Volume 11 Number 20 1983 Nucleic Acids Research

Plasmids for the cloning and expression of full-length double-stranded cDNAs under control of the SV40 early or late gene promoter

R.Breathnach and B.A.Harris

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

Received 30 August 1983; Accepted 27 September 1983

ABSTRACT

Okayama and Berg (1) have recently described a technique for the high efficiency cloning of full-length dscDNAs. We have constructed eukaryotic expression vectors compatible both with this technique (and with classical techniques) for dscDNA cloning. The vectors are such that recombinants obtained contain dscDNAs in the correct orientation downstream from a block of sequence comprising either the SV40 early or late gene promoter linked to a pair of splice sites from a rabbit β-globin gene. A sequence encoding an SV40 polyadenylation site follows the dscDNA. We have used our vectors to make a library from chicken oviduct polyA(+) RNA using the Okayama and Berg technique. Ovalbumin recombinants occur in the library at the expected frequency and a high proportion contain full length copies of the ovalbumin mRNA. However, a similar result was not obtained for conalbumin recombinants. When recombinants are introduced into eukaryotic cells by either calcium phosphate coprecipitation or protoplast fusion, expression of chicken ovalbumin or conalbumin may be detected by indirect immunofluorescence. Under optimal conditions (use of SV40 late promoter and cos 7 cells) ovalbumin protein could be detected when the ovalbumin recombinant was present in only 2 % of the protoplasts used for fusion. This suggests that colony banks obtained using our vectors could be screened in batches of 50 by protoplast fusion followed by a search for expression of a given protein using indirect immunofluorescence.

INTRODUCTION

A number of vectors are available for the expression of protein coding sequences in eukaryotic cells (2-5). We have designed a vector (pKCR) using the SV40 early gene promoter and splicing and polyadenylation signals from a rabbit β-globin gene (3). Our vector has proved useful for the expression of a variety of double-stranded cDNA (dscDNA) fragments (6-8). However, in these cases the dscDNA clones were first obtained in pBR322, and only once suitable clones were identified could fragments thereof be introduced into the expression vector. Employing an expression vector in this way is undoubtedly useful. However, the full potential of expression vectors will only be realised when dscDNA libraries can be constructed directly in them, and desired clones identified by their ability to express a part

© IRL Press Limited, Oxford, England.

icular protein in eukaryotic cells.

To achieve this aim it is necessary to have a vector in which full-length dscDNAs may be cloned in high yield in the correct orientation relative to eukaryotic promoter and processing signals. It is also necessary to have an efficient method for the introduction of such recombinants into eukaryotic cells. Okayama and Berg (1) have described a technique for high efficiency cloning of full-length dscDNAs. Briefly, a T-tailed primer fragment is used to prime cDNA synthesis from polyA(+) RNA, the cDNA-mRNA hybrid is then tailed with dC residues and annealed to a linker fragment carrying a tail of dG residues (see Fig. 3). Correct orientation of the dscDNA relative to promoter elements can thus be assured by placing these latter in the linker fragment. In contrast, using classical techniques, the annealing of dC-tailed dscDNAs to dG-tailed vector will result in 50 % of the dscDNAs in the wrong orientation relative to vector-carried promoter elements. A further advantage of the Okayama and Berg technique is that the nuclease S1 is not used, leading with this technique [as with other techniques (9)] to a greater retention of sequences representing the 5'-ends of mRNAs. Rassoulzadegan et al. (10) have modified the protoplast fusion technique pioneered by Schaffner (11) to permit very high frequency introduction of recombinant plasmids into eukaryotic cells, without the need for extensive preliminary purification of plasmid DNA. High efficiency gene transfer is also possible using a modified calcium phosphate coprecipitation technique (12).

We describe in this paper expression vectors which may be used in conjunction with the Okayama and Berg technique for dscDNA cloning. Recombinants obtained contain dscDNAs under control of either the SV40 early or late region promoter. We have obtained full length dscDNAs encoding chicken ovalbumin and conalbumin using these vectors, and these dscDNAs are expressed on introduction into eukaryotic cells. When the SV40 late promoter is used, detection of ovalbumin expression remains possible when protoplasts containing the ovalbumin dscDNA recombinant make up only 2 % of the protoplast population used for fusion to cos 7 cells, permissive for replication of the recombinant (13). This result suggests that colony libraries could be screened by expression in batches of 50 provided an antibody against the protein of interest is available. While this work was being completed, a report appeared describing the cloning of a full length hypoxanthine phosphoribosyltransferase dscDNA (14, 15) using the Okayama and Berg technique with an expression vector similar to ours.

MATERIALS AND METHODS

Construction of plasmids

pSVE (Fig. 1)

pKCR2 was obtained from pKCR (3) by eliminating one EcoRI site (shown bracketed) by partial EcoRI digestion, DNA polymerase I repair and religation. pKCR2 was cut by EcoRI, repaired with DNA polymerase I, recut with SalI and ligated to a PvuII-SalI digest of p2.40V (16) to generate pOV1. In p2.40V, the nucleotide sequence around the PvuII site reads 5'...CAGCTGCAGA-TCAAGCCAGAGAGCTC...3'. The PvuII (CAGCTG), PstI (CTGCAG) and SacI (GAGCTC) sites shown on the p2.40V map are underlined. Joining the repaired EcoRI site of pKCR2 to this PvuII site regenerates an EcoRI site (GAATTC) to give the sequence 5'...GAATTCTGCAGATCAAGCCAGAGAGCTC...3'. The underlined PstI site of pOV1 was removed by partial PstI digestion followed by DNA polymerase I treatment and religation to obtain pOV2. The sequence downstream from the EcoRI site now reads 5'...GAATTCGATCAAGCCAGAGAGCTC...3'. We also obtained a rearranged recombinant having the sequence 5'... GAATGATCAAGCC-AGAGAGCTC...3', where a BclI site (TGATCA, underlined) has been generated and the EcoRI site lost. pHB contains a HindIII-BamHI fragment of SV40 DNA [nucleotides 3476-2533, BBB system (17)] between the corresponding sites of pBR322. The BamHI site has been removed by treatment with DNA polymerase I (and therefore shown bracketed). pSVE was made by ligation of HindIII + SalI digests of pOV2 and pHB. pSVE thus contains the SV40 early region promoter (originally a HpaII-HindIII fragment of SV40 DNA, 346-5171; the HindIII site is occupied by a BamHI linker), a 640 bp BamHI-EcoRI fragment of a rabbit β-globin gene (3), a 150 bp PvuII-HindIII fragment of the chicken ovalbumin gene (16), a HindIII-BamHI fragment of SV40 DNA (3476-2533) containing a polyadenylation site, and the EcoRI-BamHI fragment of pBR322. As for pOV2, two versions of pSVE exist, one with an EcoRI site (pSVEe) and one with a BclI site (pSVEb).

pK14 (Fig. 2)

Plasmid pSV8 was obtained by introducing the large EcoRI-BamHI fragment of SV40 DNA (1782-2533) into the corresponding sites of pBR328 (18). pK14 was generated from this plasmid by introducing a KpnI linker between the HpaI sites at positions 519 and 2666. The HindIII sites of pK14 lie at nucleotides 1708, 1493 and 1046.

pSVL (Fig. 4a)

pSV327 contains a copy of SV40 DNA linearised by HpaII (346) in the ClaI site of pBR327 (18). pSVEH14 was obtained by cutting pSV327 at the

EcoRI site of its pBR327 moiety and the HpaI site at SV40 nucleotide 3733, repairing with DNA polymerase I and religating. Cutting pSVEH14 with KpnI (294) and BamHI (in pBR327 moiety), repairing with DNA polymerase I and religating regenerates in pSVLT46 a BamHI site at the position of the KpnI site in pSVEH14, i.e. n.294. pSVL was obtained by cutting pSVLT46 with AvaI, repairing, and recutting with BamHI followed by ligation to a repaired HindIII digest of pSVEe recut by BamHI. pSVL contains the AvaI-EcoRI fragment of pBR327 (18) a HpaI-KpnI fragment of SV40 DNA (3733-294) carrying the SV40 late region promoter, a 640 bp BamHI-EcoRI fragment of a rabbit β-globin gene (3) and a 150 bp fragment of the chicken ovalbumin gene (16).

pMLV (Fig. 4b)

This plasmid contains a 390 base-pair MboI-KpnI fragment [coordinates 100-490 (19)] of a Murine Moloney leukaemia virus LTR cloned between the repaired EcoRI and BamHI sites of pBR322. The KpnI site of the LTR has thus been converted into a BamHI site. pMLV contains the LTR promoter, and fragments inserted at the BamHI site should be expressed from it.

pSVLT47 (Fig. 4c)

This plasmid was prepared by ligating BamHI + SalI digests of pSVLT46 and pSVEe.

Preparation and use of primer and linker fragments.

For primer preparation (Fig. 2) pK14 was linearised by KpnI digestion and tailed with TTP and terminal deoxynucleotidyl transferase (BRL) using conditions described in Okayama and Berg (1). Tailed material was cut with EcoRI and HindIII, generating a 3.5 kb fragment and several smaller fragments. Both enzymes were used as a large excess of enzyme was needed to obtain complete digestion (excess determined by test digests of small aliquots of tailed material), possibly due to inhibition by the single-stranded tails. The 3.5 kb primer fragment was isolated on a 5-20 % sucrose gradient containing ethidium bromide using an SW41 rotor (30000 rpm, 16h, 20°C) and recovered by ethanol precipitation. Further purification of primer on an oligo dA-cellulose column (Collaborative Research) was as described (1). For linker preparation (Figs. 1 and 3), pSVE was cut with SacI and tailed with dGTP as described (1) before further digestion with PstI to generate fragments of 4.3, 1.8 and 0.45 kb. The 1.8 kb linker fragment was isolated on a 5-20 % sucrose gradient containing ethidium bromide using an SW41 rotor (35000 rpm, 16h, 20°C). For cDNA synthesis (Fig. 3), the primer fragment (1 μg) was annealed to 1 μg of chicken oviduct polyA(+) RNA previously denatured by treatment with 10 mM methylmercury hydroxide (20) and used for

cDNA synthesis as described (1). Tailing of the cDNA-mRNA hybrid with dCTP was done using the reaction conditions of Okayama and Berg (1). The material was then cut with PstI, and annealed and ligated to a two-fold excess of the linker fragment using conditions described previously (1). The mRNA strand was replaced with DNA using RNAse H (PL biochemicals) and DNA polymerase I (Biolabs). Material was transfected into E.coli strain C600 5K rendered competent for DNA uptake by overnight treatment with $CaCl_2$ (21). Ampicillin resistant colonies were selected and screened by the technique of Grunstein and Hogness (22) using nick-translated (23) chicken ovalbumin (24) or conalbumin (25) gene fragments as probes. Recombinants pSVEOV1 and 2 obtained in this manner were sequenced using the technique of Maxam and Gilbert (26) from a SacI site at position 116 on the chicken ovalbumin mRNA sequence (27).

We have found it useful to include the following tests at appropriate steps of the cloning procedure.

1) When tailing pK14 or pSVE, an estimate of tail length is first made by determination of TCA-precipitable counts. A 4 μg aliquot of tailed material is then passed over and eluted from a 100 μl bed volume oligo dA-cellulose (pK14) or oligo dC-cellulose (pSVE) column, and the relative amounts of material in the bound and non-bound fractions estimated by agarose gel electrophoresis. This ensures that conditions are found where a significant percentage of molecules are tailed. It may prove necessary to remove untailed linker fragment after the sucrose gradient step by passage over an oligo dC-cellulose column.

2) When primer fragment is made, it may be tested as follows. A 0.2 μg aliquot is incubated with reverse transcriptase in cDNA synthesis mix with increasing quantities of polyA(+) RNA. Following RNAse A digestion, the material is electrophoresed on a 0.7 % agarose gel. In this way conditions are found where all primer is converted to higher molecular weight materials. A smear will be observed.

3) The PstI digest of the dC-tailed cDNA-mRNA hybrid may be verified by electrophoresing an aliquot on a 0.7 % agarose gel. The smear discussed above should decrease in size and a discrete band at 1.2 kb become visible. Occasionally the PstI digest does not work, perhaps due to residual impurities from the tailing step. Phenol-chloroform extraction and several ethanol precipitations of the material before redigesting with PstI solves this problem. A detailed protocol for the various steps is available on request.

Cells and transfections

Maintenance of HeLa (28), LMTK$^-$ (29), NIH 3T3 (30) and cos 7 (13) cells was as described in the corresponding references. Transfection of cells by the calcium phosphate coprecipitation technique was as described by Cosaro and Pearson (12). Protoplast fusions using E.Coli strain 1106 were carried out according to Rassoulzadegan et al. (10). After 48 h cells were fixed for indirect immunofluorescence with methanol and reacted with rabbit antibodies against chicken ovalbumin or conalbumin (gift of M. Le Meur) and stained with fluorescein-conjugated antibodies using techniques described previously (31). Alternatively, total RNA was isolated by LiCl-Urea precipitation (32) and electrophoresed on a 1.5 % agarose gel containing 6 % formaldehyde (v/v) before transfer to nitrocellulose (33) and hybridization to a nick-translated (23) BamHI-XbaI fragment of pSVEOV1.

RESULTS AND DISCUSSION

Construction of a vector for cloning full-length dscDNAs under control of the SV40 early region promoter.

The expression vector pKCR (3) contains a unique BamHI site into which DNA fragments may be inserted. Such fragments lie downstream from the SV40 early gene promoter and upstream from splice sites of a rabbit β-globin gene and polyadenylation sites from this gene and the SV40 early region [pKCR has the same structure as pKCR2 (Fig. 1), except that it contains an additional EcoRI site]. While this vector may be used for the expression of fragments of dscDNAs, it is not clear whether it is suitable for the expression of full-length dscDNAs. These latter may encode their own polyadenylation signal (34). Introducing full-length dscDNAs into the BamHI site of pKCR might mean placing a polyadenylation signal upstream from the pair of β-globin gene splice sites. This would result in primary transcripts without splice sites, which in some cases do not give rise to stable cytoplasmic mRNAs (35). It would be safer to introduce full-length dscDNAs for expression downstream from a pair of splice sites. In addition, as the dscDNAs own polyadenylation signal may function inefficiently, it would be best to place a known functional polyadenylation signal downstream from it. Full-length dscDNAs for expression could therefore be placed in the environment they would enjoy if introduced into the unique EcoRI site of pKCR2 [Fig. 1; we have shown that the BamHI-EcoRI fragment of the β-globin gene contains sufficient information for splicing (B.H., unpublished). They would lie downstream from a promoter and a pair of splice sites but upstream from a

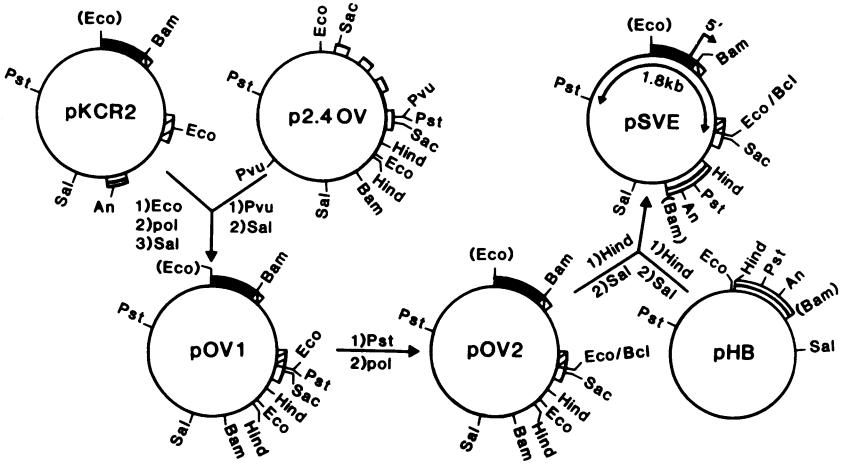

Figure 1 Construction of the linker plasmid pSVE.
Details are given in Materials and Methods. Abbreviations are as follows : Eco = EcoRI. Bam = BamHI. Pst = PstI. Sal = SalI. Pvu = PvuII. Sac = SacI. Hind = HindIII. Bcl = BclI. pol = DNA Polymerase I. The filled in box is the SV40 early gene promoter, the boxes with diagonal lines β-globin exons, the box with a horizontal line an SV40 fragment carrying the early messenger polyadenylation site (An) and the open box a small part of a chicken ovalbumin gene exon. The arrow 5' on the plasmid pSVE indicates the approximate start sites of the SV40 promoted messengers. As discussed in the text, two versions of pSVE exist. One has an EcoRI site, the other a BclI site. This is shown as Eco/Bcl on the pSVE map. Some sites have been bracketed to indicate that they no longer exist. They show the derivation of certain fragments. Drawings are not to scale.

polyadenylation site. Such a construction would generate spliced polyadenylated mRNAs where the first AUG codon would be the initiation codon carried by the dscDNA.

We wished to use the technique of Okayama and Berg (1) to clone full-length dscDNAs in the environment outlined above. In this technique linker and primer fragments become joined to the respective 5'- and 3'-ends of the dscDNA. It was thus logical to prepare a linker fragment carrying the SV40 early gene promoter and β-globin splice sites, and a primer fragment carrying an SV40 polyadenylation site (see Fig. 3). In addition, the linker and primer fragments become joined to the cDNA by annealing homopolymer tails generated by terminal deoxynucleotidyl transferase, making it desirable to have sites for enzymes which generate efficiently tailed 3'-protruding termini at the positions where tails are to be added. Also, it would be convenient to be able to use any plasmids we made for cloning dscDNAs generated by

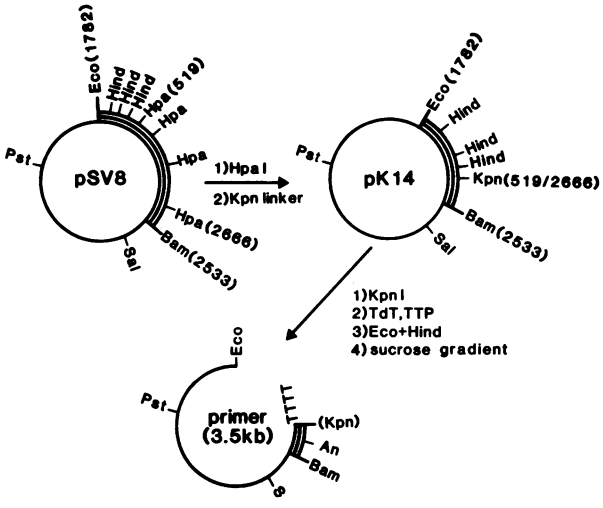

Figure 2 Construction of the primer plasmid pK14.
Details are given in Materials and Methods. Abbreviations are as in Fig. 1 with the additions : Hpa = HpaI, Kpn = KpnI, TdT = terminal deoxynucleotidyl transferase. The box with a horizontal line in the primer fragment represents the SV40 fragment containing the early messenger polyadenylation site (represented by An). The nucleotide positions of some sites on the SV40 genome (17) are shown in brackets.

classical techniques (oligo dT priming, second-strand synthesis, S1 nuclease treatment). Bearing these considerations in mind, we designed plasmids pSVE and pK14 from which suitable linker and primer fragments may be made.

The construction of pSVE is described in Fig. 1 and Materials and Methods. The aim was to introduce a SacI site (for 3'-protruding termini) a few base pairs downstream from the EcoRI site of pKCR2. This was achieved by addition of a fragment of a chicken ovalbumin gene taken from p2.40V to generate pOV1. One of the PstI sites of pOV1 was destroyed to make pOV2. Two versions were isolated, either with an EcoRI site or a BclI site (the latter site resulting from misrepair, see Materials and Methods). Finally, pOV2 was converted into pSVE (also in two versions, either with an EcoRI or BclI site, pSVEe or pSVEb respectively) in order to obtain a plasmid from which the 1.8 kb PstI-SacI linker fragment (arrowed) could be isolated by sucrose gradient ultracentrifugation. The construction of pK14 is shown in Fig. 2. A KpnI linker (for 3'-protruding termini) was introduced into plasmid pSV8 at the HpaI site (SV40 nucleotide number 2666) lying just upstream from the early messenger polyadenylation site, which is thus sandwiched

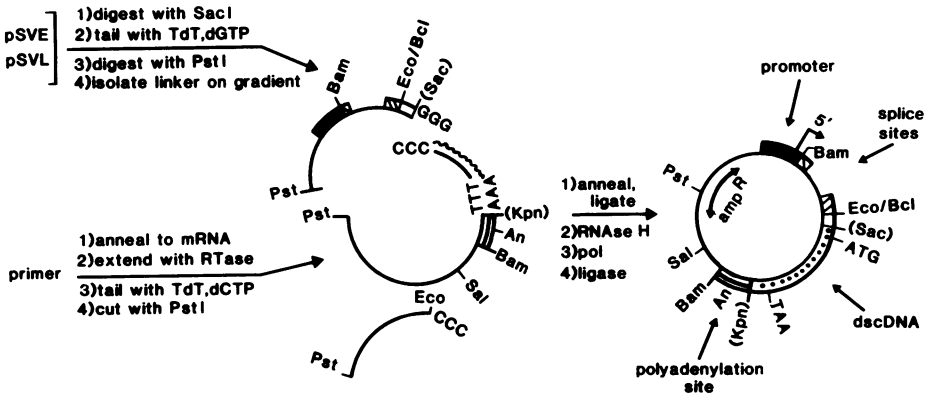

Figure 3 Structure of dscDNA recombinants obtained using the linker fragment from pSVE (or pSVL) and the primer fragment from pK14. Abbreviations are as in Figs. 1 and 2 with the addition RTase = reverse transcriptase. Symbols are as in Figs. 1 and 2. The mRNA is shown as a wavy line, and the dscDNA as a box with a dotted line. The bla gene of the plasmid pBR327 is shown as amp^R.

between the KpnI site and the BamHI site (2533) of pK14.

The generation of linker and primer fragments from pSVE (we use the term pSVE when either pSVEe or pSVEb could be used) and pK14 is shown in Figs. 2 and 3. For the linker fragment, pSVE is cut by SacI, tails of dG residues are added, and the material then cut by PstI. The 1.8 kb linker fragment is prepared by sucrose gradient ultracentrifugation. Two versions of linker fragment may be made, either with an EcoRI or a BclI site depending on whether the linker plasmid used was pSVEe or pSVEb. For the primer fragment, pK14 is cut by KpnI, tails of T residues are added and the material cut by EcoRI and HindIII. The 3.5 kb primer fragment is then isolated on a sucrose gradient.

We outline in Fig. 3 how these fragments may be used with the Okayama and Berg technique for dscDNA cloning. The resulting recombinants carry the dscDNA downstream from the SV40 early gene promoter and β-globin splice sites, but upstream from an SV40 early messenger polyadenylation site. The bacterial plasmid moiety of the recombinants is derived from pBR327 (18). Recombinants thus lack the sequences known to inhibit replication in eukaryotic systems (36). The dscDNA insert may be excised together with the SV40 polyadenylation site using the enzymes EcoRI or BclI (depending whether linker is prepared from pSVEe or pSVEb) and BamHI. Alternatively, the unit splice sites/dscDNA/polyadenylation site may be excised with BamHI, should one wish to place it under the control of a different promoter. The plasmid

pSVE may also be used for cloning dscDNAs generated by classical techniques. pSVE may be opened with SacI, tailed with dG residues and then annealed to dC-tailed dscDNAs. Alternatively, dscDNAs carrying EcoRI ends resulting from addition of linkers may be placed in the unique EcoRI site of pSVE. In either case the dscDNA will be in the environment required for expression in eukaryotic cells, provided it is in the correct orientation relative to the promoter.

Preparation of a chicken oviduct dscDNA library.

We wished to verify that our primer and linker fragments could be used for dscDNA cloning. The linker fragment derived from pSVE (in this case the plasmid pSVEb containing the BclI site was used) and the primer fragment from pK14 were used to prepare a dscDNA library from chicken oviduct mRNA as outlined in Fig. 3. We used 1 μg of polyA(+) RNA for 1 μg of primer fragment. This represents an ~ 5-fold molar excess. An excess of mRNA was preferred as any unused primer will be efficiently converted into an "empty" plasmid. We obtained 20,000 ampicillin resistant colonies under conditions where pBR322 gave 10^7 colonies/μg. About 80 % of these colonies carried plasmids harbouring inserts of the expected structure (Fig. 3). The remaining 20 % were either empty or rearranged (In further experiments with polyA(+) RNA from other sources, we have obtained 20-40,000 colonies/μg under our conditions, of which ~ 70 % contained inserts. Ocasionally, much lower efficiencies were obtained (~ 2000 colonies/μg) for no clear reason). It is difficult to compare our figures to these reported by Okayama and Berg (10^5 colonies/μg) as these authors use a smaller excess of polyA(+) RNA over primer, and do not indicate the efficiency of their transfection procedure). About 500 colonies were analysed by filter hybridisation techniques for the presence of ovalbumin or conalbumin coding sequences. Ovalbumin recombinants were detected at a frequency of ~ 20 %, while only 5 conalbumin recombinants were detected (1 %). From the known abundance of ovalbumin and conalbumin mRNAs, the expected figures are 25 % and 5 % respectively. Of 24 ovalbumin recombinants investigated, 15 gave restriction enzyme digestion patterns suggesting the presence of a full-length dscDNA copy. Two such recombinants (pSVEOV1, 2) were sequenced around the 5'-end of the cDNA by the Maxam and Gilbert (26) technique and shown to contain dscDNA copies of ovalbumin mRNA starting with the first nucleotide of the ovmRNA (27). The two recombinants differed only in the length of the GC tails (22 or 24). Of the conalbumin recombinants, only one was full length (pSVECon1), the remainder being around half-size. This result was expected, as the Okayama and

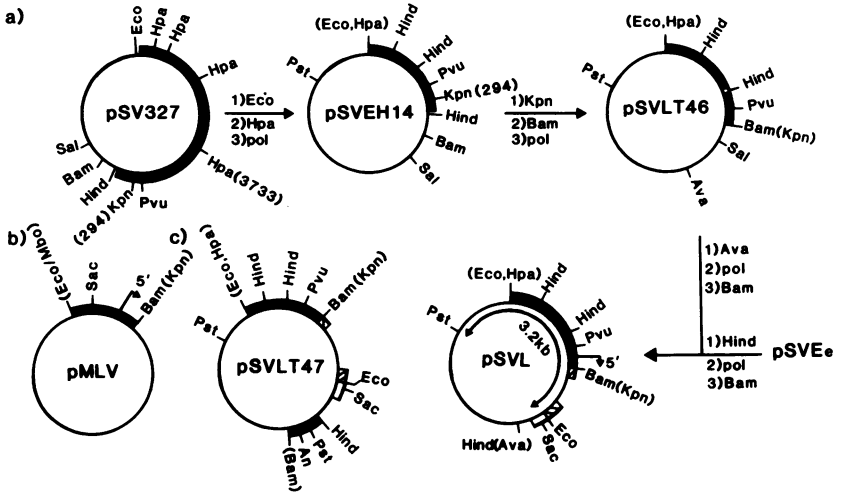

Figure 4

a, c) Construction of the linker plasmid pSVL, and plasmid pSVLT47. Details are given in Materials and Methods. Symbols and abbreviations are as in Figs. 1 and 2, with the addition Ava = AvaI. The filled-in box now represents the SV40 late gene promoter, the arrow 5' indicating the region of initiation of transcription of the SV40 promoted messengers (see text also). b) Structure of plasmid pMLV, containing a 390 bp MboI-KpnI fragment (filled-in box) of a Mo-MuLV LTR cloned in the repaired EcoRI and BamHI sites of pBR322. The arrow 5' indicates the start site of the LTR promoted messengers. Mbo = MboI. Other abbreviations are as in previous figures.

Berg technique is designed to reduce the efficiency of cloning non-full-length dscDNA copies, and it is notoriously difficult to obtain long reverse transcripts of conalbumin mRNA (20).

Expression of chicken ovalbumin and conalbumin from dscDNA recombinants.

We wished ultimately to be able to detect expression of a given protein when the corresponding dscDNA recombinant represented only a small fraction of the recombinants introduced into eukaryotic cells. It was thus important to optimize the expression of such dscDNAs; we chose to do this using as a model system the plasmids pSVEOV1 and 2 described above. We wished to introduce these recombinants into a variety of cell types using different transfection techniques and compare expression using indirect immunofluorescence. However, we decided first to construct equivalents of pSVEOV1 in which the ovalbumin dscDNA is under the control of different promoters, to whit the SV40 late gene promoter and a Moloney Murine leukaemia virus (Mo-MuLV) LTR promoter. This was achieved by excising the BamHI fragment of pSVEOV1 containing the splice sites, dscDNA and polyadenylation sites (Fig. 3) and re-

Table 1 : Transient expression of ovalbumin or conalbumin after introduction of expression plasmids into various cell lines as detected by indirect immunofluorescence. Numbers in brackets indicate percent T-antigen immunofluorescent cells obtained in these lines using pSV1, a plasmid expressing the SV40 early region [for plasmid and staining technique, see ref. (31)].

Plasmid	Protein encoded	Promoter	Host Cell	Transfection technique	Immunofluorescent cells (%)
pSVEOV1,2	ovalbumin	SV40 early	Hela	calcium phosphate	4-8 (80)
"	"	"	LMTK$^-$	"	0.05 (9)
"	"	"	NIH 3T3	"	0.1 (0.1)
pMLVOV1	"	Mo-MuLV LTR	"	"	0.1
"	"	"	Hela	"	0.1
pSVEOV1	"	SV40 early	cos 7	"	10
pSVLOV1	"	SV40 late	cos 7	"	25
"	"	"	"	fusion	25
pSVECon1	conalbumin	SV40 early	Hela	calcium phosphate	4
"	"	"	cos 7	"	10

Table 2 : Effect of dilution on transient expression of ovalbumin from expression plasmids in Hela or cos-7 cells. pSVEOV1 was diluted with increasing amounts of pBR322. pSVLOV1 was diluted with increasing amounts of pSVLH2, a plasmid of similar structure designed to express a murine histocompatibility antigen.

Plasmid	Promoter	Host Cell	Transfection technique	Dilution factor	% immunofluorescent cells
pSVEOV1	SV40 early	Hela	calcium phosphate	1	8
"	"	"	"	10	0
pSVLOV1	SV40 late	Cos 7	fusion	1	25
"	"	"	"	10	1
"	"	"	"	50	0.1

cloning it in the BamHI site of either pSVLT46 or pMLV (Fig. 4; details of these plasmids are given in Materials and Methods), generating plasmids pSVLOV1 and pMLVOV1. With pSVEOV1, these plasmids make a series where exactly the same dscDNA is under the control of three different promoters, allowing experiments to be undertaken to decide which promoter gives optimal expression. These three plasmids, together with pSVECon1, were introduced into cells either by calcium phosphate coprecipitation or the protoplast fusion technique. Expression of ovalbumin or conalbumin was detected by indirect immunofluorescence. The results are shown in Table 1. Plasmids pSVEOV1, pSVLOV1, pMLVOV1 and pSVECon1 all expressed the appropriate protein in the different eukaryotic cell lines tested. However, a number of observations may be made relative to the levels of expression observed.

1) It is clear that, where tested, the calcium phosphate and protoplast fusion techniques gave similar results.

2) Expression of ovalbumin in mouse cells is relatively inefficient, whether the promoter is the SV40 early gene promoter or the Mo-MuLV LTR. The two promoters give about the same level of immunofluorescence. This does not appear to be an effect linked to the recombinants used, as the plasmid pSV1 expressing the SV40 early region also gives a low yield of immunofluorescent cells in the mouse cell lines relative to Hela cells. A low efficiency of transfection of the mouse cells is probably responsible. Expression of ovalbumin in Hela cells is however significantly more efficient with the SV40 early region promoter compared to the Mo-MuLV promoter. This observation is consistent with recent results on the tissue specificity of activator sequences (37). The dscDNA transcripts obtained in Hela cells treated with the recombinants pMLVOV1 and pSVEOV1 were investigated by RNA blotting (Fig. 5). Both pSVEOV1 and pMLVOV1 produce transcripts (~ 2600 nucleotides) of a size consistent with that expected assuming that the β-globin intron transcript is removed by splicing and that polyadenylation occurs at the SV40 polyadenylation site. In this case the putative polyadenylation site encoded by the dscDNA itself appears not to be used. Little is known about the sequence requirements for polyadenylation; indeed, only one system, the SV40 early messenger polyadenylation site, has been studied in any detail (34). Deletions removing DNA sequences downstream from the AATAAA motif did not prevent polyadenylation of early mRNAs taking place, suggesting that sufficient information for polyadenylation resides in the motif and upstream sequences. These sequences are present in pSVEOV1, however, and do not cause polyadenylation to occur. Possibly the tract of A

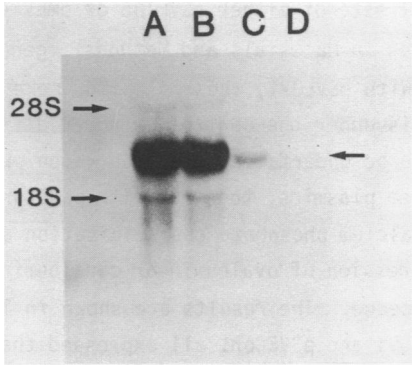

Figure 5 Detection of RNA transcripts from pSVEOV1 and pMLVOV1 after introduction into Hela cells by the calcium phosphate coprecipitation technique. Cytoplasmic RNA was prepared and electrophoresed on a formaldehyde agarose gel before transfer to nitrocellulose and hybridisation with nick-translated pSVEov1. Lanes A, B = 20 and 10 μg of RNA from cells treated with pSVE ov1. Lanes C, D = 20 and 10 μg of RNA from cells treated with pMLVOV1. Migration of 28S and 18S rRNAs are shown by arrows.

residues derived from the polyA tail of the original mRNA which already lies in an appropriate position downstream from the AATAAA sequence in pSVEOV1 inhibits polyadenylation of a transcript of this region. Alternatively, polyadenylation may in some cases by species specific.

3) Expression is most efficient in cos 7 cells using either the SV40 early or late gene promoter, greatest expression being obtained with the late gene promoter (Fig. 6). The cos 7 cells should support replication of pSVEOV1 and pSVLOV1 as both these plasmids carry SV40 origins of replication and do not have those sequences known to inhibit replication in eukaryotic cells (36).

4) Expression of pSVEOV1 in Hela cells, where replication is not expected to take place, is somewhat lower than in cos 7 cells.

We decided to investigate the combinations pSVEOV1/Hela cells and pSVLOV1/cos cells to determine whether ovalbumin expression could be detected when ovalbumin dscDNA recombinants made up only a fraction of the recombinants introduced into the cells. The results are shown in Table 2. Ten-fold dilution of pSVEOV1 before application to Hela cells by the calcium phosphate technique renders detection of ovalbumin impossible. This is not the case for pSVLOV1 introduced by fusion in cos 7 cells, however. Even at fifty-fold dilution, immunofluorescent cells may be observed. In this case around 15 clearly positive cells are reproducibly seen per 1 cm^2 area prep-

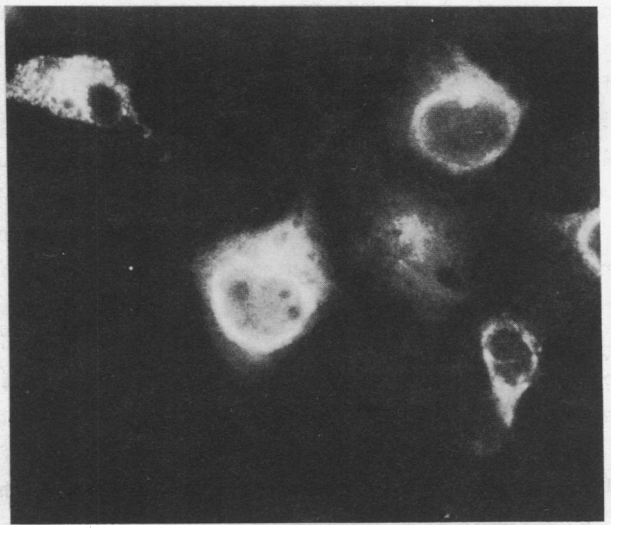

Figure 6 Immunofluorescent detection of ovalbumin produced after fusion of protoplasts carrying pSVEov1 to cos 7 cells.

ared for immunofluorescence.

The efficient expression of the full-length ovalbumin dscDNA in pSVLOV1 suggested that it would be worthwhile to create a plasmid from which a linker fragment carrying the SV40 late promoter and origin of replication joined to the pair of β-globin gene splice sites could be prepared. The construction of such a plasmid (pSVL) is shown in Fig. 4. A linker fragment may be made from pSVL using the method described for pSVE: cutting with SacI, tailing with dG residues, re-cutting with PstI and isolation of a 3.2 kb linker fragment on a sucrose gradient. We have also constructed a plasmid (pSVLT47) allowing dscDNAs prepared by classical techniques to be cloned under control of the SV40 late promoter. Thus dC-tailed dscDNAs may be cloned in the dG-tailed SacI site of pSVLT47, or dscDNAs carrying EcoRI cohesive ends may be placed in the EcoRI site of pSVLT47. It is worth noting that the plasmids pSVLT46 and 47 lack the sequences corresponding to the start sites of the major late messenger (position 325). They do however contain sequences corresponding to the majority of the remaining late messenger start sites.

CONCLUSION

We have described plasmids suitable for generating linker and primer

fragments which may be used in conjunction with the technique of Okayama and Berg (1) for the cloning of full-length dscDNAs. The plasmids may also be used for cloning dscDNAs by the classical homopolymer or linker addition techniques. In the recombinants obtained, the dscDNA lies downstream from either the SV40 early or late gene promoter and a pair of β-globin gene splice sites, and upstream from an SV40 polyadenylation site. Judging from our results with chicken ovalbumin and conalbumin dscDNAs, a dscDNA cloned in this manner will be expressed when introduced into a range of eukaryotic cell lines. Furthermore, our results suggest that colony libraries made using the SV40 late gene promoter linker may be screened by fusing protoplasts prepared from batches of up to 50 colonies at a time to cos 7 cells (13) and screening for expression of a given protein by indirect immunofluorescence. It should also be possible to use the same strategy for identifying dscDNA clones encoding proteins expressing a function which may be selected for. We estimate that around 2000 colonies could be screened per week with ease.

The success of this technique for detecting a particular dscDNA clone clearly depends on obtaining a full-length dscDNA using the vectors described. While this has proved easy for chicken ovalbumin and for some other examples [cloning of cDNA for an oestrogen-regulated breast tumor mRNA (38), and a dihydrofolate reductase mRNA (R.B., unpublished)], full-length chicken conalbumin dscDNAs were obtained at a frequency significantly lower than expected. Other examples of full-length dscDNAs hard to obtain are discussed by Okayama and Berg (1). A different problem was encountered when we tried to clone a full-length dscDNA for the SV40 large T antigen mRNA from cos 7 cells (R.B., unpublished). Screening of a library made using the pSVE linker by filter hybridization detected positive clones at a frequency of around 0.05 %. However, in every case the T antigen coding sequence had undergone extensive rearrangement. In one case examined in detail at least two inversions had occurred as well as deletions.

A number of vectors designed to permit expression of dscDNAs in bacteria are available (39, 40), and have been used to identify dscDNA clones by expression in bacteria. This technique is clearly less arduous to use than cloning by expression in eukaryotes. Nevertheless, it requires that material recognised by the antibody used be produced in the bacterial cells, and may not meet with success if the protein product is normally extensively modified. Such events would be more likely to be carried out correctly in

eukaryotic cells. Cloning by expression in eukaryotic cells may thus prove a useful complement to other available techniques.

ACKNOWLEDGEMENTS.

We thank P. Chambon for useful discussions, W. Schaffner and F. Cuzin for help with gene transfer techniques, and C. Werlé, E. Badzinski and B. Boulay for help in preparation of the manuscript. We gratefully acknowledge the excellent technical assistance of M.-C. Gesnel, M. Acker and B. Augsburger. This work was supported by grants from CNRS (ATP 3582), INSERM (PRC 124025) and MIR (83V0092). B.A. Harris was a postdoctoral fellow of the American Cancer Society.

REFERENCES

1) Okayama, H. and Berg, P. (1982) Mol. Cell. Biol. 2, 161-170.
2) Mulligan, R.C. and Berg, P. (1980) Science 209, 1422-1427.
3) O'Hare, K., Benoist, C. and Breathnach, R. (1981) Proc. Natl. Acad. Sci. USA 78, 1527-1531.
4) Gorman, C.M., Merlino, G.T., Willingham, M.C., Pastan, I. and Howard, B. (1982) Proc. Natl. Acad. Sci. USA 79, 6777-6781.
5) Gheysen, D. and Fiers, W. (1982) Journal of Molec. and Applied Genetics 1, 385-394.
6) Shaw, G.D., Boll, W., Taira, H., Mantei, N., Lengyel, P. and Weissmann, C. (1983) Nucleic Acids Res. 11, 555-573.
7) Taniguchi, T., Matsui, H., Fujita, T., Takaoka, C., Kashima, N., Yoshimoto, R. and Hamuro, J. (1982) Nature 302, 305-310.
8) Breathnach, R., Benoist, C. and O'Hare, K. (1982) Proceedings of the 9th Workshop Conference Hoechst "Genes and Tumor Genes", Raven Press, New York, Eds. Winnacker, E.L. and Schone, H.H., pp. 145-153.
9) Land, H., Grez, M., Hauser, H., Lindenmaier, W. and Schutz, G. (1981) Nucleic Acids Res. 9, 2251-2266.
10) Rassoulzadegan, M., Binetruy, B. and Cuzin, F. (1982) Nature 295, 257-259.
11) Schaffner, W. (1980) Proc. Natl. Acad. Sci. USA 77, 2163-2167.
12) Cosaro, C.M. and Pearson, M.L. (1981) Genetics 7, 603-616.
13) Gluzman, Y. (1981) Cell 23, 175-182.
14) Okayama, H. and Berg, P. (1983) Molec. Cell Biol. 3, 280-289.
15) Jolly, D.J., Okayama, H., Berg, P., Esty, A.C., Filpula, D., Bohlen, P., Johnson, G.G., Shively, J.E., Hunkapillar, T. and Friedmann, T. (1983) Proc. Natl. Acad. Sci. USA 80, 477-481.
16) Benoist, C., O'Hare, K., Breathnach, R. and Chambon, P. (1980) Nucleic Acids Res. 8, 127-142.
17) Tooze, J. (ed) (1982) DNA Tumor Viruses, Cold Spring Harbor Lab., Cold Spring Harbor, New York.
18) Soberon, X., Covarrubias, L. and Bolivar, F. (1980) Gene 9, 287-305.
19) Van Beveren, C., Rands, E., Chattopadhyay, S.K., Lowy, D.R. and Verma, I.M. (1982) J. Virol. 41, 542-556.
20) Buell, G.N., Wickens, M.P., Payvar, F. and Schimke, R.T. (1978) J. Biol. Chem. 253, 2471-2482.
21) Dagert, M. and Ehrlich, S.D. (1979) Gene 6, 23-28.
22) Grunstein, M. and Hogness, D. (1975) Proc. Natl. Acad. Sci. USA 72,

3961-3965.
23) Breathnach, R., Mandel, J.L. and Chambon, P. (1977) Nature 270, 314-319.
24) Humphries, P., Cochet, M., Krust, A., Gerlinger, P., Kourilsky, P. and Chambon, P. (1977) Nucleic Acids Res. 4, 2389-2406.
25) Perrin, F., Cochet, M., Gerlinger, P., Cami, B., Lepennec, J.P. and Chambon, P. (1979) Nucleic Acids Res. 6, 2731-2748.
26) Maxam, A. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564
27) O'Hare, K., Breathnach, R., Benoist, C. and Chambon, P. (1979) Nucleic Acids Res. 7, 321-334.
28) Banerji, J., Rusconi, S. and Schaffner, W. (1981) Cell 27, 299-308.
29) Breathnach, R., Mantei, N. and Chambon, P. (1980) Proc. Natl. Acad. Sci. USA 77, 740-744.
30) Perucho, M., Goldfarb, M., Shimizu, K., Lama, C., Fogh, J. and Wigler, M. (1981) Cell 27, 467-476.
31) Benoist, C. and Chambon, P. (1981) Nature 290, 304-310.
32) Auffray, C. and Rougeon, F. (1980) Eur. J. Biochem. 107, 303-314.
33) Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
34) Fitzgerald, M. and Shenk, T. (1981) Cell 24, 251-260.
35) Breathnach, R. and Chambon, P. (1981) Ann. Rev. Biochem. 50, 349-383.
36) Lusky, M. and Botchan, M. (1981) Nature 293, 79-81.
37) de Villiers, J., Olson, L., Tyndall, C. and Schaffner, W. (1982) Nucleic Acids Res. 10, 7965-7976.
38) Masiakowski, P., Breathnach, R., Bloch, J., Gannon, F., Krust, A. and Chambon, P. (1982) Nucleic Acids Res. 10, 7895-7903.
39) Helfman, D.M., Feramisco, J.R., Fiddes, J.C., Thomas, G.P. and Hughes, S.H. (1983) Proc. Natl. Acad. Sci. USA 80, 31-35.
40) Young, R.A. and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA 80, 1194-1198.