A small segment of polyoma virus DNA enhances the expression of a cloned β -globin gene over a distance of 1400 base pairs

Jean de Villiers and Walter Schaffner

Institut für Molekularbiologie II der Universität Zürich, Honggerberg, 8093 Zürich, Switzerland

Received 5 November 1981

ABSTRACT

The hemoglobin β -1 gene of the rabbit was linked to a 244 bp DNA fragment from the beginning of the polyoma virus late region, not including the viral origin of replication. After transfection of such recombinant DNAs into mouse 3T6 and human HeLa cells, the polyoma sequences were found to strongly enhance the level of correct β -globin gene transcripts over a distance of at least 1400 bp. These findings are similar to those obtained with a segment of DNA from the corresponding region of the SV40 genome (J.Banerji, S. Rusconi and W. Schaffner, 1981, Cell, in press) which, however, shows very limited sequence homology to the polyoma 244 bp segment. Using the same assay, a complete copy of polyoma virus DNA was found to interfere with the enhancement of globin gene expression in a cell type-specific manner which may be due to incorrect transcription. In contrast to the complete polyoma virus genome, the 244 bp DNA fragment will be particularly useful as a component of mammalian expression vectors since it almost exclusively yielded high levels of correct 3-globin gene transcripts.

INTRODUCTION

Most of the DNA sequences necessary for the transcription of eukaryotic genes by RNA polymerase II in vivo have to date been located within the first 100 base pairs (bp) upstream of the transcription initiation site of the gene $(1-6)$. In several cases however, more distant sequences have also been found to influence transcription (7-16). Grosschedl and Birnstiel (7) first reported the long range effect of a transcriptional "modulator" element more than 110 bp upstream of a sea urchin H2A gene. Upstream sequences essential for the expression of the early genes of both SV40 (9,10) and polyoma virus (15) have also been identified.

SV40 DNA sequences have been shown, in our laboratory, to enhance the expression of a linked rabbit β -globin gene in vivo by two orders of magnitude. Globin gene expression was enhanced by SV40 DNA in the absence of a viral replicon, independent of gene orientation, and most surprisingly over distances as far as 3300 bp. It was shown that the enhancing activity depended on the presence of one of two 72 bp direct repeats, positioned upstream of the viral early gene region (14). From the work of other laboratories it is known that these 72 bp repeats are required for SV40 early gene expression (9,10; M. Fromm and P. Berg, personal communication). Here we report the subcloning of the equivalent DNA fragment from the genome of a polyoma virus strain A2 derivative (15) and demonstrate that it too functions as an "enhancer" of gene expression, despite the fact that the sequences involved are not organized as direct repeats and show little homology to the SV40 "enhancer" segment. The ability of this upstream sequence to function as an "enhancer" in cell lines derived from different species is shown, emphasizing the potential of such elements in the construction of vectors for the expression of cloned eukaryotic DNA in vivo.

MATERIALS AND METHODS

Plasmid construction

Starting from the polyoma DNA recombinant plasmid p.43.34.70 (a gift from C. Tyndall and R. Kamen, London), a 244 bp DNA fragment was cloned,with XhoI and HindIII linker DNAs, into the former KpnI site between the two globin genes of the recombinant $p\beta 2x$ as shown in Fig. 1E. This 244 bp segment is ³ bp longer, in the polyoma virus strain used to construct p.43.34.70, than the corresponding segment of the A2 strain (detailed in ref. 15; for convenience the numbering system of the A2 wildtype strain is used in Fig. 1D). In addition, a 4.7 kilobase pair (kb) rabbit β -globin gene fragment was inserted into the KpnI site of the late region of the polyoma virus genome in the pBR-polyoma recombinant pPyA₂ (a gift from

Figure 1. Schematic representation of the recombinant DNAs used. A. The KpnI DNA fragment of the rabbit, which contains the hemoglobin $\beta-1$ gene (39: a gift from T. Maniatis, Harvard University). The black line

is noncoding DNA, the black bars are coding sequences and the open bars are intervening sequences (IVS1 and IVS2). B. The pBR-polyoma- β -globin recombinant pPy β late+, which has the β -globin

gene inserted in the late region of the complete polyoma virus genome. C. A map of the polyoma virus genome.

D. An expansion of the polyoma virus region between the early and the late genes. (The numbering system is according to reference 17; however, see comment in Materials and Methods).

E. The β -globin recombinant $p\beta 2x$, showing the site where the polyoma virus "enhancer" fragment was inserted to generate the recombinant $p\beta(244+)$ β . The recombinant $p\beta(244-j)\beta$, with the polyoma DNA fragment inserted in the opposite orientation was also constructed.

C. Weissmann, Zürich), generating the plasmid pPy β late+ (Fig.1B). Standard recombinant DNA techniques were employed with some modifications. The desired DNA fragments were fractionated by agarose gel electrophoresis after restriction digestion. "Low gelling temperature" agarose (Sigma type VII) containing the DNA was cut out and used in ligation reactions without further

purification. The ligated DNAs, still in low gelling agarose, were directly used to transfect competent bacteria. Clones were identified by colony hybridization or by restriction analysis of plasmid DNAs. All work involving recombinant plasmids was done under conditions conforming to the standards outlined in the U.S. National Institutes of Health Guidelines for Recombinant DNA Research.

Cell growth and transfection

The mouse fibroblast cell lines 3T6 and 3T3, the rat embryo fibroblast line Rat-2, the human carcinoma line HeLa, the monkey kidney line CV-1 and the mink lung line CCL64 were gifts from Drs. W. Topp (Cold Spring Harbor), K. O'Hare (Strasbourg), R.I. Kamen (London), U. Pettersson (Uppsala), P. Berg (Stanford) and D. Owen (Lausanne), respectively. They were grown in Dulbecco's modified Eagle's minimal medium (GIBCO) containing 2.5% fetal calf serum, 2.5% calf serum, 100 units of penicillin per ml and $100\mu q$ of streptomycin per ml (GIBCO). Plasmid DNA was purified prior to transfection of the cells either by CsCl density gradient centrifugation or else by ion-exchange chromatography, using BND-cellulose (SERVA) (18). The calcium phosphate transfections (19) were done with modifications as described previously (14).

Immunofluorescence and S1 nuclease hybridization assays

60 hrs after transfection the cells were fixed with methanol and stained with antibodies (gifts of K.D. Smith, Baltimore and F.A. Anderer, Tübingen) to reveal the presence of rabbit β -globin in the cytoplasm as described (14). The S1 nuclease mapping (20,21) was done as described (22). This procedure was modified by the use of a single-stranded DNA probe (23), except for the experiment shown in Fig.2C, lane 3.

RESULTS

For our transcription studies, we have used a 244 bp DNA fragment from the beginning of the late region of the polyoma virus genome (Figs.1C and D). This segment, referred to as the

Figure 2. Si nuclease analysis of globin gene transcripts.

A. The S1 nuclease mapping scheme, according to Weaver and Weissmann (21), using a globin gene probe lacking IVS1 $(40; a$ gift from H. Weber, Zürich). B. Enhanced expression of the β -globin gene in HeLa cells. Slot l , control hybridization to β -globin mRNA from a rabbit reticulocyte lysate. Slots 2, 3 and 4; hybridization to 60µg RNA aliquots from HeLa cells transfected with the clones $p\beta 2x$, $p\beta(244+)$ β and $p\beta(244-)$ β , respectively. $C_$. β -globin gene expression in mouse cells of a recombinant containing the complete polyoma virus genome. Slot 1, control hybridization to rabbit 3-globin mRNA. Slot 2, hybridization to RNA from 3T6 cells transfected with the clone pPyßlate+. M, endlabelled DNA fragments (pBR322 digested with HpaII). Slot 3, hybridization to RNA from 3T3 cells also transfected with the clone pPyßlate+. fl, full length input DNA (453 nucleotides); ct, correct terminus (354 nucleotides); it, incorrect terminus (about 306 nucleotides upstream of the BamHI site).

"A region" in the work of Tyndall et al.(15), neither contains the origin of DNA replication nor does it code for T-antigen, both of which are essential for viral DNA replication. However, the "A region" is required for T-antigen transcription and for viral DNA replication (15). The 244 bp DNA fragment was inserted in both orientations between the two β -globin genes of the

plasmid p β 2x, giving rise to the clones p β (244+) β and p β (244-) β (Fig.1E). Mouse 3T6 and human HeLa cells were transfected with these DNAs and after 30-40 hrs the cytoplasmic RNA was extracted. Expression of the β -globin gene was assayed by S1 nuclease mapping, using a $32P$ -end-labelled, single-stranded DNA probe (Fig.2A). Both clones, $p\beta(244+)$ β and $p\beta(244-)$ β , yielded high levels of β -globin gene transcripts, most of which had a 5' end indistinguishable from that of rabbit β -qlobin mRNA (Fig.2B, lanes 3 and 4). Transfection with the parental plasmid $p\beta 2x$, which lacks polyoma virus DNA sequences, gave a very low level of β -globin gene expression (Fig.2B, lane 2). The band indicating correct globin gene transcription in Figure 2B, lane 2, was only visible in a long reexposure of the gel (data not shown) and was estimated to be two orders of magnitude less intense than the bands in lanes 3 and 4.

An indirect immunofluorescence assay was also used to examine β -globin production. 60 hrs after transfection the cells were fixed and sequentially stained with antibodies such that the presence of β -globin in the cells could be detected, upon fluorescence microscopy, as a red cytoplasmic fluorescence (Fig.3). Four cell lines were tested for the production of β globin after transfection, namely HeLa, mouse 3T6, Rat-2 and mink lung cells. In all four lines transfection, with either $p\beta$ $(244+)$ β or p β (244-) β DNA, yielded 0.1% to 1% of the cells positive. No positive cells could be detected in any of the cell lines after transfection with $p\beta 2x$ DNA.

Time-course analysis of β -globin specific cytoplasmic RNA synthesized in mouse 3T6 and HeLa cells after transfection

Cytoplasmic RNA was harvested from mouse 3T6 and HeLa cells at different times after transfection with $p\beta(244+)$ β DNA. 120µg aliquots of RNA from each time point were hybridized to single-stranded, $32P$ -end-labelled DNA probe. The hybrids were then digested with S1 nuclease, denatured and fractionated by polyacrylamide gel electrophoresis (Fig.4). In both mouse 3T6 and HeLa cells, at all times, the only prominent band corresponded to the authentic 5' terminus of rabbit β -globin mRNA.

Figure 3. Production of β -globin monitored by immunofluorescence. A. Mouse 3T6 cells, specifically stained for rabbit β -globin with rhodamineconjugated antibody (14). B. The same cells nonspecifically stained for SV40 T-antigen (which is not present within these cells), with fluorosceine-conjugated antibody. The white bar represents a distance of 100ν m.

In the 3T6 cells the level of β -globin specific RNA was highest at the first time point (29 hrs) and steadily declined, until reaching a level less than 10% of the peak at 60 hrs posttransfection (Fig.4, lanes $3-7$). In HeLa cells β -globin-specific RNA reached a maximum at 40 hrs and declined after another 20 hrs to approximately 50% of the peak value (Fig.4, lanes 8-12).

Expression of the β -globin gene linked to the entire polyoma virus genome

Mouse 3T3, mouse 3T6, HeLa and CV-1 cells were transfected with plasmid DNA of the $pBR-polyoma-\beta-qlobin$ recombinant $pPy\beta$ late+ (Fig.1B) and the expression of β -globin was assayed by indirect immunofluorescence as before. Whereas a proportion of the mouse cells, about 0.1%, stained positive for β -globin using the immunofluorescence assay (Fig.3), no positive HeLa or CV-1 cells were seen (data not shown). SI nuclease analysis of RNA extracted from the two different mouse cell lines (Fig.2C) showed that they had different ratios of the correct to an incorrect ⁵' end, which mapped 45-50 nucleotides downstream of the mRNA cap-site. (The full length DNA band in Fig.2C, lane 3,

Figure 4. Time course of β -globin gene expression in 3T6 and HeLa cells. 120pg aliquots of cytoplasmic RNA from transfected cells were subjected to the same Sl nuclease assay outlined in Fig. 2A. Slot 1, control hybridization to carrier tRNA alone. Slot 2, control hybridization to carrier RNA plus 3-globin mRNA from a rabbit reticulocyte lysate. Slots 3-7 show the DNA probe protected by RNA extracted from 3T6 cells at 29,40,44,48 and 60 hrs post-transfection, respectively. Slots 8-12 show the probe protection by HeLa cell RNA extracted at the same times as above. (Lane 10 suffered the loss of some material). M indicates the marker DNA fragments generated by BclI and DdeI double-digestion of polyoma virus DNA.

indicatesthe presence of unspliced transcripts that initiated upstream of the PstI site and/or contamination by the double-stranded DNA probe which was used for this one experiment).

DISCUSS ION

We have shown that linkage of a 244 bp polyoma virus DNA fragment, devoid of the replication origin, to a cloned rabbit β -globin gene yielded high levels of β -globin-specific RNA in transfected cells. RNA from both mouse 3T6 and human HeLa cells was assayed by S1 nuclease mapping and in both cases the β -qlobin gene transcripts were found to have the correct 5'-terminus. There exists the formal possibility that these 5' ends are generated from longer transcripts by nucleolytic cleavage, but

there is evidence from many laboratories which suggests that the ⁵' end of a mRNA is the site of transcription initiation (24-29). This enhancement effect was found to be independent of the orientation of the polyoma DNA fragment with respect to the direction of β -globin gene transcription. It also functions over a distance of at least 1400 bp, which was the distance from the "enhancer" to the cap-site of the closest β -qlobin gene. In addition, linkage of the cloned gene to the polyoma virus "enhancer" sequence allowed detection of β -globin protein in the cytoplasm of mouse 3T6, human HeLa, Rat-2 and mink lung cells by indirect immunofluorescence.

The phenomenon of enhanced gene expression is neither unique to polyoma virus DNA nor to the rabbit β -qlobin gene. Our group has shown that a DNA fragment from the equivalent region of the SV40 genome, i.e. the viral DNA between the early and the late genes, functions as an "enhancer" of β -globin gene expression (14) and that linkage of a sea urchin histone gene cluster to a pBR-SV40 vector enhanced the expression of three histone genes (M. Bendig, C. Hentschel and W. Schaffner, unpublished data). The SV40 "enhancer" element has subsequently also been used to get efficient transcription of the human β -globin gene (30; R. Treisman and T. Maniatis, personal communication), and of chimeric genes containing promoters from the conalbumin gene, or from the adenovirus major late region (16). An increased transformation frequency of the herpesvirus thymidine kinase gene, after linkage to SV40 DNA (31,32), may be due to an enhanced transcription of the thymidine kinase gene. The 73 bp repeat found within the "long terminal repeat" sequence of a retrovirus (mouse Moloney sarcoma virus) can replace the SV40 72 bp repeat as an essential component for expression of the SV40 early genes (P. Gruss and G. Khoury, personal communication), and an "enhancer"-bearing fragment of SV40 appears to complement its counterpart in the polyoma virus genome (C. Tyndall and J. de Villiers, unpublished data). These elements, whatever their mode of action, have potential as components of vector systems to study the expression of cloned

eukaryotic DNA in vivo, by virtue of the following properties:

- 1) The "enhancers" are relatively short sequences.
- 2) They seem to function independently of their orientation.
- 3) They are effective over long distances.
- 4) The effect does not appear to be restricted to a particular linked gene.
- 5) Transcription of linked genes initiates at the mRNA cap sites.
- 6) High levels of transcripts (and protein) are detectable.
- 7) The enhancement effect is not confined to a single cell line.

The polyoma virus "enhancer" fragment used in this study does not include all the sequences required for viral DNA replication (15). Vectors have been constructed which include a viral replicon (but lack "enhancer" sequences) to increase the copy number of the DNA template (6). The enhancement phenomenon and viral replication can be combined in a "high expression" vector by using a slightly larger DNA fragment from either SV40 or polyoma virus, since both these functions are encoded within the 450 bp between the early and the late genes (R. Treisman and T. Maniatis, personal communication; S. Rusconi and W. Schaffner, unpublished results). Special cell lines which synthesize viral T-antigens (33; C. Tyndall and R. Kamen, personal communication) make the requirement for the early viral genes obsolete in such a vector system.

The time-course analysis of RNA from mouse 3T6 and human HeLa cells has demonstrated that differences exist in the expression of the transfecting DNA's and these should be considered if comparative studies, with a number of cell lines, are undertaken. The peak level of β -globin RNA in the 3T6 cells is reached much earlier than in the HeLa cells, which may be a consequence of the shorter doubling time of the mouse cells.

It has been reported that neighboring plasmid sequences can interfere with the expression of the rabbit β -globin gene (22,14). The present studies indicate that polyoma virus DNA sequences outside the "enhancer" element also affect the correct expression of the linked β -qlobin gene. High levels of incorrect transcripts were obtained in mouse cells when the β -globin gene

was linked to the whole polyoma genome (Fig.2C), but not when it was linked to the 244 bp "enhancer" fragment (Fig.4). After transfection with the same recombinant containing whole polyoma DNA, 3T6 and 3T3 cells showed different ratios of correct to incorrect β -globin gene transcripts (Fig.2C) indicating that interference from polyoma sequences has a cell type-specific host range. Even more striking is the observation that HeLa and monkey CV-1 cells, after transfection with β -globin recombinants containing the whole polyoma genome, do not produce detectable amounts of β -globin protein. This cannot be explained by the fact that polyoma DNA replicates in mouse but not in primate cells, since the 244 bp fragment alone enhances β -globin production in all the cell lines tested, including HeLa cells (see above). It remains to be seen whether the failure of HeLa and CV-1 cells to produce β -globin after transfection with recombinants containing the complete polyoma virus genome is due to the production of incorrect transcripts or whether other factors are involved.

The major transcript with an incorrect 5' end in our Si nuclease assay mapped 45-50 nucleotides downstream of the authentic 3-globin mRNA cap site. This transcript is not peculiar to our assay: Rabbit β -1 globin genes from different sources have been introduced into a variety of vertebrate cells. The same incorrect terminus has been observed, in most of the cases together with the correct 5' terminus, in these other systems (4,5,14,22,34) including a mouse cell-polyoma vector system (5). This shorter DNA fragment,in our S1 nuclease assay, could result from protection by RNA which is either initiated 45-50 nucleotides downstream of the cap site or is generated by splicing of remote transcripts to this same site. We believe the latter to be the case since a primer extension with reverse transcriptase failed to produce a "strong-stop" band mapping at this position downstream of the cap site (J.de Villiers and R.I. Kamen, unpublished data). Examination of the β -globin gene sequence (35) indeed reveals the presence of an AG dinucleotide, preceding the position 48 nucleotides downstream of the cap

 $\bar{\mathcal{L}}$

Nucleic Acids Research

site, which could function as a splice acceptor site (36).

The mode of action of viral "enhancer" sequences has yet to be elucidated. Trivial explanations, such as a copy number effect due to amplification of the globin gene via a viral replicon or readthrough from viral promoters and the generation of globin mRNA by processing, were rendered highly unlikely by the experiments of Banerji, Rusconi and Schaffner (14) with SV40 and our experiments with polyoma: The 244 bp polyoma "enhancer" fragment lacks both a functional replication origin and sequences coding for T-antigen, and the transcripts of the globin gene map at the same position as authentic rabbit β globin mRNA. Nevertheless it is possible that the enhancement of globin gene expression involves the binding, to the viral 244 bp segment, of factors required for DNA replication and/or transcription.

An intriguing observation is that the "enhancer" elements of polyoma and SV40 do not show extended homologies in their DNA sequence. Even the 72 bp direct repeat of the SV40 "enhancer" segment and the corresponding 68 bp repeat of the closely related human papovavirus BKV (37) have very limited homology. Differences in the primary structure between "enhancer" elements should, however, not be overemphasized since such differences do not preclude the conservation of higher order structures (38; J. Banerji, personal communication). The possible lack of extensive sequence homology should be considered when eukaryotic genomes are screened for the presence of cellular counterparts to viral enhancers, which may activate genes within chromosomal domains.

ACKNOWLEDGEMENTS

We are indebted to R.I. Kamen (Imperial Cancer Research Fund Laboratories, London) and his team, especially C. Tyndall, G. La Mantia and P. Jat, for help and for valuable discussions during the visit of J.d.V. to London. We are also grateful to all the colleagues mentioned in the text for gifts of materials, to S. Rusconi and P. Dierks for technical advice and to B.

Wieringa and E. Mullen (Zürich) for help with cell culture. We thank S. Oberholzer and F. Ochsenbein for typing the manuscript and for preparing the graphic figures, respectively, and J. Banerji and M.L. Birnstiel for critical reading of the manuscript. This work has been supported by the Swiss National Research Foundation, grant No.3.687.0.80 and by the Kanton of Zürich, by a Swiss Federal Exchange Stipend to $J.d.V.$ and by a short term fellowship of the Imperial Cancer Research Fund (London) to J.d.V.

REFERENCES

- 1. Grosschedl, R. and Birnstiel, M.L.(1980) Proc.Natl.Acad.Sci. USA 77, 1432-1436
- 2. Gluzman, Y., Sambrook, J.F. and Frisque, R.J. (1980) Proc. Natl.Acad.Sci.USA 77, 3898-3902
- 3. Benoist, C. and Chambon, P. (1980) Proc.Natl.Acad.Sci.USA 77, 3865-3869
- 4. Dierks, P., van Ooyen, A., Mantei, N. and Weissmann, C. (1981) Proc.Natl.Acad.Sci.USA 78, 1411-1415
- 5. Dierks, P., Wieringa, B., Marti, D., Reiser, J., van Ooyen, A., Meyer, F., Weber, H. and Weissmann, C. (1981) ICN-UCLA Symp. on Molecular and Cellular Biology, XXXIII, D.D. Brown and C.F. Fox, eds. (New York: Academic Press), in press
- 6. Mellon, P., Parker, V., Gluzman, Y. and Maniatis, T. (1981) Cell, in press
- 7. Grosschedl, R. and Birnstiel, M.L. (1980) Proc.Natl.Acad. Sci.USA 77, 7102-7106
- 8. Grosschedl, R. and Birnstiel, M.L. (1981) Proc.Natl.Acad. Sci.USA, in press
- 9. Benoist, C. and Chambon, P. (1981) Nature 290, 304-310
- 10. Gruss, P., Dhar, R. and Khoury, G. (1981) Proc.Natl.Acad. Sci.USA 78, 943-947
- 11. Faye, G., Leung, D.W., Tatchell, K., Hall, B.D. and Smith, M. (1981) Proc.Natl.Acad.Sci.USA 78, 2258-2262
- 12. Guarente, L. and Ptashne, M. (1981) Proc.Natl.Acad.Sci.USA 78, 2199-2203
- 13. Struhl, K. (1981) Proc.Natl.Acad.Sci.USA 78, 4461-4465
- 14. Banerji, J., Rusconi, S. and Schaffner, W. (1981) Cell, in press
- 15. Tyndall, C., La Mantia, G., Thacker, C.M., Favaloro, J. and Kamen, R. (1981) Nucl.Acids Res., in press
- 16. Moreau, P., Hen, R., Wasylyk, B., Everett, R., Gaub, M.P. and Chambon, P. (1981) Nucl.Acids Res., in press
- 17. Soeda, E., Arrand, J., Smolar, N., Walsh, J.E. and Griffin, B.E. (1980) Nature 283, 445-453
- 18. Schlegel, R.A., Pyeritz, R.E. and Thomas, C.A. (1972) Analyt.Biochem. 50, 558-568

 $\ddot{}$

 \bar{z}