Volume 9 Number 22 1981 Nucleic Acids Research

The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants

P.Moreau, R.Hen, B.Wasylyk, R.Everett, M.P.Gaub and P.Chambon

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

Received 16 October 1981

ABSTRACT

By introduction of recombinant plasmids into monkey CV1 cells, we have unambiguously demonstrated that sequences entirely within the 72 bp repeat, which is located upstream of the SV40 early region, are crucial for T-antigen expression *in vivo*. We have also shown that a DNA fragment containing the 72 bp repeat, inserted directly before chicken conalbumin or adenovirus-2 major late promoter sequences in chimeric plasmids where these promoters replace that of the SV40 early genes, caused a dramatic increase in the expression of T-antigen *in vivo*. This effect was independent of the orientation of the 72 bp repeat, but was sensitive to its location within the plasmid. When the 72 bp repeat was separated from the promoter sequences, T-antigen expression was reduced. Insertion of the 72 bp repeat into equivalent plasmids containing no known eukaryotic promoter sequences (plasmids which were not detectably expressed *in vivo*) gave rise to a measurable, but smaller level of expression. The stimulation of expression by the 72 bp repeat is cis-acting : it required covalent linkage to the recombinant. We discuss the possibility that the 72 bp repeat region in SV40 may act as a bi-directional entry site for RNA polymerase B such that promoter sequences linked to the repeat are more efficiently utilised.

INTRODUCTION

A prokaryotic promoter is the genetic element necessary for expression of a particular structural gene. It is defined by cis-dominant mutations that affect the basal level of expression of this gene, and it is different from the structural gene itself and from elements that regulate its expression (1, 2). The molecular analysis of E.coli promoters has shown that promoters are regions of DNA that specifically interact with RNA polymerase before correct initiation of transcription of the structural gene can occur (for reviews, see 3, 4). All of the necessary promoter elements, such as the Pribnow box and the -35 sequence are located within 50 bp upstream from the site of mRNA initiation (for reviews, see 3, 4).

Recent studies on the eukaryotic promoters for RNA polymerase B (II) have suggested that it is not a simple RNA polymerase binding site, analogous to

prokaryotic promoters. The role of sequences located more than 100 bp upstream before the mRNA initiation sites has been demonstrated for both viral (5, 6, 7) and cellular (8, 9) genes (see also ref. 10 for a review). For instance, the results obtained with a number of deletion mutants have suggested that the most important elements for SV40 early gene expression lie within the 72 bp repeat region which starts about 115 bp upstream from the mRNA startsites (6, see also Fig. 1). The eukaryotic promoter could therefore be quite large when compared with its prokaryotic counterpart, but in addition, gene expression can still occur when only the upstream element is present, and when other promoter elements (such as the TATA box) are deleted (6). This striking finding, with no known counterpart in prokaryotes, has prompted us to study in more detail the nature of the SV40 sequences involved and to analyze their "promoter" properties.

Using deletion mutants and indirect immunofluorescence as an assay for T-antigen expression, we here demonstrate unambiguously that sequences lying within the 72 bp repeat are crucial for SV40 early gene expression. Moreover, the SV40 72 bp repeat region has a promoter-like activity by itself. We have also constructed chimeric recombinant plasmids containing the SV40 72 bp repeat region and sequences from known chicken conalbumin or adenovirus 2 major late promoter regions linked to the coding sequences of SV40 T-antigens. Using transfection or nuclear microinjection techniques to introduce the DNA of these recombinants into monkey cells, and indirect immunofluorescence as an assay for T-antigen expression, we show not only that the SV40 72 bp repeat region can act in either orientation with heterologous promoter sequences, but also that it retains some of its activity even when it is located far from the promoter and T-antigen structural sequences. Mechanisms which could possibly account for the puzzling effect of the SV40 72 bp repeat region are discussed in the light of our present observations.

MATERIALS AND METHODS

A) SV40 Nucleotide sequence numbering system.

The sequence for SV40 is from (11) as modified in (12) and numbered in (13), BBB system.

B) Recombinant plasmids.

pSV1 [containing the SV40 early gene wild type sequence from the HpaII (346) to the BamHI (2533) sites], AS (derived from pSV1 by deletion of 9 bp at the BglI site), pEMP (containing all of the T-antigen structural sequences from 5227 to 2533), and pHS102 have been described (6). DNA sequencing analysis (B.W. unpublished data) has shown that pHS102 has a deletion between position 5227 and 113 (see Fig. 1) with a BamHI site reconstructed at position 113.

DB14 was constructed by cloning the pHS102 box fragment A (Fig. 1) between the BamHI and PvuII sites of pBR322.

1) Construction and sequence of the SV40 mutants (TB series)

TB0 was constructed by digesting pSV1 with SphI which cuts once in pBR322 (position 565) and once in each 72 bp repeat (see Fig. 1). The two largest fragments were purified on sucrose gradients and ligated. This resulted in TB0 with only one 72 bp repeat (sequence in Fig. 2). TB101, TB202 and TB208 were constructed by isolating the pSV1 HinfI fragment (position 631 in pBR322 to position 5135 in SV40) which was cleaved by BstNI (EcoRII). The DNA fragments were treated with Bal31 nuclease, repaired with DNA polymerase I and ligated. After cleavage with BglI and ClaI, the fragments were inserted between the ClaI and BglI sites of pSV1 (partial digest). The resulting mutants have deletions of one 72 bp sequence and 21 (TB208), 22 (TB202) or 32 nucleotides (TB101) in the remaining 72 bp sequence (Fig. 2). [The sequence of these mutants was deduced (14) after labelling at the pBR322 HindIII site]. TB101 and TBBI101 were obtained by blunt ligation of the pHS102 box fragment A (Fig. 1) into the

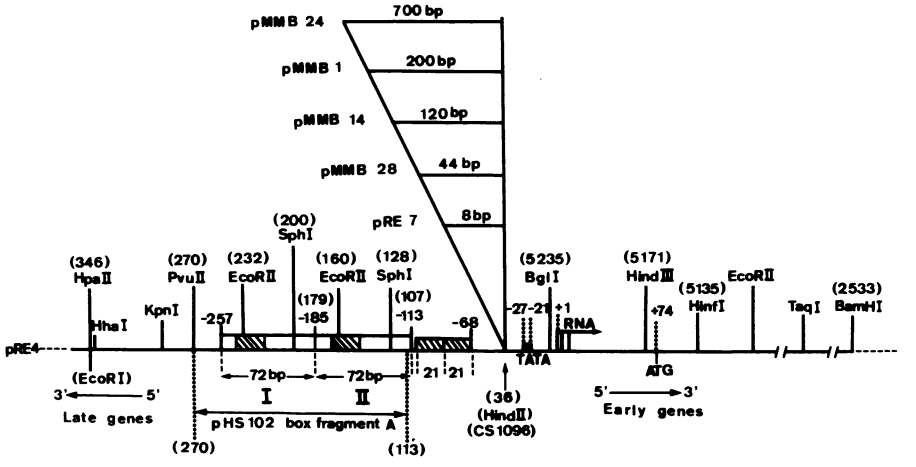

Fig. 1 : The early promoter region of SV40, pRE and pMMB. The construction of pRE4, pRE7 and the pMMB series is described in "Materials and Methods". Numbers in parentheses indicate the nucleotide position in the SV40 genome (Materials and Methods). +1 corresponds to one of the cap sites of the early mRNAs and the negative numbers count backwards from this +1 position (6). The two 72 bp sequences directly repeated (numbered I and II) are shown as open boxes between nucleotides 179 and 250, and 107 and 178. The hatched areas indicate GC-rich regions within these repeats. A further GC-rich exact 21 bp direct repeat (62-82 and 83-103) is shown by hatched boxes. The extra HindII site in cs1096 and pRE4 is at position 36. pRE7 has an 8 bp insert of a BamHI linker at this position and the pMMB series plasmids have further insertions of the indicated sizes in this BamHI site. The pHS102 box fragment A indicates the BamHI-PvuII fragment which is excised from pHS102 and used for construction of various recombinants (Materials and Methods). TATA = the SV40 early region TATA box (6). ATG = the translation initiation codon for T-antigens. (EcoRI) = the pBR322 EcoRI site in which the early HpaII-BamHI fragment of SV40 cs1096 was cloned to give pRE4.

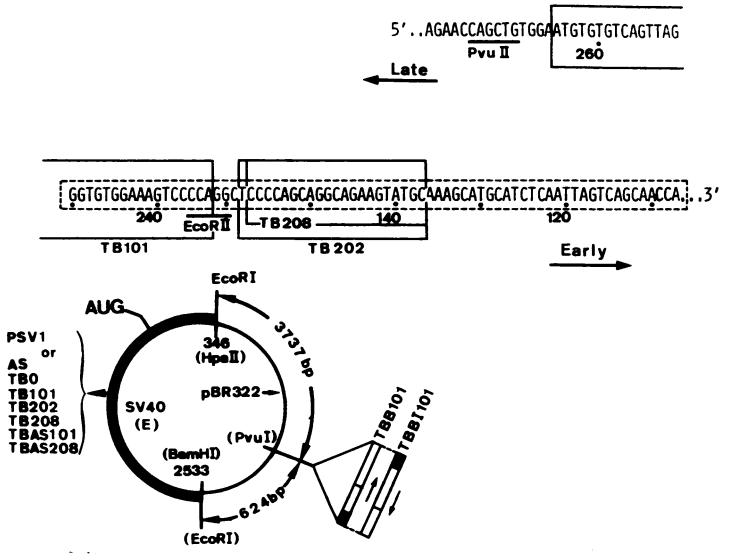

Fig. 2: The TB and TBB SV40 recombinants. The sequence (early non-coding strand) shows the unique 72 bp sequence (dashed line box) which is present in the TB0 deletion mutant (Materials and Methods). The solid line boxes indicate the extent of the additional deletions present in TB101, TB202 and TB208. The general structure of the various recombinants described in Materials and Methods is shown below the sequence. The thick line represents the SV40 sequences and the thin line corresponds to pBR322. SV40 coordinates of restriction sites used for cloning are shown inside the circle. Sites in parentheses were not regenerated (Materials and Methods). The pHS102 box fragment A (Fig. 1) is represented by two open boxes (the 72 bp repeat) and one filled box, and the arrows indicate its natural orientation with respect to SV40 early genes (E).

PvuI site of pBR322. Recombinants TBAS101 and TBAS208 were constructed by insertion in the correct orientation of the SphI fragment which contains the mutated SV40 early region in TB101 and TB208 between the SphI sites of mutants AS (see above).

2) Construction of pRE and pMMB series.

The SV40 early gene HpaII-BamHI fragment of the SV40 mutant cs1096, [which contains two point mutations one of which results in extra HindII site at position 36 - see Fig. 1 (15)] was cloned into the EcoRI site of pRE3 in an analogous fashion to the construction of pSV1 (5). pRE3 is a mutant of pBR322 obtained with bisulphite treatment (16) which has lost the BamHI site, but retains tetracycline resistance (R.E. unpublished results). The resultant plasmid, pRE4 (Fig. 1), was partially digested with HindII, and linear full length molecules were then isolated by sucrose gradient centrifugation. These molecules were ligated at 4°C in the presence of a 20-fold molar excess of BamHI linker (Collaborative Research). One of the resultant recombinants, pRE7, was identical to pRE4 except that a single BamHI linker had been inserted into the HindII site at positon 36, giving an unique BamHI site. The insertions in

the pMMB series (Fig.1) were obtained by blunt-end ligation of a total AluI digest of pX3 in the BamHI site of pRE7 [pX3 contains cs1096 cloned into the BamHI site of pBR322 (D. DiMaio, personal communication)]. This gives BamHI sites at the boundaries of the inserted fragments. The sizes of these fragments were determined by restriction mapping. None of them was derived from the AluI fragment (5226 to 271) which includes the SV40 origin of replication.

pRE7BS, a derivative of pRE7 which has lost the BglI site of SV40, was constructed by ligation of S1 nuclease treated full-length linear pRE7 isolated from a sucrose gradient following partial BglI digestion. The insert fragments from pMMB1 and pMMB28 were ligated into the BamHI site of pRE7BS to create pMMB1BS and pMMB28BS.

3) Recombinants containing either no known promoter (pEMP series and pHSX102) or a chicken conalbumin gene "promoter" fragment (pTCT and pSVCT series).

pTCT : the conalbumin +62 to -102 AluI fragment (17) was cloned in the repaired unique BamHI site of pEMP (5).This regenerated BamHI sites at both ends of the conalbumin fragment. pTCT has the conalbumin fragment in the same orientation with respect to transcription as the SV40 T-antigen coding sequences (Fig. 3). pTCTI has the conalbumin fragment in the opposite orientation (not shown). pTCTB, pTCTBI, pTBCT, pTBICT, pEMPB and pEMPBI : the repaired fragment A from pHS102 (Fig. 1) was cloned in pTCT which had been partially digested with BamHI and repaired (see Fig. 3). pTCTB3, pTCTBI3, pEMPB3 and pEMPBI3 : the repaired pHS102 fragment A was cloned in the repaired SalI site of pTCT (to give pTCTB3 and pTCTBI3, Fig. 3) or pEMP (to give pEMPB3 and pEMPBI3 Fig.3). pTCTB37, pTCTBI37, pEMPB37 and pEMPBI37 (Fig. 3) : first pTCT (tet^+) and pEMP (tet^+) (ie. pTCT and pEMP with intact tetracycline genes) were constructed by ligating the EcoRI (repaired) to SalI (non repaired) 650 bp fragment from pBR322 to pTCT which had been partially digested with BamHI, repaired and cleaved with SalI. The PstI 1423 bp fragment (from 3612 in pBR322 to 3204 in SV40) was replaced with the equivalent fragment from pSVPBA500 or pSVPBIA500 (see below) (Fig. 3). pSVCT and pSVCT : the +113 (SV40) to +375 (pBR) BamHI fragment from pHS102 (which contains the 72 bp repeat) was cloned into partially BamHI digested pTCT or pTCTI (pSVCTI) (Fig. 3). pHSX102 : the EcoRI (position 440) to HinfI (position 564) fragment from the chicken conalbumin double-stranded cDNA (18) was inserted by blunt-end ligation into the BamHI site of pHS102 (SV40 position 113).

4) Recombinants containing an adenovirus 2 major late (Ad2 MLP) promoter fragment (pSVA series).

The Ad2 MLP promoter fragments with deletions extending from upstream to positions -500, -40 and -29, (19) were excised with SalI (position 650 in pBR322) and PvuII (position +33 of the Ad2 MLP, see Ref. 19). The resulting fragments were inserted between the repaired BamHI site and the SalI site of pEMP (see above), yielding pSVA500, pSVA40 and pSVA29 (Fig. 4). The repaired pHS102 box fragment A (Fig. 1) was then inserted in both orientations in plasmids pSVA500, pSVA40 and pSVA29 at two different positions : the SstI site (present at positions -65 and -43, with respect to the mRNA start site in pSVA40 and pSVA29) or the PvuI site of pBR322, both made blunt-end with DNA polymeraseI (see Fig. 4). Insertions of two tandemly repeated fragments A in both orientations were also obtained (pSVPB2A500 and pSVPB2IA500).

The identity of all plasmids constructions shown in Fig. 1-4 was checked by extensive restriction enzyme analysis and by DNA sequencing (14) when necessary.

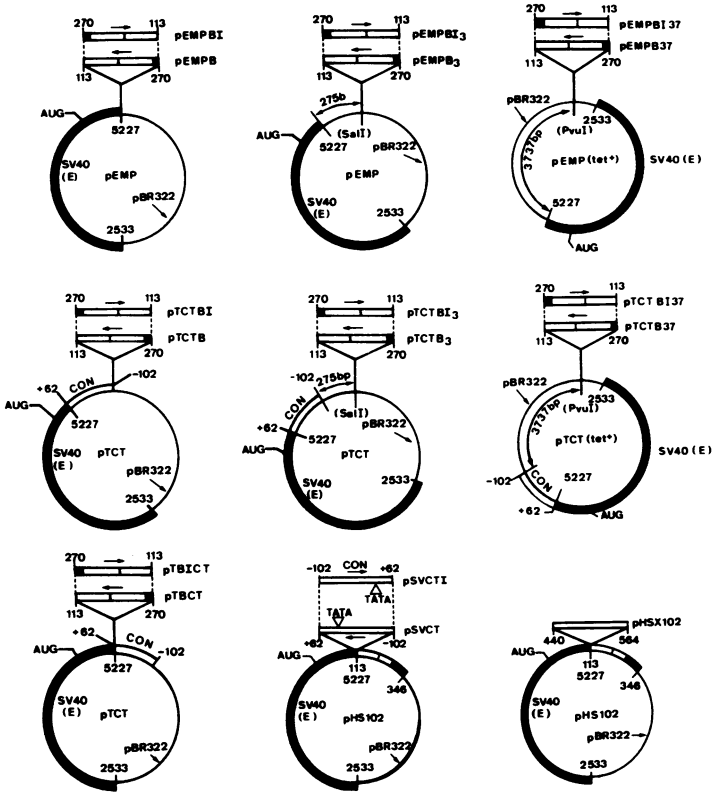

Fig. 3: Recombinants containing the SV40 72 bp repeat and either a heterologous promoter sequence (coordinates outside of the circle) from the chicken conalbumin gene (the pTCT series and pSVCT series) or no known equivalent sequence (the pEMP series and pHSX102). For the construction of recombinants see Materials and Methods. SV40 (E) = SV40 early region from nucleotide 5227 to 2533. Con = conalbumin gene promoter fragment from +62 to -102 (19) (the arrow indicates the direction of transcription of the conalbumin gene). TATA = "TATA box" homology sequence (or Goldberg-Hogness box) found in most genes transcribed by RNA polymerase B (or II) (10). For other symbols see Fig. 2.

C) Transfection of DNA into CV1 cells and indirect immunofluorescence assay.

The DEAE-dextran procedure was as described in (5,6) except that we use 1 µg plasmid DNA per 60 mm dish. After 36-48 hours expression of the SV40 early region was estimated by indirect immunofluorescence detection of large T-antigen as described (5, 6). 5 mm diameter circles were marked on each plate under areas of confluent cells. The cells with fluorescent nuclei within the marked areas were counted. Approximately 1% of the cells were fluorescent when the positive SV40 wild type recombinant pSV1 was used. The response varied with the amount of DNA used, but a plateau level was reached between 0.2 and 1 µg of DNA. The results were expressed in percentages of the value obtained for

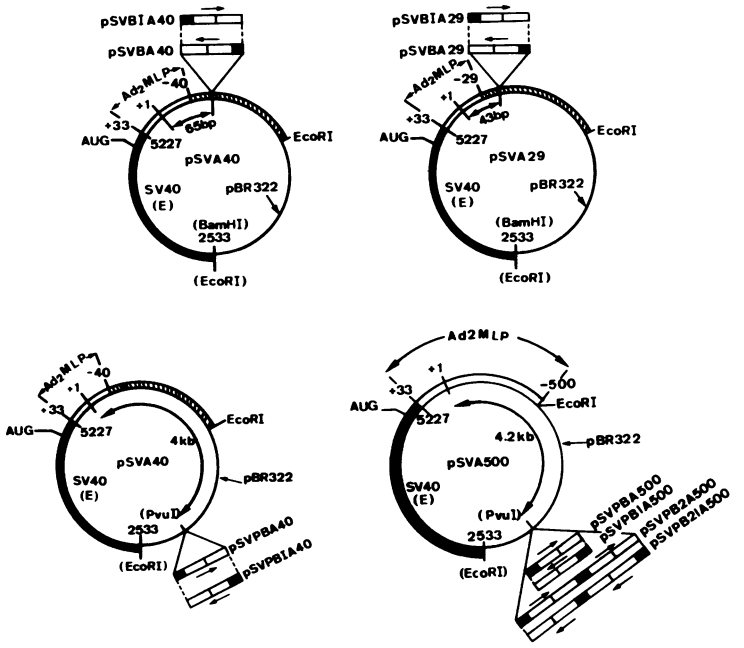

Fig. 4: Recombinant containing the SV40 72 bp repeat (fragment A, Fig. 1, open boxes and filled box outside the circle) and different sequences from the adenovirus 2 major late promoter (Ad2 MLP) (numbers outside the circle). Coordinates of the Ad2 MLP segments are given with respect to the mRNA startsite (position +1, see Ref. 19). Other numbers and symbols are as described in legends to Figs. 2 and 3.

pSV1 or AS recombinant in that particular experiment. In most cases numbers quoted in the tables are averages of the results from several independent experiments using at least two different preparations of plasmid DNA. The individual percentages could vary by as much as 50 % of the mean value (however the usual variation was less than this). Therefore we consider there to be little difference between plasmids which gave mean values of, for example 50 % and 70 % ; on the other hand the difference between plasmids giving 2 % and 10% was marked and we consider this to be a significant difference. Further we do not know the quantitative relationship between the percentage of positive cells and the actual level of T-antigen expression. We suspect that a 20-50 % reduction in positive cells reflects a far greater decrease in T-antigen levels (see Discussion).

D) Microinjection in CV1 cell nuclei.

We used the technique of Graesmann (20) as modified by W. Ansorge (21). Approximately 120 cells, in two different places on the dish, were microinjected on average with 1-2 x 10^{-11} ml (about 150-300 DNA molecules per cell). However it is likely that in practice there are much wider variations in the actual number of plasmid molecules which enter a nucleus. 24 h. later the number of dead cells were counted and the cells were fixed and stained for T-anti-

gen (6). The results were expressed as the percentage of surviving cells (about 80 %) which were positive for T-antigen relative to cells microinjected with pSV1. For a given DNA recombinant the variation of the percentage of positive cells from one experiment to another was about the same as with the transfection assay.

All other Materials and Methods were as previously described (5, 6).

RESULTS

A) Effect of deletions in the 72 bp repeats on SV40 early gene expression.

Starting from the SV40 wild-type early region recombinant pSV1 (5) we constructed a series of deletion mutants in the 72 bp repeat (Fig. 2). As shown in Table 1, removal of just one 72 bp sequence (TB0) resulted in only a minor reduction of T-antigen expression, in agreement with the previous observation of Gruss et al. (22) and with the existence of natural variants of SV40 with only one 72 bp sequence (12, 23, 24). However, deletions within this sequence (TB208 and TB202) gave a drastic reduction in expression. The microinjection results confirm that T-antigen expression is strongly decreased but not abolished in TB202. These results extend our previous observations (6) and demonstrate unambiguously that sequences within the 72 bp sequence, on the early side of the EcoRII site (Fig. 2), are crucial for SV40 early gene expression. However, other sequences on the late side of the EcoRII site are also important (TB101, Table 1). Whether additional deletions located exclusively within the 72 bp sequence would result in a complete shut-off of T-antigen expression is unknown at the present time.

To demonstrate that the above decreases were not due to a defect in replication, we constructed a series of derivatives of plasmid AS (6) which contains a 9 bp deletion in the origin of replication and therefore cannot replicate (25) but still gives roughly the same level of T-antigen expression as pSV1 (Table 1). As shown in Table 1, the effect of the deletions in the 72 bp sequence was very similar with recombinants containing a deletion in the origin of replication (TBAS101 and TBAS208).

B) The introduction into various recombinants of the 72 bp repeat region in either orientation stimulates gene expression.

1) Effect of introduction of the 72 bp box region directly upstream from the T-antigen coding sequence.

We have previously shown that deleting all of the sequences located upstream from the SV40 early gene mRNA startsites (pEMP mutant, see Ref. 5) abolished T-antigen expression. However, a low, but definite, level of express-

TABLE 1

Immunofluorescent CV1 cells (%)*

SV40 PROMOTER RECOMBINANTS	TRANS-FECTION	MICRO-INJECTION
pSV1, AS	100	100
TB0	85 (5)	ND
TB101	<1 (10)	13 (5)
TBAS101	<1 (2)	ND
TB208	4 (5)	ND
TBAS208	10 (3)	ND
TB202	4 (4)	32 (1)
TBB101	52 (4)	86 (4)
TBBI101	51 (1)	ND
pRE4	63 (12)	ND
pRE7	45 (12)	ND
pMMB28	37 (6)	ND
pMMB14	20 (6)	ND
pMMB1	8 (5)	ND
pMMB24	4 (3)	ND
pRE7BS	36 (4)	ND
pMMB28BS	27 (4)	ND
pMMB1BS	4 (4)	ND

PROMOTERLESS RECOMBINANTS	TRANS-FECTION	MICRO-INJECTION
pHS102	3 (6)	48 (4)
pEMP	0 (7)	0 (5)
pEMPB	2 (5)	ND
pEMPBI	3 (5)	ND
pEMPB3	3 (12)	ND
pEMPBI3	4 (13)	ND
pEMPB37	<1 (2)	ND
pEMPBI37	<1 (2)	18 (3)
pHSX102	3 (2)	50 (1)
pEMP+DB14	ND	0 (1)

*see Materials and Methods. Numbers in parentheses correspond to the number of independent assays which were carried out for a given recombinant. ND = not determined. Recombinants are as described in Materials and Methods and in Figs. 1 and 2.

ion was obtained in a mutant (pHS102 in Ref. 6 - see also Table 1, microinjection results) which retained the sequence located between the beginning of the 72 bp region (nucleotide 113) and the HpaII site (nucleotide 346) (see Fig. 1). The same level of expression was obtained with the mutant pEMPB (Fig. 3) in which the SV40 113 to 270 nucleotide segment ("pHS102 box fragment A" - Fig. 1) was inserted directly before the SV40 early gene coding region. The 72 bp repeat region can therefore "promote", to some extent, T-antigen production in the absence of any additional known promoter sequences. Moreover, its effect is apparently not related to its orientation with respect to the coding sequences, since identical results were obtained with pEMPB and pEMPBI (Table I and Fig. 3). In addition the 72 bp repeat region must be covalently linked to the structural gene sequences for the stimulation of expression to be observed, i.e. it is cis-acting. This was shown by experiments in which the plasmids pEMP and DB14 (the "pHS102 box fragment A" cloned in pBR322) were coinjected in CV1 cells (Table 1).

2) Stimulation of gene expression by the 72 bp repeat region in chimeric plasmids containing sequences of the conalbumin or Adenovirus 2 major late promoters.

Since the level of expression of the T-antigen was strongly decreased in pEMPB which lacks all of the SV40 promoter sequences located between the early mRNA startsites and nucleotide 113, particularly the TATA box region which is known to play a promoter role both *in vitro* (26) and *in vivo* (6, 27), we decided to construct recombinants in which the corresponding sequences of the conalbumin (nucleotides +62 to -102, Ref. 17) or Adenovirus 2 major late (Ad2 MLP) (nucleotide +33 to -40, Ref. 19) promoter regions were inserted (in the correct orientation with respect to the coding sequence) between the SV40 T-antigen structural sequence and the "pHS102 box fragment A" (Fig. 1) (see Figs. 3 and 4). In the absence of the 72 bp repeat, there was no detectable production of T-antigen in the conalbumin recombinant pTCT and only a very low level in the Adenovirus recombinant pSVA40 (Table 2). We do not know whether this difference, which is more obvious with the microinjection assay, reflects a difference in the properties of the two promoter regions or rather differences in the actual construction of the recombinants. However, in both cases, adding the 72 bp region in either orientation resulted in a striking increase of

TABLE 2

Immunofluorescent CV1 cells (%)*

CONALBUMIN PROMOTER RECOMBINANTS	TRANS-FECTION	MICRO-INJECTION	ADENOVIRUS-2 MAJOR LATE RECOMBINANTS	TRANS-FECTION	MICRO-INJECTION
pSV1, AS	100	100	PSV1, AS	100	100
pTCT	0 (4)	0 (2)	pSVA40	1 (12)	10 (3)
pTCTB	78 (3)	100 (2)	pSVBA40	82 (12)	119 (2)
pTCTBI	76 (4)	100 (1)	pSVBIA40	79 (12)	96 (2)
pTCTB3	9 (9)	83 (2)	pSVA29	<1 (12)	3 (3)
pTCTBI3	10 (11)	87 (5)	pSVBA29	30 (12)	90 (2)
pTCTB37	<1 (3)	77 (2)	pSVBIA29	27 (9)	74 (2)
pTCTBI37	<1 (2)	80 (3)	pSVPBA40	14 (5)	88 (3)
pTBCT	4 (5)	ND	pSVPBIA40	10 (5)	70 (3)
pTBICT	8 (5)	ND	pSVA500	1 (12)	33 (3)
pSVCT	98 (10)	98 (2)	pSVPBA500	23 (10)	93 (3)
pSVCTI	21 (9)	100 (2)	pSVPBIA500	15 (8)	83 (3)
pTCT+pHS102	4 (2)	ND	pSVPB2A500	41 (6)	ND
pTCT+DB14	0 (1)	0 (1)	pSVPB2IA500	36 (6)	ND
			pSVA40+DB14	<1 (2)	7 (1)

* see Materials and Methods. Number in parentheses and ND, see Table 1. Recombinants are as described in Materials and Methods and in Figs. 3 and 4.

T-antigen expression to a level very close to that of pSV1 (compare pTCTB, pTCTBI, pSVBA40 and pSVBIA40, Table 2 and Figs. 3 and 4). Sequences present in the 72 bp repeat were actually involved in the stimulatory effect : replacement of the 72 bp repeat region in pSVBA40 with a fragment containing the deleted 72 bp sequence present in TB208 (Fig. 2) almost abolished T-antigen expression, whereas its replacement with a fragment containing the non-deleted 72 bp sequence of TB0 (Fig. 2) did not affect noticably the original level of expression (R.H., data not shown). In addition this stimulatory effect was cis-acting like that discussed above for the pEMP derivatives (see in Table 2 the results of transfections or microinjections with pTCT + pHS102, pTCT + DB14 and pSVA40 + DB14).

Several lines of evidence indicate that it is not the mere insertion of conalbumin or Adenovirus fragments, with lengths approximately equal to that of the deleted SV40 sequences, which is responsible for the efficient expression of T-antigen. Firstly, insertion of an Ad2 MLP segment in which the first two bases of the TATA box, which are important for promotion of transcription *in vitro* (19), had been deleted (+33 to -29 segment, recombinants pSVA29 in Fig. 4) resulted in a significant decrease of T-antigen production (Table 2 - compare the results of transfection experiments with pSVBA40, pSVBIA40, pSVBA29, pSVBIA29). Secondly, inversion of the conalbumin segment (+62 to -102) which contains the TATA box which has been shown to play a promoter role both *in vitro* (28-30) and *in vivo* (31), led (at least after transfection) to a strong decrease in the percentage of immunofluorescent cells (compare pSVCT and pSVCTI, Fig. 4 and Table 2). Thirdly, restoring the initial distance between the 72 bp repeat and the SV40 T-antigen coding sequences by insertion of a fragment of 125 bp isolated from the conalbumin double-stranded cDNA (recombinant pHSX102, Fig. 3, see Materials and Methods) gave a level of expression (Table 1) which was similar to that of the parental deleted plasmid (pHS102). In addition, these results may suggest that the recombinant plasmids are actually transcribed *in vivo* by RNA polymerase class B (II), since the TATA box is a specific element of genes known to be transcribed by this enzyme (for reviews, see Refs. 10 and 19).

C) The location of the 72 bp repeat region in the various recombinants affects its ability to stimulate gene expression.

The 72 bp region can act in SV40 not only when it is in its natural location,but also when it is displaced by about 4.0 kb (Fig. 2), as shown by the resultsobtained with recombinants TBB101 and TBBI101, in which the 72 bp re-

peats were inserted in either orientation in the PvuI site of the 72 bp deletion mutant TB101 (Table 1). In both cases, there was a strong stimulation of T-antigen expression, irrespective of the orientation of the inserted repeat.

A similar, but reduced stimulation was obtained in the Adenovirus series where the 72 bp region was inserted at about 4.0 kb upstream from either the +33 to -40 or the +33 to -500 Ad2 MLP sequences (pSVPBA40, pSVPBIA40, pSVPBA500, pSVPBIA500 - Fig. 4 and Table 2). It is interesting to observe that insertion of four 72 bp sequences instead of two increased significantly the level of T-antigen expression (compare the pSVPBA500 and pSVPB2A500 series - Fig. 4 - in Table 2). The reduced effect of the insertion of the 72 bp region in the Adenovirus derivatives (when compared to the corresponding SV40 TBB101 recombinants) may be due to the fact that only a fraction of the 72 bp region has been deleted in mutant TB101 (see below). It is striking that in the case of the conalbumin derivatives, insertion of the 72 bp region in either orientation at about 3.7 kb from the conalbumin promoter sequences (pTCTB37, pTCTBI37 - Fig. 3 and Table 2) resulted in a much lower level of T-antigen expression than for the SV40 and Adenovirus corresponding derivatives. The results obtained with conalbumin derivatives where the 72 bp region is inserted in either orientation in the SalI site (pTCTB3 and pTCTBI3 - Fig. 3 and Table 2) suggest that the effect of the 72 bp repeat region decreased with increasing distances from the promoter sequences. It is possible that this "effect of distance" is more noticable with the conalbumin than with the Adenovirus derivatives (see above), because the promoter region of the latter is more efficient (pSVA40 expression is higher than that of pTCT, see microinjection results in Table 2). A similar "effect of distance" also seems to occur with the "promoterless" pEMP series (pEMPB, pEMPBI, pEMPB3, pEMPBI3 , pEMPB37, pEMPBI37 - Table 1 and Fig. 3). However, the microinjection results indicate that the 72 bp region has an effect on T-antigen expression in the conalbumin and "promoterless" derivatives even when the effect is not detectable by transfection (compare in Table 1 pEMP and pEMPBI37, and in Table 2 pTCT, pTCTB37 and pTCTBI37).

From the above results it appears that 1) the effect at distance of the 72 bp repeat can be observed in the absence of any known promoter sequence ; 2) there is a decrease in the effect of the 72 bp region with increasing distances from the promoter sequences ; 3) the same effects were obtained irrespective of the orientation of the inserted repeat region. It is important to stress that the nature of the assay which is used for measuring the variation of gene expression with respect to distance of the 72 bp region is of paramount importance. Clearly, the "effect of distance" would not have been noticed, if

we had used the microinjection assay alone. This is due to the difference in the sensitivity of the assays as discussed below.

The extent of stimulation of gene expression by the 72 bp region is influenced not only by the distance between the promoter sequences and the inserted region, but also by the relative order of the components present in the recombinants. This is shown by the results of recombinants pTBCT and pTBCTI (Fig. 3 and Table 2) in which the 72 bp region has been inserted in either orientation between the conalbumin promoter sequence and the SV40 T-antigen coding sequence. It appears that in this particular configuration the level of expression is not influenced by the presence of the conalbumin promoter region (compare pEMPB and pEMPBI with pTBCT and pTBCTI in Tables 1 and 2).

D) Effect of separation of the 72 and 21 bp repeat containing region from the TATA box region in SV40 recombinants.

The results described above indicate that, as the distance between the 72 bp repeat and the promoter sequences is increased, there is a decrease in the levels of T-antigen expression. To investigate this effect in more detail, we constructed a recombinant, pRE7, with a BamHI linker at position 34. A large number of insertion derivatives in the BamHI site of pRE7 were then constructed (Materials and Methods and Fig. 1). T-antigen expression obtained with these recombinant plasmids is shown in Table 1. The parent recombinant, pRE4, which is derived from the SV40 mutant cs1096 (15, Materials and Methods and Fig. 1) showed a small but reproducible decrease in expression when compared with the wild-type pSV1. This can only be due to the two single base changes in cs1096 and may be allied to the cold-sensitivity of the mutant. pRE7, with an 8 bp insertion, gave a further small fall in expression which could be due either to disruption of an important sequence in the -40 region, or to alteration of a critical distance separating two domains located on either side of this position. However, this function is not essential since the SV40 deletion mutant dl892 (23) which has lost 19 bp including those mutated in cs1096, is viable.

The insertion of increasingly longer DNA fragments in the BamHI site of pRE7 resulted in further decreases in expression which were particularly marked in the cases of pMMB1 (200 bp inserted) and pMMB24 (700 bp inserted) (Table 1). It is possible that the inhibition was due to the DNA sequences of the insertions themselves, rather than to the separation of the 72 bp and 21 bp repeat containing segment from the TATA box region. However, all of the 16 pMMB series plasmids studied gave reduced expression compared to pRE7, and of the 12 plas-

mids with insertions larger than 120 bp, 9 gave values of 12 % or less (R.E., unpublished results). Therefore it is likely that the inhibition is related to the insertion of a segment of DNA, and not to inhibitory DNA sequences.

This fall in expression could be due to the distance effect of moving the 72 bp repeat and the 21 bp GC-rich repeat away from the TATA box region, or it could be due to decreased replication of the plasmid constructs *in vivo*. To test this latter possibility, we constructed derivatives of pRE7, pMMB1 and pMMB28 (pRE7BS, pMMB1BS and pMMB28BS) in which the BglI site at the SV40 origin of replication was destroyed (Materials and Methods). These plasmids gave expression values in CV1 cells similar to their parents (Table 1). In addition, pRE4, pRE7 and the pMMB series plasmids gave the same pattern of results after transfection of mouse 3T34E cells (R.E., unpublished data), which are not permissive for SV40 DNA replication. These data exclude the possibility that the reduction in expression from the pMMB series plasmids was due to a defect in replication in the transfected cells.

In general therefore, increases in distance between the 72 and 21 bp repeat containing region and the TATA box region led very rapidly to marked reduction in expression of T-antigen. This could be due to the altered position of the 72 bp repeat region alone and therefore similar to some of the "effect of distance" described above. However the comparison of pMMB1 and pMMB24 with TBB101 (Figs. 1 and 2) indicates that although the position of the inserted 72 bp boxes in the latter is much further from the TATA box region than is the case in the pMMB plasmids, its expression level was significantly higher (Table 1). In the TBB101, the 21 bp repeats are in their natural position. Therefore the reduction of expression seen with pMMB plasmids could be due not only to moving the 72 bp boxes away, but also to the altered position of the 21 bp repeats, which suggests that the latter could play an important role in the SV40 early promoter. In fact, a deletion between nucleotides 36 and 113 (R.E., unpublished data) and other deletions within this region (6, R.E., unpublished data) do cause important reductions in T-antigen production in CV1 cells.

DISCUSSION.

A) The validity of the immunofluorescence assay.

The T-antigen immunofluorescence assay is an indirect method of measuring the efficacity of gene expression of the various recombinants. In particular, since we have scored only the number and not the fluorescence intensity of the CV1 positive cells obtained after transfection, some justification is required

to support our assumption that this number reflects the relative efficiencies of gene expression in the recombinants. It is reasonable to assume that in any one experiment the same proportion of cells in each culture dish used will be capable of DNA uptake and that the amount of DNA taken up by the nucleus of these competent cells would not be constant, but would rather follow a normal distribution varying from perhaps one to several tens of plasmid copies per cell. (This is certainly an upper limit because the results obtained by transfection and microinjection were markedly different in the case of poorly expressing recombinants - see below and Tables 1 and 2). It is also reasonable to assume that for a particular recombinant there is a minimum amount of DNA which is required in a cell to give, in total, enough T-antigen production for a nucleus to become visibly fluorescent. It follows that, when in a particular recombinant plasmid, the efficiency of T-antigen production is decreased, the number of plasmid copies required for a cell to become fluorescent increases and the number of positive cells in a constant area of confluent cell growth decreases relative to the wild-type situation. The justifications of these assumptions are empirical. Firstly, the results obtained were reproducible : the same relative efficiency was repeatedly observed for the different recombinants. Secondly, the SV40 wild-type recombinant pSV1 gave the complete range of fluorescent intensity (illustrating that different cells take up different amounts of DNA), but poorly expressed recombinants gave only a low number of weakly fluorescent positive cells. This implies that even cells which had taken up the maximum amount of DNA gave a fluorescence intensity that was weaker than those that had taken up the maximum amount of control DNA.

The results obtained after microinjection of DNA can be interpreted along the same lines, except that the amount of recombinant DNA entering the nucleus was on average much higher. Therefore even for weakly expressing T-antigen recombinants, this could now be a saturating amount of DNA which would result in a high proportion of fluorescent cells. Therefore the two methods are complementary rather than confirmatory and it is very likely that they can detect wide variations in T-antigen production. However it should be stressed that, since we have not taken into account the variations in fluorescence intensity of the T-antigen positive cells, our results may afford a minimal estimate of the variations of gene expression, and that significant decreases might be missed.

B) Characteristics of the effect of the SV40 72 bp sequence on gene expression.

Our results support and extend our (6) and other (22) previous conclusions.

With plasmids containing one 72 bp sequence and deletions within it, we demonstrate unequivocally that the 72 bp element contains sequences which are essential for SV40 early gene expression. In agreement with our previous suggestion (6) the GC-rich block present within the 72 bp sequence appears to be important for this function, since deletions on both sides of the EcoRII site result in a marked decrease of early gene expression (Fig. 2 and Table 1).

We have reported previously that a SV40 early region deletion mutant(pHS102, Fig. 2 in Ref.6) in which the "promoter" sequences located between the early mRNA startsites and the beginning of the second (II) 72 bp sequence had been deleted, but which still contained the 72 bp repeat and the adjacent sequences up to the HpaII site (position 346, see Fig. 1), expressed the early genes although with a low efficiency (see also the present paper, Table 1). The present results indicate that a DNA fragment which contains most of the 72 bp repeat and only 20 additional bp on their 5' side (from nucleotides 113 to 270) can also "promote" in a cis-acting fashion the expression of the T-antigen when inserted upstream from the SV40 early structural sequences (recombinant pEMPB and pEMPBI, Fig. 3 and Table 1). The effect was particularly noticeable when the DNA was introduced by microinjection where more than 30 % of the cells were T-antigen positive when the 72 bp repeat was present in the recombinant pEMPBI37, whereas no positive cells were ever observed in the absence of the repeat (pEMP).

Although the 72 bp repeat can act in the absence of the 117 bp segment located immediately upstream from the SV40 early mRNA startsites, it is clear that it functions best when closely associated with sequences which are known to belong to promoter regions active *in vitro* or/and *in vivo*. This is obviously the case for SV40, but it is striking that the SV40 sequences can be replaced by "promoter" sequences from the Adenovirus major late transcription unit (19, 32) (Fig. 4 and Table 2) or from the chicken conalbumin gene (19, 28-31) (Fig. 3 and Table 2). These are markedly inefficient in "promoting" T-antigen expression in recombinants lacking the 72 bp repeat. Several observations indicate that it is the promoter sequences of the Adenovirus or conalbumin segments themselves, rather than the mere insertion of DNA fragments with lenghts approximately equal to that which has been removed from the SV40 early region, which are important for the cis-acting effect of the 72 bp repeat (see Result section).However,the 72 bp repeat do not have to be positioned at exactly the same distance from the Adenovirus or conalbumin TATA boxes as they are from the SV40 TATA box to be fully efficient, at least with our assay.

It was surprising that in all of our recombinants, the 72 bp repeat

functions equally efficiently in either orientation with respect to Adenovirus or conalbumin promoter DNA segments and to SV40 early structural sequences. This may reflect the *in vivo* situation. The 72 bp repeat is located between the early and late SV40 genes, which are transcribed in opposite directions. It has been suggested (22) that the 72 bp repeat region could also affect the efficiency of SV40 late gene expression. However, the order of the three DNA segments is important to achieve efficient expression : to be fully active (as judged by our transfection assay) the 72 bp repeat has to be located upstream (with respect to transcription of the SV40 early genes) from the inserted promoter segment (compare recombinants pTBCT and pTBICT with pTCTB and pTCTBI, and with pEMP and pEMPBI - Figs. 3 and 4, Tables 1 and 2).

It is very striking that, even when the 72 bp repeat is 4 kb away, it can stimulate the expression of the SV40 early structural sequences in the absence (the pEMP series, Fig. 3 and Table 1) and in the presence of the SV40 early (the TBB series, Fig. 2 and Table 1), the Adenovirus major late (the pSVPBA40 and 500 series, Fig. 4 and Table 2), or conalbumin (the pTCTB series, Fig. 3 and Table 2) promoter regions. However, the effect of the 72 bp repeat as assayed after transfection decreased with distance, although for the same distance this decrease was greater for the conalbumin than for the Adenovirus major late promoter recombinants. As the pBR322 sequences are identical in the two cases, this difference must be due to the Adenovirus or conalbumin promoter sequences (see above). It is remarkable that these variations in the "effect at distance" are almost erased when the DNA is introduced by microinjection, most likely because the number of recombinant copies per nucleus is too high to allow the immunofluorescence assay to detect differences in the efficiency of T-antigen expression by the various recombinants.

The insertion of four rather than two 72 bp sequences in tandem, in either orientation, resulted in a more efficient expression (recombinants of the pSVA500 series, Fig. 4 and Table 2). In the SV40 case, a recombinant with one 72 bp sequence (TBO, Fig. 2 and Table 1) appears to be slightly less efficient in early gene expression than the corresponding wild-type recombinant with the 72 bp repeat (pSV1). These observations raise the question whether the basic functional unit is composed of one or two 72 bp sequences and whether adjacent 72 bp sequences are interacting or acting independently.

C) Possible mechanisms for the effect of the 72 bp repeat region.

As our estimate of SV40 early gene expression is based on an assay of T-antigen, there are in principle three basic levels (pre-transcriptional,

transcriptional and post-transcriptional) at which one can imagine a role for the 72 bp repeat. Although we cannot definitely exclude it, it is unlikely that the 72 bp sequence is involved in mechanisms related to transport of the recombinant DNA to the nucleus or its stabilization therein (by any mechanism including its integration in the cellular genome), since 1) recombinants (for instance pEMP and pTCT - Tables 1 and 2) which are fully negative by transfection are also fully negative by microinjection ; 2) it would be difficult with such mechanisms to account for the pattern of the "effect of distance". It is also unlikely that the main role of the 72 bp sequence would be post-transcriptional, resulting in a more efficient processing or/and transport, or in a greater stabilization of SV40 early gene transcripts. Since our experiments demonstrate clearly that the function of the 72 bp repeat is cis-acting, a post-transcriptional mechanism would imply that the 72 bp repeat is transcribed (at least in part) and their transcript linked to the T-antigen coding sequence. It would then be difficult to understand how the same effects could be obtained irrespective of the orientation of the inserted 72 bp box region and why the 72 bp box is not efficient when placed between the conalbumin promoter sequence and the structural T-antigen sequence (see above). It has been reported (33) that nuclear matrix preparations of transformed 3T3 cells are preferentially enriched in SV40 DNA sequences. Since the nuclear matrix could be involved in RNA processing and/or transport (for Refs. see 33), one could imagine that the role of the 72 bp sequence would be to provide a DNA attachment site to the matrix. However, this possibility would not readily account for the "effect of distance". Therefore, although we cannot exclude that the 72 bp sequence could act indirectly at the post-transcriptional level, we consider this hypothesis improbable. A more reasonable interpretation of all of our data is that the 72 bp repeat acts in some way at the transcriptional level and possesses some promoter-like activity. In fact, the same conclusion was derived from an extensive study of variations in the RNA initiation sites on the SV40 early genes of a series of deletion mutants around and within the 72 bp repeat (6).

At the transcriptional level the 72 bp sequence could act indirectly by facilitating the formation of an "active" template or play a more direct role by interacting with some components of the transcription machinery. The chromatin structure of at least a fraction of the SV40 minichromosomes is known to be in a very "open" conformation in the region which encompasses the 72 bp repeat and the origin of replication (for Refs., see 6, 34, 35). Therefore one could speculate that the 72 bp region somehow induces this peculiar chromatin

structure, which in turn allows the transcription machinery to operate. However, experiments recently performed in our laboratory suggest that the 72 bp repeat cannot by itself induce the formation of the nucleosome gap nor the nuclease sensitivity pattern characteristic of the "open" minichromosomes (C.B. Chae, D. Mathis, J. Jongstra, C. Benoist and P. Chambon, in preparation). Alternatively the 72 bp repeat could be involved in other mechanisms known to activate transcription at the template level, for instance they could interact in some way with the topoisomerase machinery controlling DNA supercoiling, which is known to affect gene expression in prokaryotes (36-38). Although such a mechanism cannot be excluded at the present time, it would not readily account for the "effect of distance". The same remark applies to another possible mechanism, in which the 72 bp sequence would direct the DNA template into a specialized nuclear compartment containing all factors required for efficient transcription.

Could the 72 bp repeat interact directly with some of the components of the transcription machinery ? As discussed above, it "promotes" gene expression, although not efficiently, in the absence of any known promoter sequence, which suggests that it could be in itself an element of a promoter region. A plausible hypothesis which could account for all of our observations is that the 72 bp sequence corresponds to a particularly efficient bi-directional entry site for a component of the transcription machinery, for instance and for the sake of simplicity, RNA polymerase B (II). The RNA polymerase would then track the DNA to find an initiation site. This would explain the high efficiency of expression when the 72bp repeat is close to elements known to have promoter functions, like the Adenovirus and conalbumin sequences which have been inserted in some of our recombinants, and the lower efficiency, when the 72 bp repeat is moved away. Since it appears that in the absence of the TATA box region initiation takes place *in vitro* (26) as well as *in vivo* (6) from multiple cryptic startsites, this hypothesis would also explain that gene expression can take place with a low efficiency in the presence of the 72 bp repeat, but in the absence of any known promoter sequence. This possibility is consistent with results from M. From, P. Berg and M. Singer (personal communications) who have also noted the enhancing effect of the 72 bp repeat region and found that in pBR322 recombinants containing the 72 bp region of SV40 inserted at distance from a known promoter or promoter-like region, initiation of transcription can also take place at some sites of the intervening pBR322 sequences. *In vitro* transcription studies and detailed analyses of the RNA synthesized *in vivo* from our recombinants are currently in progress to test

the validity of the bi-directional "super" entry site hypothesis.

However, it should be clear that several of the possible mechanisms that we have discussed could operate concomitantly. It is also important to stress that the effect of the 72 bp sequence does not exhibit striking species specificity, as similar results were obtained in the present study using SV40 permissive monkey CV1 cells or non-permissive mouse 3T3 or L cells (A. Dierich, B.W. and R.E., unpublished observations). In this regard, we note that Capecchi (39) reported that the insertion of a SV40 DNA fragment which contains the 72 bp repeat into a HSV-thymidine kinase gene recombinant enhances the transformation frequency of mouse LMTK$^-$ cells, and W. Schaffner (personal communication) found that the expression of the rabbit β-globin gene is enhanced in HeLa cells when it is inserted in a recombinant vector which contains SV40 sequences which include the 72 bp repeat.

D) Are there functional equivalents of the SV40 72 bp repeats in other viral or cellular genes ?

There is no obvious sequence similarity between the 72 bp repeat sequence in SV40 and the corresponding region in polyoma virus DNA whose general genome organization is similar to that of SV40. However, some sequences located upstream from the polyoma early mRNA startsites play a role analogous to that of the SV40 72 bp repeat (C. Tyndall, G. La Mantia, C. Thacker and R. Kamen, personal communication). The long terminal repeats (LTR) of some retroviruses could contain sequences with a role similar to that of the SV40 72 bp repeat, but again in this case there are no obvious sequence similarities between the SV40 repeat and LTR sequences (H. Varmus, P. Gruss, personal communications).

Whether functional equivalents of the 72 bp sequence are present in cellular genes is not clear. Upstream sequences crucial for gene expression and located at about the same distance from the mRNA startsites as the SV40 72 bp repeat have been characterized in several cellular gene promoter regions, notably in the sea urchin H2A (8) and herpes virus thymidine kinase (40) genes. Whether they share the main properties of the 72 bp repeat is however unknown. Short homologies have been found (41) between the 150-166 GC-rich sequence present in the 72 bp repeat and sequences present in monkey DNA fragments selected by hybridization to a SV40 DNA region including the origin of replication but it is unknown whether these monkey sequences are functionally equivalent. The question is therefore open whether the function exerted by the SV40 72 bp repeat, which has no precedent in prokaryotic cells, is restricted to certain viruses (to divert the transcriptional machinery for their own use) or whether

it corresponds to a widely spread mechanism to control gene expression in eukaryotic cells.

ACKNOWLEDGEMENTS.

We thank Drs. D. Di Maio and D. Nathans for a gift of SV40 mutants cs1096 and pX3, Dr. R. Wilsnack for a gift of antiserum and Drs. C. Benoist, J. Corden and J.M. Jeltsch for gifts of several recombinants. We gratefully acknowledge the help of Dr. C. Benoist for the construction of some recombinants and the helpful advice of Drs. P. Oudet and W. Ansorge for microinjections. We thank K. Dott, J.M. Garnier, M.C. Gesnel, D. Schmitt and C. Wasylyk for expert technical assistance. This work was supported by grants from the INSERM (ATP 72.79.104 and PRC 124.026), the CNRS (ATP 006520/50), the Fondation Française pour la Recherche Médicale, The Association pour le Développement de la Recherche sur le Cancer and the Fondation S. et C. Del Duca. P. Moreau and R. Everett were respectively recepeients of post-doctoral fellowships from the Gouvernement de la province de Québec and from the Royal Society (G.B.). Preliminary accounts of this work have been presented at the Symposium on Promoters : Structure and Function at Monterey, USA (May 1981) and at the IX Congress of the International Society of Developmental Biologists in Basel, Switzerland (September 1981).

REFERENCES.

1. Jacob, F., Ullman, A. and Monod, J. (1964) C.R. Hebd. Séances Acad. Sci. Ser. X 258, 3125.
2. Scaife, J. and Beckwith, J.R. (1966) Cold Spring Harbor Symp. Quant. Biol. 31, 403-408.
3. Rosenberg, M. and Court, D. (1979) Annu. Rev. Genet. 13, 319-353.
4. U. Siebenlist, R. Simpson and W. Gilbert (1980) Cell 20, 269-281.
5. Benoist, C. and Chambon, P. (1980) Proc. Natl. Acad. Sci. USA 77, 3865-3869.
6. Benoist, C. and Chambon, P. (1981) Nature 290, 304-310.
7. Gluzman, Y., Sambrook, J. and Frisque, R. (1980) Proc. Natl. Acad. Sci. USA 77, 3898-3902.
8. Grosschedl, R. and Birnstiel, M. (1980) Proc. Natl. Acad. Sci. USA 77, 7102-7106.
9. Struhl, K. (1981) Proc. Natl. Acad. Sci. USA 78, 4461-4465.
10. Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349-383.
11. Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voorde, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G. and Ysebaert, M. (1978) Nature 273, 113-119.
12. Van Heuverswyn, H. and Fiers, W. (1979). Eur. J. Biochem. 100, 51-60.
13. Tooze, J. (ed.) (1980) DNA Tumor viruses, Cold Spring Harbor Lab., Cold Spring Harbor, New York, p. 61-338.
14. Maxam, A. and Gilbert, W. (1980) Meth. Enzym. 65, 499-580.
15. DiMaio, D. and Nathans, D. (1980) J. Mol. Biol. 140, 129-142.
16. Shortle, D. and Nathans, D. (1978) Proc. Natl. Acad. Sci. USA 75, 2170-2174.
17. Cochet, M., Gannon, F., Hen, R., Maroteaux, L., Perrin, F. and Chambon, P. (1979) Nature 282, 567-574.
18. Jeltsch, J.M. and Chambon, P. (1981) Eur. J. Biochem., in press.
19. Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kédinger, C. and Chambon, P. (1980) Science 209, 1406-1414.
20. Graessmann, M. and Graessmann, A. (1976) Proc. Natl. Acad. Sci. USA 73, 366-370.

21. Ansorge, W. (1981) Exp. Cell Research, in press.
22. Gruss, P., Dhar, R. and Khoury, G. (1981) Proc. Natl. Acad. Sci. USA 78, 943-947.
23. Subramanian, K. and Shenk, T. (1978) Nucleic Acids Res. 5, 3635-3642.
24. Waldeck, W. and Sauer, G. (1976) FEBS Lett. 71, 313-315.
25. Van Heuverswyne, H. and Fiers, W. (1979) Eur. J. Biochem. 100, 51-60.
25. Myers, R.M. and Tjian, R. (1980) Proc. Natl. Acad. Sci. USA 77, 6491-6495.
26. Mathis, D.J. and Chambon, P. (1981) Nature 290, 310-315.
27. Ghosh, P.K., Lebowitz, P., Frisque, R.J. and Gluzman, Y. (1981) Proc. Natl. Acad. Sci. USA 78, 100-104.
28. Wasylyk, B., Kédinger, C., Corden, J., Brison, O. and Chambon, P. (1980) Nature 285, 388-390.
29. Wasylyk, B., Derbyshire, R., Guy, A., Molko, D., Roget, A., Téoule, R. and Chambon, P. (1980) Proc. Natl. Acad. Sci. USA 77, 7024-7028.
30. Wasylyk, B. and Chambon, P. (1981) Nucleic Acids Res. 9, 1813-1824.
31. Grosschedl, R., Wasylyk, B., Chambon, P. and Birnstiel, M. (1981) Nature, in press.
32. Sassone-Corsi, P., Corden, J., Kédinger, C. and Chambon, P. (1981) Nucleic Acids Res. 9, 3941-3958.
33. Nelkin, B.D., Pardoll, D.M. and Vogelstein, B. (1980) Nucleic Acids Res. 8, 5623-5633.
34. Saragosti, S., Moyne, G. and Yaniv, M. (1980) Cell 20, 65-73.
35. Jakobovits, E.B., Bratosin, S. and Aloni, Y. (1980) Nature 285, 262-265.
36. Liu, L.F. and Miller, K.G. (1981) Proc. Natl. Acad. Sci. USA 78, 3487-3491.
37. Sternglanz, R., DiNardo, S., Voelkel, K.A., Nishimura, Y., Hirota, Y., Becherer, K., Zumstein, L., Wang, J.C. (1981) Proc. Natl. Acad. Sci. USA 78, 2747-2751.
38. Hsieh, T. and Brutlag, D. (1980) Cell 21, 115-125.
39. Capecchi, M.R. (1980) Cell 22, 479-488.
40. S.L. McKnight, E.R. Gavis, R. Kingsbury and R. Axel (1981) 25, 385-398.
41. Queen, C., Lord, S.T., McCutchan, T.F. and Singer, M.F. (1981) Cell. and Molec. Biol., in press.