.*, 1983; Hen *et al.*, 1983; Nordheim and Rich, 1983; Lusky *et al.*, 1983; Hearing and Shenk, 1983; Veldman *et al.*, 1985). However, such enhancer consensus sequences occur in DNA segments without enhancer properties, and enhancer activity can be generated by duplicating DNA sequences without enhancer activity on their own (Weber *et al.*, 1984; Swimmer and Shenk, 1984), thus questioning the real functional significance of these consensus sequences.

Therefore, we decided to determine, at the nucleotide level, the DNA sequences essential for the activity of the prototype SV40 enhancer. Systematic deletions and point mutations have been constructed throughout the enhancer region and their effect on SV40 early transcription investigated *in vivo*, using a transient expression assay in HeLa cells. We show here that the SV40 enhancer encompasses a large DNA segment of ~100 nucleotides containing the 72-bp sequence, but also extending further upstream. Furthermore, the present study reveals that the enhancer is composed of at least two distinct DNA domains which exhibit very little enhancing activity on their own. However, their association results in a dramatic 400-fold enhancement of transcription, virtually irrespective of their relative orientation and, to some extent, of the distance between them. Enhancer activity can also be generated by duplication of either domain. In addition, we show that the activity of each domain is due to the presence of several specific sequence motifs. Various assortments of these motifs are found in other viral and cellular enhancers, suggesting that enhancers are mosaics of a limited number of basic evolutionary related sequence motifs.

Results

Assay for enhancer activity

We have studied the activity of the SV40 enhancer in its natural position within the early promoter region and in selected cases when separated by 600 bp from the other promoter elements. However, as the activity of the SV40 early promoter is autoregulated by one of the products of the SV40 early transcription unit (large T-antigen, see Tooze, 1982 for references), a SV40/β-globin recombinant has been constructed (pAO, Figure 1A) in which the T-antigen coding sequences have been replaced by the coding sequences of the rabbit β-globin gene. In this construction, initiation of transcription occurs exclusively from the SV40 early promoter and the amount of β-globin RNA synthesized in a transient expression assay directly reflects the transcriptional activity of the SV40 early promoter in the absence of DNA replication and repression by T-antigen (Wasylyk *et al.*, 1983a; Baty *et al.*, 1984 and data not shown). Unless otherwise stated, the wild-type sequence of the early promoter region was modified in the pA-type recombinants by deleting exactly one of the two 72-bp repeated sequences (from positions 107 to 178, see Figures 1A and 2A) and by generating a *Bam*HI site precisely at the junction (position 101) between the 72-bp sequences and the 21-bp repeat region (pAO in Figures 1A and 2A). This '*Bam*HI' mutation does not affect the activity of the early promoter and the presence of one copy of the 72-bp sequence con-

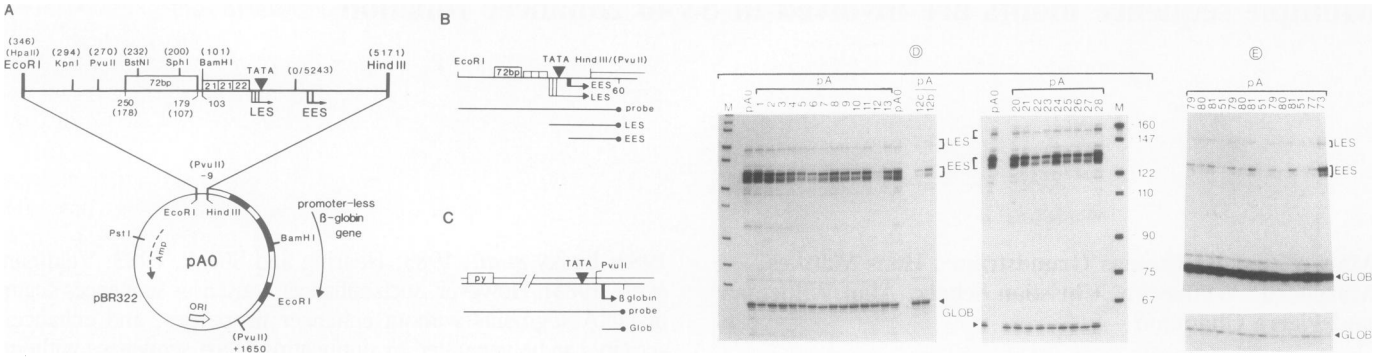


Fig. 1. (**A**) Structure of the basic recombinant pA0 which contains the SV40 early promoter region from *Hpa*II (346) to *Hin*dIII (5171) cloned in front of the coding sequences of the rabbit β-globin gene [*Pvu*II fragment, from position −9 to +1650 (van Ooyen *et al.*, 1979), double line, intron and exon sequences are indicated by open or black boxes, respectively] in pBR322 (single line) between *Eco*RI (4361) and *Pvu*II (2066) sites. Cloning restored the *Hin*dIII site at SV40 position 5171 (BBB system, see references Tooze, 1982) and destroyed the *Pvu*II sites at both ends of the β-globin gene. Some SV40 DNA restriction sites, the 72-bp sequence, and 21-bp repeat region, the TATA box, the EES and LES (Wasylyk *et al.*, 1983a; Baty *et al.*, 1984), are indicated. A *Bam*HI site was engineered at position 101, at the junction between the 72-bp sequence and the 21-bp repeat region (see Grundström *et al.*, 1985 and also Figure 2A); this *Bam*HI site is contained in all pA-type recombinants (unless otherwise indicated) and does not affect the SV40 early promoter activity (data not shown). All recombinants of the pA series with a number beyond 100 (see below) have the *Bam*HI site in the rabbit β-globin gene (position 476) destroyed by 'filling in' with *Escherichia coli* DNA polymerase (Klenow fragment) which leaves transcription from the SV40 early promoter and mRNA stability unaffected (data not shown). (**B**) and (**C**) Determination of transcriptional activity of the SV40 early promoter. The amount of RNA synthesized from the SV40 early promoter in the pA-type recombinants after transient expression in HeLa cells was determined by quantitative S1 nuclease mapping. (**B**) Part of the recombinant pA0 with the SV40 promoter region (*Eco*RI to *Hin*dIII, single line) cloned in front of rabbit β-globin coding sequences (double line). EES and LES represent single-stranded probe fragments (see below) of 130−137 and 164−178 nucleotides in length, respectively, protected against S1 nuclease digestion by RNA initiating at the SV40 EES or LES, respectively. (**C**) Part of the recombinant pβ(244$^+$)β (De Villiers and Schaffner, 1981) used as reference gene in the transient expression assay. The TATA box and start site of the β-globin gene (double line) are indicated, as well as the polyoma virus enhancer (py). The same single-stranded DNA probe as in (**B**) was used for quantitative S1 nuclease mapping of RNA initiating at the globin cap site, yielding a protected DNA fragment of 60 nucleotides in length (Glob). (**D**) and (**E**) Quantitative S1 nuclease analysis of RNA synthesized by various enhancer mutants. (**D**) Effect of the scanning point mutations present in pA1−pA13 and pA20−pA28 (see text and Figure 2A). EES and LES correspond to RNA transcribed from the SV40 early promoter and GLOB to RNA transcribed from the co-transfected reference recombinant (see text and **B** and **C**). **Lane M**: size markers (in nucleotides), ^{32}P 3′ end-labeled *Msp*I digest of pBR322. (**E**) Effect of point mutations when the enhancer has been transposed 600 nucleotides away from the remaining part of the SV40 early promoter. Three independent experiments are shown. pA73 has been constructed by inserting 600 nucleotides of pBR322 sequence (*Bam*HI-*Nru*I fragment, position 375−971) in the *Bam*HI site of pAO (position 101, **A**). This was carried out by adding a *Bgl*II linker at position 973 of pBR322, inserting the resulting *Bam*HI-*Bgl*II fragment at the *Bam*HI site of pAO, and selecting clones which restore the *Bam*HI site proximal to the enhancer (a *Xho*II site is created at the other end of the insert). pA77, pA79, pA80 and pA81 (Table I) are derivatives of pA73 containing the deletion of pA62 or the point mutations of pA6, pA12 or pA26, respectively (Figure 2A). pA51 is described in Figure 2A. Autoradiography was for 2 weeks. The lower panel shows a shorter exposure of the same gel used for measuring the signal from the reference gene. EES, LES and GLOB are as in **D**.

siderably facilitates the mutagenesis of the enhancer, while decreasing its activity by only ~3-fold (see below).

The effect of the various enhancer mutations was measured by quantitative S1 nuclease analysis of RNA synthesized during transient expression in HeLa cells (Figure 1B, C, D and E; in all cases the amount of specific RNA synthesized was proportional to the amount of transfected recombinant). The protected probe DNA sequences EES and LES correspond to mRNA initiating at the early-early and late-early start sites of the SV40 early-early and late-early overlapping promoters, respectively (see Wasylyk *et al.*, 1983a; Baty *et al.*, 1984). By densitometric scanning of autoradiograms corresponding to different exposures of the same gel, signals three orders of magnitude lower than those of wild-type transcription could be measured (see Figures 1D, E and 3). Transcription from an internal reference recombinant [rabbit β globin gene in pβ(244+)β (De Villiers and Schaffner, 1981), Figure 1C], co-transfected in all experiments, yielded a signal (GLOB in Figures 1D, E and 3) whose intensity was used to correct for variations in transfection efficiencies. The effect of enhancer mutations on variations in RNA initiated at the LES have not been systematically studied here, but appear to follow closely those of RNA initiated at the EES (see Figures 1D, E and 3), as expected from previous studies (Wasylyk *et al.*, 1983a).

The size of the enhancer is at least 100 bp

To delineate the DNA sequences required for enhancer activity, we first constructed a series of deletion mutants using both natural and engineered restriction sites (Figure 2A). When sequences in the 5′ moiety and upstream of the 72-bp sequence are deleted (pA223, pA233, Figure 2A), transcription is dramatically decreased (0.5% and 2%, respectively). Deleting sequences in the 3′ portion of the 72-bp sequence (pA62, pA234, Figure 2A) also results in a marked decrease in transcription (7% and 3.5% respectively), demonstrating the importance of both moieties of the 72-bp sequence for efficient early transcription.

To investigate whether sequences located further upstream from the 72-bp sequence are required for enhancer function, a *Bgl*II site was engineered at the 5′ boundary of this sequence (pA35, Figure 2A), which by itself diminishes transcription by 3-fold. Deletion of all SV40 sequences 5′ to the *Bgl*II site (pA104, Figure 2A) results in a 30-fold reduction of transcription, which unambiguously demonstrates that sequences beyond the 72-bp sequence are required for enhancer function. Transcription is not significantly affected by deleting the SV40 sequences upstream from the *Kpn*I site (pA53), but is reduced by 2-fold when the deletion extends further downstream to position 284 (pA260, Figure 2A). Thus by comparing pA104, pA260 and pA53, we conclude that sequences located between positions 256 and 284 are very important for enhancer activity, with sequences further upstream (position 284−300) stimulating transcription an additional 2-fold. Surprisingly, when sequences located between positions 255 and 298 are removed in the internal deletion mutant pA54 (*Bgl*II-*Kpn*I deletion), transcription is not more affected

than in pA35, which contains the engineered *Bgl*II site used to generate the deletion. It appears, therefore, that sequences located upstream from position 298, which have very little enhancer activity by themselves (see the deletion mutants pA56, pA270 and pA103 in Figure 2A), can functionally substitute for the 256−284 sequences when these are deleted. However, in the absence of deletions within the main body of the enhancer, these potential enhancer sequences are clearly not required for full enhancer activity (see pA53 in Figure 2A), at least under the present assay conditions.

Thus, the present analysis of a series of deletion mutants (Figure 2A and data not shown) localizes the essential SV40 enhancer sequence between positions 179 (107) and 284, whereas auxiliary DNA sequences located further upstream between positions 284 and 346 increase enhancer activity by an additional 2-fold [compare pAO (100%) with pA270 (0.2%) and pA260 (46%) in Figure 2A]. Transcription is stimulated ~1000-fold when all SV40-specific sequences located upstream from the 21-bp repeat region are present up to position 346 (see pA51, pA52 and pA101 in Figure 2A and pA51 in Figure 1E). Since deleting one 72-bp sequence results in an ~3-fold decrease in enhancer activity (see below), the overall stimulation of transcription brought about by the enhancer element in its natural SV40 early promoter environment is ~3000-fold.

The enhancer contains multiple sequence motifs

A systematic series of clustered point mutations was generated throughout the essential enhancer region (position 179−284) to determine at the nucleotide level which sequences are required for enhancer activity. In view of their auxiliary role (see above) the sequences located further upstream from position 284 were not considered in the present point mutation scanning analysis. Clusters of three neighbouring nucleotides were mutated simultaneously *in vitro* by generating non-complementary transversions (A⇋C, G⇋T), using either M13 oligonucleotide-directed mismatch-primer mutagenesis or the 'shot-gun' ligation technique (Grundström *et al.*, 1985). After sequencing, the mutated enhancer sequences were cloned into the expression vector pAO, yielding the recombinants pA1−pA34 (Figure 2A), which were analyzed for transcriptional activity. Two autoradiograms corresponding to such quantitative S1 nuclease analyses are shown in Figure 1D; the results are summarized in Figure 2A, where transcription originating from the EES of the SV40 early promoter in the pA1−pA34 mutants is plotted relative to wild-type pAO transcriptional activity. For each mutant the bar represents the average value of at least three independent transfection experiments using at least two different plasmid DNA preparations, after correction for transcription from the co-transfected reference recombinant (GLOB in Figure 1D, see above).

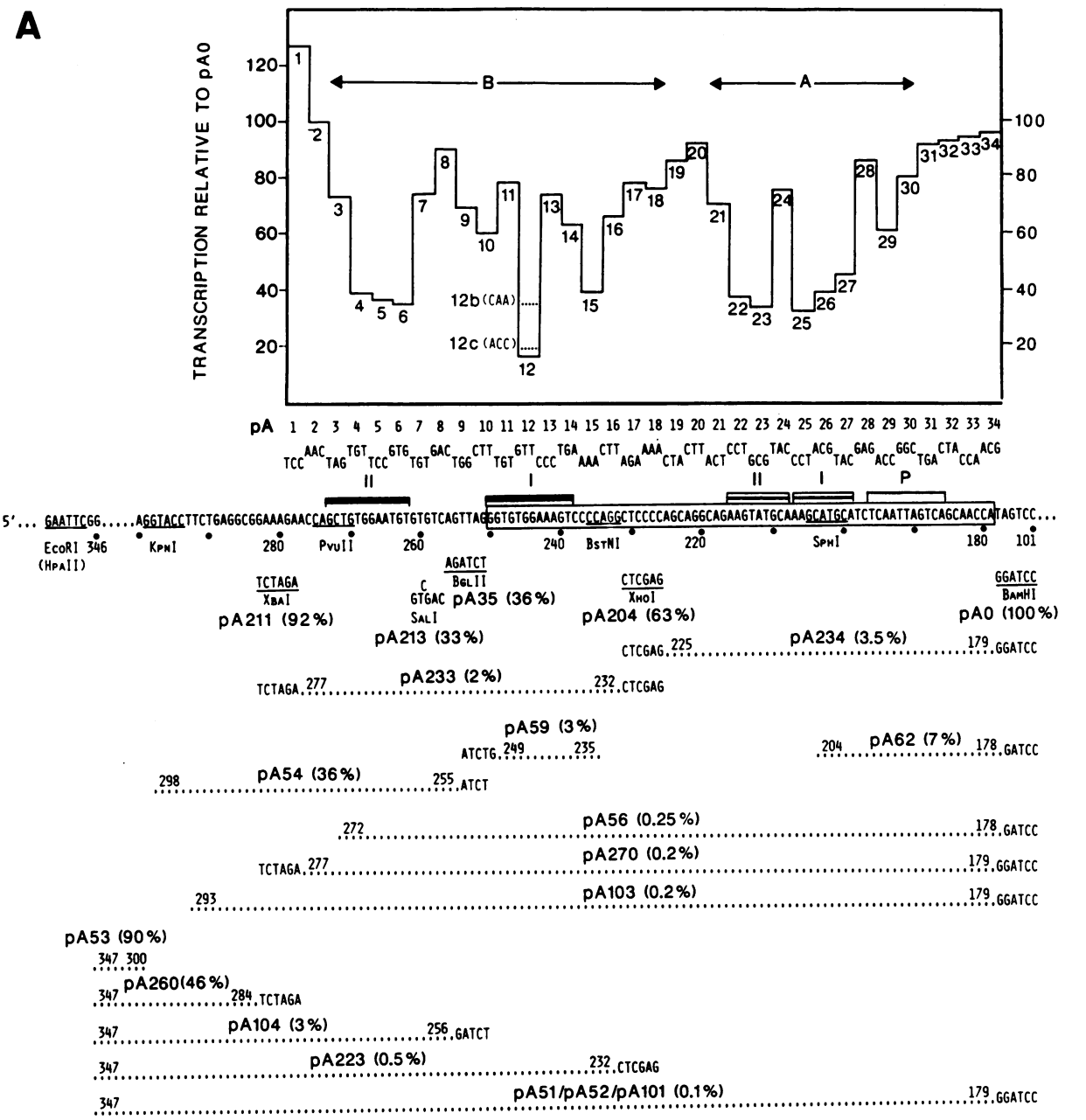
Mutations detrimental to enhancer activity are scattered throughout the entire DNA segment analysed and none of the scanning mutations abolishes enhancer function; no single scanning mutation reduces transcription by >6-fold (pA12). It is unlikely that a sequence important for enhancer activity has not been detected because the appropriate mutation has not been made. For example, two alternative mutations at positions 246−244 changing the sequence TGG to CAA (pA12b) and ACC (pA12c) decrease transcription to 35% and 15%, respectively, compared with 12% for pA12. Furthermore, the mutations generating a *Xho*I site at position 226 (pA204, 63%, Figure 2A), which change a cluster of three C residues to GAG, affect transcription to the same degree as the corresponding scanning mutation pA18 (CCC→AAA, 75%). Additional examples are discussed below (Figure 2B). The scanning mutant pA11, which is located in the center of an essential enhancer element (GT-motif I, see below) and does not severely affect transcription, may be an exception, because this mutation generates the sequence GGTG^{T}TGG, leaving TG-motif I (see below) virtually unaffected.

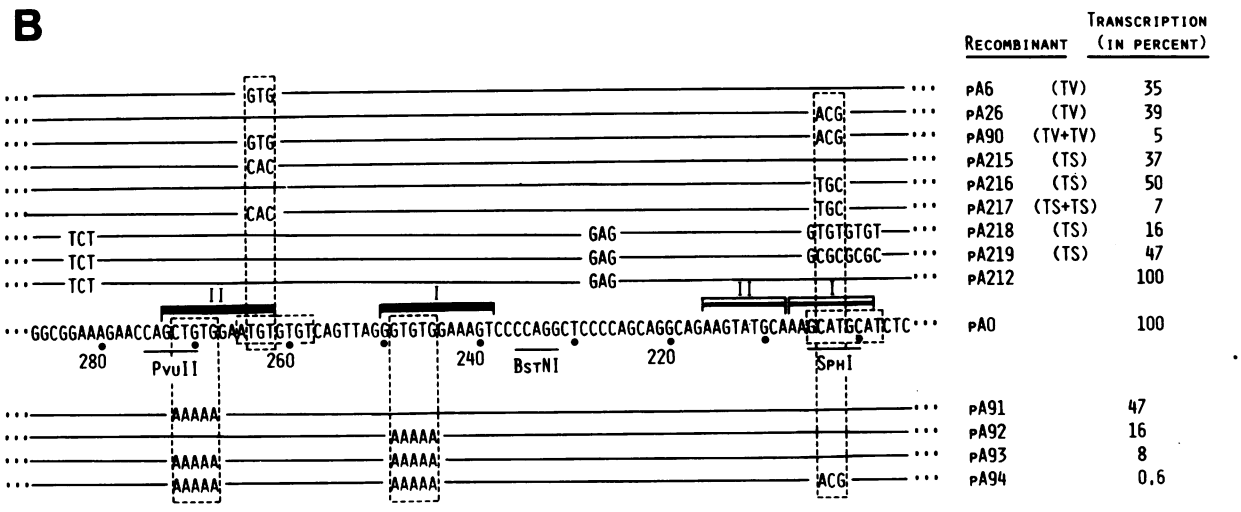
To investigate whether the various scanning mutations identify sequences with genuine enhancer properties, we have also studied their effect when the enhancer element is separated from the remaining SV40 early promoter elements by 600 nucleotides of pBR322 sequence. In agreement with a previous report (Wasylyk *et al.*, 1984), this 'spacing' results in a 30-fold decrease in enhancer activity (Figure 1E and Table I, pA73 versus pA0). Three point mutations (pA6, pA12, pA26) situated at key positions within the SV40 enhancer and mutant pA62, in which the 3′ end portion of the 72-bp sequence is deleted (Figure 2A), were chosen for this study. As shown in Figure 1E and summarized in Table I, these mutations have the same relative detrimental effect on transcriptional activity, irrespective of the enhancer location, confirming that they affect enhancer sequences.

The point mutation scanning analysis reveals several regions important for enhancer function. One of these is defined by mutants pA21−pA30 (positions 219−190, Figure 2A) and contains a direct repetition of the sequence 5′-AAG$^{C}/_{T}$ATGCA-3′ at positions 207−199 and 216−208 (Figure 2A and B), called hereafter the 'Sph-motifs I and II', since the 3′ repeated sequence harbours the *Sph*I site. Equivalent bases of these repeated DNA sequences are located nine nucleotides apart which, assuming a B-DNA-like DNA structure, corresponds to approximately one DNA helical turn (by using different plasmid DNA preparations in several independent transfection experiments, we have repeatedly found that pA24 has a transcriptional activity close to wild-type, see Figures 1D and 2A).

A second region important for enhancer function, which is defined by mutants pA3−pA18 (positions 273−226), consists of sequences located both within and upstream from the 72-bp sequence. The repeated sequence 5′-G$^{C}/_{G}$TGTGG-3′, which belongs to a longer repeated sequence 5′-G$^{C}/_{G}$TGTGGAA$^{A}/_{T}$GT-3′ (called hereafter the GT-motifs I and II, Figure 2A and B), appears to be an essential motif of this enhancer region (see mutants pA3−pA6, pA10 and pA12). These two GT-motifs are located 23 nucleotides apart, which corresponds to approximately two DNA helical turns. We also note that the GT-motif I is followed by a repetition of the motif 5′-TCCCCAG-3′ (TC-motif), whose equivalent bases are nine nucleotides apart, but that only mutations within the more upstream (positions 239−233) of these two motifs appear to efficiently decrease transcriptional activity (pA15). Note that the mutation present in pA1 generates an additional TC-motif, which is apparently accompanied by an increase in transcription.

It has been suggested that two clusters of alternating purine and pyrimidine residues (positions 265−258 and 205−198, Figure 2B) are important sequence features for SV40 enhancer function, because they have the potential to form Z-DNA structures (for references, see Nordheim and Rich, 1983; Azorin and Rich, 1985). The present scanning mutation analysis indicates that these alternating purine-pyrimidine motifs do in fact belong to important sequence elements of the SV40 enhancer (see mutants pA5, pA6, pA7, pA25, pA26 and pA27, Figure 2A). To address the question of whether the alternation of purine/pyrimidine residues *per se* is important for enhancer activity, transition mutants were constructed in which this alternation is conserved (pA215, pA216, Figure 2B). It is clear that transition mutations in either motif, as well as a combination of


A
TRANSCRIPTION RELATIVE TO pA0
B
II
I
P
pA211 (92%)
pA213 (33%)
pA35 (36%)
pA204 (63%)
pA234 (3.5%)
pA0 (100%)
pA233 (2%)
pA59 (3%)
pA62 (7%)
pA54 (36%)
pA56 (0.25%)
pA270 (0.2%)
pA103 (0.2%)
pA53 (90%)
pA260(46%)
pA104 (3%)
pA223 (0.5%)
pA51/pA52/pA101 (0.1%)


B
RECOMBINANT
TRANSCRIPTION (IN PERCENT)
pA6 (TV) 35
pA26 (TV) 39
pA90 (TV+TV) 5
pA215 (TS) 37
pA216 (TS) 50
pA217 (TS+TS) 7
pA218 (TS) 16
pA219 (TS) 47
pA212 100
pA0 100
pA91 47
pA92 16
pA93 8
pA94 0.6

transition mutations in both motifs (pA217), are as detrimental to enhancer function as the corresponding transversion mutations (pA6, pA26 and pA90, Figure 2B). Furthermore, when the sequence at positions 205–198 is mutated to either alternating GC or GT residues (pA218 and pA219 respectively, Figure 2B) the enhancer activity is also decreased. These results clearly demonstrate that the contribution of the two motifs containing the alternating purine-pyrimidine residues to enhancer activity cannot be simply ascribed to their potential to form Z-DNA structures.

As each individual cluster of point mutations affects the SV40 enhancer only to a moderate extent (Figure 2A), combinations of clusters of point mutations situated at key positions within the enhancer were constructed (pA91–pA94, Figure 2B). Because the GT-motifs may not be sufficiently modified by the G⇋T transversions which were used to create the scanning mutants, we constructed the point mutants pA91 and pA92, in which the sequences $^{G}/_{C}$TGTG (positions 272–268 and 249–245) are replaced by a cluster of five A residues (Figure 2B). Analysed individually, both mutations (pA91 and pA92) diminish enhancer activity (Figures 2B and 3A) to the same extent as the corresponding scanning mutations (pA4, pA5 and pA12, respectively, in Figure 2A). However, a combination of the mutations present in pA91 and pA92 (pA93, Figures 2B and 3A) affects the transcriptional activity much more dramatically than either of them. It is also important to stress that the enhancer activity is decreased to 8% in pA93, even though the two alternating purine-pyrimidine motifs are intact. Note that a similar decrease was obtained by combining clusters of point mutations located exclusively in these two motifs: pA6 and pA26 in recombinant pA90 (Figures 2B and 3), and pA215 and pA216 in recombinant pA217 (Figure 2B). The most dramatic decrease in transcription (~200-fold reduction) was obtained by combining the cluster of point mutants present in pA91, pA92 and pA26 in recombinant pA94 (Figures 2B and 3). This finding is particularly striking, considering that pA94 retains the entire SV40 enhancer sequence except for the point mutations introduced, but yet has a transcriptional activity reduced to approximately the same extent as recombinants having the entire enhancer element deleted (pA56 and pA270 in Figure 2A). Thus, abolishing enhancer activity requires multiple mutations at specific positions within the enhancer, which supports the idea that the SV40 enhancer represents a large DNA segment composed of multiple important sequence motifs, all of them being required for maximal enhancer activity.

Two independent enhancer domains

The deletion and point mutation scanning analysis reveals two broad regions A and B (positions 219–190 and 273–226, respectively) important for enhancer function, each of which on its own exhibits a low, but still significant, enhancing activity (pA233 and pA234, Figures 2A and 4). To investigate whether the dramatic increase in transcriptional activity generated by their juxtaposition depends critically on their relative location, we have examined the effect of inserting DNA segments of various lengths between these two regions (at the *Xho*I site engineered at position 226, pA204, Figures 2A and 3B). The insertion of 4 bp (pA205, Figure 3B and Table II), which separates regions A and B by nearly half a DNA helical turn, has little effect on enhancer activity (77% of the parental pA204). Similarly, insertions of either 12, 18 and 58 nucleotides (pA206, pA207 and pA249, respectively, Figure 3B and Table II) or a deletion of 9 bp in this region (positions 224–232) (data not shown) have small effects on the enhancer activity. Thus enhancer function apparently does not require a close juxtaposition nor a precise stereospecific alignment between the two enhancer regions A and B, which appear to behave as individual domains, even though they functionally cooperate to generate the enhancer activity. Insertions of longer fragments (pA250, pA251, pA252, in Figure 3 and Table II) result in a strong decrease in enhancer activity. This observation excludes the possibility that domain B may represent the enhancer, whose activity would be insensitive to 'distance', and that domain A corresponds to an 'upstream' promoter element with an activity sensitive to 'distance'. Note that the residual enhancer activity is still significantly higher in the mutant containing a 325-bp insert (pA251) than in a recombinant which is devoid of all SV40 sequences upstream from region A (pA223 in Figures 2, 3B and 4).

To investigate further the physical independence of domains A and B, we used the *Xba*I site engineered at position 278 (pA211, Figures 2 and 4), which either alone or in combination with the *Xho*I site at position 226 (pA212, Figures 2B and 4) does not affect enhancer activity. Isolated domains A (pA223, Figure 4, see also Figure 3B) and B (pA224, Figure 4) stimulate transcription by 4- to 8-fold over the basal level of transcription observed in the absence or any enhancer sequence (pA101 in

Fig. 2. Effect of deletion and point mutations within the enhancer region on transcription initiating from the EES of the SV40 early promoter. All results are expressed relative to pA0, taken as 100%. The sequence of the SV40 enhancer region of wild-type SV40 containing only one copy of the 72-bp sequence position 107–178 or 179–250, boxed in (**A**) cloned in the recombinant pA0 (Figure 1A) is shown, as well as the engineered restriction sites. Some naturally occurring restriction sites, the GT-motifs I and II (solid lines, I and II), the Sph-motifs I and II (double lines I and II) and the P-motif [single line, P in (**A**)] are indicated (see text). (**A**) In the upper part of the figure the mutations present in mutants pA1–pA34 are shown and lined up with their relative transcriptional activity as measured by quantitative S1 nuclease mapping (Figure 1B and C, e.g., 1D) after transient expression in HeLa cells. The lower part shows various deletion mutants with the extent of the deletions indicated by dots (the position of the first and last deleted nucleotide is given); their transcriptional activity is indicated in parentheses relative to pA0. These values which represent the average (±20%) of at least three independent transfection experiments with different plasmid DNA preparations were obtained by densitometric scanning of autoradiograms similar to those shown in Figures 1D, E and 3 and exposed for various times. pA35, pA54 and pA59 contain a *Bgl*II site at position 251 generated by changing wild-type SV40 sequence at position 254–251 to 5′-ATCT-3′ using mismatch-primer mutagenesis (Grundström *et al.*, 1985). All the other recombinants contain a *Bam*HI site at position 101 as described above (Figure 1A). pA52, pA101, pA104, pA223, pA233, pA260 and pA270 contain part of the polylinker region of M13mp12 (Kahn *et al.*, 1984) from *Eco*RI to *Bam*HI (pA52, pA101, pA104), from *Eco*RI to *Xho*I (pA223), from *Xba*I to *Xho*I (pA233), from *Eco*RI to *Xba*I (pA260), and from *Xba*I to *Bam*HI (pA270), respectively, in place of the deletion. The small polylinker region replacing the SV40-specific DNA sequences deleted in pA103 (position 179–293) is identical to that present in the corresponding M13 derivative M13mp9A102 as described elsewhere (Grundström *et al.*, 1985). pA51 combines the deletions of pA53 and pA103 in one construct. The scanning point mutants pA1 to pA34, pA12b and pA12c as well as recombinants pA204, pA211 and pA213 have been constructed by *in vitro* mutagenesis using either mismatch-primer mutagenesis or the 'shot-gun ligation' technique essentially as described elsewhere (Grundström *et al.*, 1985). The sequence of all recombinants was confirmed by DNA sequencing. (**B**) Effect of multiple point mutations in the SV40 enhancer. The two stretches of alternating purine/pyrimidine (positions 198–205 and 258–265) are boxed with dashed lines. The various mutants have been constructed by 'shot-gun ligation' as described (Grundström *et al.*, 1985) and contain the mutations as indicated. TV and TS represent transversion and transition mutants, respectively. pA218 and pA219 contain the mutations of pA204 and pA211 (see **A**) which generate a *Xho*I site (position 226) and a *Xba*I site (position 278), in addition to the mutations at position 198–205. The transcription efficiency, as determined from the amount of RNA initiated at the EES on autoradiograms similar to that shown in Figure 3A, is expressed relative to pA0.

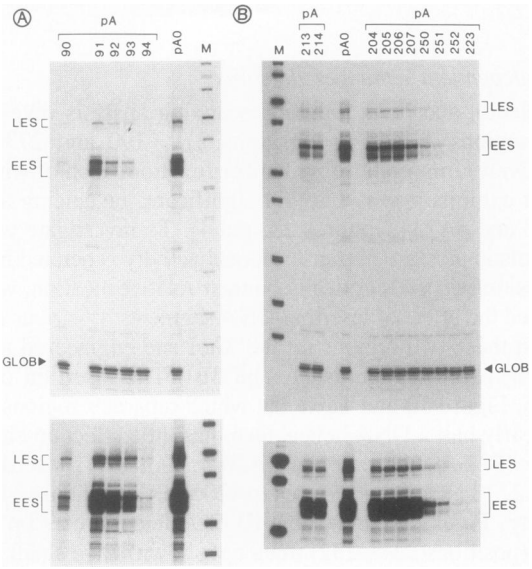
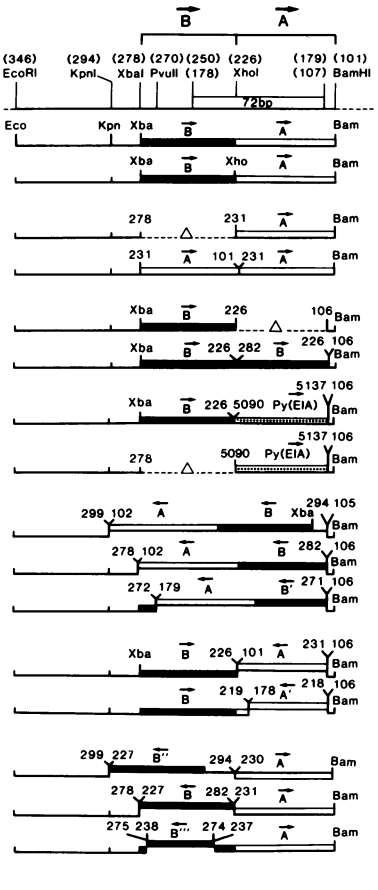


Fig. 3. (**A**) Quantitative S1 nuclease analysis of RNA synthesized by the multiple point mutants pA90 to pA94 (Figure 2B). (**B**) Effect of inserting spacers of increasing length between enhancer domains A and B. The structure of the recombinants pA204, pA213 and pA223 is shown in Figure 2A; recombinants pA205 – pA207 and pA250 – pA252 are described in Table II. pA214 was obtained by 'filling in' the *SalI* site (position 257) in pA213 (Figure 2A) using DNA polymerase I (Klenow fragment). EES, LES and GLOB are as indicated in legend to Figure 1D and E. The lower panels show a longer exposure of the same autoradiograms.

Table I. The relative effect of mutations within the enhancer is independent of 'distance'

Natural position		At a distance of 600 bp		
Recombinant	Relative transcription (%)	Recombinant	Relative transcription (%)	
pAO	100	pA73	3.3 [2.8,3.2,4.0]	(100)
pA6	35 [28,34,42]	pA79	0.52 [0.26,0.76,0.54]	(16)
pA12	17 [18,18,14]	pA80	0.23 [0.23,0.19,0.28]	(7)
pA26	40 [45,40,34]	pA81	0.77 [0.51,0.74,1.06]	(23)
pA62	7 [6.5,6.0,8.0]	pA77	0.30 [0.37,0.25,0.28]	(9)

The effects of mutations affecting enhancer activity are compared when the enhancer is situated either in its natural position within the SV40 early promoter (pA0, pA6, pA12, pA26 and pA62, see Figure 2A) or separated from the 21-bp repeat region by a spacer of 600 nucleotides of pBR322 sequence (pA73, pA79, pA80, pA81 and pA77, respectively; for the construction of these recombinants see Figure 1E). The experiment was performed in triplicate (see Figure 1E) using three different DNA preparations (the individual values are bracketed). The values correspond to the amount of RNA initiated at the EES (determined by quantitative S1 nuclease analysis), expressed relative to pA0, taken as 100%. The numbers in parentheses on the right hand side correspond to the transcription of pA79, pA80, pA81 and pA77, expressed relative to pA73, taken as 100%.



Transcription relative to pA0(EES)			
with Eco-Xba region		without Eco-Xba region	
	percent		percent
pA211	92	pA260	46
pA212	102	pA261	38
pA233	2.3	pA223	0.4
pA235	5.2	pA225	1.6
pA234	3.3	pA224	0.8
pA237	15.5	pA227	8
pA411	51	pA401	26
pA410	3.0	pA400	0.6
pA301	79		
pA306	62	pA308	45
pA201	62		
pA302	42	pA310	18
pA203	53		
pA305	41		
pA303	44	pA311	8.5
pA202	33		

Fig. 4. Effect of rearrangement of enhancer domains A and B. The SV40 enhancer region of pA0 from *Eco*RI (position 346) to *Bam*HI (position 101) is depicted at the top of the figure. Enhancer domains A and B encompass sequences from *Bam*HI to *Xho*I (position 101 – 226, open box) and from *Xho*I to *Xba*I (position 226 – 278, filled box), respectively (see Figure 2A). The arrows indicate the relative orientation of the two domains within the wild-type and rearranged enhancer. The various recombinants were constructed using synthetic linkers and adaptors, taking advantage of restriction sites engineered in the SV40 enhancer region: *Xho*I site (position 226, pA204); *Xba*I site (position 278, pA211); combination of both sites (pA212) (Figure 2A and B). Two classes of recombinants with or without the sequences further upstream from the *Xba*I site up to the *Eco*RI site (position 278 – 346) were constructed (with *Eco-Xba* and without *Eco-Xba* region, respectively). The coordinates of the extremity of the rearranged segments are indicated. For all recombinants the predicted sequence was confirmed by DNA sequencing; these data and a detailed description of the cloning procedure are available upon request. The inversions in recombinants pA201, pA202 and pA203 (position 179 – 271, 238 – 274 and 178 – 218, respectively) were constructed by using 'shot-gun ligation' (Grundström *et al.*, 1985). The same technique was employed to construct the *Xho*I/*Bam*HI DNA segment harboring part of the polyoma virus enhancer [Py(E1A), position 5090 – 5137, see Tyndall *et al.*, 1981] in recombinants pA400, pA401, pA410 and pA411. Transcription initiated at the EES of each recombinant was determined by quantitative S1 nuclease analysis as indicated in legend to Figure 1 and was expressed relative to pA0, taken as 100%. In view of the requirement for stereospecific alignment between the 21-bp repeat region and the enhancer (Takahashi *et al.*, 1986), the relative enhancer activities of the various constructions must be considered as minimal values.

Figure 2A, whose transcription is 0.1% of wild-type). Note that there is an additional 4- to 6-fold increase of this activity when the *Eco*RI-*Xba*I region is present (pA233 and pA234, Figure 4), which supports the conclusion that this region possesses some potential enhancer activity (see above).

The fact that the SV40 enhancer is known to stimulate transcription irrespective of its orientation (Moreau *et al.*, 1981; Banerji *et al.*, 1981; Wasylyk *et al.*, 1983b; see also pA301, pA306, pA201 and pA308 in Figure 4), prompted us to investigate whether domains A and B can also function bidirectionally, both in the presence and absence of the potential enhancer sequences located upstream from the *Xba*I site (see above). As shown in Figure 4, inversion of the entire domain B (pA305, pA303 and pA311) or of a sequence containing its essential motifs (pA202) does not result in a dramatic decrease of its activity. Domain B in its reversed orientation stimulates transcription by ~20-fold (compare pA305, pA303 and pA202 with pA233, and

Table II. Effect of DNA insertions between enhancer domains A and B

Recombinant	Size of insertion (in bp)	Relative transcription (%)		
pAO		100		
pA204		62	[63,61,62]	(100)
pA205	4	48	[45,54,44]	(77)
pA206	12	60	[51,60,67]	(97)
pA207	18	55	[49,61,54]	(88)
pA249	58	42	[35,48]	(68)
pA250	91	9	[9,7,12]	(15)
pA251	325	2.8	[2.3,3.4]	(4.5)
pA252	780	0.9	[1.0,0.8]	(1.5)

pA205 was constructed by 'filling in' the *Xho*I site of pA204 (Figure 2A) using DNA polymerase I (Klenow fragment). In pA206, 12 nucleotides (5'-TCGAGTCTAGAC-3') were inserted between positions 230 and 231 of pA204 using synthetic oligonucleotides. In pA207, a segment of 18 nucleotides (a polylinker with restriction sites for *Xho*I, *Nar*I and *Xba*I, 5'-CTCGAGGGCGCCTCTAGA-3') was inserted between positions 227 and 228 using 'shot-gun ligation' (Grundström *et al.*, 1985). Recombinant pA249, which contains an insert of 58 nucleotides between position 227 and 228, was constructed by cloning a 46-bp (random DNA sequence) between the *Xho*I and *Xba*I site of the polylinker of pA207 using 'shot-gun ligation' (Grundström *et al.*, 1985). The constructs pA250, pA251 and pA252 contain pBR322 sequences of 91, 325 and 780 nucleotides in length, inserted in the *Xho*I site of pA204, respectively. This was performed by cutting pBR322 with either *Sph*I (561), *Nru*I (971) or *Ava*I (1424), treating with DNA polymerase I (Klenow fragment), adding *Sal*I linkers (8-mers), excising the corresponding *Sal*I fragments and inserting them into the *Xho*I site of pA204.

pA311 with pA223), whereas the same domain in its natural orientation enhances transcription by 40- to 100-fold (compare pA233 with pA212, and pA223 with pA261). Similarly, the enhancer activity is not dramatically affected when the entire domain A (pA302, pA310) or a sequence containing its essential motifs (pA203) are inverted. Transcription is stimulated by ~15- to 20-fold (compare pA302 and pA203 with pA234, and pA310 with pA224), whereas the stimulation brought about by domain A in its natural orientation is ~30- to 40-fold (compare pA234 with pA212, and pA224 with pA261). These results support the notion that the ability of domains A and B to cooperate in generating the enhancer activity does not critically depend on their relative orientation.

To investigate whether one enhancer domain can functionally substitute for the other, recombinants containing a dimer of either domain were constructed (pA235, pA225, pA237 and pA227 in Figure 4). It is clear that dimers of either domains A or B are more efficient than the corresponding monomers in stimulating transcription. However, wild-type enhancer activity is not achieved, even with multimers of domains A or B (data not shown). It is worth noting that a recombinant in which the promoter-proximal B domain of pA237 has been replaced by a random sequence of identical length is not more efficient in stimulating transcription than pA234, which contains a single B domain (data not shown), indicating that the increase in transcription observed when domain B is duplicated is not related to an optimal distance requirement between domain B and the 21-bp repeat region. Thus, the enhancer activity, which is considerably decreased when either domain A or B is deleted, can be restored, at least to some extent, by duplicating either domain A or B. Furthermore domain A can be replaced by a 48-bp segment of the polyoma virus enhancer (positions 5090–5137, see Tyndall *et al.*, 1981; polyoma enhancer domain A, see Herbomel *et al.*, 1984) which contains a sequence motif in common with the adenovirus E1A enhancer element and is duplicated in several

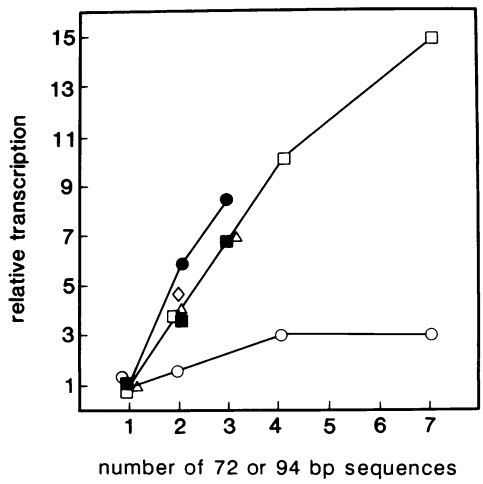


Fig. 5. Effect of multiple copies of the 72-bp sequence and of the SV40 enhancer on enhancer activity. Transcription initiating at EES relative to pA0 (taken as 1) as a function of increasing number of 72-bp sequences (two, four and seven copies, pAW2, pAW4 and pAW7, respectively) using different amounts of recombinant DNA for transfection (—○—, 1.5 μg or —□—, 0.15 μg DNA/6 × 10^6 cells). The transcriptional activity of recombinants containing one to three copies of the 72-bp sequence (—■—, pA71, pAW22, pAW23, respectively) or of the 94-bp sequence (position 179–272, cloned from pCWori, a gift of H.Kopecka and M.Girard; —●—, pAW32, pAW33, respectively) is also shown when the enhancer is separated from the 21-bp repeat region by an insertion of 43 nucleotides of pBR322 sequence in the *Bam*HI site of pA0 (position 101). This 43-nucleotide insertion (from *Bam*HI to *Nar*I, pBR322 coordinates 375–413) was obtained by adding a *Bgl*II linker at the *Nar*I site of pBR322, isolating the resulting *Bam*HI-*Bgl*II fragment, inserting it at the *Bam*HI of pA0, and selecting for clones which restore the *Bam*HI site proximal to the enhancer. pAW52 (◇) (Wildeman *et al.*, 1984) is a derivative of pAW22 containing a 8-bp deletion in the 72-bp repeat (position 174–181, see Benoist and Chambon, 1981). The transcriptional activity of pAW22, pAW23, pAW32, pAW33 and pAW52 is expressed relative to pA71 (itself 30% of pA0), taken as 1. —△—, transcription with increasing number of 72-bp sequences (one, two and three copies, pA73, pA82 and pA83, respectively) when the enhancer has been separated from the 21-bp repeat region by inserting 600 nucleotides of pBR322 sequence, essentially as described in the legend to Figure 1E. The transcriptional activity of pA82 and pA83 is expressed relative to pA73 (itself 3.3% of pA0, Table I), taken as 1.

polyoma virus variants (reviewed in Ruley and Fried, 1983; Melin *et al.*, 1985). The chimeric enhancer (pA411 and pA401) appears to be 50–60% as efficient as the wild-type pA0 SV40 enhancer, which is particularly striking, since the isolated polyoma virus enhancer domain A (pA400, pA410) is as inefficient in stimulating transcription as the isolated SV40 domains A or B (pA223, pA224, pA233, pA234). Thus, it appears that transcriptional enhancement requires the presence of at least two enhancer domains, which by themselves have very little activity and whose synergy is, to a large extent, independent of their spacing and relative orientation.

A linear increase of transcription with multiple copies of the SV40 enhancer

Most SV40 isolates contain a duplication of the 72-bp sequence (the 72-bp repeat), although variants with only one copy (Waldeck and Sauer, 1976; van Heuverswyn and Fiers, 1979; Subramanian and Shenk, 1978) or with various deletions within the repeat (Reddy *et al.*, 1978; Benoist and Chambon, 1981) have been found. This prompted us to investigate the enhancing activity of multiple copies of the 72-bp sequence. There is a modest increase up to four copies (Figure 5, open circles), when transfection is performed under our standard conditions. However, when 10

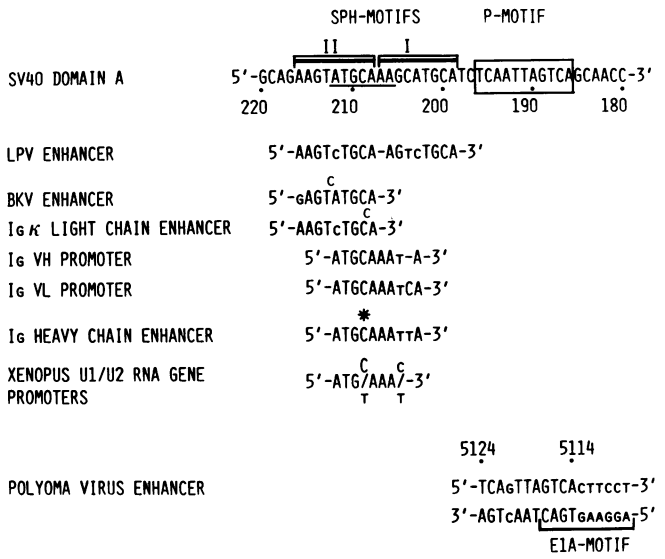


Fig. 6. Sequence elements homologous to the motifs present in the SV40 enhancer domain A. The two Sph-motifs I and II (positions 216 – 208 and 207 – 199) are indicated as in Figure 2A and B, the sequence 5'-ATGCAAA-3' homologous to the immunoglobulin heavy chain enhancer (position 212 – 206) is underlined, and the polyoma virus enhancer homology (P-motif, position 196 – 186) is boxed. References for the various sequence elements are as follows (see also text): LPV enhancer (Pawlita *et al.*, 1985; Mosthaf *et al.*, 1985); BK virus (BKV) enhancer (Rosenthal *et al.*, 1983); Igκ light chain enhancer (Picard and Schaffner, 1984; Queen and Stafford, 1984); consensus sequences of the upstream promoter elements of the immunoglobulin heavy (VH) and light (VL, opposite strand) chain genes (Falkner and Zachau, 1984; Parslow *et al.*, 1984); Ig heavy chain enhancer (opposite strand, Ephrussi *et al.*, 1985 and references therein; the star indicates the C complementary to the G protected against dimethylsulfate modification, see text); *Xenopus* U1/U2 RNA genes (Mattaj *et al.*, 1985; Ciiberto *et al.*, 1985; Krol *et al.*, 1985); polyoma virus enhancer (opposite strand, Veldman *et al.*, 1985; Ruley and Fried, 1983). For comparison the E1A-like motif of the polyoma virus enhancer (Hearing and Shenk, 1983; Herbomel *et al.* 1984) is shown.

times less specific DNA is transfected, transcription increases more dramatically, almost linearly up to seven copies of the 72-bp sequence, with an approximate 3.5-fold increase between the monomer and the dimer (Figure 5, open squares). The same relative increases in transcription are observed when enhancers containing multiple copies of the 72-bp sequence are separated from the 21-bp repeat region by a 43-bp (Figure 5, filled squares) or a 600-bp (Figure 5, open triangles) insert in the *Bam*HI site at position 101. An enhancer variant, in which there is an 8-bp deletion in GT-motif I of the 72-bp sequence proximal to the 21-bp repeat region (from positions 174 to 181 of the 72-bp repeat, see Benoist and Chambon, 1981), stimulates transcription as efficiently as an intact 72-bp repeat (Figure 5, open diamond). Enhancers containing multiple copies of not only the 72-bp sequence, but of the entire 94-bp region (position 179 – 272) are even more efficient in stimulating transcription (Figure 5, filled circles). Thus, increasing the number of 72-bp sequences, or even better the entire enhancer 94-bp sequence, results in a linear increase of transcription with respect to that achieved with the monomeric enhancer form. This linear increase is in marked contrast to the dramatic non-linear increase brought about by juxtaposing two enhancer domains (see above).

Discussion

The SV40 enhancer domains

The present study shows clearly that the 72-bp sequence is not the SV40 enhancer, since sequences located further upstream are essential for enhancer activity (pA104, Figure 2). Both deletion and point mutation analyses indicate that the activity of the minimal enhancer (containing one copy of the 72-bp sequence) is spread over a region of ~100 bp (185 – 275, Figure 2A). Two domains A and B, which by themselves exhibit a very low transcriptional activity, have been identified within this region. The existence of a third domain C (position 298 – 347) located upstream from the *Kpn*I site, whose activity is barely apparent when domains A and B are intact, is suggested because of the ability of this DNA segment to rescue the truncated enhancer present in the deletion mutant pA104 (see Figure 2A). That this segment contains a potential enhancer activity is further supported by the observations of Weber *et al.* (1984) and Swimmer and Shenk (1984), who reported that duplication of sequences from 298 to 376 or 357, respectively, generates a functional enhancer.

Both domains A and B have genuine enhancer properties, since mutations in either domain are equally detrimental to enhancer activity when the enhancer is either in close apposition to the 21-bp repeat or moved 600 bp away (Figure 1 and Table I). Since isolated domains A and B exhibit very little enhancer activity (Figures 2 and 4), the 400-fold stimulation of transcription brought about by their juxtaposition must involve some efficient synergistic mechanism. Surprisingly, this synergy does not appear to depend critically on the relative orientation of either domain (i.e. each domain can function bidirectionally like the entire enhancer) (Figure 4) or the exact distance between them (Table II). Thus, the two enhancer domains A and B appear to behave as individual domains, even though they functionally cooperate to generate enhancer activity. However, maximal enhancer activity, which is achieved when the enhancer is close to the remaining part of the SV40 early promoter, requires a stereospecific alignment between some element(s) of domain A and the 21-bp repeat region (Takahashi *et al.*, 1986).

Enhancer activity can be generated by association of domains A and B, by duplication of either domains A or B, or even by replacing SV40 domain A by the polyoma virus enhancer domain A (Figure 4). Generation of enhancer activity by duplication of a non-functional enhancer domain has been previously reported (Veldman *et al.*, 1985; Weber *et al.*, 1984; Swimmer and Shenk, 1984; Herr and Gluzman, 1985). It is interesting to note that, once the basic SV40 enhancer activity has been achieved by associating domains A and B, increasing the number of 72-bp sequences within the enhancer or polymerization of the entire enhancer results in a further linear increase of enhancer activity. This observation suggests that the underlying mechanism is different from that responsible for the burst of enhancer activity generated by juxtaposing domains A and B.

Multiple sequence motifs in the SV40 enhancer

Perhaps one of the most striking properties of the SV40 enhancer is that no individual point mutation decreases its activity by >8-fold (irrespective of the presence of enhancer domain C, data not shown), which still represents an ~50-fold enhancement relative to an enhancerless promoter. This suggests that many sequence elements are involved in generating full enhancer activity. Although there are no strong overall sequence homologies between different enhancers, several short and degenerate consensus sequences have been identified, but their functional significance has been studied only in a limited number of cases. On the basis of sequence homology between various enhancers and the results of a random mutagenesis, Weiher *et al.* (1983) have suggested that the so-called 'core' sequence $GTGG^{A}/_{T}{}^{A}/_{T}{}^{A}/_{T}G$ plays a key role in enhancer activity. Our

present systematic deletion and point mutation analysis confirms their suggestion, but reveals that there are in fact several additional sequence elements which, within domains A and B, are as critical for enhancer activity as the 'core' sequence.

Domain B contains two directly repeated sequence elements (the GT-motifs) 5′-G$^{C}/_{G}$TGTGGAA$^{A}/_{T}$GT-3′ (Figure 2), parts of which are crucial for enhancer activity. The downstream GT-motif I includes the 'core' sequence. The sequence TGG (positions 269 − 267 and 246 − 244, Figure 2) appears particularly important. Mutations of the other residues of this sequence element seem better tolerated in GT-motif I than in GT-motif II. Additional sequences located downstream from each motif are also involved in enhancer activity. GT-motif II is followed by an alternating purine/pyrimidine cluster, whose alteration has only a moderate effect on transcription (pA7, Figure 2), whereas GT-motif I is followed by the repetition of the sequence 5′-TCCCCAG-3′ (the TC-motif, Figure 2), of which only the upstream one is important for enhancer activity. It is striking that the tandem duplications of Herr and Gluzman (1985), which resuscitate SV40 by restoring the activity of enhancers mutated in their alternating purine/pyrimidine motifs, have the sequence 5′-TGTGGAAAGTCCCCA-3′ in common. This sequence encompasses exactly the residues which, within and downstream from GT-motif I, are important for enhancer activity in our transient assay in HeLa cells (see Figure 2A). It is noteworthy that

246 236

the SV40 sequence 5′-ṪGGAAAGTCCĊ-3′ is found in four 18-bp direct repeats present in the human cytomegalovirus enhancer (Boshart *et al.*, 1985) and that the SV40 sequence

236 228

5′-ĊCAGGCTCĊ-3′is present in the opposite orientation as a direct repeat in the promoter region of the hypoxanthine phosphoribosyltransferase gene (Melton *et al.*, 1984). Both GT-motifs are clearly required for enhancer activity, since deleting either of them abolishes its function (pA59, and compare pA104 with pA260 in Figure 2A). That the sequence 5′-GGTGTGG-3′ present in the two GT-motifs is also repeated in other enhancers [bovine papillomavirus (Lusky *et al.*, 1983; Weiher and Botchan, 1984), adenovirus E1A (Hen *et al.*, 1983), immunoglobulin κ-light chain (Queen and Stafford, 1984)] further supports this conclusion. While the centres of the two GT-motifs are separated by 23 bp, there appears to be no stringent distance requirement between them, since inserting 5 bp at position 257 (pA214, 23% of pA0; see also Figure 3B and pA213 in Figure 2A) does not abolish domain B activity. Furthermore, deleting GT-motif II can be compensated for by juxtaposing domain C to GT-motif I (compare pA104 with pA54 in Figure 2A). Thus domain B may be composed of two subdomains B1 and B2, both of which are required to generate an active domain B, but which can function, at least to some extent, independently of one another.

The directly repeated Sph-motifs 5′-AAG$^{T}/_{C}$ATGCA-3′, whose equivalent residues are separated by 9 bp, clearly represent one of the key features of domain A (Figure 2). Both motifs are required to generate an active domain A (see deletions pA62, Figure 2A and data not shown). Sequence elements closely related to the Sph-motifs are found as a single motif in the BK virus (Rosenthal *et al.*, 1983) and as a repeat in the lymphotropic papovavirus (LPV) (Furuno *et al.*, 1984; Pawlita *et al.*, 1985; Mosthaf *et al.*, 1985) (see Figure 6), where their function has not yet been studied. The Sph-motif is also present in the immunoglobulin κ-light chain enhancer (Picard and Schaffner, 1984; Queen and Baltimore, 1983), where its deletion is detrimental to enhancer activity (Queen and Stafford, 1984). In addition, the juxtaposition of the two Sph-motifs in the SV40 enhancer generates a sequence element which is present in the immunoglobulin heavy chain enhancer (Banerji *et al.*, 1983; Gilles *et al.*, 1983; Neuberger, 1983) and in the upstream elements of the promoters of the VH and VL immunoglobulin genes (Falkner and Zachau, 1984; Parslow *et al.*, 1984) (Figure 6). A homologous motif, which exhibits enhancer properties (Mattaj *et al.*, 1985), has also been described in a conserved sequence of a distal element of the *Xenopus* U1 and U2 RNA genes (Mattaj *et al.*, 1985; Ciliberto *et al.*, 1985; Krol *et al.*, 1985). It is interesting to note that the G residue complementary to the C present in this sequence in the immunoglobulin heavy chain enhancer (see Figure 6) is protected against dimethylsulfate methylation in lymphoid B cells, but not in other cell lines, which suggests that this sequence interacts with a protein specifically in lymphoid B cells (Ephrussi *et al.*, 1985). Since a mutation affecting this G is apparently not detrimental to SV40 enhancer activity in HeLa cells (pA24 in Figure 2A), it will be interesting to determine whether the same mutation will affect its activity in cells of the B lineage. This homology to the immunoglobulin heavy chain enhancer is not present in the polyoma virus enhancer domain A, which can efficiently substitute for SV40 domain A in HeLa cells, which raises the question as to whether the SV40-polyoma virus chimeric enhancer will efficiently activate transcription in lymphoid cells. The downstream Sph-motif (Sph-motif I), but not the upstream one (Sph-motif II), contains an 8-bp stretch of alternating purine/pyrimidine residues, within which mutations are detrimental to enhancer activity (Figure 2). As already discussed, the contribution of this sequence to enhancer activity cannot be simply ascribed to its potential to form a Z-DNA structure. In this respect, we note also that replacing SV40 domain A by polyoma virus domain A in the chimeric enhancer of pA411 (Figure 6) does not strongly decrease the enhancer activity, although the polyoma virus domain A does not contain any potential Z-DNA structure.

Domain A contains an additional sequence motif which is defined by mutants pA29 − pA30. The sequence 5′-TCAATTAG-TCA-3′ (called hereafter the P-motif) is found (in the opposite strand) in the polyoma virus enhancer domain A, where it partially overlaps the adenovirus E1A-like enhancer motif (see Veldman *et al.*, 1985; Herbomel *et al.*, 1984 and Figure 6). A

258 251

P-motif related sequence ṪCAGTTAĠ is also present in SV40 domain B where its mutation is detrimental to enhancer activity (see Figure 2). The P-motif is duplicated in a variety of naturally occurring polyomavirus isolates and in mutants which are selected for growth in PCC4 teratocarcinoma cells (reviewed in Ruley and Fried, 1983; Melin *et al.*, 1985), suggesting that it corresponds to an important element of the polyoma enhancer. Since a 26-bp polyoma virus segment which contains both the P-motif and the E1A-like enhancer motif acquires enhancer activity when polymerized (Veldman *et al.*, 1985), it would be interesting to determine which of the two motifs (E1A-like or P-motif or both) of the polyoma virus domain A is functional in the chimeric SV40-polyomavirus enhancer. In this respect it is worth noting that although the wild-type polyoma virus enhancer is almost inactive in embryonal carcinoma (EC) cells (F9 and PCC3), our chimeric enhancer is as active in these cells as in HeLa cells (our unpublished results), indicating that exchanging domain B of the polyoma virus enhancer for domain B of the SV40 enhancer results in an enhancer active in EC cells.

Modular organisation of enhancers

In conclusion, we have demonstrated that the SV40 enhancer is composed of several sequence motifs whose integrity is required for full enhancer activity in HeLa cells. These motifs belong to two major domains which have very little stimulatory activity by themselves, but generate a very potent enhancer when associated, suggesting the involvement of some type of cooperativity, the mechanism of which is unknown at present. All known viral and cellular enhancers contain short sequence stretches which exhibit some homology to one or several of the SV40 motifs, and when investigated, their importance for enhancer activity has indeed beeen demonstrated (Weiher *et al.*, 1983; Hen *et al.*, 1983; Weiher and Botchan, 1984; Queen and Stafford, 1984). It appears, therefore, that enhancers are made up of several sequence motifs which have been combined in various ways during evolution. Results reported elsewhere (Wildeman *et al.*, in preparation) strongly suggest that the role of the various SV40 motifs is to bind specific *trans*-acting factors involved in enhancer function (Schöler and Gruss, 1984; Mercola *et al.*, 1985; Wildeman *et al.*, 1984; Sergeant *et al.*, 1984; Sassone-Corsi *et al.*, 1985). Thus, the redundancy of enhancer motifs in SV40 may only be apparent if one assumes that some of the *trans*-acting factors are host- or cell-type specific, and that the ubiquitous activity of the SV40 enhancer is related to the multiplicity of its enhancer motifs. In this respect we note that the activity of the polyoma virus enhancer which lacks the Sph-motif and the duplication of the GT-motif, is restricted to certain cell types (see Ruley and Fried, 1983 for references), but that mutants active in some of these cells contain duplications of either polyoma virus enhancer domain A or B (reviewed in Ruley and Fried, 1983; Melin *et al.*, 1985). Similarly, the cell-type specificity of the human polyomavirus JC is modified by multiplication of a GT-motif-like sequence (Miyamura *et al.*, 1985). In this respect, it is also worth recalling that our chimeric SV40-polyoma virus enhancer, in which the polyoma virus enhancer domain B is replaced by the SV40 domain B, is active in EC cells.

Thus, host-range and cell-type specificity of enhancer activity may be achieved both by the presence (either singly or duplicated) or absence of a given motif and the presence or absence of the corresponding specific *trans*-acting factor. Absence of a given motif or of its cognate *trans*-acting factor may be compensated by duplication of another motif. Obviously, such a modular organization of enhancers would enormously increase the combinatorial possibilities of transcriptional control of gene expression. The set of SV40 mutants described in the present study will be very useful to test these hypotheses.

Materials and methods

Recombinant plasmids were introduced into HeLa cells by the calcium phosphate co-precipitation technique (Banerji *et al.*, 1981). Semiconfluent HeLa cells ($\sim 6 \times 10^6$ cells per 150 cm^2 Petri dish) were transfected with 1.5 μg of pA plasmid DNA, 4.5 μg of the reference gene p$\beta(244^+)\beta$ (De Villiers and Schaffner, 1981) and 21 μg of pBR322 DNA. After 24 h at 37°C the medium was removed, the cells washed with phosphate-buffered saline (PBS) and fed again with fresh culture medium [Dulbecco's modified Eagle's minimal essential medium (Gibco), containing 2.5% fetal calf serum, 2.5% calf serum, 500 U/ml penicillin, 40 μg/ml gentamycin and 100 μg/ml streptomycin]. Cytoplasmic RNA was prepared 40–50 h after transfection by lysing the cells with 0.5% Nonidet P-40. 5–25 μg of RNA were hybridized with an excess of single-stranded ^{32}P 5′ end-labeled DNA probe (see below) at 42°C in 50% (v/v) formamide, 400 mM NaCl, 40 mM Pipes, pH 6.5, 1 mM EDTA. After S1 nuclease digestion (BRL, 170 units at 25°C for 2 h) the protected DNA sequences were resolved in 8% polyacrylamide/8.3 M urea/salt gradient gels (Biggin *et al.*, 1983). The single-stranded DNA probe was prepared by first hybridizing a synthetic oligonucleotide, 5′-GCACCATTCTGTCTGTTTTGGG-3′, complementary to positions +39 to +60 of the rabbit β-globin gene to single-stranded M13mp8βSV phage DNA which contains the *Eco*RI/*Taq*I fragment of pA0 (SV40 coordinate 346 to β-globin coordinate +307) cloned between the *Eco*RI and *Acc*I sites of M13mp8. After primer extension, performed essentially as described elsewhere (Grundström *et al.*, 1985), and digestion with *Eco*RI, the single-stranded DNA (SV40 coordinate 346 to β-globin coordinate +60) was purified by polyacrylamide gel electrophoresis (6% polyacrylamide, 8.3 M urea) and recovered by electroelution. Aliquots of this single-stranded DNA preparation were then 5′ end-labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase.

Acknowledgements

We thank H.Kopecka and M.Girard for a generous gift of plasmid pCWori. We are grateful to Drs B.Wasylyk, I.Davidson and G.Albrecht for critical readings of the manuscript. The help of A.Staub for oligonucleotide synthesis is gratefully acknowledged. We thank F.Ruffenach and the cell culture group for excellent technical assistance. We are most grateful to C.Werlé and B.Boulay for the illustrations and to the secretarial staff for their patience in typing the manuscript. This work was supported by the CNRS (ATP 6984), the INSERM (PRC 124026), the Fondation pour la Recherche Médicale, the Ministère de la Recherche et de la Technologie (82V1283) and the Association pour le Développement de la Recherche sur le Cancer. T.Grundström, A.Wildeman and M.Zenke were recipients of fellowships from MFR-INSERM, Natural Sciences and Engineering Research Council of Canada and the CNRS/Deutsche Forschungsgemeinschaft, respectively.

References

Azorin,F. and Rich,S. (1985) *Cell*, **41**, 365–374.
Baty,D., Barrera-Saldana,H.A., Everett,R.D., Vigneron,M. and Chambon,P. (1984) *Nucleic Acids Res.*, **12**, 915–932.
Banerji,J., Rusconi,S. and Schaffner,W. (1981) *Cell*, **27**, 299–308.
Banerji,J., Olson,L. and Schaffner,W. (1983) *Cell*, **33**, 729–740.
Benoist,C. and Chambon,P. (1981) *Nature*, **290**, 304–310.
Biggin,M.D., Gibson,T.J. and Hong,G.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3963–3965.
Boshart,M., Weber,F., Jahn,G., Dorsch-Häsler,K., Fleckenstein,B. and Schaffner,W. (1985) *Cell*, **41**, 521–530.
Chambon,P., Dierich,A., Gaub,M.P., Jakowlev,S.B., Jongstra,J., Krust,A., LePennec,J.P., Oudet,P. and Reudelhuber,T. (1984) *Recent Prog. Horm. Res.*, **40**, 1–42.
Ciliberto,G., Buckland,R., Cortese,R. and Philipson,L. (1985) *EMBO J.*, **4**, 1537–1543.
De Villiers,J. and Schaffner,W. (1981) *Nucleic Acids Res.*, **9**, 6251–6264.
Ephrussi,A., Church,G.M., Tonegawa,S. and Gilbert,W. (1985) *Science*, **227**, 134–140.
Falkner,F.G. and Zachau,H.G. (1984) *Nature*, **310**, 71–74.
Furuno,A., Miyamura,T. and Yoshiike,K. (1984) *J. Virol.*, **50**, 451–456.
Gilles,S.D., Morrison,S.L., Oi,V.T. and Tonegawa,S. (1983) *Cell*, **33**, 717–728.
Grundström,T., Zenke,M., Wintzerith,M., Matthes,H.W.D., Staub,A. and Chambon,P. (1985) *Nucleic Acids Res.*, **13**, 3305–3316.
Gruss,P. (1984) *DNA*, **3**, 1–5.
Hearing,P. and Shenk,T. (1983) *Cell*, **33**, 695–703.
Hen,R., Borrelli,E., Sassone-Corsi,P. and Chambon,P. (1983) *Nucleic Acids Res.*, **11**, 8747–8760.
Herbomel,P., Boucharot,B. and Yaniv,M. (1984) *Cell*, **39**, 653–662.
Herr,W. and Gluzman,Y. (1985) *Nature*, **313**, 711–714.
Kahn,J., Matthes,H.W.D., Gait,M.J. and Brenner,S. (1984) *Gene*, **32**, 217–224.
Khoury,G. and Gruss,P. (1983) *Cell*, **33**, 313–314.
Kriegler,M. and Botchan,M. (1983) *Mol. Cell. Biol.*, **3**, 325–339.
Krol,A., Lund,E. and Dahlberg,J.E. (1985) *EMBO J.*, **4**, 1529–1535.
Laimins,L.A., Khoury,G., Gorman,C., Howard,B. and Gruss,P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6453–6457.
Lusky,M., Berg,L., Weiher,H. and Botchan,M. (1983) *Mol. Cell. Biol.*, **3**, 1108–1122.
Mattaj,I.W., Lienhard,S., Jiricny,J. and De Robertis,E.M. (1985) *Nature*, **316**, 163–164.
Melin,F., Pinon,H., Reiss,C., Kress,C., Montreau,N. and Blangy,D. (1985) *EMBO J.*, **4**, 1799–1803.
Melton,D.W., Konecki,D.S., Brennand,J. and Caskey,C.T. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2147–2151.
Mercola,M., Goverman,J., Mirell,C. and Calame,K. (1985) *Science*, **227**, 226–270.
Miyamura,T., Furuno,A. and Yoshiike,K. (1985) *J. Virol.*, **54**, 750–756.
Moreau,P., Hen,R., Wasylyk,B., Everett,R., Gaub,M.P. and Chambon,P. (1981) *Nucleic Acids Res.*, **9**, 6047–6068.
Mosthaf,L., Pawilta,M. and Gruss,P. (1985) *Nature*, **315**, 597–600.
Neuberger,M.S. (1983) *EMBO J.*, **2**, 1373–1378.

Nordheim,A. and Rich,A. (1983) *Nature*, **303**, 674–679.
Parslow,T.G., Blair,D.L., Murphy,W.L. and Granner,D.K. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2650–2654.
Pawlita,M., Clad,A. and zur Hausen,H. (1985) *Virology*, **143**, 196–211.
Picard,D. and Schaffner,W. (1984) *Nature*, **307**, 80–82.
Queen,C. and Baltimore,D. (1983) *Cell*, **33**, 741–748.
Queen,C. and Stafford,J. (1984) *Mol. Cell. Biol.*, **4**, 1042–1049.
Reddy,V.B., Thimmappaya,B., Dhar,R., Subramanian,K.N., Zain,S., Pan,J., Gosh,P.K., Celma,M.L. and Weissman,S.M. (1978) *Science*, **200**, 494–502.
Rosenthal,N., Kress,M., Gruss,P. and Khoury,G. (1983) *Science*, **222**, 749–755.
Ruley,H.E. and Fried,M. (1983) *J. Virol.*, **47**, 233–237.
Sassone-Corsi,P., Dougherty,J., 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.
Schöler,H.R. and Gruss,P. (1984) *Cell*, **36**, 403–411.
Serfling,E., Jasin,M. and Schaffner,W. (1985) *Trends Genet.*, **1**, 224–230.
Sergeant,A., Bohmann,D., Zentgraf,H., Weiher,H. and Keller,W. (1984) *J. Mol. Biol.*, **180**, 577–600.
Subramanian,K.N. and Shenk,T. (1978) *Nucleic Acids Res.*, **5**, 3635–3642.
Swimmer,C. and Shenk,T. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6652–6656.
Takahashi,K., Vigneron,N., Matthes,H., Wildeman,A. and Chambon,P. (1986) *Nature*, **319**, 121–126.
Tooze,J. (1982) *Molecular Biology of Tumor Viruses, 2nd Ed., Part 2, Revised. DNA Tumor Viruses.* Cold Spring Harbor Laboratory Press, New York.
Tyndall,C., LaMantia,G., Thacker,C.M., Favaloro,J. and Kamen,R. (1981) *Nucleic Acids Res.*, **9**, 6231–6250.
van Heuverswyn,H. and Fiers,W. (1979) *Eur. J. Biochem.*, **100**, 51–60.
van Ooyen,A., van der Berg,J., Mantei,N. and Weissmann,C. (1979) *Science*, **206**, 337–344.
Veldman,G., Lupton,S. and Kamen,R. (1985) *Mol. Cell. Biol.*, **5**, 649–658.
Waldeck,W. and Sauer,G. (1976) *FEBS Lett.*, **71**, 313-315.
Wasylyk,B. (1986) *Genetic Engineering*. Academic Press, New York, in press.
Wasylyk,B., Wasylyk,C., Matthes,H., Wintzerith,M. and Chambon,P. (1983a) *EMBO J.*, **2**, 1605-1611.
Wasylyk,B., Wasylyk,C., Augereau,P. and Chambon,P. (1983b) *Cell*, **32**, 503-514.
Wasylyk,B., Wasylyk,C. and Chambon,P. (1984) *Nucleic Acids Res.*, **12**, 5589-5608.
Weber,F., DeVilliers,J. and Schaffner,W. (1984) *Cell*, **36**, 983-992.
Weiher,H. and Botchan,M.R. (1984) *Nucleic Acids Res.*, **12**, 2901-2916.
Weiher,H., Koenig,M. and Gruss,P. (1983) *Science*, **219**, 626-631.
Wildeman,A.G., Sassone-Corsi,P., Grundström,T., Zenke,M. and Chambon,P. (1984) *EMBO J.*, **3**, 3129-3133.
Yaniv,M. (1984) *Biol. Cell*, **50**, 203-216.

Received on 4 November 1985