Avian Retroviral Vectors Derived from Avian Defective Leukemia Virus: Role of the Translational Context of the Inserted Gene on Efficiency of the Vectors

M. BENCHAIBI,* F. MALLET,† P. THORALVA,* P. SAVATIER,† J. H. XIAO,†
G. VERDIER,† J. SAMARUT,* and V. NIGON†

*Laboratoire de Biologie, Ecole Normale Supérieure de Lyon, 46 Allée d’Italie, 69364 Lyon Cédex 07, France, and †Laboratoire de Biologie Cellulaire, Institut de Chimie, Biologie Moléculaire et Cellulaire, Université Claude Bernard Lyon I, 43 Boulevard du 11 Novembre 1918, 69622 Villeurbanne Cédex, France

Received May 2, 1988; accepted September 21, 1988

We have constructed retroviral vectors derived from the genome of avian erythroblastosis virus ES4 (AEV ES4). The neo selectable gene was substituted for the original v-erbA or v-erbB oncogenes of AEV, either in the same or in a different reading frames. Recombinant retrovirus were rescued and used to infect chicken embryo fibroblasts or quail QT6 cells. When the neo gene was inserted in the same reading frame as the original oncogene, we obtained (1) a high level of expression of the neo gene, (2) a balanced ration of both genomic and subgenomic RNAs, and (3) high titer recombinant viruses. Conversely, when the neo gene was inserted in a reading frame different from that of the original oncogene, we observed (1) a very low level of expression of the neo protein, (2) a predominance of the viral transcript used as translational template for the neo protein synthesis, and (3) low titer recombinant viruses. One of the vectors was used to transfer a human β-globin gene into avian cells in culture without detectable rearrangement of this gene, but exhibited a deletion within the conserved noncoding region located between the two original oncogenes. Our data provide information for further construction of double expression vectors. Furthermore, three of the vectors would provide helpful tools to identify genetic elements of the virus genome involved in splicing regulation. © 1989 Academic Press, Inc.

INTRODUCTION

Retroviruses are now widely used to transfer foreign DNA sequences into eukaryotic cells in vivo and in vitro. Retrovirus-mediated gene transfer into cells in vitro offers several advantages over naked DNA transfection. First, it is much more efficient since a high fraction of cells in a culture can be infected and then stably integrate the retrovirus genome in their DNA, whereas only a small fraction of cells (10^{-3} to 10^{-6}) can be stably transformed by DNA transfection. Second, expression of genes introduced into cells by retroviral infection is more efficient than that of transfected genes (Hwang and Gilboa, 1984). Whereas the transfected genes usually integrate into host genome at multiple copies reaching a few hundred per cell, the number of integrated retrovirus genomes per infected cell is limited to one to 20 copies (Varmaus and Swanstrom, 1982). Finally, retrovirus-mediated gene transfer is the most efficient technique to transform somatic and germ cells in vivo (reviewed in Gilboa et al., 1986; Stewart et al., 1986). The efficiency and stability of the integration of provirus DNA into the host cell genome are provided by defined viral genome sequences that play specific functions in the molecular machinery of the virus replication cycle (Varmaus and Swanstrom, 1982). The construction of recombinant retroviruses by substituting foreign genes for viral genes implicates preserving those genetic elements involved in the machinery of virus replication. These essential elements include the long terminal repeats (LTRs) that provide the promoter and terminator signals (Coffin, 1985), the packaging sequences (Watanabe and Temin, 1982; Mann and Baltimore, 1985; Katz et al., 1986; Bender et al., 1987), short priming sequences involved in the synthesis of proviral DNA (Sorge and Hughes, 1982; Varmaus and Swanstrom, 1982), and other cis-acting sequences whose specific functions are still unclear but which enhance virus recovery (Sorge and Hughes, 1982; Armentano et al., 1987). The structure of the foreign genes inserted into retrovirus genomes is also determinant for the replication of recombinant viruses. Furthermore the level of expression of the inserted gene depends on the number of AUGs and the nucleotide sequence around the AUGs that precede the initiation codon of the foreign sequence (Bandyopadhyay and Temin, 1984a). Internal polyadenylation signals may prevent the synthesis of full-length viral genomes and thereby abort the replication cycle (Shimotohno and

1 Present address: Laboratoire de Génétique Moléculaire des Eucaryotes, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg Cédex, France.
2 To whom requests for reprints should be addressed.

0042-6822/88 $3.00
Copyright © 1989 by Academic Press, Inc.
All rights of reproduction in any form reserved.
Temin, 1981). Inserted sequences containing direct repeats induce instability of the vector viruses (Rhode et al., 1987). Many retroviral vectors have been developed to transfer foreign genes into mammalian and more specifically human cells. These vectors usually derived from murine leukemia viruses (Cepko et al., 1984; and review in Gilboa et al., 1986; Stewart et al., 1986). In avian species, retrovirus vectors were derived mainly from spleen necrosis virus (SNV), an avian virus that can infect some mammalian cells (Bangopadhyay and Temin, 1984b; Dougherty and Temin, 1987; Emerman and Temin, 1986a, Miller and Temin, 1986; Shimotohno and Temin, 1981). Avian replication competent vectors were also developed from Rous sarcoma virus (RSV), (Hughes and Kosik, 1984; Souza et al., 1984). We have explored the possibility of using an avian defective leukemia retrovirus as vector.

As a prototype virus, we chose the avian defective leukemia virus, avian erythroblastosis virus (AEV), that carries and expresses independently two oncogenes, respectively, v-erbA and v-erbB (Vennström et al., 1980). The two oncogenes of AEV, v-erbA and v-erbB, are expressed through two distinct viral transcripts (Anderson et al., 1980). AEV induces sarcomas and erythroleukemias in chicken in vivo and transforms fibroblasts and erythocyte progenitor cells in vitro (Gazollo et al., 1980; Graf and Beug, 1978). We first constructed recombinant retroviruses by substituting the selectable marker gene neo (Colbere-Garapin et al., 1981), for the v-erbB oncogene. The construction was optimized to produce high titer viruses by inserting the neo gene in the same coding phase as the original v-erbB gene. In a second step, we constructed recombinant retroviruses free of oncogenes and substituted the neo gene for the v-erbA oncogene. The construction was optimized by inserting the neo gene in the same reading frame as the original v-erbA sequence. We show that the efficiency of the vectors depends on the translational context of the inserted selectable gene. In a third step, we inserted a human β-globin gene, as a model gene, into the v-erbA oncogene. One of the recombinant viruses developed in this work was recently used to transfer the v-erbA oncogene into chicken cells in vitro (Gandrillon et al., 1987).

MATERIALS AND METHODS

Plasmids

The plasmid pAG50 (Colbere-Garapin et al., 1981) containing the neo gene was obtained from Dr. A. C. Garapin (Institut Pasteur, Paris). pAEV11, a plasmid containing a permuted AEV ES4 genome (Vennström et al., 1980), was obtained from J. M. Bishop (UCSF, San Francisco). One set of LTRs was added at both ends of the AEV ES4 genome in pAEV11 to give plasmid pAEV2LTR. The plasmid pBR3Pst, provided by Dr. Maniatis (Harvard University), carries a 2.3-kb PstI–PstI fragment which contains the β-globin gene flanked upstream by 450 bp and downstream by 340 bp of noncoding human DNA.

Enzymes

Enzymes used to modify DNA were purchased from Boehringer and used as recommended by the supplier.

Cells and media

Chicken embryo fibroblasts were prepared from C/O Spafas embryos (Rhone Mérieux, Lyon) and grown as previously described (Gandrillon et al., 1987). QT6 cells (Moscovici et al., 1977) were grown in the same medium. G418 (GIBCO) was dissolved in 100 mM HEPES buffer, pH 7.3, as a 10 mg/ml solution.

Construction of the recombinant virus genomes

Construction of the virus genomes were performed using regular techniques (Maniatis et al., 1982). Only the principles of the constructions are reported. Details of the intermediate steps and protocols can be provided upon request.

Construction of X11 and X12 genomes. Most of the v-erbB sequence was removed from pAEV2LTR after cutting out with appropriate restriction enzymes, and further recurrent digestion with Bsr31. The open space resulting from v-erbB elimination was filled up with the BglII–BglII fragment containing the neo gene of pAG50.

In X1, the region upstream from the neo coding frame contains residual Tn5 sequences linked to 28 bp of residual coding v-erbB sequence. To construct the X12 genome, we created NdeI site sequences through oligonucleotide site-directed mutagenesis (Zoller and Smith, 1983), respectively, on the initiator ATG of the neo gene and the ATG at the beginning of the v-erbB coding sequence. Ligation between both NdeI sites removed residual v-erbB and Tn5 sequences and put the neo coding sequence in the same frame as the original v-erbB.

Construction of TXN3' and TXN3'gag'. The TXN3' genome was constructed from X12 by deleting all the v-erbA sequence using Bsr31 digestion. The original gag sequence was preserved. TXN3'gag'- was constructed from TXN3' by deleting the gag sequence downstream of the XhoI site (Thoraval et al., 1987).

Construction of TSN, TXhoN, and TXN5'. These plasmids were constructed from an intermediate plasmid pBRgaglenv2LTR that derived from pAEV2LTR after removing most of the v-erbA sequences and keeping the region between both oncogenes which contains a
splice acceptor site. In pTSN the neo gene isolated from pAG50 was inserted downstream of the residual gag sequence. In pTXhoN the NdeI–BglII neo sequence isolated from pXJ12 was inserted at the XhoI site of gag by blunt end ligation. In this construct the gag and neo respective initiation codons are not in the same phase. In pTXN5' the XhoI site in gag was converted to a NdeI site by linker insertions. The neo gene isolated from pXJ12 was then linked through its NdeI site to the gag NdeI site. In this construct the initiation codons of gag and neo are in the same phase.

Construction of the genome of XJIA16 virus. Recombinant retrovirus XJIA16 was constructed by inserting the human δ-globin gene into the place occupied by the 0.5-kb PstI–PstI v-erbB fragment in XJ12 vector (see Fig. 5). The δ-globin gene sequence was a PstI–HpaI fragment of 1.8 kb, whose HpaI site in the second intron was converted into a PstI site by linker addition. This fragment does not contain the third exon and polyadenylation signal. The globin gene was inserted in anti-sense orientation with respect to retroviral transcription orientation.

Recovery of infectious virus particles

Plasmid DNAs (1–5 µg) carrying the virus genomes were cotransfected with pRAV1 helper virus DNA on secondary CEFs according to the method described by Kawai and Nishizawa (1984). G418 was added 24 hr later at a concentration of 200 µg/ml. Foci of resistant CEFs usually became visible after 10 days. After 10–12 days the drug was removed and the cells were then passaged in regular CEF medium. Viruses were usually collected 1–2 weeks after drug removal from subconfluent cultures.

Titration of viruses carrying the neo gene

Viruses carrying a neo gene were assayed for their ability to induce resistance to G418 of infected CEFs or QT6 cells. The titration assay was performed in 100-mm petri dishes. Secondary cells (2 × 10⁵) were seeded in 7–8 ml of regular CEF medium. Sixteen hours later the medium was removed and replaced by 4 ml of virus suspension diluted in CEF medium. After 4–12 hr at 37° the infection medium was discarded and the cells were covered with 7–8 ml of regular CEF medium containing 200 µg/ml G418. Medium with the drug was then changed every 3–4 days. The foci were scored 8–12 days after infection.

Analysis of neomycin phosphotransferase (NPT II) protein

The dosage of NPTII activity in infected CEFs was performed according to Reiss et al. (1984). The procedure includes separation of the enzyme on a nondenaturing electrophoresis gel followed by an indicator agarose gel overlay containing kanamycin and P32-ATP as substrates. NPTII isolated from Escherichia coli was used as a standard. The size of the protein was assigned using molecular weight standard proteins for nondenaturing gels (Sigma).

Preparation and analysis of cellular RNAs

Total cellular RNAs were extracted from infected CEFs according to Auffray and Rougeon (1980). Poly(A)+ RNAs were isolated by the technique of Long-agre and Rutter (1977), by chromatography on oligo(dT)-cellulose T3 (Collaborative Research). RNAs were analyzed on Northern blots using the standard procedures (Maniatis et al., 1982).

Analysis of proviral DNA

High-molecular-weight cellular DNA was prepared as described by Emerman and Temin (1986b). For DNA blot analysis 20–25 µg of DNA were digested with restriction endonucleases and fractionated on 0.8% agarose gels. Transfer of the fractionated DNA onto nitrocellulose filters and probing with nick-translated DNA fragments were done by standard procedures (Maniatis et al., 1982).

Preparation of labeled probes

The neo probe was isolated from pAG50 as a 1.2-kb BglII–BglII fragment. The δ-globin probe was the PstI–PstI fragment inserted into the vector. The env probe was the EcoRI–Accl fragment from the env residue of the AEV ES4. All probes were isolated by agarose electrophoresis and electroelution. The probes were labeled by nick-translation or random primer with [³²P]-dCTP using standard procedures.

RESULTS

I. Structure of the genome of the recombinant retroviruses

In the AEV ES4 genome, the v-erbB coding sequence is linked on its 5’ side to a residual c-erbB intron sequence termed J, that provides the splice acceptor site of the virus genome between v-erbA and v-erbB genes (Henry et al., 1985, Figs. 1 and 2A). An ATG codon is located 16 bp downstream this splice acceptor site in the v-erbB coding sequence. This codon will hence be referred to as the v-erbB ATG (or AUG). In the viral subgenomic RNA, the v-erbB coding sequence is joined to a RNA sequence transcribed from the leader region upstream from the δ-gag sequence. This leader sequence contains the gag initiation codon hereafter
called the leader ATG (or AUG, Fig. 2A). In the subgenomic RNA, the leader and the v-erbB AUGs are in the same reading frame. We decided to substitute either the v-erbA or v-erbB sequence by the neo sequence from the Tn5 transposon. The structures of the various constructs are shown in Fig. 1. In pTSN, pTXhoN, and pTXN5′ the neo sequence was inserted in place of the v-erbA gene, whereas in pXJ1, pXJ12, pTXN3′, and pTXN3′gag- it was inserted in place of v-erbB. In the first group of viruses the NPTII protein encoded by neo should be translated on the genomic RNA. In the second group it should be translated on the subgenomic transcript. The partial sequences of the transcripts used as templates for NPTII are shown in Fig. 2B. In pTSN, pTXhoN, and pXJ1 the leader and NPTII respective initiation codons are not in the same phase. Moreover, in pTSN and pXJ1, an additional out of frame AUG codon persisted between the initiation codons of gag and neo, respectively, and the neo coding sequence contains a terminator codon in phase with the leader AUG. The additional AUG derived from residual non-coding Tn5 sequence upstream of the neo coding sequence. In pXJ1, in addition, the v-erbB AUG was conserved. In pTXN5′, pXJ12, pTXN3′, and pTXN3′gag- the leader and NPTII initiation codons are in the same phase, and only the NPTII coding sequence was inserted. The phase and the distance between the leader and neo initiation codons on the template transcript are summarized in Table 1 for the various viruses. All the viruses were rescued on chicken embryo fibroblasts by cotransfection of the respective plasmid with a plasmid carrying the genome of the RAV1 helper virus.

II. Transcription of the proviruses in infected cells

Expression of the viral RNAs was analyzed in QT6 cells infected with the respective viruses. The cells were selected with G418 prior to RNA analysis. To identify the genomic and subgenomic RNAs of viruses XJ1, XJ12, TXN3′, and TXN3′gag- carrying the neo gene
Fig. 2. (A) Partial structure of the AEV ES4 provirus. DS and AS refer respectively to splice donor and acceptor sites. gag initiation codon (underlined) and v-erbB ATG (italic underlined) are represented on both provirus and subgenomic mRNA. / fusion between the splice donor site and the acceptor site. (B) Nucleotide sequences of genomic RNA (TSN, TXhoN, TXN5') of provirus vectors. The reading frame begins at the initiator codon of the gag message. Underlined AUG, gag AUG. Bold face AUG, neo AUG. Italic underlined AUG, one AUG of v-erbB corresponding to an internal AUG of c-erbB. Italic AUG, AUG from Tr16 transposon. Twice underlined, nonsense codon. / fusion between the splice donor site and the acceptor site.

at the v-erbB site we used a neo probe. Conversely, this probe should reveal only the genomic RNAs of viruses TSN, TXhoN, and TXN5' containing the neo gene inserted in the v-erbB site. In this latter case, an env probe was used to detect the subgenomic and genomic transcripts. This probe should also detect the genomic and subgenomic transcripts of the helper virus. Results are shown in Fig. 3A. All the vector viruses expressed a genomic and a subgenomic RNA of nearly the expected sizes. In cells infected with TSN, TXhoN, TXN5', or TXN3' additional unexpected transcripts of unknown origin were observed. Moreover a great difference could be observed between the amounts of genomic and subgenomic RNAs. For the virus XJ1 the subgenomic RNA was estimated to be 5- to 10-fold more abundant than the genomic RNA. Conversely XJ12, TXN3', and TXN3'gag- showed a nearly balanced ratio between both types of transcripts, and this ratio was similar to that observed with wild-type AEV (data not shown). Most of the transcripts of TSN and TXhoN vectors were under the genomic form whereas the genomic and subgenomic transcripts of TXN5' showed a balanced ratio. Analysis by Southern blot of the structure of the provirus in TSN, TXhoN, and TXN5' did not
TABLE 1

<table>
<thead>
<tr>
<th>Provirus</th>
<th>Frame AUG</th>
<th>Distance (bp) between AUGs</th>
<th>NPTII CEF ng/100 μg</th>
<th>NPTII QT6 protein</th>
<th>Initiator codon</th>
<th>Neo protein size (AA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSN</td>
<td>No</td>
<td>800</td>
<td>&lt;0.4&lt;sup&gt;a&lt;/sup&gt;</td>
<td>&lt;0.4&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—&lt;sup&gt;b&lt;/sup&gt;</td>
<td>—&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>TXhoN</td>
<td>No</td>
<td>253</td>
<td>&lt;0.4&lt;sup&gt;a&lt;/sup&gt;</td>
<td>&lt;0.4&lt;sup&gt;a&lt;/sup&gt;</td>
<td>—&lt;sup&gt;b&lt;/sup&gt;</td>
<td>—&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>TXN5&lt;sup&gt;+&lt;/sup&gt;</td>
<td>Yes</td>
<td>264</td>
<td>15.4</td>
<td>28</td>
<td>gag</td>
<td>353</td>
</tr>
<tr>
<td>X11</td>
<td>No</td>
<td>95</td>
<td>8.7</td>
<td>2.6</td>
<td>neo</td>
<td>264</td>
</tr>
<tr>
<td>X112</td>
<td>Yes</td>
<td>30</td>
<td>55.8</td>
<td>63</td>
<td>gag</td>
<td>275</td>
</tr>
<tr>
<td>TXN3</td>
<td>Yes</td>
<td>30</td>
<td>55.3</td>
<td>71</td>
<td>gag</td>
<td>275</td>
</tr>
<tr>
<td>TXN3 gag</td>
<td>Yes</td>
<td>30</td>
<td>54.8</td>
<td>86</td>
<td>gag</td>
<td>275</td>
</tr>
</tbody>
</table>

Column 2 indicates whether the neo gene and the gag gene are in the same reading frame. Column 3 indicates distance (bp) between the initiator codon for gag message and the neo AUG in mature RNA. Columns 4 and 5 show the concentration of NPT II in CEF and QT6 cells infected with the corresponding vectors. The concentration was estimated by comparing activity of the NPT II retroviral product with that of standard NPT II produced in E. coli transformed by pAG 50, assuming identical phosphotransferase activities. Should the altered neo proteins have decreased activity, then the values for TXN5<sup>+</sup>, X11, TXN3, and TXN3 gag would be underestimated. Column 6 indicates which AUG was used to obtain a NPT II functional protein whose size (number of amino acid) is given in column 7.

<sup>a</sup> Undetectable.
<sup>b</sup> Not assessable.

show any detectable rearrangements of the virus genomes (data not shown).

These data show, therefore, that the pattern of expression of the vector virus transcripts depends on the position of the neo gene in the virus genome. In particular, it appears that only those vectors in which the neo coding sequence was inserted in phase with the gag initiation codon (see Table 1) displayed a balanced expression of genomic and subgenomic transcripts. In all other vectors, we observed a predominance of the RNA used as template for the NPTII protein.

III. Effects of drug selection on the pattern of expression of the viral transcripts

We checked whether the pattern of expression of the viral RNAs in infected cells would depend on the selective pressure imposed by G418.

CEF s were infected with either X11 or X112 and then selected in medium containing G418. When sufficient numbers of G418-resistant CEFs were obtained the selection was raised and the cells were further amplified by passing three to four times in drug-free medium. Poly(A)<sup>+</sup> RNAs were extracted from the cells were taken either during or after the selection. Poly(A)<sup>+</sup> RNAs were also extracted from CEFs freshly infected with X11 prior to G418 selection. The RNAs were analyzed on Northern blots by hybridization with a neo probe to detect genomic and subgenomic RNAs (Fig. 3B).

In CEFs infected with X11 two transcripts were revealed with the neo probe. In size these transcripts were similar to the expected genomic and subgenomic viral RNAs, respectively. The ratio between the amounts of genomic and subgenomic transcripts, about 1:4, was identical in X11-infected CEFs taken prior, during, or after G418 selection. The same patterns were observed with CEFs infected with two different batches of X11 derived from two independent transfections of pXJ1 DNA.

In CEFs infected with X112 the subgenomic viral transcript was nearly of the same size as that observed in X11-infected CEFs. However, the X112 genomic transcript was shorter by 100–200 bases than that transcribed from X11. This difference in size is greater than predicted (85 nucleotides) from the constructed X112 genome and results from an unexpected deletion within the provirus (see below). In CEFs infected by X112 the ratio between the amounts of genomic and subgenomic transcripts was about 2:1. The same pattern was observed in CEFs taken either during or after selection.

IV. Analysis of NPTII protein in infected cells

The expression of NPTII recombinant protein was analyzed in CEFs or QT6 cells infected by the respective viruses. The cells were selected with G418 prior to the analysis. The sizes of the protein are shown in Fig. 4, compared with that of the native NPTII protein encoded by the Tn5 transposon.

No functional NPTII protein could be detected by this biochemical assay in cells infected with TSN or TXhoN viruses, presumably because of the poor sensitivity of the assay, suggesting that low expression of NPTII in
Fig. 3. Northern blot analysis of viral transcripts of the different constructs in infected cells. (A) Viral genomic and subgenomic mRNAs extracted from QT6 were revealed by hybridization with random primed neo or env probe. The 28 S (4.6 kb) and 18 S (1.8 kb) ribosomal RNAs were used as molecular weight markers as well as the genomic and subgenomic transcripts of RAV1 revealed by the env probe. Genomic transcript of RAV1 (•) or vectors (○). Subgenomic transcript of RAV1 (-•) or vectors (○). •, unexpected transcripts. (B) Viral genomic and subgenomic mRNAs of XJ1 and XJ12 were revealed by hybridization with a nick-translated neo probe. Poly(A)^+ mRNAs were isolated from infected CEFs collected before (○), during (+), or after (−) G418 selection.

the cells is sufficient to induce their resistance to G418. In cells infected with TXN6', the size of the NPTII protein was estimated at 353 amino acids, i.e., 89 amino acids longer than the natural NPTII which demonstrates an initiation of the protein at the leader AUG.

In cells infected with XJ1 the protein exhibited an estimated size of 264 amino acids, identical to that of the protein encoded by the Tn5 transposon, and was therefore initiated at its original initiation codon. In cells infected with XJ12, TXN3', or TXN3'gag- the NPTII protein showed a size of nearly 275 amino acids that is consistent with an initiation at the leader AUG. The amount of functional NPTII in cells infected by each virus was estimated by reference to purified NPTII pro-
tein isolated from bacteria expressing the Tn5 transpo-
son. The data are shown in Table 1. Functional NPTII protein could not be quantified in cells infected by TSN or TXhoN. Cells infected with XJ12, TXN3' or TXN3'gag-
expressed similar amounts of NPTII, at a level 3- to 4-
fold higher than cells infected by TXN5' and 6- to 25-
fold higher than cells infected by XJ1.

**Titers of the recombinant viruses**

The titers of the recombinant viruses were estimated as the numbers of G418-resistant CEF or QT6 foci in-
duced by 1 ml of virus suspension. The data are shown in Table 2. The highest titers were obtained with the viruses TXN5', XJ12, TXN3', and TXN3'gag-. These tit-
ers were in the same range as those of AEV ES4 esti-
ated in a transformation assay (data not shown). All viruses carrying a neo coding sequence out of frame with the gag sequence exhibited a titer 5- to 40-fold lower.

**VI. Use of XJ12 as a retroviral vector for foreign sequences**

As the XJ12 virus could be recovered at a high titer and displayed a pattern of transcription very similar to that of natural retroviruses, we investigated whether it could be used as an efficient vector to transfer foreign sequences into avian cells.

The human β-globin gene was inserted in reverse ori-
entation into the genome of XJ12 in place of the Psrl–
Psrl β-erbA sequence (Fig. 5A). The inserted β-globin fragment, 1.8 kb, was derived from a human genomic clone. The recombinant virus XJ516 was recovered at a titer of $4 \times 10^4$ rfu/ml. To analyze the stability of the vector, CEFs were infected with the virus, and then se-
lected with G418 and analyzed for the presence of inte-
grated proviruses. CEFs infected with XJ12 were used as a control.

Analysis of the DNA of XJ12-infected CEFs is shown in Fig. 5B after hybridization with a neo probe. The di-
gestion with Psrl gave the expected 0.9-kb band. How-
ever the restrictions by either EcoRI + Sall or BamHI + EcoRI gave a doublet at 1.7–1.8 and 3.1–3.2 kb, re-
spectively. Within each doublet the difference between the size of the bands was estimated between 100 and 150 nucleotides. Further restriction of these bands by Apat in both cases resulted in a unique band of 1.2 kb hybridizable to the neo probe. The simplest interpreta-

---

**TABLE 2**

<table>
<thead>
<tr>
<th>Virus</th>
<th>Average titer rfu/ml&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Batch 1</th>
<th>Batch 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSN</td>
<td>$3.0 \times 10^4$</td>
<td>$4.0 \times 10^3$</td>
<td></td>
</tr>
<tr>
<td>TXhoN</td>
<td>$9.0 \times 10^3$</td>
<td>$8.0 \times 10^3$</td>
<td></td>
</tr>
<tr>
<td>TXN5'</td>
<td>$3.6 \times 10^3$</td>
<td>$5.0 \times 10^3$</td>
<td></td>
</tr>
<tr>
<td>XJ1</td>
<td>$2.4 \times 10^4$</td>
<td>$2.7 \times 10^4$</td>
<td></td>
</tr>
<tr>
<td>XJ12</td>
<td>$3.2 \times 10^5$</td>
<td>$5.8 \times 10^5$</td>
<td></td>
</tr>
<tr>
<td>TXN3'</td>
<td>$8.0 \times 10^4$</td>
<td>$1.7 \times 10^5$</td>
<td></td>
</tr>
<tr>
<td>TXN3'gag-</td>
<td>$2.0 \times 10^5$</td>
<td>$2.1 \times 10^5$</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>The helper virus used to rescue the recombinant virus is Rav-1.<br>
<sup>b</sup>The titers were determined from the number of G418-resistant foci and expressed as resistance forming unit (rfu)/ml.
tion for these doublets is the occurrence of two kinds of XJ12 proviruses that differ by 100–150 bp within their 5' half.

The size of the double bands was therefore determined by comparison with restriction fragments of the original plasmid pXJ12 run on the same gels (not shown). We concluded that the shortest band within each doublet resulted from a deletion in the virus genome. This deletion was therefore located between the SalI and Apal sites upstream of the splice acceptor site (Fig. 5A). The same deletion was observed in another batch of XJ12 virus issued from an independent transfection (data not shown), which shows that deletion of this region is a characteristic feature of XJ12 virus.

Analysis of the DNA of XJ816 infected CEFs is shown in Fig. 5C. The digestions with either PstI or BamHI gave bands of expected sizes after hybridization with neo or δ-globin probes. On the other hand double digestions with either BamHI and EcoRI or SalI and EcoRI followed by hybridization with the neo probe gave fragments smaller by 100–150 bp than expected in the region spanning the viral sequence located between the SalI and EcoRI sites. As the δ-globin and neo sequence have been shown to be full length we conclude that this reduction of size of the provirus is the same as that shown in XJ12 genome and is more accurately located between the SalI and Apal sites.

**DISCUSSION**

The purpose of this work was to investigate the possibility of using avian defective leukemia viruses as retroviral vectors to carry several foreign genes. AEV was chosen as prototype defective leukemia virus since it is one of the few leukemia viruses that carries and independently expresses two oncogenes (Vennström et al., 1980). Insertion of a dominant selectable marker gene into a retroviral vector is necessary to ensure selection of the vector at a high titer. We therefore substituted the bacterial gene neo that encodes the NPTII protein for either the v-erbA or v-erbB oncogene of AEV. We analyzed the effects of sequences involved in the con-
trol of protein translation on the efficiency of expression of the NPTII protein and production of recombinant vector viruses.

In AEV, the p75gag-erbA protein is translated from the leader initiation codon upstream of the gag sequence. It is likely that this same initiation codon is used on the subgenomic RNA to translate the gp74v-erbB protein. Indeed this initiation codon is the only one to match the consensus structure proposed by Kozak (1981). Moreover, the nearest AUG codon brought by the v-erbB coding sequence, although in frame, represents an internal codon in the c-erbB proto-oncogene from which v-erbB originated (Henry et al., 1985). In the constructed virus genomes XJ1, XJ12, TXN3', and TXN3'gag- the neo gene was substituted for v-erbB. In the constructs TSN, TXhoN and 'TXN5' the marker gene was inserted in place of v-erbA. Only in the viruses TXN5', XJ12, TXN3', and TXN3'gag- were the NPTII and leader initiation codons in the same frame.

Our results clearly show that the transcription strategies used by the recombinant viruses were different according to the site of insertion of the neo gene. It appeared that each virus in which the neo gene was in frame with the leader initiation codon, expressed, in infected cells, balanced steady-state amounts of genomic and subgenomic RNAs. This situation was very similar to that observed for AEV. Conversely, in cells infected with TSN, TXhoN, or XJ1, we observed a predominance of the transcript used as translation template for the NPTII protein. The unbalanced expression of the genomic and subgenomic transcripts might result either from differential splicing efficiency or from selective degradation of some transcripts. It should be noticed in Fig. 3 that the total amounts of viral transcripts in XJ1- and XJ12-infected cells are nearly the same, suggesting that there was no preferential degradation of the genomic transcripts in XJ1-infected cells. However, a definitive conclusion would require a precise analysis of the life span of each transcript in cells infected by each virus. The mechanisms that regulate the splicing of viral transcripts in infected cells are not clearly known (Varmus and Swanstrom, 1982). Involved sequences have been described in the Moloney murine leukemia virus intron (Hwang et al., 1984b). A noncoding sequence between the env and src genes of Rous sarcoma virus influences splicing efficiency (Stoltzfus et al., 1987). It has been shown that the insertion of foreign sequences in retroviral genome may reduce mRNA splicing (Miller and Temin, 1986).

Our observation might be interpreted by considering that competition between several potential initiation AUGs on the viral transcripts could decrease the efficiency of translation of the NPTII protein (Kozak, 1984; Liu et al., 1984, Bandyopadhyay and Temin, 1984a). As a consequence, the growth of the infected cells in presence of G418 would require overproduction of the neo RNA template. This hypothesis would then assume that the selective pressure of the drug could modulate the strategy of transcription of the proviruses. This is not consistent with our observation that selection with G418 of CEFs infected by XJ1 had no effect on the ratio between genomic and subgenomic RNAs transcribed from XJ1 proviruses. One explanation of this discrepancy might be that the strategy of transcription of the virus genomes was imposed during the first drug selection following transfection of the constructed genomes into CEFs to rescue the recombinant viruses. This procedure, might have selected viruses with genome mutations that would change the pattern of transcription to enhance expression of the NPTII protein. Checking this hypothesis would require cloning and sequencing the rescued virus genomes. Should this hypothesis be valid, then the viruses TSN, TXhoN, and XJ1 would provide helpful tools to identify genetic elements of the virus genome that regulate the splicing of viral transcripts in retrovirus-infected cells. However, we may notice that a single mutation in consensus splice site could drastically reduce the splicing efficiency as it was observed in some cases of β-thalassemia (Treisman et al., 1983). Our hypothesis on the origin of the abnormal transcription strategy of viruses TSN, TXhoN, and XJ1 is strongly supported by the data of the analysis of the NPTII protein in infected cells. The NPTII protein was found in high amounts only in the cells infected with viruses showing a balanced expression of their transcripts. In all cases the NPTII protein was initiated at the leader initiation codon. In cells infected by XJ1, the NPTII protein was barely detectable and initiated at its own initiation codon. This observation is interesting because the subgenomic RNA of XJ1 contains one termination codon, in frame with the leader initiation codon, partially overlapping the NPTII initiator codon (see Fig. 2). According to Thomas and Cappechi (1986), such a structure might induce the scanning back of the ribosome and reinitiation at the internal AUG with a very low efficiency. In cells infected with either TSN or TXhoN, the NPTII protein was expressed at an even lower level than in XJ1-infected cells, suggesting that reinitiation at the neo AUG was poorly efficient in this case. The reason for this low efficiency is unknown but might possibly be due to the long gag sequences.

The XJ12 and its derivative XJ816 carrying a δ-globin gene showed a similar spontaneous deletion in their genome. This deletion occurred upon successive passages of the viruses in cell cultures (data not shown). The deleted sequence (100–200 bp) was located im-
mediatey upstream the splice acceptor site of the virus. Since X12 encodes a functional p75gagerbA protein (Gandrillon et al., 1987), it is likely that the deletion is localized within the conserved noncoding region that separates the v-erbA and v-erbB oncogenes in AEV ES4. This region contains a 200-bp sequence derived from an intron of the c-erbB proto-oncogene (Henry et al., 1985). The reason for the deletion of this region in exclusively X12 and X16 is unknown.

Our data show that X12 provides an efficient vector to transfer foreign sequences inserted into the v-erbA gene. This vector was able to transfer a human globin gene into avian cells in culture without detectable rearrangement of this gene.

In conclusion our data provide helpful information for the further construction of double expression retroviral vectors. Production of vectors with a high titer requires that the selectable marker sequence inserted into the vector must be in frame with the gag initiation codon. Moreover, they suggest that expression of a correct protein from an inserted non selectable sequence will require that this sequence carry its own initiation codon and that it is separated from the gag initiation codon by an in frame terminator codon. Furthermore, the eukaryotic translational mechanism of initiation at an internal downstream AUG seems to require a spacer between the stop codon and the reinitiator AUG estimated around 79 bp (Kozak, 1987); such a distance is observed in the case of the src oncogene where 63 bp separate the TGA in frame with the gag AUG and the src initiator codon (Schwartz et al., 1983).

ACKNOWLEDGMENTS

We thank all members of our laboratories for continuous helpful discussions. This work was supported by grants from the Institut National de la Recherche Agronomique, the Centre National de la Recherche Scientifique, and Association pour la Recherche contre le Cancer.

REFERENCES


Miller, C. K., and Temin, H. M. (1986). Insertion of several different DNAs in reticuloendotheliosis virus strain T suppresses transformation by reducing the amount of subgenomic mRNA. J. Virol. 58, 75–80.


