JOURNAL OF VIROLOGY, Aug. 2005, p. 10408–10419 Vol. 79, No. 16
0022-538X/05/$08.00+0 doi:10.1128/JVI.79.16.10408–10419.2005

The Receptor for the Subgroup C Avian Sarcoma and Leukosis Viruses, Tvc, Is Related to Mammalian Butyrophilins, Members of the Immunoglobulin Superfamily

Daniel Elleder,[1]† Volodymir Stepanets,[1] Deborah C. Melder,[2] Filip Šenigl,[1] Josef Geryk,[1] Petr Pajer,[3] Jiří Plachý,[1] Jiří Hejnar,[1] Jan Svoboda,[1]* and Mark J. Federspiel[2]*

Department of Cellular and Viral Genetics[1] and Department of Molecular Virology,[3] Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, Prague 166 37, Czech Republic, and Molecular Medicine Program, Mayo Clinic College of Medicine, Rochester, Minnesota 55905[2]

Received 2 February 2005/Accepted 9 May 2005

The five highly related envelope subgroups of the avian sarcoma and leukosis viruses (ASLVs), subgroup A [ASLV(A)] to ASLV(E), are thought to have evolved from an ancestral envelope glycoprotein yet utilize different cellular proteins as receptors. Alleles encoding the subgroup A ASLV receptors (Tva), members of the low-density lipoprotein receptor family, and the subgroup B, D, and E ASLV receptors (Tvb), members of the tumor necrosis factor receptor family, have been identified and cloned. However, alleles encoding the subgroup C ASLV receptors (Tvc) have not been cloned. Previously, we established a genetic linkage between *tvc* and several other nearby genetic markers on chicken chromosome 28, including *tva*. In this study, we used this information to clone the *tvc* gene and identify the Tvc receptor. A bacterial artificial chromosome containing a portion of chicken chromosome 28 that conferred susceptibility to ASLV(C) infection was identified. The *tvc* gene was identified on this genomic DNA fragment and encodes a 488-amino-acid protein most closely related to mammalian butyrophilins, members of the immunoglobulin protein family. We subsequently cloned cDNAs encoding Tvc that confer susceptibility to infection by subgroup C viruses in chicken cells resistant to ASLV(C) infection and in mammalian cells that do not normally express functional ASLV receptors. In addition, normally susceptible chicken DT40 cells were resistant to ASLV(C) infection after both *tvc* alleles were disrupted by homologous recombination. Tvc binds the ASLV(C) envelope glycoproteins with low-nanomolar affinity, an affinity similar to that of binding of Tva and Tvb with their respective envelope glycoproteins. We have also identified a mutation in the *tvc* gene in line L15 chickens that explains why this line is resistant to ASLV(C) infection.

Retroviruses require an interaction between the viral glycoproteins and a specific cell surface protein (receptor) to initiate entry into a cell (reviewed in references 32 and 56). The envelope glycoproteins of retroviruses are composed of trimers of two glycoproteins: the surface glycoprotein (SU), which contains the domains responsible for interaction with the host receptor, and the transmembrane glycoprotein (TM), which anchors SU to the membrane and mediates fusion of the viral and host membranes. The interaction of the SU glycoprotein with the host receptor usually involves multiple, noncontiguous determinants in both proteins that specify receptor choice and binding affinity and trigger a conformational change in the envelope glycoproteins that initiates the fusion process. Despite the complexity and specificity of the interaction between the viral glycoproteins and host receptors, closely related retroviruses carry envelope glycoproteins with mutations that alter receptor usage. The natural selection of retroviral subgroups with altered receptor usage may help the virus overcome host resistance and promote coinfection and may lead to heterotransmission.

The five highly related envelope subgroups of the avian sarcoma and leukosis viruses (ASLVs), subgroup A [ASLV(A)] to ASLV(E), are thought to be an example of the evolution of receptor usage by an ancestral retrovirus (reviewed in references 6 and 54). The ASLV(A) to ASLV(E) SU glycoproteins are almost identical except for five hypervariable regions, vr1, vr2, hr1, hr2, and vr3 (12, 13, 20, 21). Analyses of ASLV *env* genes have identified the hr1 and hr2 domains as the principal receptor interaction determinants; vr3 also plays a role in determining the specificity of receptor recognition (20, 21, 51, 52). In experiments that mimic the evolutionary forces of natural selection in cells resistant to ASLV entry, new viral variants with mutations in the ASLV envelope glycoproteins that altered receptor usage and binding affinity were selected; these mutations were in hr1 and vr3 (29, 30, 38, 42, 50).

ASLVs have been especially useful for studying the early events of retrovirus infection, not only because they have members with closely related SU glycoproteins that use different cellular receptors but also because several ASLV receptors

* Corresponding author. Mailing address for Mark J. Federspiel: Molecular Medicine Program, Mayo Clinic College of Medicine, 200 First Street, SW, Rochester, MN 55905. Phone: (507) 284-8895. Fax: (507) 266-2122. E-mail: federspiel.mark@mayo.edu. Mailing address for Jan Svoboda: Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo nam. 2, Prague 166 37, Czech Republic. Phone: 420-2-243-10-238. Fax: 420-2-243-10-955. E-mail: svoboda@img.cas.cz.

† Present address: The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037-1099.

have been cloned and soluble forms of these receptors have been developed. In chicken cells, three genetic loci determine the susceptibility and resistance to subgroup A to E ASLVs: *tva* [susceptibility to ASLV(A)], *tvb* [susceptibility to ASLV(B), ASLV(D), and ASLV(E)], and *tvc* [susceptibility to ASLV(C)] (53, 54). Susceptibility to ASLV infection is dominant, and therefore it is likely that the tva^r, tvb^r, and tvc^r ASLV resistance alleles contain defects that either block expression or alter the protein so that it is not an efficient ASLV receptor. The ASLV(A) receptors (Tva) are related to the family of low-density lipoprotein receptors (LDLR) (8, 57). Two *tva* genes have been cloned, from quail (8, 57) and chicken (22), that encode closely related Tva receptors but with critical differences in the conserved 40-amino-acid LDLR ligand binding domain important for ASLV(A) SU interaction (9, 30, 38, 45–47, 58, 59). Three highly related Tvb receptors have been identified and cloned, all of which are related to the tumor necrosis factor receptor (TNFR) family. There are two different *tvb* susceptibility alleles in chickens. The tvb^{s1} allele confers susceptibility to subgroups B, D, and E; the tvb^{s3} allele confers susceptibility only to subgroups B and D (1, 3). These alleles encode the chicken Tvb^{S1} (3) and Tvb^{S3} (14) receptors, respectively, and differ by a single amino acid change that presumably alters the structure of the Tvb^{S1} protein so that it no longer functions as an ASLV(E) receptor. A third cloned *tvb* receptor, the turkey Tvb^{T} receptor (2), confers susceptibility to only subgroup E ASLV. A 15-amino-acid region of the chicken Tvb receptors, residues 32 to 46, can serve as a minimal receptor for ASLV(B) viruses. However, the Tvb determinants required for ASLV(E) entry appear to be conformation dependent and require residues distinct from ASLV(B) (1, 3, 35).

To date, alleles encoding the ASLV(C) receptor, Tvc, have not been cloned. We have cloned the *tvc* gene and identified the Tvc receptor. Previously, two studies reported that the genetic loci controlling the susceptibility to subgroup A and subgroup C ASLVs are closely linked in chickens (40, 41). We refined the linkage of the *tva* and *tvc* genes and mapped these loci to a specific region of chicken chromosome 28 (23). This information allowed us to identify a bacterial artificial chromosome (BAC) containing a portion of chicken microchromosome 28 that conferred susceptibility to ASLV(C) infection. The region containing the *tvc* gene was located on a 15-kb fragment by using subclones constructed from this BAC. The *tvc* gene was identified on this genomic DNA fragment and encodes a 488-amino-acid protein related to mammalian butyrophilins, members of the immunoglobulin protein family (18, 39, 43).

MATERIALS AND METHODS

Chicken lines. The inbred White Leghorn chicken lines H6 and L15 were originally developed at the Northern Poultry Breeding Station (Reaseheath, Cheshire, United Kingdom) and imported to Prague, Czech Republic, in 1989 and 1977, respectively. Line H6 is sensitive and line L15 resistant to ASLV(C) infection (23). The inbred White Leghorn chicken line M was developed at the Institute of Molecular Genetics (Prague, Czech Republic) and is sensitive to ASLV(C) infection (26).

Cell culture and virus propagation. Chicken embryo fibroblasts (CEFs) were prepared from 10-day-old embryos of chicken lines H6, L15, M, and Brown Leghorn as described previously (24). CEFs and DF-1 cells, a continuous fibroblastic cell line derived from line 0 CEFs (28, 48), were grown in Dulbecco's modified Eagle's medium–nutrient mixture F-12 Ham (Sigma, St. Louis, Missouri) supplemented with 5% calf serum, 1% fetal calf serum, 1% chicken serum, 10% tryptoso-phosphate broth, and antibiotics (100 units/ml penicillin and 100 μg/ml streptomycin) in a 5% CO_2 atmosphere at 37°C. For some experiments, DF-1 cells were maintained in Dulbecco's modified Eagle's medium (GIBCO/BRL) supplemented with 10% fetal bovine serum (GIBCO/BRL), 100 units of penicillin per ml, and 100 μg of streptomycin per ml (Quality Biological, Inc., Gaithersburg, MD) at 39°C and 5% CO_2. The immortalized fibroblastoid Syrian hamster cell line NIL-2 (19) was cultivated in Dulbecco's modified Eagle's medium–nutrient mixture F-12 Ham supplemented with 5% calf serum and antibiotics under the same conditions as for CEFs. The B lymphoid chicken $DT40^{Cre1}$ cells (4, 16) and the DT40$tvc^{-/-}$ derivative were maintained in RPMI 1640 medium supplemented with 10% fetal bovine serum, 1% chicken serum, glutamine, antibiotics, and 0.1 mM β-mercaptoethanol in a 5% CO_2 atmosphere at 41°C. They were passaged by seeding 10^5 cells/ml every 3 to 4 days.

Two methods were used to generate the replication-competent ASLV long terminal repeats (LTR) with a splice acceptor, equipped with Bryan polymerase (RCASBP). In one method, 3×10^5 DF-1 cells were seeded per well in a 24-well multidish (Nunclon, Roskilde, Denmark) and transfected the next day with 1 μg of the RCASBP(A)GFP, RCASBP(B)GFP, or RCASBP(C)GFP plasmid (24) by using Lipofectamine 2000 (Invitrogen, Carlsbad, Calif.) according to the manufacturer's procedure. Cells were subcultured until they reached confluence on 100-mm culture dishes. The supernatant was collected and centrifuged at $2,000 \times g$ for 10 min at 4°C, filtered through 0.2-μm syringe filter (Nalgene, Rochester, NY), and stored in aliquots at −80°C. The virus titer was determined by infecting CEFs with diluted virus stock. Both RCASBP(A)GFP and RCASBP(C)GFP virus stocks regularly reached a titer of 10^7 IU/ml. Stocks of RCASBP viruses were also generated by transfection of plasmid DNA by the calcium phosphate precipitation method (24). In standard transfections, 5 μg of purified plasmid DNA was introduced into DF-1 cells by the calcium phosphate precipitation method (33). Viral spread was monitored by assaying culture supernatants for ASLV capsid protein (CA) by enzyme-linked immunosorbent assay (ELISA) (49). Virus stocks were generated from cell supernatants cleared of cellular debris by centrifugation at $2,000 \times g$ for 10 min at 4°C and stored in aliquots at −80°C.

DNA constructs. The pTvc expression plasmid was constructed using PCR with the sequence encoding the Tvc receptor truncated after the glycine residue (Gly335) of the cytoplasmic B30.2 domain (see Fig. 2). The PCR product was generated using primers TVC6 and TVC3 (see below) and line H6 cDNA and was subsequently cloned into the EcoRI and BglII sites of the pSG5 eukaryotic expression vector under the control of the simian virus 40 early promoter (25). Note that pTvc was derived from the *tvc* mRNA splice variant that lacks the last of the five small exons downstream from the transmembrane domain, amino acid residues 305 to 315 in Tvc. A similar cloning strategy was used to construct pTvc-F, an expression plasmid that contains the entire *tvc* cDNA coding region.

A gene encoding a soluble form of the chicken Tvc receptor (*stvc-mIgG*) and a soluble form of the chicken Tvb^{S3} receptor (*stvb*^{S3}*-mIgG*) (1) was constructed as described previously for soluble forms of the chicken and quail Tva receptors (22, 30, 31). These genes encode the extracellular domain of the particular ASLV receptor fused to the constant region of a mouse immunoglobulin G (mIgG) heavy chain and are in the CLA12NCO adaptor plasmid (24). The *stvc-mIgG* and *stvb*^{S3}*-mIgG* gene cassettes were isolated as ClaI fragments and subcloned into the ClaI site of the RCASBP(A) vector. DF-1 cells were infected with each virus, and infected cell supernatants that contained either the sTvc-mIgG protein or the sTvbS3-mIgG protein were collected. The chicken sTva-mIgG protein was collected in supernatants from a stable DF-1 cell line expressing the *tvaSstva-mIgG* gene in the TFANEO expression vector (22). The supernatants were cleared by centrifugation at $2,000 \times g$ for 10 min at 4°C and stored in aliquots at −80°C.

A gene encoding the SU glycoprotein from RCASBP(C) fused to the constant region of a rabbit immunoglobulin G (rIgG), SU(C)-rIgG, was constructed as described previously for SU(A)-rIgG (29) and SU(B)-rIgG (3). The *SU(C)-rIgG* gene was subcloned into the CLA12NCO adaptor plasmid, recovered as a ClaI fragment, and subcloned into the ClaI site of the RCASBP(A) vector (24). DF-1 cells were infected with RCASBP(A)*SU(C)-rIgG* virus, and supernatant from the infected cells containing the SU(C)-rIgG protein was collected. Supernatant containing the SU(A)-rIgG protein was collected from a stable DF-1 cell line expressing the *SU(A)-rIgG* gene in the TFANEO expression vector (29). The supernatants were cleared by centrifugation at $2,000 \times g$ for 10 min at 4°C and stored in aliquots at −80°C.

The replication-competent ASLV recombinant viruses RCASBP(A)AP, RCASBP(B)AP, and RCASBP(C)AP, containing the human heat-stable alkaline phosphatase (AP) gene, and RCASBP(A)GFP and RCASBP(C)GFP, containing the green fluorescent protein (GFP) gene, have been described previ-

TABLE 1. Expression of Tvc confers susceptibility to ASLV(C) infection

Cells	ASLV receptor(s)	ASLV titer (avg ± SD)[a]		
		RCASBP(C)AP	RCASBP(A)AP	RCASBP(B)AP
DF-1	Tva, Tvb, Tvc	$(2.4 \pm 0.6) \times 10^4$	$(4.5 \pm 3.3) \times 10^6$	$(4.2 \pm 1.2) \times 10^4$
NIL-Tvc	Tvc	$(7.9 \pm 1.5) \times 10^3$	<1	5.3 ± 2.5
NIL-Tva	Tva	<1	$(5.4 \pm 2.3) \times 10^4$	3.0 ± 2.6
NIL	None	<1	<1	1.0 ± 1.0

[a] Cells were infected with 10-fold serial dilutions of ASLV stocks of RCASBP (C)AP, RCASBP(A)AP, and RCASBP(B)AP viruses, and the titer was determined by the AP assay and presented as AP-positive foci per milliliter. The results are the averages and standard deviations from three experiments.

ously (24). The receptor subgroup of the viral envelope glycoprotein is given in parentheses in these virus designations.

The *tvc* knockout constructs were generated by insertion of isogenic homology regions from the 5′ and 3′ ends of the DT40 *tvc* gene into the multiple cloning sites of the pLoxPuro and pLoxBsr vectors bearing the puromycin or blasticidin S resistance genes, respectively, driven by the chicken β-actin promoter (4). Homology region 1 was obtained as a PCR product from DT40 genomic DNA with primers KO1L and KO1R (Table 1), shortened at the 5′ end by KpnI and at the 3′ end by SalI digestion, and ligated into unique KpnI and SalI sites in the pLoxBsr vector. Homology region 1 is 1,907 bp in length and spans nucleotides −2923 to −1017 5′ from the initiation ATG of *tvc*. Homology region 2 was obtained by PCR using primers KO2L and KO2R (Table 1), digested by SpeI in the noncomplementary part of the primers, and ligated into the unique SpeI site in the pLoxBsr vector. Homology region 2 is 2,788 bp in length and spans nucleotides 3033 to 5820 3′ from the ATG. After the 5′ and 3′ target locus sequences were cloned into pLoxBsr, creating pLOX*tvc*Bsr, the blasticidin S resistance gene was replaced with the puromycin resistance gene from pLoxPuro by using the SalI and XhoI sites to create pLOX*tvc*Puro (4).

ALV AP assay. For AP assays, cell cultures (~30% confluent) were incubated with 10-fold serial dilutions of the appropriate RCASBP/AP virus stocks for 36 to 48 h. The assay for alkaline phosphatase activity was described previously (31).

BACs. BACs containing chicken genomic DNA inserts were obtained from BACPAC Resources (Oakland, Calif.), from the Wageningen University Chick-FPC collection (http://www.zod.wau.nl/vf), and from Jerry Dodgson (Michigan State University, East Lansing). A human BAC containing the GFP gene that was used as an experimental control was a gift from Matt Cotten (Axxima Pharmaceuticals) (5). BAC DNAs were amplified and purified with the QIAGEN large-construct kit (QIAGEN Inc., Valencia, Calif.). BAC CH261-100I9 DNA was digested with XhoI and SfuI, and a 15-kb DNA fragment containing the *tvc* locus was purified by agarose gel electrophoresis and subcloned into the SalI and NarI sites of plasmid pUC18 (55).

Reverse transcription-PCR (RT-PCR) and PCR. The following oligonucleotide primers were used in this study (all oligonucleotides are written 5′ to 3′): TVC1, CTGACCCTGTGGCCGTGGCTGTG; TVC2, TGGGGATCTCCTCA TCCACGCTG; TVC3, ctagagatcTGCCATCATCAGTCAGGACC; TVC4, CT CCTGGTCCCCATAGTGGTCCTG; TVC5, GGGTCGCTCTGGATTATGA AGTGG; TVC6, atccgaatTCCATGgAGA*c*GATGTTTTTTGGCTG; TVC7, CT CGCTGGCAGAGCCAGGAC; TVC8, AGGCTGGTTGGTTTAGCAGTAG; TVA1, CATGGTGCGGTTGTTGGAG; TVA2, GGGATCGCGCGGCTCCG AAC; TVB1, CAGACCTCCAGAAGCCAGAC; TVB2, CCCAGGCACTTGG GAAAG; GAPDH1, CCATGACAACTTTGGCATTG; GAPDH2, TCCCCAC AGCCTTAGCAG; KO1L, agtcgtcgacCAGGCAAAGAAGGGAATGAA; KO1R, agtcgtcgacACAGTTGGGCAAGCCTGTAA; KO2L, agtactagtTTCTCT GCAGGTGACTGTGG; KO2R, agtactagtCTGCTGGCTGCTGTAGTCTG; HR, ATACACTGGGGAGGTTCAGAAGC; BS1, CGATTGAAGAACTCAT TCCACTCAAATATACCC; PU4, CAGCGCCCGACCGAAAGGAGCGCAC GACC; and IGL, AGGCAGGTATAACGCCCTCT. In these sequences, the lowercase letters designate sequences at the 5′ ends of the primers that are not complementary to the templates, and the restriction enzyme sites used for cloning of the PCR products are underlined (EcoRI in TVC3, BglII in TVC6, SalI in KO1L and KO1R, and SpeI in KO2L and KO2R). The two lowercase letters in italic in primer TVC6 indicate mismatches with the *tvc* sequence in inbred line H6. The mismatches arose by designing this oligonucleotide based on the *tvc* sequences from expressed sequence tag data, which differ at these two positions from the *tvc* sequence in inbred line H6.

Total RNA was isolated from CEFs or from chicken tissues by using the TRI reagent (Sigma-Aldrich, St. Louis, Mo.) according to the manufacturer's instructions. RNA samples (1 μg) were converted to cDNA by using the Moloney murine leukemia virus reverse transcriptase (Promega, Madison, Wis.) with an oligo(dT) primer. PCR amplifications of regions of the *tvc* cDNA were performed using the Accu Taq polymerase (Sigma-Aldrich, St. Louis, Mo.) according to the manufacturer's instructions. The following PCR conditions were used: 2 min at 94°C; 33 cycles of 15 s at 94°C, annealing for 40 s at primer-specific temperatures (see below), and 2 min at 68°C; and a final extension of 7 min at 68°C. The initial amplification of a 1,257-bp fragment containing most of the *tvc* open reading frame used primers TVC1 and TVC2 and an annealing temperature (T_A) of 58°C. The 5′ and 3′ end regions of the *tvc* transcript were determined from rapid amplification of cDNA ends (RACE) reactions (see below). The entire coding region of *tvc* was amplified using primers TVC7 and TVC8 located in the 5′ and 3′ untranslated region sequences (T_A = 56°C). The truncated form of Tvc contained in the pTvc plasmid was generated using primers TVC6 and TVC3 (T_A = 56°C).

To compare the expression levels of *tva*, *tvb*, and *tvc* transcripts in various chicken tissues, the cDNA samples were amplified with specific primers by using *Taq* DNA polymerase (Takara, Kyoto, Japan) in the presence of 1% dimethyl sulfoxide and 1 M betaine. PCR amplifications were performed using the following conditions: 2 min at 94°C; 30 cycles of 15 s at 94°C, annealing for 40 s at primer-specific temperatures (see bellow), and 40 s at 72°C; and a final extension of 3 min at 72°C. Using these PCR conditions, *tvc*, *tva*, *tvb*, and *gapdh* PCR products were sampled after different numbers of cycles to determine the exponential phase of the reactions. To be within the exponential phase of the reactions, 30 amplification cycles were used to generate receptor-specific PCR products, and 24 amplification cycles were used to amplify the highly expressed *gapdh* transcripts in individual samples. The primer combinations and annealing temperatures used for this analysis were as follows: for amplification of *tvc* cDNAs, primers TVC6 and TVC4 and a T_A of 55°C; for amplification of *tvb* cDNAs, primers TVB1 and TVB2 and a T_A of 58.5°C; for amplification of *tva* cDNAs, primers TVA1 and TVA2 and a T_A of 57°C; and for amplification of *gapdh* cDNAs, primers GAPDH1 and GAPDH2 and a T_A of 58°C.

RACE. 5′- and 3′-RACE reactions were used to generate a complete *tvc* sequence. The SMART RACE (Clontech, Palo Alto, CA) procedure was employed according to the manufacturer's instructions. One microgram of total RNA from line H6 CEFs was converted to cDNA by using Moloney murine leukemia virus reverse transcriptase (Promega, Madison, Wis.) and used as a template for both RACE reactions. The *tvc*-specific antisense primers TVC2, TVC4, and TVC6 were used for the 5′-RACE. The *tvc*-specific sense primers TVC1 and TVC5 were used for the 3′-RACE. Both 5′- and 3′-RACE reactions yielded single distinct PCR products whose nucleotide sequences were determined directly by sequencing with the TVC4 and TVC5 primers, respectively.

Assays for Tvc receptor function. Line L15 CEFs were seeded at 3×10^5 per 30-mm dish. The next day, the cells were transiently transfected using Lipofectamine 2000 with 2 μg of DNA. The DNA samples tested included different BAC clones, plasmid constructs containing DNA subcloned from BACs, the pTvc plasmid, and DNA from control BAC or plasmid clones. Two days after transfection, the cells were infected with RCASBP(C)GFP at a multiplicity of infection of 4. The number of GFP-positive cells was determined 2 days later by fluorescence-activated cell sorting (FACS) with a Coulter Epics Elite ESP apparatus (Coulter Corporation, Hialeah, Florida) and analyzed with WinMDI software (J. Trotter, The Scripps Research Institute, San Diego, Calif.).

For receptor analysis in stably transfected hamster cells, the NIL-2 cell line was transfected using Lipofectamine 2000 with 2 μg of plasmid pTva or pTvc together with 0.2 μg of plasmid pMC1 NEO poly(A) (Stratagene, La Jolla, Calif.), which contains the neomycin resistance gene. The transfected cells were grown for 10 days in G418 (500 μg/ml) to select for neomycin resistance. Cell clones were isolated from soft agar, expanded, and challenged with RCASBP(A)GFP and RCASBP(C)GFP at a multiplicity of infection of 4. The GFP-positive cells were quantitated 2 days after infection as described above.

Sequence prediction programs. The SignalP 3.0 program (10) was used to predict the location of the signal peptide cleavage site in the amino acid sequence of Tvc. The transmembrane region was predicted by using the TMHMM 2.0 server (http://www.cbs.dtu.dk/services/TMHMM). The IgV and B30.2 domains in Tvc were predicted by searching the Conserved Domain database with the BLAST program (37) or by performing the search against the Pfam protein families database (7). The IgC domains in Tvc and other butyrophilins were not predicted by these domain prediction programs, in part because these domains exhibit considerable sequence variation. The presence and position of the IgC domain in Tvc were determined by performing a multiple-sequence alignment with butyrophilins from three species, which shows the conserved residues in IgC.

Knockout of *tvc* in chicken DT40 cells. To knock out one *tvc* allele, DT40^{Cre1} cells (5×10^6 in 400 μl of medium) were electroporated (25 μF and 700 V applied in a 4-mm cuvette) with 20 μg of linearized pLOX*tvc*Bsr by using the Gene Pulser Xcell (Bio-Rad). After electroporation, 5×10^4 cells were plated per microtiter well in 100 μl of growth medium. Twenty-four hours later the medium was changed to growth medium supplemented with15 μg/ml blasticidin S (Invitrogen). After 10 days of drug selection, drug-resistant colonies were cloned, expanded, and checked for homologous recombination. To knock out the second *tvc* allele, DT40 *tvc*$^{+/-}$ cells were electroporated with 20 μg of linearized pLOX*tvc*Puro as described above and selected for puromycin resistance in medium supplemented with 1 μg/ml puromycin (Sigma). Puromycin-resistant clones were checked for resistance to blasticidin S and for homologous recombination of the pLOX*tvc*Puro construct. Homologous recombination of pLOX*tvc*Bsr was detected using primer HR, complementary to nucleotides 4181 to 4159 5′ from the *tvc* ATG, and primer BS1, complementary to sequence in the blasticidin S resistance gene (4). Homologous recombination of pLOX*tvc*Puro was detected using primer HR and primer PU4 within the puromycin resistance gene sequence. Nontargeted *tvc* genes were detected using primer HR and primer IGL 3′ (the end of the Ig-like domain-coding sequence and complementary to nucleotides 460 to 479). The primers are listed in the "RT-PCR and PCR" section above. One DT40 *tvc*$^{-/-}$ clone was used for testing the sensitivity or resistance to RCASBP(B)GFP and RCASBP(C)GFP viruses.

ELISA. The levels of the mIgG fusion proteins were quantitated in culture supernatants by ELISA for the mouse IgG tag as previously described (31). The levels of the rIgG fusion proteins were quantitated in culture supernatants by ELISA for the rabbit IgG tag as previously described (29). The linear range for a standard experiment was between 0.5 and 50 ng of ImmunoPure IgG Fc fragment per ml.

Immunoprecipitations and Western immunoblot analysis. The sTvc-mIgG, sTva-mIgG, and sTvbS3-mIgG proteins were immunoprecipitated separately with anti-mouse IgG-agarose beads (Sigma) for ≥1 h at 4°C, and the SU(C)-rIgG, SU(A)-rIgG, and SU(B)-rIgG proteins were immunoprecipitated separately with anti-rabbit IgG agarose beads (Sigma) for ≥1 h at 4°C. The protein-antibody agarose bead complexes were collected by centrifugation and washed twice in dilution buffer (50 mM Tris-buffered saline, 1% Triton X-100, 1 mg/ml bovine serum albumin), once in 50 mM Tris-buffered saline, and once in 0.05 M Tris-Cl, pH 6.8. The washed complexes were collected by centrifugation, resuspended in 50 μl Laemmli sample buffer (2% sodium dodecyl sulfate [SDS], 10% glycerol, 0.0625 M Tris-Cl, pH 6.8, 0.1% bromophenol blue, 5% β-mercaptoethanol), and heated for 5 min at 100°C. The precipitated and denatured proteins were separated by SDS-12% polyacrylamide gel electrophoresis, transferred to a nitrocellulose membrane, blocked, and washed as described previously (31, 36). The immunoblots were probed with a either a peroxidase-conjugated goat anti-mouse IgG (heavy plus light chains [H+L]) (50 ng/ml) or a peroxidase-conjugated goat anti-rabbit IgG (H+L) (50 ng/ml) (Kirkegaard and Perry Laboratories, Gaithersburg, MD). After extensive final washing, immunodetection of the protein-antibody-peroxidase complexes was performed with the Western blot chemiluminescence reagent (DuPont NEN, Boston, MA). The immunoblots were then exposed to Kodak X-Omat film.

Binding affinity analyzed by FACS. DF-1 cells or DF-1 cells infected with either RCASBP(A), RCASBP(B), or RCASBP(C) were removed from culture with trypsin de Larco (Quality Biological, Inc.) and washed with Dulbecco's phosphate-buffered saline (PBS). The cells were fixed with 4% paraformaldehyde in PBS at room temperature for 15 min and then washed with PBS. Approximately 1×10^6 cells in PBS supplemented with 1% calf serum (PBS-CS) were incubated with supernatant containing one of the receptor-mIgG or SU-rIgG fusion proteins on ice for 30 min. The stable DF-1 cell lines TF/sTvaS-18 (expressing chicken sTva-mIgG), TF/sTvb^{S3}-8 (expressing sTvb^{S3}-mIgG), TF/SU(A)-19 [expressing SU(A)-rIgG], and TF/SU(B)-12 [expressing SU(B)-rIgG]; RCASBP(A)*stvc-mIgG*-infected DF-1 cells (expressing sTvc-mIgG); and RCASBP(A)*SU(C)-rIgG*-infected DF-1 cells [expressing SU(C)-rIgG] were the sources of the receptor-mIgG or SU-rIgG fusion proteins. The DF-1 cells were then washed with PBS-CS and incubated with either 5 μl of goat anti-mouse IgG (H+L) linked to phycoerythrin or 5 μl of goat anti-rabbit IgG (H+L) linked to phycoerythrin (Kirkegaard & Perry Laboratories, Gaithersburg, MD) in PBS-CS (1-ml total volume) on ice for 30 min. The cell–soluble receptor-IgG–Ig-phycoerythrin complexes were washed with PBS-CS, resuspended in 0.5 ml PBS-CS, and analyzed with a Becton Dickinson FACSCalibur using CELLQuest 3.1 software.

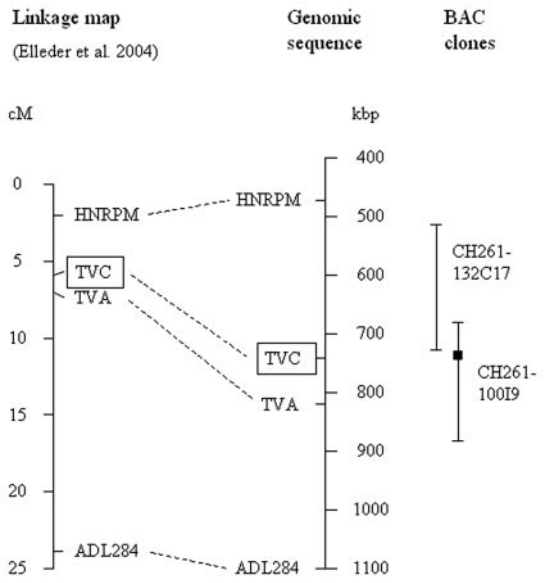


FIG. 1. Alignment of the genetic linkage map and the genome sequence of a region of chicken chromosome 28. A portion of chicken chromosome 28 is shown, depicting the positions of *tva*, *tvc*, and the two nearest genetic markers. Corresponding loci are connected by dotted lines. The positions of BACs CH261-100I9 and CH261-132C17, which tested positive and negative for the presence of *tvc*, respectively, are shown on the right. The small black square depicts the position of the 15-kbp genomic fragment that contained the *tvc* gene.

***K_D* calculations.** The maximum possible fluorescence and apparent dissociation constant (K_D) for each data set obtained from the FACS binding assays were estimated by fitting the data via nonlinear least squares to a log logistic growth curve function, $f(y) = M/[1 + e^{-r(\log x - \log K_D)}]$, where y is the mean fluorescence, M is the maximum fluorescence, r is the rate, x is the concentration of the receptor-mIgG or SU-rIgG fusion protein, and K_D is the dissociation constant, which is defined as the concentration of the receptor-mIgG or SU-rIgG fusion protein at half-maximal binding (30).

Nucleotide sequence accession number. The complete nucleotide sequence of the *tvc* cDNA has been submitted to GenBank and assigned accession number AY847576.

RESULTS

Isolation of the *tvc* gene by positional cloning. Our previous genetic linkage analysis mapped the location of the *tvc* locus to within 1.1 centimorgans of the *tva* locus on chicken chromosome 28 (Fig. 1) (23). To clone the *tvc* gene, we obtained overlapping BAC genomic clones from this region of chromosome 28 from the Wageningen University ChickFPC collection (http://www.zod.wau.nl/vf) and from Jerry Dodgson (Michigan State University, East Lansing). Each BAC DNA was transiently transfected into line L15 CEFs, which are resistant to ASLV(C) infection, and challenged with RCASBP(C)GFP, an ASLV(C) virus containing the GFP gene (24, 48). The challenged culture was scored for infection by the RCASBP (C)GFP virus by fluorescence microscopy. The transfection

efficiency of BAC DNA was extremely low by either the lipofection or calcium phosphate precipitation methods (data not shown). The transfection efficiencies were estimated using a control BAC clone containing a 130-kb human DNA genomic insert and a GFP expression cassette (5) and averaged 20 GFP-positive cells out of 10^6 transfected CEFs. Despite the extremely low BAC DNA transfection efficiencies, there were GFP-positive cells observed in RCASBP(C)GFP-challenged line L15 CEF cultures transfected with BAC CH261-100I9 but not in those transfected with the other chicken BAC clones, including the adjacent BAC CH261-132C17 (data not shown). The draft sequence of the chicken genome from Washington University (http://pre.ensembl.org/Gallus_gallus) enabled the mapping of the BAC clones to their locations on chromosome 28. The BAC CH261-100I9 clone contains ~200 kb of chicken genomic chromosome 28 DNA (draft positions ~681000 to 883000) and the *tva* gene (~818000 to 819500).

We analyzed the ~90-kb chicken DNA insert region of the BAC CH261-100I9 between the *tva* gene and the overlap with the BAC CH261-132C17 insert. Fragments of this region were subcloned into a plasmid expression vector by using convenient restriction enzyme sites, and each fragment was screened for the ability to confer susceptibility to RCASBP(C)GFP infection by transient transfection into line L15 CEFs as described above. This analysis mapped the *tvc* gene to a 15-kb fragment of chromosome 28 (positions 735137 to 750326 in the draft sequence) (data not shown). Further deletions narrowed the location of the *tvc* gene to the 10-kb region between positions 740000 and 750000 (data not shown). The Ensemble browser (http://www.ensembl.org/) predicted only one gene in this region, encoding a homologue to mammalian butyrophilins.

Cloning of cDNAs encoding the putative Tvc receptor. A tentative nucleotide sequence of the putative *tvc* coding region was assembled using the Ensemble gene prediction software and sequences in the chicken expressed sequence tag databases (11, 17). Using RT-PCR and primers based on the assembled *tvc* sequence, a ~1,200-bp cDNA product was amplified from total RNA isolated from the ASLV(C)-susceptible line H6 CEFs (data not shown). To obtain a complete cDNA, additional primers were designed from the nucleotide sequence of this initial cDNA clone and were used to amplify the 5′ end (5′-RACE) and 3′ end of the *tvc* cDNA (data not shown). The resulting sequences were assembled into a *tvc* cDNA consensus sequence (Fig. 2). This *tvc* cDNA consensus sequence was verified by PCR amplification of full-length *tvc* cDNAs by using primers based on the ends of the consensus sequence (data not shown). The *tvc* mRNA is 1,875 nucleotides long with a single open reading frame encoding 488 amino acids. The deduced amino acid sequence of Tvc was used in a protein-protein BLAST search for homologous protein sequences in the National Center for Biotechnology Information protein databases. Figure 3 compares the Tvc protein sequence to the two most similar proteins identified in this search, human and bovine butyrophilin BTN1A1 (butyrophilin, subfamily 1, member A1) and mouse butyrophilin BTN1A1.

The butyrophilins are members of the immunoglobulin superfamily and are type I, single-transmembrane proteins with several conserved features (18, 39, 43). The extracellular regions of most butyrophilin proteins contain two immunoglobulin-like domains, IgV and IgC, which are related to CD80 and CD86 costimulatory molecules of the immune system. The cytoplasmic domain contains a B30.2 domain, a domain present in a large number of proteins, that may function as a protein-binding domain. Some members of the butyrophilin family are highly expressed in secretory epithelium of the mammary gland during lactation; other butyrophilin homologues are expressed predominately in skeletal muscle, intestine, or erythroid cells; while still other homologues are widely expressed in many tissues. To date, the function of any of the butyrophilin proteins is not understood, but the conserved structural domains and diverse expression profiles suggest that these proteins may have important general and tissue-specific functions within and outside the immune system.

We used computer predictions and the homology between Tvc and the mammalian butyrophilins to identify the putative signal peptide, the IgV domain, the IgC domain, the transmembrane domain, and the B30.2 domain in the Tvc protein sequence and the 5′ and 3′ noncoding regions of the *tvc* cDNA (Fig. 2). The extracellular region of Tvc contains 227 amino acids with two potential N-linked glycosylation sites and four cysteine residues, the transmembrane region is 24 amino acids long and the cytoplasmic region contains 215 amino acids. The exon boundaries of the *tvc* gene were determined by comparing the *tvc* cDNA nucleotide sequence to the draft chicken genome sequence (Fig. 2). All *tvc* intron-exon boundaries are phase 1 junctions, the major domains are each encoded on a single exon, and the sequence between the transmembrane and B30.2 domains contains five short exons, all characteristics shared by other butyrophilin family genes.

Expression of the *tvc* cDNA confers susceptibility to ASLV(C) infection. To test the ability of the *tvc* cDNA to confer susceptibility to ASLV(C) infection, the expression plasmid pTvc was constructed from a fragment of the *tvc* cDNA encoding the extracellular domain, transmembrane domain, and a truncated cytoplasmic domain of Tvc (Fig. 4A). Line L15 CEFs were transiently transfected with either pTvc DNA or a control plasmid DNA, pTva, encoding the extracellular domain, transmembrane domain, and a truncated cytoplasmic domain of the chicken Tva receptor (22), and subsequently challenged with RCASBP(C)GFP. The challenged cultures were analyzed by phase-contrast and fluorescence microscopy (Fig. 4B), and the GFP-positive infected cells were quantitated by flow cytometry (Fig. 4C and D). CEFs transfected with pTvc were efficiently infected by RCASBP(C)GFP; on average 15% of the cells were GFP positive. In contrast, CEFs transfected with the pTva control plasmid were not efficiently infected by RCASBP(C)GFP and yielded only rare GFP-positive cells (<0.05%). These experiments demonstrate that the expression of Tvc confers susceptibility to ASLV(C) infection in chicken cells normally resistant to ASLV(C) infection. We initiated these experiments with this truncated *tvc* cDNA at a time when a cDNA encoding the complete B30.2 cytoplasmic domain was not yet cloned. This *tvc* cDNA did contain the complete extracellular domain that is important for its function as a ASLV(C) receptor. We have since repeated these experiments with the expression plasmid pTvc-F, which contains the entire *tvc* cDNA coding region, and obtained similar results (data not shown). In addition, no toxicity was observed in cells expressing either the full-length or truncated Tvc protein.

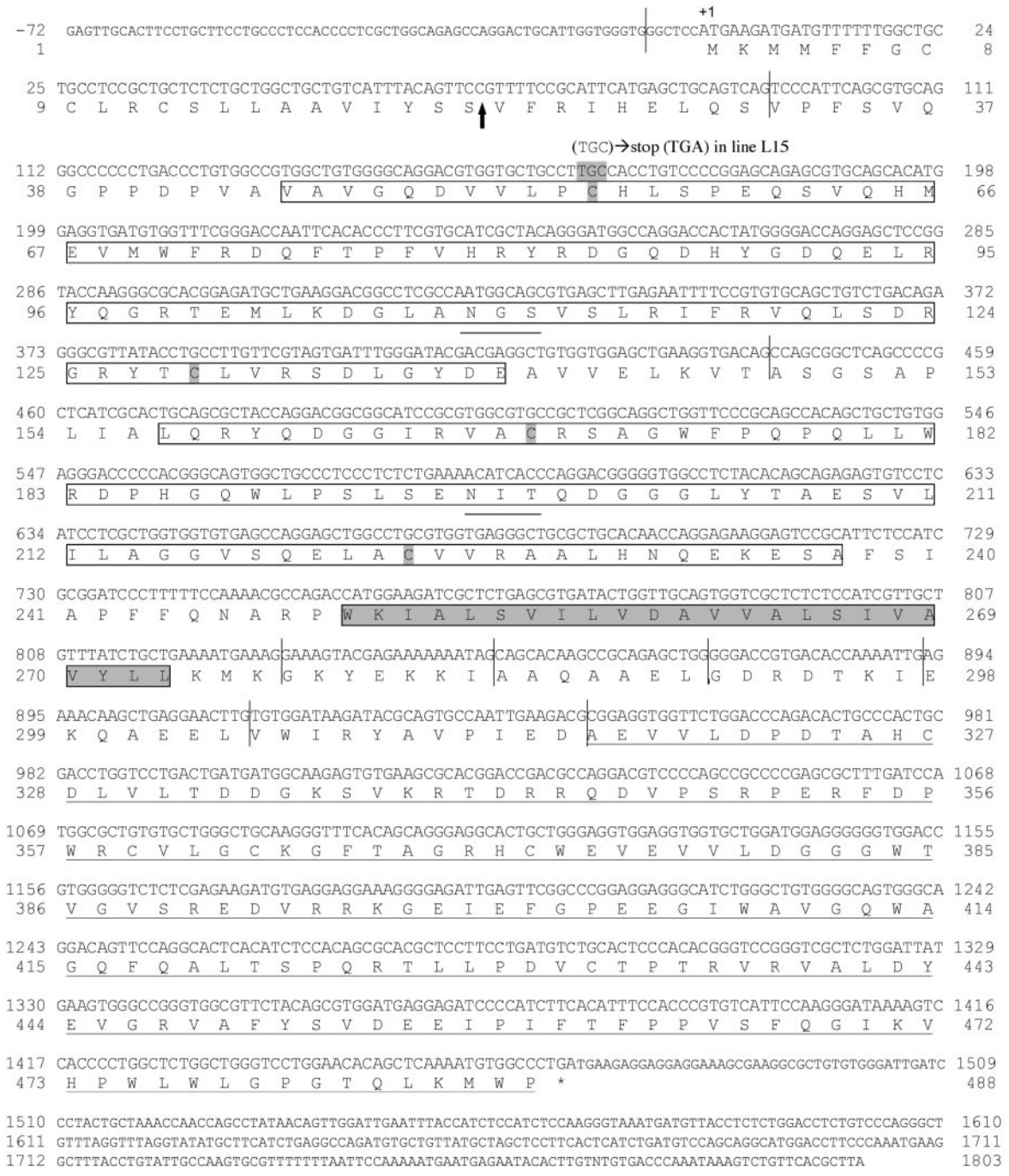


FIG. 2. The nucleotide sequence of the *tvc* cDNA. The nucleotide and deduced amino acid sequences of Tvc are shown, with the smaller nucleotide letters indicating the 5′ and 3′ untranslated regions of the *tvc* mRNA. The open boxes in the extracellular region of Tvc are the IgV and IgC domains, the shaded box is the transmembrane domain, and the long underlined region in the intracellular domain is the region of Tvc homologous to the B30.2 domain. The predicted leader peptidase cleavage site is marked by an arrow and the exon junctions by vertical lines. The cysteine residues in both Ig domains are shaded, and the site of the mutation in the line L15 *tvc* cDNA is highlighted. The two potential N-linked glycosylation sites are also underlined.

Mammalian cells do not normally express functional receptors for any of the ASLVs. To further characterize the specificity of ASLV susceptibility conferred by Tvc, stable mammalian cell lines that express either a truncated form of the Tvc receptor or the chicken Tva receptor were generated. Hamster NIL cells were transfected with either the pTvc or pTva plasmid and stable lines generated that express each receptor, named NIL-Tvc and NIL-Tva, respectively. NIL-Tvc cells and NIL-Tva cells were challenged with RCASBP(C)GFP or RCASBP(A)GFP (the same ASLV but with a subgroup A

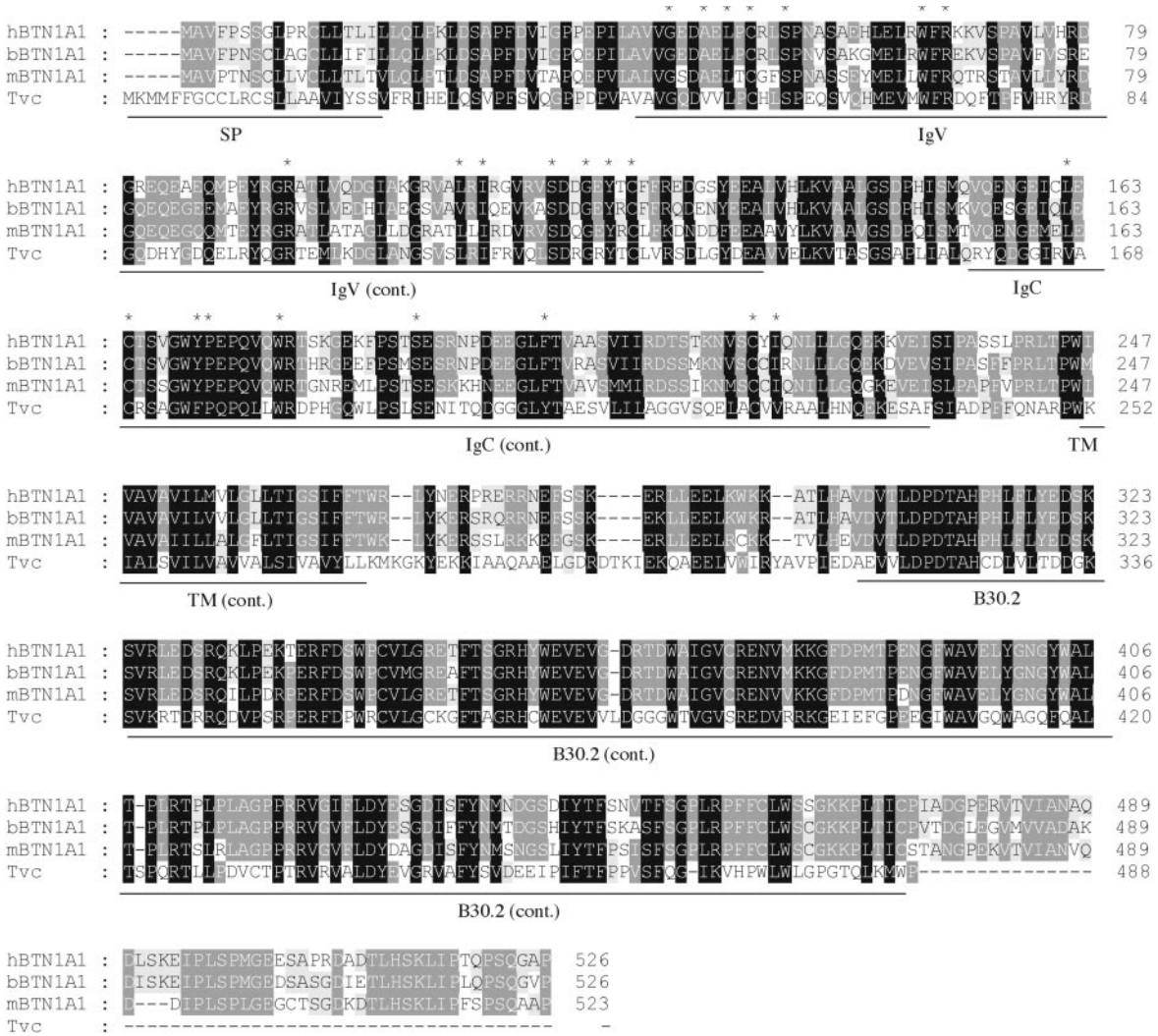


FIG. 3. Alignment of Tvc with human, bovine, and murine butyrophilins. The amino acid sequence of Tvc was aligned with those of the three proteins with the highest homology from the BLAST search, three butyrophilins classified as subfamily 1 member 1 (BTN1A1): human (hBTN1A1, accession number NP_001723.1), bovine (bBTN1A1, accession number NP_776933.1), and mouse (mBTN1A1, accession number AAH31459.1). From this analysis, homologous domains were predicted in Tvc (underlined); SP indicates the signal peptide and TM the transmembrane domain. Amino acid residues that are conserved within the IgV and IgC1 folds are indicated with asterisks (18). The protein alignment was performed using the ClustalW program and then highlighted in gray scales using the Genedoc program.

env). The challenged cultures were analyzed by phase-contrast and fluorescence microscopy, and the GFP-positive cells were quantitated by flow cytometry. The NIL-Tvc cells were efficiently infected with RCASBP(C)GFP but not with RCASBP (A)GFP, while the NIL-Tva cells were efficiently infected with RCASBP(A)GFP but not with RCASBP(C)GFP (data not shown). The titers of RCASBP(C)AP, RCASBP(A)AP, and RCASBP(B)AP virus stocks on the NIL cell lines and chicken DF-1 cells were compared. As expected, DF-1 cells were efficiently infected by viruses of all three envelope subgroups, since DF-1 cells express Tvc, Tva, and Tvb receptors (Table 1). NIL-Tvc cells were efficiently infected only by RCASBP(C)AP, and NIL-Tva cells were efficiently infected only by RCASBP(A)AP. The parental NIL cells, which do not express functional ASLV receptors, were not efficiently infected by any of these ASLV viruses.

Targeted deletion of *tvc* renders DT40 cells resistant to ASLV(C) infection. The *tvc* gene was deleted in chicken DT40 cells, a B-cell line with high rates of homologous recombination (4, 16). The 5′ and 3′ genomic regions that flank the *tvc* gene in DT40 cells were cloned and used for homologous recombination to target integration to completely delete the *tvc* coding sequence (Fig. 5A). A DT40 $tvc^{-/-}$ cell clone with a cell morphology, viability, and growth rate similar to those of parental DT40 cells was chosen for further study. Parental DT40 cells and DT40$tvc^{-/-}$ cells were challenged with RCASBP(C)GFP and RCASBP(B)GFP viruses and analyzed by phase-contrast and fluorescence microscopy (Fig. 5B). Parental DT40 cells were infected by both RCASBP(C)GFP (Fig. 5B, panel b) and RCASBP(B)GFP (Fig. 5B, panel f). In contrast, DT40$tvc^{-/-}$ cells were highly resistant to RCASBP (C)GFP (Fig. 5B, panel d) but still susceptible to infection by

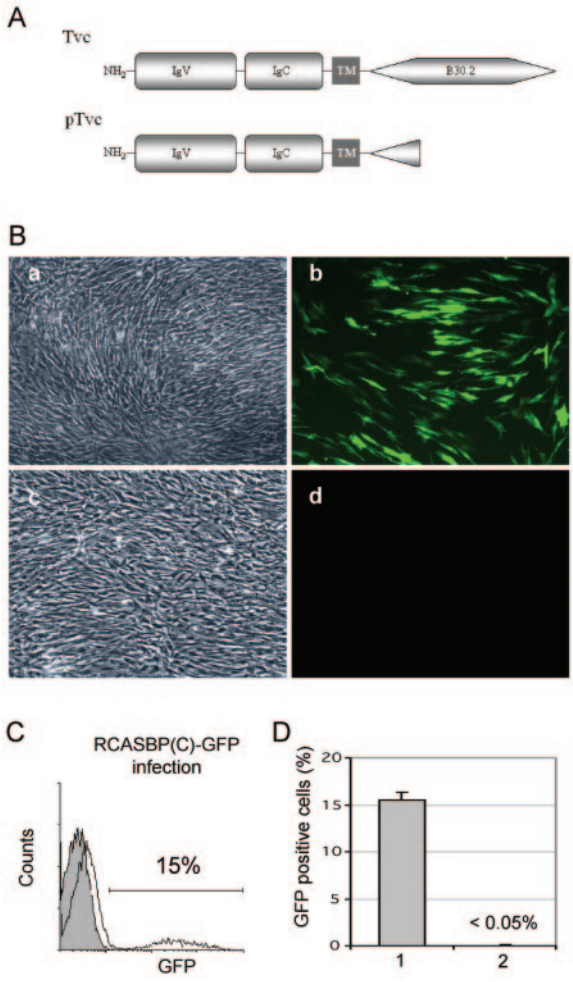


FIG. 4. Expression of the *tvc* cDNA allows ASLV(C)s to infect resistant chicken cells. Line L15 CEFs were transfected with the plasmid pTvc and subsequently infected with RCASBP(C)GFP virus. As a negative control, the pTva plasmid was transfected into line L15 CEFs and challenged with RCASBP(C)GFP virus. The infected GFP-positive cells were quantified by flow cytometry. (A) Schematic diagrams of the full-length Tvc protein and of the truncated product expressed from the pTvc plasmid. (B) Images from phase contrast microscopy (a and c) and fluorescence microscopy (b and d) of L15 CEFs transfected with pTvc (a and b) or the control pTva plasmid (c and d). (C) A representative histogram depicting flow cytometry analysis of cells transfected with pTvc (open curve) or a control plasmid (shaded curve). The relative fluorescence is plotted against the cell count; the percentage of cells in which the fluorescence activity was above that measured in the control is indicated. (D) Percentage of GFP-positive cells from FACS analysis; the averages and standard deviations from two experiments performed in triplicate are shown. L15 CEFs were transfected with pTvc (bar 1) or control plasmid (bar 2).

RCASBP(B)GFP (Fig. 5B, panel h). These data confirm that *tvc* is the ASLV(C) receptor.

Tvc and ASLV(C) glycoproteins bind with low-nanomolar affinity. Two approaches were used to estimate the binding affinities of ASLV receptors for ASLV envelope glycoproteins. In one approach, the ASLV envelope glycoproteins expressed on the surface of DF-1 cells infected with ASLV(C), ASLV(A), or ASLV(B) were assayed for their ability to bind soluble forms of three ASLV receptors, sTvc-mIgG, sTva-

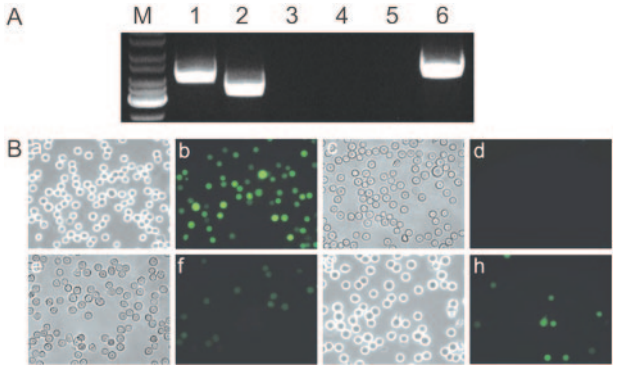


FIG. 5. Targeted deletion of *tvc* renders DT40 cells resistant to ASLV(C) but not to ASLV(B) infection. (A) PCR detection of homologous recombination of knockout constructs and *tvc* in DT40*tvc*$^{-/-}$ (lanes 1 to 3) and parental (lanes 4 to 6) DT40 cells. Homologous recombination of the pLOX*tvc*Bsr construct was detected using primers HR and BS1 (lanes 1 and 4), homologous recombination of the pLOX*tvc*Puro knockout construct was detected using primers HR and PU4 (lanes 2 and 5), and intact *tvc* was detected using primers HR and IGL (lanes 3 and 6). M, DNA marker ladder. (B) Phase-contrast (a, c, e, and g) and fluorescence (b, d, f, and h) microscopy of infected cells. Parental DT40 cells (a, b, e, and f) and DT40*tvc*$^{-/-}$ cells (c, d, g, and h) were infected by RCASBP(C)GFP (a to d) or RCASBP(B)GFP (e to h) virus.

mIgG, and sTvb^{S3}-mIgG, by FACS as described previously (22, 30, 38). In a second approach, the ASLV receptor proteins expressed by DF-1 cells were assayed for their ability to bind soluble forms of ASLV SU glycoproteins, SU(C)-rIgG, SU(A)-rIgG, and SU(B)-rIgG, by FACS. The integrity of the three soluble receptor proteins and the three soluble SU proteins was determined by immunoprecipitation and Western analysis (Fig. 6A). The concentration of each protein stock was quantitated by ELISA for either the mouse or rabbit IgG (29, 31).

For the soluble receptor protein concentrations assayed, sTvc-mIgG bound only to ASLV(C)-infected cells (Fig. 6B), sTva-mIgG bound only to ASLV(A)-infected cells (Fig. 6C), and sTvb^{S3}-mIgG bound only to ASLV(B)-infected cells (Fig. 6D). All three soluble receptor forms bound their respective envelope glycoproteins with subnanomolar affinity (Table 2). The estimated binding affinity of the sTva-SU(A) interaction (0.05 nM) was ~10-fold higher than the sTvc-SU(C) (0.55 nM) and sTvbS3-SU(B) (0.9 nM) binding affinities in these experiments. We were able to detect the binding of SU(C)-rIgG and SU(A)-rIgG, but not SU(B)-rIgG, to uninfected DF-1 cells (Fig. 6E). The binding affinities measured for the SU(C)-Tvc interaction (3.93 nM) and the SU(A)-Tva interaction (0.84 nM) by using this approach were both ~10-fold lower than those we obtained using the soluble receptor approach (Table 2).

The *tvc*^{r} gene in the ASLV(C)-resistant line L15 contains a mutation that introduces a premature stop codon. Line L15 is highly resistant to ASLV(C) infection, presumably due to a mutation in the *tvc* gene. To test this hypothesis, we cloned *tvc* cDNAs from line L15 total RNA by RT-PCR using primers that amplified the complete transcript. When the nucleotide sequences of the line L15 *tvc* cDNA and the line H6 *tvc* cDNA were compared, one nucleotide difference was found that changed codon 55 (TGC, cysteine) to a termination codon

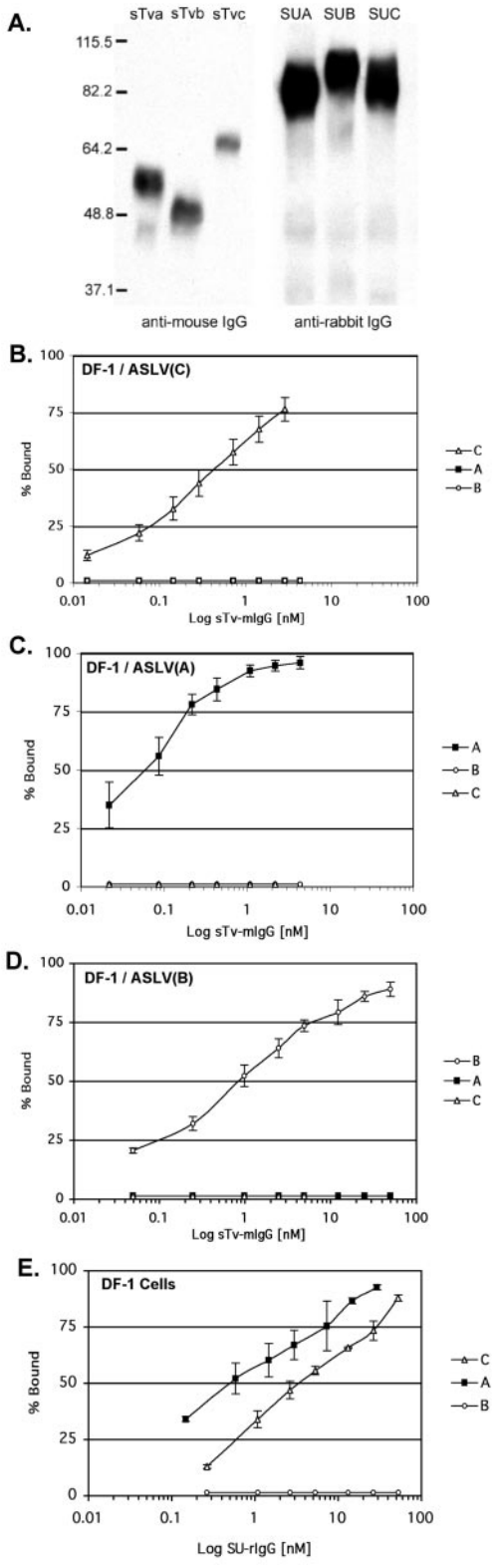


FIG. 6. Binding affinity of the ASLV envelope glycoproteins for ASLV receptors. (Panel A) Western immunoblot analysis of the soluble forms of the Tvc receptor sTvc-mIgG (sTvc) the chicken Tva receptor TvaSsTva-mIgG (sTva), and the Tvb receptor sTvb^{S3}-mIgG (sTvb) immunoprecipitated with anti-mouse IgG-agarose beads and of the secreted forms of the SU glycoproteins SU(C)-rIgG (SUC), SU(A)-rIgG (SUA), and SU(B)-rIgG (SUB) immunoprecipitated with anti-rabbit IgG-agarose beads. The precipitated proteins were denatured, separated by SDS-12% polyacrylamide gel electrophoresis, and transferred to nitrocellulose. The filters were probed with either peroxidase-conjugated goat anti-mouse IgG or goat anti-rabbit IgG, and the bound protein-antibody complexes were visualized by chemiluminescence using Kodak X-Omat film. Molecular masses (in kilodaltons) are given on the left. (Panels B to E) DF-1 cells chronically infected with either ASLV(C) (panel B), ASLV(A) (panel C), or ASLV(B) (panel D) and uninfected DF-1 cells (panel E) were fixed with paraformaldehyde and incubated with different amounts of each soluble receptor (panels B to D) or each secreted SU-rIgG (panel E). The receptor-viral glycoprotein complexes were bound to either goat anti-mouse IgG or goat anti-rabbit IgG linked to phycoerythrin. The amount of phycoerythrin bound to the cells was determined by FACS, and the maximum fluorescence was estimated (see Materials and Methods). The data were plotted as percent maximum fluorescence bound versus soluble receptor sTvc-mIgG (C), sTva-mIgG (A), or sTvb-mIgG (B) concentration (panels B to D) or secreted SU(C)-rIgG (C), SU(A)-rIgG (A), or SU(B)-rIgG (B) concentration (panel E). The values shown are averages and standard deviations from three experiments.

(TGA) (Fig. 2). This would produce a severely truncated Tvc receptor, the signal peptide plus 32 amino acid residues of the mature protein, and abolishes its use as an ASLV(C) receptor. We are designating the chicken line L15 *tvc* gene *tvc^r*. There are now five different mutations identified in the subgroup A to E ASLV receptors that result in resistance to infection by specific ASLV envelope subgroups in inbred White Leghorn chickens (1, 2, 22, 34; this study). The molecular defects encoded by these mutations either alter the structure of the receptor and reduce the binding affinity to the ASLV glycoprotein (e.g., *tva^r* and *tvb^{S3}*) or eliminate the expression of the receptor (e.g., *tvc^r*, *tva^{r2}*, and *tvb^r*). These mechanisms are consistent with the recessive nature of the ASLV-resistant phenotypes.

Distribution of the ASLV receptor transcripts in chicken tissues. To characterize the distribution of Tvc, Tva, and Tvb receptors in chickens, the receptor mRNA expression levels were analyzed by semiquantitative RT-PCR. A fragment of each receptor transcript was amplified from total RNA isolated from a variety of tissues from outbred Brown Leghorn chickens susceptible to ASLV subgroups A, B, and C. The pattern of receptor mRNA expression differs for each receptor (Fig. 7). The *tvc* mRNA is preferentially expressed in thymus, spleen, and bursa, organs involved in immune function. The *tva* mRNA levels are more abundant in ovary and testes, while *tvb* mRNA is more broadly expressed. Although there are differences in receptor mRNA expression levels, *tvc*, *tva*, and *tvb* mRNAs can be detected in all tissues if the cycles of amplification are increased, with the exception that *tva* mRNA was not detected in breast muscle.

DISCUSSION

It is likely that the related subgroup A to E ASLV *env* genes evolved from a single ancestral gene. The ability to use different cellular proteins as receptors would help the virus counter the development of resistance and host receptor variation. Three cell surface proteins have been identified as receptors for the subgroup A to E ASLVs. Tvc, the receptor for sub-

TABLE 2. Estimated binding affinities of soluble forms of the ASLV receptors for ASLV envelope glycoproteins expressed on the surface of infected DF-1 cells and of soluble forms of the ASLV surface glycoproteins for endogenous levels of the ASLV receptors expressed on DF-1 cells

Cells	Apparent K_D (nM)[a]					
	Receptor			Surface glycoprotein		
	sTvc-mIgG	sTva-mIgG	sTvb^{S3}-mIgG	SU(C)-rIgG	SU(A)-rIgG	SU(B)-rIgG
DF-1	—[c]	—	—	3.93 ± 0.26	0.84 ± 0.28	NDB[b]
DF-1/ASLV(C)	0.55 ± 0.20	NDB	NDB	—	—	—
DF-1/ASLV(A)	NDB	0.05 ± 0.02	NDB	—	—	—
DF-1/ASLV(B)	NDB	NDB	0.90 ± 0.25	—	—	—

[a] Apparent K_D values were estimated by fitting the data via nonlinear least squares to a log logistic growth curve function as described in Materials and Methods. Each result is the average and standard deviation from three experiments.
[b] NDB, no detectable binding.
[c] —, binding reaction not performed.

group C ASLVs, is related to a member of the immunoglobulin protein Ig family; Tva, the receptor for subgroup A ASLVs, is related to the LDLR family; and Tvb, the receptor for the subgroup B, D, and E ASLVs, is related to the TNFR family. Except for the fact that all three proteins are ASLV receptors, the proteins have no obvious sequence or structural homology, nor do they appear to be functionally related. However, the ASLV SU glycoproteins have evolved the ability to bind each of the proteins and carry out the multistep process leading to fusion of the viral and cellular membranes. Although it has not yet been proven experimentally, all of the subgroup A to E ASLVs likely require both receptor binding and low pH to complete the fusion process. This dual requirement is, as far as is now known, found only in the ASLV *env* proteins (6). It is likely that the requirement for low pH will require ASLV virions to enter through a particular cellular compartment, which could limit the choice of cellular proteins that could be acceptable ASLV receptors.

While the Ig, LDLR, and TNFR protein families are not homologous, they share several basic characteristics. First, these proteins are all type I, single-transmembrane glycoproteins, which may be important or required for the receptors to interact with ASLV SU. Most other retroviral *env* proteins use receptors that have multiple membrane-spanning domains; the SUs interact with several of the extracellular loops of the receptor to initiate the fusion process, which does not involve a low-pH step. Second, the Ig, LDLR, and TNFR protein families have one or more conserved extracellular domains that may be important for their interactions with the ASLV SU glycoproteins (e.g., the 40-amino-acid LDLR ligand binding domain in Tva and the three cysteine-rich domains in Tvb). These conserved protein domains, which appear to be very different, may present some as-yet-unidentified combination of receptor determinants that enables the evolution of ASLV SU/receptor usage. Testing of this hypothesis will require a more complete understanding of the structures of the *env* proteins and the receptors and their interactions. Third, each ASLV receptor protein is a member of a family of homologous proteins with related functions. Retrovirus infection causes, in infected cells, the synthesis of viral envelope glycoproteins that can down-regulate and/or block the normal functions of the receptor. The related family members may compensate for functional loss of the ASLV receptor protein in infected cells. The normal functions of Tvc, Tva, and Tvb proteins in chickens are not known; however, viral infection presumably does not involve the normal function of the receptor, so a physiologically functional receptor protein is probably not required.

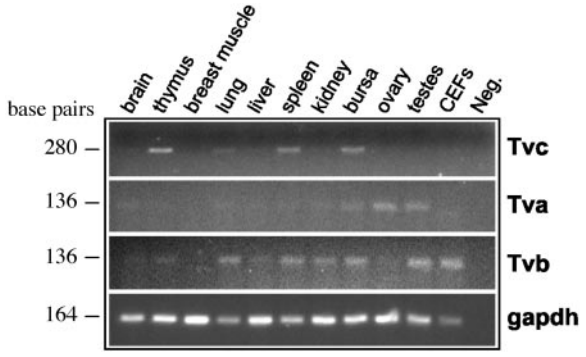


FIG. 7. RT-PCR amplification of *tva*, *tvb*, and *tvc* transcripts from various chicken tissues. The cDNA samples were prepared from various tissues of outbred Brown Leghorn chickens or from CEFs from inbred line M. PCR primers were designed to amplify short regions in the N-terminal regions of the *tva*, *tvb*, and *tvc* genes (see Materials and Methods); the lengths of the expected products are indicated on the left. The figure shows photographs of PCR products after 30 cycles of amplification, separated on agarose gels and stained with ethidium bromide. The 30 PCR cycles were within the exponential phase of product amplification. By increasing the number of amplification cycles, the receptor transcripts were detectable in all tissues (except for *tva* in breast muscle). GAPDH was included as a standard. Neg., no-cDNA PCR control.

The levels of the endogenous Tva and Tvb receptors expressed on avian cells appear to be extremely low. Because the receptor levels are low and the immunological reagents that can be used to detect ASLV SU glycoproteins and their receptors are few, the binding affinity of the ASLV SU/receptor interactions has been estimated using IgG-tagged forms of SU and/or receptor. Using these approaches, we estimate that subgroup C, A, and B ASLV SU glycoproteins bind their respective receptors with low-nanomolar affinities (Fig. 6; Table 2). This affinity may be important for optimal infection efficiency. ASLV(A) variants carrying mutations in SU that significantly reduced the binding affinity for the quail Tva receptor infected quail cells inefficiently (30, 38). Using the SU-rIgG reagents, we were able to detect endogenous Tva and Tvc on DF-1 cells, but not endogenous Tvb. The two experimental approaches used to measure ASLV SU receptor binding affin-

ities produced slightly different estimates of the affinities of these protein-protein interactions (Table 2). This apparent discrepancy may not be surprising, since in each approach, an IgG fusion protein that likely forms a dimer is used as one component of the SU-receptor interaction. However, the natural SU-receptor interaction involves trimeric ASLV glycoproteins binding to monomeric receptors. Despite these caveats, we believe that the estimates of both the relative and the absolute binding affinities of ASLVs for their receptors are reasonable.

The tissue distribution of retroviral receptors may be a determinant of pathogenicity. The ability of ASLVs to efficiently infect cells that express the receptors at extremely low levels has made a systematic characterization of ASLV subgroup A and B tissue tropisms and their possible effect(s) on pathogenicity difficult. We were able to detect *tvc*, *tva*, and *tvb* mRNAs by RT-PCR in almost all the tissues we tested, although there were differences in the expression profiles (Fig. 7). This result supports previous reports that if the bird is genetically susceptible, most chicken tissues can be infected by all three ASLV subgroups (15, 44). If most or all chicken cells express low levels of all three ASLV receptors, the tissue tropism of virus infection could be influenced by the subtle differences in receptor level. However, tropism could also be affected by other factors (sites of integration, long terminal repeat structure, and promoter/enhancer specificity, for example), and receptor expression may not necessarily be the major factor determining tropism or pathology.

The genetic linkage of the *tva* gene and the *tvc* gene in the chicken genome is striking. This close genetic linkage made it possible to use a positional approach to clone Tvc. It had been suggested that because of this linkage that the Tvc receptor would be related to the Tva receptor and that the two loci would derive from gene duplication. However, as this study shows, Tvc and Tva belong to very different protein families. At present, the close proximity of the *tvc* and *tva* genes on chicken microchromosome 28 appears to be a fortuitous coincidence. The *tvb* gene is located on chicken microchromosome 22 (27). This strengthens the argument that the proximity of the *tva* and *tvc* loci on chromosome 28 is an evolutionary accident.

The availability of three distinct receptor-SU pairs makes the ASLV system useful for additional evolutionary studies and structure/function analyses. The fact that all three of the receptors are type I membrane proteins has made it possible to express soluble versions of the receptors; this will make it easier to study the interactions between the receptors and their cognate SUs biochemically and, we hope, structurally. The fact that the structural changes in the ASLV *env* proteins required to cause the fusion of the viral and cellular membranes involve both receptor binding and low pH should make it possible to dissect the molecular details of the events that lead to membrane fusion.

ACKNOWLEDGMENTS

This work was supported in part by grant 523/04/0489 from the Grant Agency of the Czech Republic, project AV0Z50520514 awarded by the Academy of Sciences of the Czech Republic, and National Institutes of Health grant AI48682 and the Mayo Foundation (to M.J.F.).

We thank Stephen Hughes and Jerry Dodgson for discussions and reading of the manuscript, Richard Crooijmans and Martien Groenen for an initial PCR screen of the Wageningen BAC library, Jerry Dodgson and Michael Romanov for providing BAC library filter sets and BAC clone DNAs, Jean-Marie Buerstedde and Hiroshi Arakawa for training in the DT40 knockout technology, and Lisa Stubbs for sharing unpublished chicken genomic sequence data. We thank all the personnel of the Koleč animal farm and Audelia Munguia for excellent technical assistance.

REFERENCES

1. **Adkins, H. B., S. C. Blacklow, and J. A. T. Young.** 2001. Two functionally distinct forms of a retroviral receptor explain the nonreciprical receptor interference among subgroups B, D, and E avian leukosis viruses. J. Virol. **75:**3520–3526.
2. **Adkins, H. B., J. Brojatsch, J. Naughton, M. M. Rolls, J. M. Pesola, and J. A. T. Young.** 1997. Identification of a cellular receptor for subgroup E avian leukosis virus. Proc. Natl. Acad. Sci. USA **94:**11617–11622.
3. **Adkins, H. B., J. Brojatsch, and J. A. T. Young.** 2000. Identification and characterization of a shared TNFR-related receptor for subgroup B, D, and E avain leukosis viruses reveal cysteine residues required specifically for subgroup E viral entry. J. Virol. **74:**3572–3578.
4. **Arakawa, H., D. Lodygin, and J.-M. Buerstedde.** 2001. Mutant loxP vectors for selectable marker recycle and conditional knock-outs. BMC Biotechnology **1:**7.
5. **Baker, A., and M. Cotten.** 1997. Delivery of bacterial artificial chromosomes into mammalian cells with psoralen-inactivated adenovirus carrier. Nucleic Acid Res. **25:**1950–1956.
6. **Barnard, R. J., and J. A. T. Young.** 2003. Alpharetrovirus envelope-receptor interactions. Curr. Top. Microbiol. Immunol. **281:**107–136.
7. **Bateman, A., L. Coin, R. Durbin, R. D. Finn, V. Hollich, S. Griffiths-Jones, A. Khanna, M. Marshall, S. Moxon, E. L. Sonnhammer, D. J. Studholme, C. Yeats, and S. R. Eddy.** 2004. The Pfam protein famalies database. Nucleic Acid Res. **32:**D138–D141.
8. **Bates, P., J. A. T. Young, and H. E. Varmus.** 1993. A receptor for subgroup A Rous sarcoma virus is related to the low density lipoprotein receptor. Cell **74:**1043–1051.
9. **Belanger, C., K. Zingler, and J. A. Young.** 1995. Importance of cysteines in the LDLR-related domain of the subgroup A avian leukosis and sarcoma virus receptor for viral entry. J. Virol. **69:**1019–1024.
10. **Bendtsen, J. D., H. Nielson, G. von Heijne, and S. Brunak.** 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. **340:**783–795.
11. **Boardman, P. E., J. Sanz-Ezquerro, I. M. Overton, D. W. Burt, E. Bosch, W. T. Fong, C. Tickle, W. R. Brown, S. A. Wilson, and S. J. Hibbard.** 2002. A comprehensive collection of chicken cDNAs. Curr. Biol. **12:**1965–1969.
12. **Bova, C. A., J. P. Manfredi, and R. Swanstrom.** 1986. Env genes of avian retroviruses: nucleotide sequence and molecular recombinants define host range determinants. Virology **152:**343–354.
13. **Bova, C. A., J. C. Olsen, and R. Swanstrom.** 1988. The avian retrovirus *env* gene family: molecular analysis of host range and antigenic variants. J. Virol. **62:**75–83.
14. **Brojatsch, J., J. Naughton, M. M. Rolls, K. Zingler, and J. A. T. Young.** 1996. CAR1, a TNFR-related protein, is a cellular receptor for cytopathic avain leukosis-sarcoma viruses and mediates apoptosis. Cell **87:**845–855.
15. **Brown, D. W., and H. L. Robinson.** 1988. Influence of *env* and long terminal repeat sequences on the tissue tropism of avian leukosis viruses. J. Virol. **62:**4828–4831.
16. **Buerstedde, J.-M., and S. Takeda.** 1991. Increased ratio of targeted to random integration after transfection of chicken B cell lines. Cell **67:**179–188.
17. **Cogburn, L. A., R. Morgan, and J. Burnside.** 2003. Expressed sequence tags, DNA chip technology and gene expression profiling, p. 629–645. *In* W. M. Muir and S. E. Aggrey (ed.), Poultry genetics, breeding and biotechnology. CABI Publishing, Wallingford, United Kingdom.
18. **Compte, E., P. Pontarotti, Y. Collette, M. Lopez, and D. Olive.** 2004. Characterization of BT3 molecules belonging to the B7 family expressed immune cells. Eur. J. Immunol. **34:**2089–2099.
19. **Diamond, L.** 1967. Two spontaneously transformed cell lines derived from the same hamster embryo culture. Int. J. Cancer **2:**143–152.
20. **Dorner, A. J., and J. M. Coffin.** 1986. Determinants for receptor interaction and cell killing on the avian retrovirus glycoprotein gp85. Cell **45:**365–374.
21. **Dorner, A. J., J. P. Stoye, and J. M. Coffin.** 1985. Molecular basis of host range variation in avian retroviruses. J. Virol. **53:**32–39.
22. **Elleder, D., D. C. Melder, K. Trejbalova, J. Svoboda, and M. J. Federspiel.** 2004. Two different molecular defects in the Tva receptor gene explain the resistance of two tva^r lines of chickens to infection by subgroup A avain sarcoma and leukosis viruses. J. Virol. **78:**13489–13500.
23. **Elleder, D., J. Plachy, J. Hejnar, J. Geryk, and J. Svoboda.** 2004. Close linkage of genes encoding receptors for subgroups A and C of avian sarcoma/leucosis virus on chicken chromosome 28. Anim. Genet. **35:**176–181.
24. **Federspiel, M. J., and S. H. Hughes.** 1997. Retroviral gene delivery. Methods Cell Biol. **52:**179–214.
25. **Green, S., I. Issemann, and E. Sheer.** 1988. A versatile in vivo and in vitro

eukaryotic expression vector for protein engineering. Nucleic Acids Res. **16:**369.
26. **Hala, K., and J. Plachy.** 1997. Inbred strains of chickens, p. 2286–2292. *In* I. Lefkovitz (ed.), Immunology methods manual. Academic Press, London, United Kingdom.
27. **Hillier, L. W., W. Miller, E. Birney, et al.** 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertabrate evolution. Nature **432:**695–716.
28. **Himly, M., D. N. Foster, I. Bottoli, J. S. Iacovoni, and P. K. Vogt.** 1998. The DF-1 chicken fibroblast cell line: transformation induced by diverse oncogenes and cell death resulting from infection by avian leukosis viruses. Virology **248:**295–304.
29. **Holmen, S. L., and M. J. Federspiel.** 2000. Selection of a subgroup A avian leukosis virus [ALV(A)] envelope resistant to soluble ALV(A) surface glycoprotein. Virology **273:**364–373.
30. **Holmen, S. L., D. C. Melder, and M. J. Federspiel.** 2001. Identification of key residues in subgroup A avian leukosis virus envelope determining receptor binding affinity and infectivity of cells expressing chicken or quail Tva receptor. J. Virol. **75:**726–737.
31. **Holmen, S. L., D. W. Salter, W. S. Payne, J. B. Dodgson, S. H. Hughes, and M. J. Federspiel.** 1999. Soluble forms of the subgroup A avian leukosis virus [ALV(A)] receptor Tva significantly inhibit ALV(A) infection in vitro and in vivo. J. Virol. **73:**10051–10060.
32. **Hunter, E.** 1997. Viral entry and receptors, p. 71–120. *In* J. M. Coffin, S. H. Hughes, and H. E. Varmus (ed.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
33. **Kingston, R. E., C. A. Chen, and H. Okayama.** 1989. Introduction of DNA into eukaryotic cells, p. 911–919. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology, vol. 1. John Wiley & Sons, Inc., New York, N.Y.
34. **Klucking, S., H. B. Adkins, and J. A. T. Young.** 2002. Resistance of infection by subgroup B, D, and E avian sarcoma and leukosis viruses is explained by a premature stop codon within a resistance allele of the *tvb* receptor gene. J. Virol. **76:**7918–7921.
35. **Klucking, S., and J. A. T. Young.** 2004. Amino acid residues Tyr-67, Asn-72, and Asp-73 of the Tvb receptor are important for subgroup E avian sarcoma and leukosis virus interaction. Virology **318:**371–380.
36. **Kvaratskhelia, M., P. K. Clark, S. Hess, D. C. Melder, M. J. Federspiel, and S. H. Hughes.** 2004. Identification of glycosylation sites in the SU component of the avian sarcoma/leukosis virus envelope glycoprotein (subgroup A) by mass spectrometry. Virology **326:**171–181.
37. **Marchler-Bauer, A., J. B. Anderson, C. DeWeese-Scott, N. D. Fedorova, L. Y. Geer, S. He, D. I. Hurwitz, J. D. Jackson, A. R. Jacobs, C. J. Lanczycki, C. A. Liebert, C. Liu, T. Madej, G. H. Marchler, R. Mazumder, A. N. Nikolskaya, A. R. Panchenko, B. S. Rao, B. A. Shoemaker, V. Simonyan, J. S. Song, P. A. Thiessen, S. Vasudevan, Y. Wang, R. A. Yamashita, J. J. Yin, and S. H. Bryant.** 2003. CDD: a curated Entrez database of conserved domain alignments. Nucleic Acids Res. **31:**383–387.
38. **Melder, D. C., V. S. Pankratz, and M. J. Federspiel.** 2003. Evolutionary pressure of a receptor competitor selects different subgroup A avian leukosis virus escape variants with altered receptor interactions. J. Virol. **77:**10504–10514.
39. **Ogg, S. L., A. K. Weldon, L. Dobbie, A. J. H. Smith, and I. H. Mather.** 2004. Expression of butyrophilin (Btn1a1) in lactating mammary gland is essential for the regulated secretion of milk-lipid droplets. Proc. Natl. Acad. Sci. USA **101:**10084–10089.
40. **Pani, P. K.** 1974. Close linkage between genetic loci controlling response to fowl subgroup A and subgroup C sarcoma viruses. J. Gen. Virol. **22:**187–195.
41. **Payne, L. N., and P. K. Pani.** 1971. Evidence for linkage between genetic loci controlling response of fowl to subgroup A and subgroup C sarcoma viruses. J. Gen. Virol. **13:**253–259.
42. **Rainey, G. J. A., A. Natonson, L. F. Maxfield, and J. M. Coffin.** 2003. Mechanisms of avian retroviral host range extension. J. Virol. **77:**6709–6719.
43. **Rhodes, D. A., M. Stammers, G. Malcherel, S. Beck, and J. Trowsdale.** 2001. The cluster of BTN genes in the extended major histocompatibility complex. Genomics **71:**351–362.
44. **Robinson, H. L., L. Ramamoorthy, K. Collart, and D. W. Brown.** 1993. Tissue tropism of avian leukosis viruses: analyses for viral DNA and proteins. Virology **193:**443–445.
45. **Rong, L., and P. Bates.** 1995. Analysis of the subgroup A avian sarcoma and leukosis virus receptor: the 40-residue, cysteine-rich, low-density lipoprotein receptor repeat motif of Tva is sufficient to mediate viral entry. J. Virol. **69:**4847–4853.
46. **Rong, L., A. Edinger, and P. Bates.** 1997. Role of basic residues in the subgroup-determining region of the subgroup A avian sarcoma and leukosis virus envelope in receptor binding and infection. J. Virol. **71:**3458–3465.
47. **Rong, L., K. Gendron, B. Strohl, R. Shenoy, R. J. Wool-Lewis, and P. Bates.** 1998. Characterization of determinants for envelope binding and infection in Tva, the subgroup A avian sarcoma and leukosis virus receptor. J. Virol. **72:**4552–4559.
48. **Schaefer-Klein, J., I. Givol, E. V. Barsov, J. M. Whitcomb, M. VanBrocklin, D. N. Foster, M. J. Federspiel, and S. H. Hughes.** 1998. The EV-0-derived cell line DF-1 supports efficient replication of avian leukosis-sarcoma viruses and vectors. Virology **248:**305–311.
49. **Smith, E. J., A. M. Fadly, and W. Okazaki.** 1979. An enzyme-linked immunoabsorbant assay for detecting avian leukosis sarcoma viruses. Avian Dis. **23:**698–707.
50. **Taplitz, R. A., and J. M. Coffin.** 1997. Selection of an avian retrovirus mutant with extended receptor usage. J. Virol. **71:**7814–7819.
51. **Tsichlis, P. N., and J. M. Coffin.** 1980. Recombinants between endogenous and exogenous avian tumor viruses: role of the C region and other portions of the genome in the control of replication and transformation. J. Virol. **33:**238–249.
52. **Tsichlis, P. N., K. F. Conklin, and J. M. Coffin.** 1980. Mutant and recombinant avian retroviruses with extended host range. Proc. Natl. Acad. Sci. USA **77:**536–540.
53. **Weiss, R.** 1982. Experimental biology and assay of RNA tumor viruses, p. 209–260. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses, vol. 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
54. **Weiss, R. A.** 1992. Cellular receptors and viral glycoproteins involved in retrovirus entry, p. 1–108. *In* J. A. Levy (ed.), The retroviruses, vol. 2. Plenum Press, New York, N.Y.
55. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
56. **Young, J. A. T.** 2001. Virus entry and uncoating, p. 87–103. *In* D. M. Knipe and P. M. Howley (ed.), Fields virology, vol. 2. Lippincott Williams & Wilkins, Philadelphia, Pa.
57. **Young, J. A. T., P. Bates, and H. E. Varmus.** 1993. Isolation of a chicken gene that confers susceptibility to infection by subgroup A avian leukosis and sarcoma viruses. J. Virol. **67:**1811–1816.
58. **Zingler, K., C. Belanger, R. Peters, D. Agard, and J. A. T. Young.** 1995. Identification and characterization of the viral interaction determinant of the subgroup A avian leukosis virus receptor. J. Virol. **69:**4261–4266.
59. **Zingler, K., and J. A. T. Young.** 1996. Residue Trp-48 of Tva is critical for viral entry but not for high-affinity binding to the SU glycoprotein of subgroup A avian leukosis and sarcoma viruses. J. Virol. **70:**7510–7516.

REPRODUCTION

RESEARCH

Retrovirus-mediated *in vitro* gene transfer into chicken male germ line cells

Jiří Kalina, Filip Šenigl[1], Alena Mičáková, Jitka Mucksová, Jana Blažková[1], Haifeng Yan[2], Martin Poplštein, Jiří Hejnar[1] and Pavel Trefil

BIOPHARM, Research Institute of Biopharmacy and Veterinary Drugs Ltd, 254 49 Jílové u Prahy, Czech Republic, [1]Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Vídeňská 1083, CZ-14220 Prague 4, Czech Republic and [2]HIAVS, Hunan Institute of Animal and Veterinary Science, Quantang, Changsha 410131, Hunan, China

Correspondence should be addressed to J Hejnar; Email: hejnar@img.cas.cz
P Trefil; Email: trefil@bri.cz

Abstract

Chicken testicular cells, including spermatogonia, transplanted into the testes of recipient cockerels sterilized by repeated γ-irradiation repopulate the seminiferous epithelium and resume the exogenous spermatogenesis. This procedure could be used to introduce genetic modifications into the male germ line and generate transgenic chickens. In this study, we present a successful retroviral infection of chicken testicular cells and consequent transduction of the retroviral vector into the sperm of recipient cockerels. A vesicular stomatitis virus glycoprotein G-pseudotyped recombinant retroviral vector, carrying the enhanced green fluorescent protein reporter gene was applied to the short-term culture of dispersed testicular cells. The efficiency of infection and the viability of infected cells were analyzed by flow cytometry. No significant CpG methylation was detected in the infected testicular cells, suggesting that epigenetic silencing events do not play a role at this stage of germ line development. After transplantation into sterilized recipient cockerels, these retrovirus-infected testicular cells restored exogenous spermatogenesis within 9 weeks with approximately the same efficiency as non-infected cells. Transduction of the reporter gene encoding the green fluorescent protein was detected in the sperms of recipient cockerels with restored spermatogenesis. Our data demonstrate that, similarly as in mouse and rat, the transplantation of retrovirus-infected spermatogonia provides an efficient system to introduce genes into the chicken male germ line.

Reproduction (2007) **134** 445–453

Introduction

The availability of genetic models created through transgenesis and gene targeting is the major advantage of using the mouse over all other experimental animals. Due to the lack of immortal and pluripotent embryonic stem (ES) cells in species other than the mouse, efficient gene transfer and gene knockout technology is not feasible. In the chicken, particularly, this task is further complicated by specific features of the reproductive system, which make the newly fertilized zygote relatively inaccessible and difficult to handle.

Because of limited success in cultivation and genetic modification of chicken ES cells (Pain *et al.* 1996, Petitte *et al.* 2004), the current methods of chicken transgenesis (Sang 2004) concentrate on blastodermal cells of stage X embryo (Koo *et al.* 2004, Kwon *et al.* 2004, McGrew *et al.* 2004, Lillico *et al.* 2007) and primordial germ cells (PGCs; Vick *et al.* 1993, Naito *et al.* 1999, van de Lavoir *et al.* 2006). Spermatogonial stem cells (SSCs) might serve as an alternative to ES cells provided that they could be genetically modified *in vitro* and transplanted into recipient cockerels. This approach of germ line gene transfer could eliminate the background of mosaic animals and produce hemizygously transgenic animals in F1 generation.

Mouse SSCs are seldomly found among freshly explanted testicular cells, but expand *in vivo* in a functional transplantation assay (Ogawa *et al.* 2000) as well as *in vitro* (Nagano *et al.* 2003) and can be infected with a replication-defective ecotropic retroviral vector (Nagano *et al.* 2000). Growth conditions essential for self-renewal and expansion of SSCs in the mouse and even in rats have been elaborated (Kanatsu-Shinohara

ISSN 1470–1626 (paper) 1741–7899 (online)

DOI: 10.1530/REP-06-0233
Online version via www.reproduction-online.org

et al. 2003, Kubota *et al.* 2004, Hamra *et al.* 2005 respectively). In such a way, genetically modified mouse SSCs restore fertility after long-term *in vitro* culture (Kanatsu-Shinohara *et al.* 2003, 2004, Kubota *et al.* 2004). In contrast, avian spermatogenesis is poorly understood (Thurston & Korn 2000) and SSCs as well as other germ cells of the seminiferous epithelium have yet to be well characterized in birds. Recently, we have demonstrated that the transfer of dispersed testicular cells from fertile donor cockerels resulted in the colonization of infertile recipient testes with final production of viable and fertilization-competent spermatozoa (Trefil *et al.* 2006). Therefore, efficient gene delivery into the crude mix of testicular cells may target the rare SSCs and pave the way to chicken transgenesis. The exact requirements for the maintenance of chicken SSCs in culture are not yet known. Under such circumstances, infection with a high-titer virus vector seems to be the only choice for gene delivery because virus entry needs only a short time to displace the SSCs from their protected niche in the seminiferous tubule. In addition, infection with retrovirus in comparison with transfection, lipofection, or electroporation techniques minimizes the harsh chemical or physical influences leading to the decrease in the transplantation capacity of SSCs. Initially, replication-competent avian sarcoma and leukosis virus (ASLV)-derived and reticuloendotheliosis virus-derived vectors (Salter *et al.* 1987, Bosselman *et al.* 1989) were used, and since that time, retroviral vectors have become a common gene transfer vehicle for chicken cells. Recent studies show that replication-defective retroviral vectors can be used to introduce foreign genes into the avian germ line and that the major approaches to avian transgenesis have been made through retroviral and lentiviral vectors (Kwon *et al.* 2004, McGrew *et al.* 2004, Lillico *et al.* 2007).

In this report, we demonstrate that chicken testicular cells can be efficiently infected *in vitro* with a replication-defective reporter retrovirus vector pantropized by vesicular stomatitis virus envelope glycoprotein (VSV-G). All cell types present in the mix of testicular cells express the transduced reporter enhanced green fluorescent protein (EGFP) after *in vitro* cultivation and freshly explanted infected testicular cells restore spermatogenesis in sterilized recipient cockerels. The presence of the transduced reporter gene in the sperm of transplanted recipients suggests that the application of these techniques might be useful for transgenesis in chicken.

Materials and Methods

Animals

Inbred Black Minor (BM; genotype ii, EE, b/b) and White Leghorn (WL; genotype II) cockerels at the age of 32 weeks were used as donors and recipients of testicular cells respectively, as successful transplantation of testicular cells from BM into WL cockerels was demonstrated previously in our chicken germ cell transplantation model (Trefil *et al.* 2003, 2006). Three males of the same genetic origin as recipients were retained as positive controls and other three males were subjected to the same sterilizing treatment as recipient males but kept untransplanted as negative controls. All animals used in the experiments were obtained from the Experimental Animal Farm of the Institute of Molecular Genetics (Prague). Only animals of standard weight for these breeds and known fertility were used in the experiments. Cockerels were kept in individual cages (4200 cm^2) fitted with perches under standard husbandry conditions. Feed and water were provided *ad libitum* and a 12 h light:12 h darkness photoperiod was arranged. Experimental animals were killed by cervical dislocation and decapitation. All experiments were performed in accordance with Czech legal requirements for animal handling.

Irradiation treatment

The radiation treatment unit Theratron T1000 (Theratronics, Kanata, Canada) was used to irradiate testes of recipient WL cockerels according to the protocol described previously by Trefil *et al.* (2003).

Construction of the retroviral vector

The retroviral vector was assembled by insertion of the EGFP coding sequence from the pEGFP plasmid into the pLPCX vector designed for retroviral gene delivery and expression (both plasmids obtained from BD Biosciences Clontech, Palo Alto, CA, USA). The pEGFP was digested with EcoRI and HindIII and the 791 bp fragment was ligated into the multiple cloning site of pLPCX cleaved with EcoRI and HindIII. To eliminate the puromycin resistance gene and the internal P_{CMVIE} promoter, the resulting vector was partially digested with AgeI, the single cut fragment was isolated and digested with XhoI. The overhangs were removed by treatment with mung bean nuclease and the 5658 bp fragment was isolated and self-ligated. The resulting pLG vector (Fig. 1) was used for virus production. The 5′ LTR of the pLG vector is derived from Moloney murine sarcoma virus (MoMuSV), whereas the 3′ LTR comes from Moloney murine leukemia virus (MoMuLV).

GP-293 cell culture, virus propagation, and fluorescence microscopy

The HEK 293-based cell line GP-293 (purchased from BD Biosciences Clontech) was maintained in F-12 and MEM-D mixed in the ratio of 1:1 (Sigma) supplemented

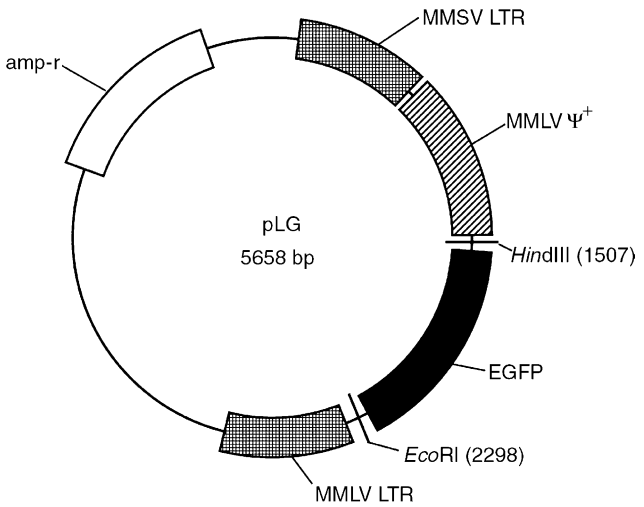


Figure 1 Schematic structure of the reporter retrovirus. The construct is based on the parental vector pLPCX, in which the internal P_{CMVIE} promoter was deleted and puromycin resistance gene is replaced by the EGFP gene. The positions of HindIII and EcoRI sites used for the EGFP cloning are depicted. Not in the appropriate scale.

with 5% calf serum and antibiotics in 3% CO_2 atmosphere at 37 °C.

For virus production, the GP-293 cells were cultivated on a 150 mm Petri dish. The cells were calcium phosphate co-transfected with 50 μg pLG and 10 μg pVSV-G (BD Biosciences Clontech). The medium containing the virus was collected 24, 36, and 48 h after transfection. The virus stock was clarified by centrifugation (150 ***g*** at 4 °C for 10 min). The supernatant was centrifuged in the SW28 rotor (Beckman Instruments, Fullerton, CA, USA) at 73 000 ***g*** at 4 °C for 2 h to pellet the virus. After centrifugation, the supernatant was aspirated and the pelleted virus was resuspended in a small volume of cultivation medium and stored in aliquots at −80 °C. The virus titer was determined by infecting chicken embryo fibroblasts with diluted virus stock and counting GFP-positive cell clusters in the infected culture under a fluorescence microscope (Olympus IX 51) with excitation wavelength 488 nm and emission wavelength 507 nm. Typically, virus stocks containing 5×10^7 infection units (IU) per ml were obtained.

Retrovirus infection of testicular cells

Dispersed testicular cells from BM male donors were prepared using a two-step enzymatic digestion as described by Bastos *et al.* (2005). Testicular cell cultures were infected with concentrated virus stocks at the multiplicity of infection (MOI) as high as 90. Infections were carried out in a small volume of media for 60–120 min at 41 °C in 5% CO_2 atmosphere. Afterwards, complete cultivation medium was added to the normal volume without removing the virus. Infected testicular cells were either used for further cultivation and expression analysis or transferred immediately into sterilized recipient cockerels.

Flow cytometry and classification of cell ploidy

The infected suspension of testicular cells was subjected to fluorescence activated cell analysis (FACS) using a FACSVantage SE apparatus (BD Biosciences) at 48 and 120 h post-infection. We assessed the proportion of EGFP-positive cells, measured the ploidy classes of testicular cells by staining with Hoechst 33342 (H42; Sigma) and verified their viability by counterstaining with propidium iodide (PI; 2 μg/ml).

Cells were harvested using trypsin/EDTA solution (0.05% trypsin in 0.5 mM EDTA) and resuspended in HBBS. For detection of the cell ploidy, testicular cells were further stained with H42 at a final concentration of 10 μM for 30 min at room temperature. H42 and PI were excited by u.v. helium/neon laser 50 mW (Coherent Enterprise, Santa Paula, CA, USA). H42 blue was collected using a combination of 485/58 nm long pass and 505 nm short pass filters in front of the first detector. PI and H42 red fluorescences were detected with a 660 nm/20 nm band pass filter in front of the second detector. GFP was excited with a second 488 nm laser and fluorescence was collected with a 530/30 nm band pass filter.

Genomic DNA analysis

Genomic DNA was extracted from 5-day-old culture of testicular cell mixture and from ejaculates of recipient cockerels after transplantation of infected donor testicular cells using a genomic DNA purification kit (DNeasy Tissue Kit, Qiagen) or using a standard process of phenol purification. The N-terminal primer (5′-TGACCCT-GAAGTTCATCTGCA-3′) and the C-terminal primer (5′-ACGAACTCCAGCAGGACCATGT-3′) correspond to the pEGFP nucleotide sequences that yield a 546 bp DNA fragment. Each reaction mixture contained 1 μg genomic DNA extract, 50 pmol of each primer, 0.2 mM dNTP, 5 μl complete 10× PCR buffer, 2.5 U Taq polymerase (Top Bio), and the reaction volume was adjusted to 50 μl with ddH_2O. The samples were denatured at 95 °C for 5 min and then subjected to 36 cycles of PCR amplification. Each cycle was as follows: denaturation at 94 °C for 30 s, primer annealing at 60 °C for 30 s, and extension at 72 °C for 30 s.

Bisulfite cytosine methylation analysis

DNA samples from virus-infected testicular cells were isolated by phenol–chloroform extraction and digested by HindIII restriction endonuclease. Bisulfite treatment of DNA was performed according to Hájková *et al.* (2002). We amplified the 5′ LTR of integrated proviruses as a 525 bp DNA fragment encompassing 21 CpGs from

bisulfite-treated DNA by nested PCR: forward external primer (A) 5′-GAATAGATGGAATAGTTGAATATGGG-3′ (nucleotides –342 to –316 upstream to the transcription start), forward internal primer (B) 5′-GGGTTAAATA-GGATATTTGTGGTAAGTAG-3′ (nucleotides –319 to –290 upstream to the transcription start), reverse internal primer (C) 5′-AACCCCCAAATAAAAAACCC-3′ (nucleotides 133–153 downstream to the transcription start), and reverse external primer (D) 5′-CCTAAAC-AAAAATCTCCAAATCC-3′ (nucleotides 160–183 downstream to the transcription start). The first round of PCR contained ~25 ng DNA. Taq polymerase (Takara, Shiga, Japan) was used following manufacturer's recommendation with 5 μM $MgCl_2$, 0.9 M betaine, and 0.9% dimethylsulfoxide. Conditions for PCR were as follows: 95 °C for 1 min, 58 °C for 2 min, and 72 °C for 1 min (40 cycles). The second round of PCR started with 1 μl of the first-round PCR product and the PCR conditions were identical. PCR products were subsequently cloned using the pGEM-T vector cloning system (Promega). Individual PCR clones were sequenced using the AmpliTaq FS L′ Big Dye Terminator sequencing kit (Applera, Norwalk, CT, USA) with universal pUC/M13 forward and reverse primers. Only PCR clones with at least 95% conversion of Cs outside CpGs were taken into account. Only PCR clones containing the hybrid MoMuSV/MoMuLV LTR sequence were used for further analysis in order to discern between reversely transcribed and integrated proviral DNA and original plasmid DNA persisting in the infected cells.

Transplantation of retroviral-infected testicular cells

Approximately, 300 μl of the dispersed virus-infected testicular cells with densities varying from 10^5 to 10^6 cell/ml were injected into both testes of anesthetized recipient cockerels WL. The injection needle was inserted through the tunica albuginea directly into the testes and donor cell suspension was applied. Recipient cockerels were anesthetized by i.m. injection of 15 mg/kg ketamine (Narkamon, Spofa, Czech Republic) and 4 mg/kg xylazine (Rometar, Spofa, Czech Republic) prior to transfers.

Sperm quality assessment

The repopulation of the seminiferous tubules and the onset of spermatogenesis in the testes of recipient WL cockerels were assessed by light microscopy. Ejaculates from transplanted and control cockerels were collected at 1-week-interval using the conventional abdominal massage technique (Burrows & Quinn 1937). Progressive sperm motility was assessed and rated in the five-grade scale, where a rating of five means very good motility and a rating of one is a very poor motility score. Sperm concentration was determined after centrifugation in microcapillaries (Bonitz 1970).

Statistical analysis

The best regression equation for the adjustment of the time course of spermatozoa concentration from recipient males transplanted with adult dispersed testicular cells was achieved using a linear model with polynomial interpolation (Meloun *et al.* 1994). Comparison between the regression equations was performed using the Barlett and Chow test.

Results

Infection of dispersed testicular cells in vitro *with a reporter retrovirus*

We have developed an experimental system to determine whether chicken SSCs could be infected with a pantropic retrovirus *in vitro* after being displaced from their protected niche in the seminiferous epithelium. The basic features of this system are: (i) non-replicative character of the virus vector, which makes it a suitable model for approaching transgenesis in chicken; (ii) transduction of the *EGFP* reporter gene, which serves for simple detection of retroviral expression; (iii) pantropic character of infectious virus particles that can enter cells without any receptor restrictions due to the VSV-G (Emi *et al.* 1991); and (iv) differences between DNA sequences of 5′ and 3′ LTRs, which make it possible to discern the integrated proviruses versus episomal DNA of the cloned retrovirus. The schematic structure of the reporter retrovirus cloned in pLG construct is shown in Fig. 1. The infectious retrovirus was propagated in GP293 packaging cells lacking any envelope sequence in their helper retroviral construct. After co-transfection of pLG together with a vector expressing VSV-G, we obtained stocks of infectious virus, which was concentrated 50 times by ultracentrifugation. The titer of the concentrated viral stocks was determined by infection of chicken embryo fibroblasts and typically reached 10^7–5×10^7 IU/ml.

The testicular cell suspension was infected by the retroviral vector at the MOI reaching 90 infectious virus particles per cell. We followed the EGFP fluorescence in *in vitro* cultivated testicular cells from five donor cockerels for 5 days with the first inspection 8 h post-infection. A substantial fraction of cells emitted green light at that time. The percentage of EGFP-positive cells increased during the *in vitro* cultivation, with gradual overgrowth of EGFP-positive colonies (Fig. 2A and B). We counted the percentage of EGFP-positive cells and assessed the transduction efficiency by flow cytometry analysis 48 h after infection. We were able to identify at least 20% of EGFP-positive cells in the whole testicular cell suspension. After 5 days of cultivation, we were able

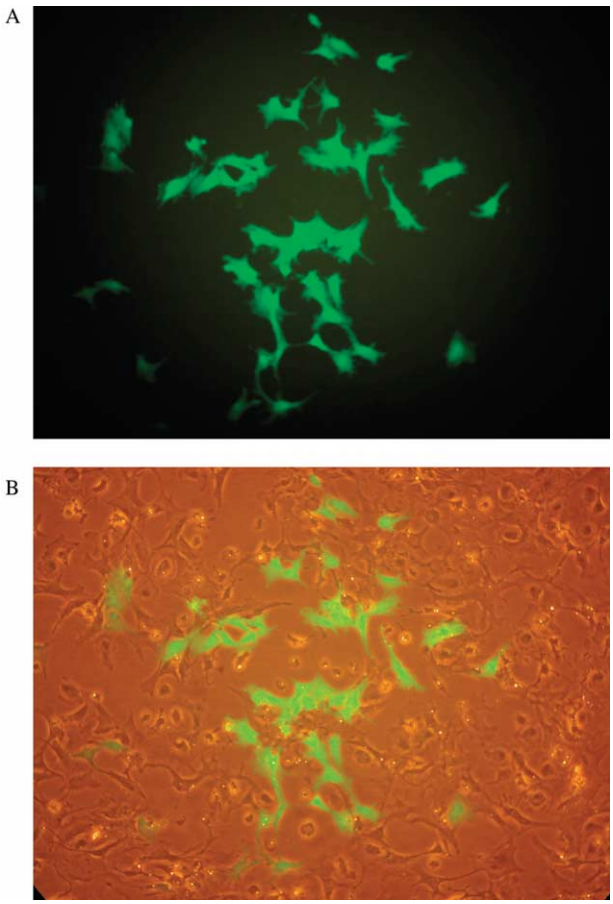


Figure 2 Four-day-old testicular cell culture 48 h after infection with the EGFP-transducing retroviral vector. Images from fluorescence microscopy (A) and from phase contrast with low level of fluorescence (B) illustrate the percentage of EGFP-positive cells (magnification 200×).

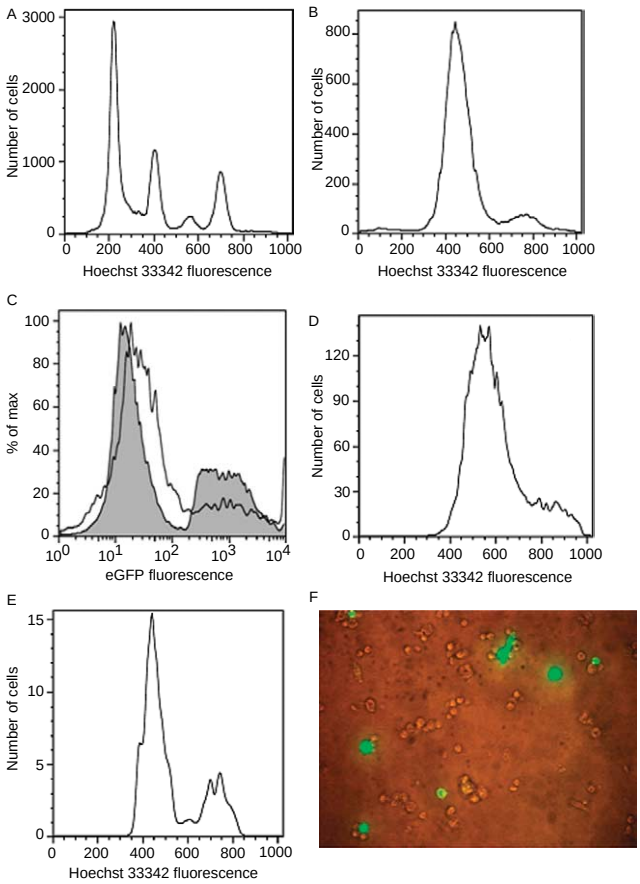


Figure 3 EGFP expression and ploidy classification in populations of infected testicular cells. Using the H42 dye, the ploidy of (A) 2-day and (B) 5-day cell cultures was analyzed by flow cytometry. (C) The percentage of EGFP-positive cells in the entire population of infected testicular cells (the gray part of the graph represents 5-day culture and the white part represents 2-day culture). The ploidy of selected EGFP-positive cells after (E) 2-day and (D) 5-day culture was classified separately by flow cytometry and H42 staining. (F) The image of cells FACS-sorted by H42 2 days after infection from phase contrast with low level of fluorescence.

to detect ∼37% EGFP-positive cells (Fig. 3C). The viability of these suspension cells was very high as measured by the PI exclusion. As much as 99.7% of cells from the infected culture of testicular cells were PI-negative. This clearly demonstrates that EGFP expression is non-toxic for chicken testicular cells. Using H42 fluorescence, cell cycles of 2-day (A) and 5-day (B) cell cultures were analyzed.

During the second day of cultivation of testicular cells (Fig. 3A), we were able to detect 20% of diploid, 16% of tetraploid, and 44% of haploid cells. After the fifth day, however, the character of the cell populations changed (Fig. 3B). The haploid population represented by spermatides was no longer detected. On the contrary, the proportion of diploid cells increased significantly up to 85%.

In 20% (2-day culture) or 37% (5-day culture) of EGFP-expressing testicular cells (Fig. 3C), that were stained with H42, we were able to identify changes mainly in the number of diploid cells. In the 2-day culture, it was possible to identify various classes of ploidy (Fig. 3E), while in the 5-day culture, we were able to detect mainly diploid cells (Fig. 3D). When the fraction of haploid cells from the infected 2-day culture stained with H42 was sorted by FACS, we have observed round cells representing the stage of round spermatids. EGFP-positive cells were found among these cells (Fig. 3F), but the percentage is clearly lower in comparison with the primary testicular cells.

CpG methylation of the retroviral promoter region in cultivated testicular cells

To analyze the CpG methylation of integrated proviruses, we performed the bisulfite sequencing of proviral DNA in testicular cells cultivated *in vitro* for 5 days. We focused on the 5′ LTR, which contains 21 CpG dinucleotides scattered along the whole LTR, with

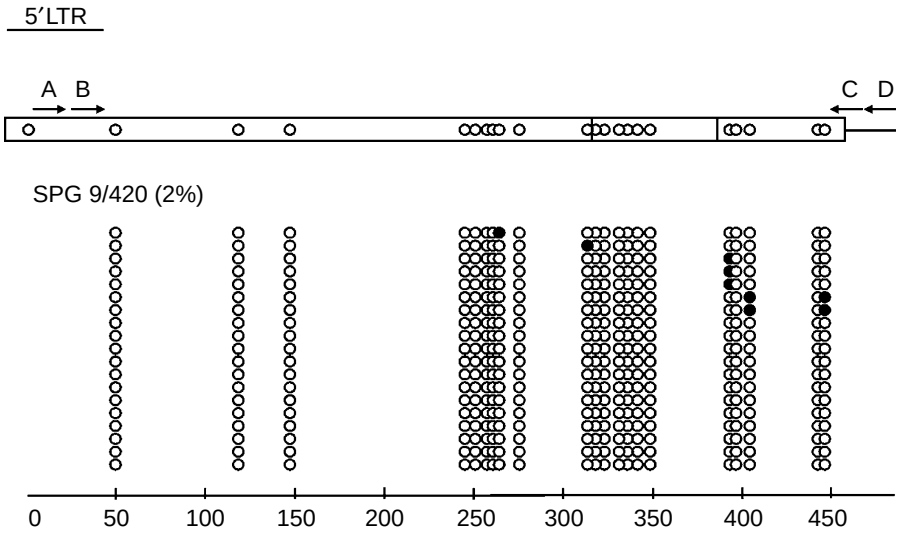


Figure 4 CpG methylation of the retroviral promoter region in cultivated testicular cells. We performed bisulfite sequencing to analyze CpG methylation of the integrated proviruses. We have focused our attention to the 5′ LTR, which contains 21 CpG dinucleotides scattered along the entire LTR with a majority of them concentrated in a circa 100 bp region around the transcription start. Localization of CpG dinucleotides (vertical bars) and PCR primers (arrows). Methylated CpGs are depicted by filled circles, non-methylated CpGs by empty circles.

a majority of them concentrated in a circa 100 bp region around the transcription start. In the 19 analyzed clones of the PCR product, we have not found any densely methylated sequence pointing to transcriptionally inactive silenced provirus. The level of CpG methylation was extremely low; only 2% out of 240 CpGs analyzed were methylated (Fig. 4). None of the 21 CpGs analyzed was found to be methylated with a strikingly higher frequency than others and no non-CpG methylation was detected. These results are in agreement with our previous observation that retroviral sequences integrated in chicken cells are methylated with low efficiency even after prolonged cultivation (Jana Blažková, unpublished results) and correspond with the transcriptionally active state of proviruses. Among the majority of hybrid MoMuSV/MoMuLV LTR sequences indicating reversely transcribed and integrated proviral DNA, we have found the MoMuSV LTR sequences as well. This demonstrates the persistence of plasmid DNA in the retroviral vector preparations and in the infected testicular cells.

Production of spermatozoa from the recipient testis

Seven WL males were transplanted with dispersed infected testicular cells from five adult BM donors. Three males of the same genetic origin as recipients were retained as positive controls and other three males were subjected to the same sterilizing treatment as recipient males but kept untransplanted as negative controls. Two of the WL recipients restored spermatogenesis within 2–3 months after transplantation and started to produce ejaculates.

A small volume of semen was first collected 7 weeks after transplantation. Sperm concentration increased gradually but with high variation and reached 500×10^6 sperm/ml at 23 week post-transplantation (Fig. 5). The control group of irradiated but non-transplanted WL cockerels failed to produce any semen during the same period of time. One of the three positive control WL males transplanted with non-infected testicular cells restored spermatogenesis and produced ejaculates 2 months after transplantation.

The presence of the viral vector in spermatozoa of recipient males transplanted with retrovirus-infected testicular donor cells was demonstrated by PCR of the specific EGFP DNA sequence (Fig. 6) and determination the nucleotide sequence of this PCR product (data not shown). This EGFP was not detected in spermatozoa produced by recipients transplanted with non-infected testicular cells.

Discussion

Efficient transplantation of donor testicular cells was widely described in mammals (Brinster & Avarbock 1994, Ogawa *et al.* 2000). Trefil *et al.* (2006) have shown previously that this transplantation in chickens resulted in

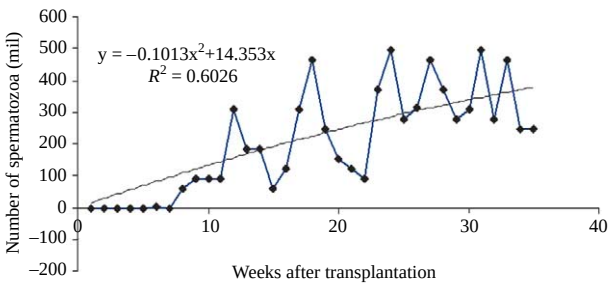


Figure 5 The time course of spermatozoa production in one WL recipient transplanted with dispersed testicular cells from adult BM donors (A). Only recipient cockerels with restored spermatogenesis were included into this calculation. Regression lines were constructed to demonstrate the rate of increase in sperm concentration.

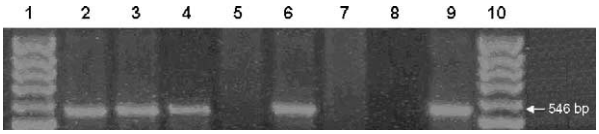


Figure 6 PCR analysis of the EGFP gene in spermatozoa isolated from a recipient cockerel to whom retrovirus-infected testicular cells carrying EGFP (lanes 3 and 4) were transplanted and PCR analysis of retrovirus-infected testicular cells (lane 6). The expected size of the amplified band of the *EGFP* gene was 546 bp, which corresponds with the 100 bp marker of molecular size (lanes 1 and 10). For positive (lanes 2 and 9) and negative (lanes 5, 7, and 8) controls, plasmid DNA (pEGFP) and genomic DNA isolated from non-infected spermatozoa (lane 5) and non-infected cells (lane 7) were used. Lane 8, the negative PCR control without DNA.

the re-colonization of the seminiferous epithelium with partial resumption of spermatogenesis and production of viable and fertilization-competent spermatozoa. In this report, we have continued our studies and observed the functional recovery of exogenous spermatogenesis after transplantation of *in vitro* infected testicular cell mixture. We have clearly demonstrated that infection with a high-titer reporter retrovirus can be achieved during short-term *in vitro* cultivation of donor testicular cells. Furthermore, these cells can be transplanted into sterilized recipients without any substantial loss of their capacity to repopulate recipient testes and resume spermatogenesis. Finally, we have detected the reporter DNA sequences in ejaculates of transplanted recipient cockerels, which suggest that this procedure might be a promising way to produce transgenic chickens.

Two factors are obviously important for retrovirus infection of the chicken male germ line cells. First, a pantropic VSV-G-pseudotyped retrovirus vector was used in our experiments to overcome the strict requirement for specific receptors. ASLV-derived replication-defective vectors could be used for this purpose as well; however, the sensitivity of particular chicken breeds to individual ASLV subgroups (A–E, J; Barnard *et al.* 2006) should be tested. Expression of *Tva* and *Tvb* but not *Tvc* receptor genes was shown in the testes (Elleder *et al.* 2005), but their expression in spermatogonia remains to be characterized. Secondly, the recovery of spermatogenesis by donor cells can be established only from rare SSCs. This cell population has not yet been unequivocally identified in chicken and, as a consequence, the precise number of stem spermatogonia transferred into each male remained unknown. To target them, a reasonable titer of retrovirus vector is necessary. We have used the titer of 5×10^7 IU/ml VSV-pseudotyped reporter vector, which was applied at MOI reaching 90. Even at this MOI, only a fraction of infected testicular cells was EGFP-positive, but the colony proliferation and overgrowth of these cells is an evidence that undifferentiated cells with the potential to repopulate recipients' testes are targeted by virus infection. The retrovirus titer in our study was significantly higher than those used in the previous study in mice (Nagano *et al.* 2001), which was clearly beneficial for the increase in number of infected testicular cells.

Two out of seven (28.6%) cockerels transplanted with infected testicular cells produced sperm after 2–3 months. This efficiency is comparable with animals transplanted in parallel directly with uninfected testicular cells, but a bit lower in comparison with our previous transplantation experiments (Trefil *et al.* 2006), where the efficiency reached 50%. Small numbers of birds used in these experiments might cause such differences. Clearly, in addition to the transient removal from their specific microenvironment, the survival and maintenance of proliferation capacity of germ line stem cells might be further affected by infection of the virus that is equipped with the toxic VSV-G envelope. Our partial success in re-colonization of testes in sterilized recipients evidences the retrovirus-mediated transduction of a reporter gene into male germ line stem cells. Further examination of the progeny of these recipients will show whether this approach may be of use in chicken transgenesis.

Of particular importance in retrovirus gene delivery experiments is the long-term expression of the transduced gene. After integration, proviruses are often subjected to epigenetic silencing by CpG methylation and histone modifications. Transcription of MLV and MLV-derived vectors is efficiently suppressed in embryonic carcinoma cells, an *in vitro* model of pluripotent stem cells. CpG methylation probably plays a role in epigenetic silencing of retroviruses infecting heterologous hosts. For example, Rous sarcoma virus (RSV) proviruses are often densely methylated after infection of rodent cells (Hejnar *et al.* 1996) and RSV LTR-driven transcription is more sensitive to CpG methylation in mammalian than in chicken cells (Hejnar *et al.* 1999). We have not detected any significant CpG methylation within the LTR of our vector integrated into testicular cells, which were cultivated in parallel with cells transplanted in recipient cockerels. The non-methylated status of the integrated vector correlates with expression of the transduced EGFP marker gene. We have not followed the EGFP expression in recipients' testes or in ejaculates after resumption of spermatogenesis, but the analogous study in mouse (Nagano *et al.* 2000) demonstrated that the β-galactosidase activity in frequent spermatogenic colonies derived from transplanted, retrovirus-infected germ line stem cells. In future studies, we will check the F1 progeny of fertile recipient cockerels for the presence of integrated retroviral vector to see if we are able to produce transgenic chicks and detect the expression of EGFP in the tissues of positive animals. In the case of silenced proviral copies, we plan to continue experiments with modified LTRs containing sequences blocking CpG methylation, for example, core sequences of CpG island (Hejnar *et al.* 2001).

To date, there is no evidence of a successful infection of chicken testicular cells that would have led to the production of transgenic spermatozoa. The optimized irradiation protocol (Trefil *et al.* 2003) eliminating the original acceptor germ cells enabled us to perform the efficient transplantation of dispersed testicular cells (Trefil *et al.* 2006). In this report, we have demonstrated that retrovirus infection of these testicular cells before transplantation could be used for gene delivery into the male germ line and for transgenesis in chicken.

Acknowledgements

The authors would like to thank Alena Kudějová for her skilled technical assistance and Dr Daniela Kotalová for her skilled technical assistance in the treatment of cockerels as well as Zdeněk Cimburek (Institute of Microbiology, Academy of Sciences of the Czech Republic) for FACS analysis. They also thank Jasper Manning for the critical reading of our manuscript. This research was supported by the Ministry of Education, Youth and Sports of the Czech Republic (Grant no. 1P05ME722 to Pavel Trefil), the Grant Agency of the Czech Republic (Grant no. 523/04/0569 to Jiří Hejnar), and the Academy of Sciences of the Czech Republic (project AV0Z50529514). The authors declare that there is no conflict of interest that would prejudice the impartiality of this scientific work.

References

Barnard RJ, Elleder D & Young JA 2006 Avian sarcoma and leukosis virus-receptor interactions: from classical genetics to novel insights into virus-cell membrane fusion. *Virology* **344** 25–29.

Bastos H, Lassalle B, Chicheportiche A, Rioul L, Testart J, Allemand I & Fouchet P 2005 Flow cytometric characterization of viable meiotic and post meiotic cells by Hoechst 33 342 in mouse spermatogenesis. *Cytometry* **65A** 40–49.

Bonitz W 1970 Die Ermittlung des Spermazellvolumens in Ejakulaten von Hähnen mit Hilfe der Mikro-Kapillarmethod. *Archiv für Geflügelzucht und Kleintierkunde* **19** 193–201.

Bosselman RA, Hsu RY, Boggs T, Hu S, Bruszweski J & Ou S 1989 Germline transmission of exogenous genes in the chicken. *Science* **243** 533–535.

Brinster RL & Avarbock MR 1994 Germ line transmission of donor haplotype following spermatogonial transplantation. *PNAS* **91** 11303–11307.

Burrows WH & Quinn JP 1937 Collection of spermatozoa from domestic fowl and turkey. *Poultry Science* **16** 19–24.

Elleder D, Stepanets V, Melder DC, Šenigl F, Geryk J, Pajer P, Plachý J, Hejnar J, Svoboda J & Federspiel M 2005 The receptor for the subgroup C avian sarcoma and leukosis viruses, Tvc, is related to mammalian butyrophilins, members of the immunoglobulin superfamily. *Journal of Virology* **79** 10408–10419.

Emi N, Friedemann T & Yee JK 1991 Pseudotype formation of murine leukemia virus with the G protein of vesicular stomatitis virus. *Journal of Virology* **65** 1202–1207.

Hájková P, El-Maarri O, Engemann S, Oswald J, Olek A & Walter J 2002 DNA-methylation analysis by the bisulfite-assisted genomic sequencing method. *Methods in Molecular Biology* **200** 143–154.

Hamra FK, Chapman KM, Nguyen DM, Williams-Stephens AA, Hammer RE & Garbers DL 2005 Self renewal, expansion, and transfection of rat spermatogonial stem cells in culture. *PNAS* **102** 17430–17435.

Hejnar J, Svoboda J, Geryk J, Fincham VJ & Hák R 1996 High rate of morphological reversion in tumor cell line H-19 associated with permanent transcriptional suppression of the LTR, v-src, LTR provirus. *Cell Growth and Differentiation* **5** 277–285.

Hejnar J, Plachý J, Geryk J, Machoň O, Trejbalová K, Guntaka RV & Svoboda J 1999 Inhibition of the Rous sarcoma virus long terminal repeat-driven transcription by *in vitro* methylation: different sensitivity in permissive chicken cells versus mammalian cells. *Virology* **255** 171–181.

Hejnar J, Hájková P, Plachý J, Elleder D, Stepanets V & Svoboda J 2001 CpG island protects Rous sarcoma virus-derived vectors integrated into nonpermissive cells from DNA methylation and transcriptional suppression. *PNAS* **98** 565–569.

Kanatsu-Shinohara M, Ogonuki N, Inoue K, Miki H, Ogbura A, Toyokuni S & Shinohara T 2003 Long-term proliferation in culture and germ line transmission of mouse male germline stem cells. *Biology of Reproduction* **69** 612–616.

Kanatsu-Shinohara M, Toyokuni S & Shinohara T 2004 Transgenic mice produced by retroviral transduction of male germ line stem cells *in vivo*. *Biology of Reproduction* **71** 1202–1207.

Koo BC, Kwon MS, Choi BR, Lee HT, Choi HJ, Kim J-H, Kim N-H, Jeon I, Chang W & Kim T 2004 Retrovirus-mediated gene transfer and expression of EGFP in chicken. *Molecular Reproduction and Development* **68** 429–434.

Kubota H, Avarbock MR & Brinster RL 2004 Growth factors essential for self-renewal and expansion of mouse spermatogonial stem cells. *PNAS* **101** 16489–16494.

Kwon MS, Koo BC, Choi BR, Lee HT, Kim YH, Ryu W-S, Shim H, Kim J-H, Kim N-H & Kim T 2004 Development of transgenic chickens expressing enhanced green fluorescent protein. *Biochemical and Biophysical Research Communications* **320** 442–448.

van de Lavoir MC, Diamond JH, Leighton PA, Mather-Love C, Heyer BS, Bradshaw R, Cerchner A, Hooi LT, Gessaro TM, Swanberg SE, Delany ME & Etches RJ 2006 Germline transmission of genetically modified primordial germ cells. *Nature* **44** 766–769.

Lillico SC, Sherman A, McGrew MJ, Robertson CD, Smith J, Haslam C, Barnard P, Radcliffe PA, Mitrophanous KA, Elliot EA & Sang HM 2007 Oviduct-specific expression of two therapeutic proteins in transgenic hens. *PNAS* **104** 1771–1776.

McGrew MJ, Sherman A, Ellard FM, Lillico SG, Gilhooley HJ, Kingsman AJ, Mitrophanous KA & Sang H 2004 Efficient production of germline transgenic chickens using lentiviral vectors. *EMBO Reports* **5** 728–733.

Meloun M, Militky J & Forina M 1994 PC-Aided regression and related methods, In *Chemometrics for Analytical Chemistry,* 2nd edn. Chichester: Ellis Horwood.

Nagano M, Shinohara T, Avarbock MR & Brinster RL 2000 Retrovirus-mediated gene delivery into male germ line stem cells. *FEBS Letters* **475** 7–10.

Nagano M, Brinster CJ, Orwig KE, Ryu B-E, Avarbock MR & Brinster RL 2001 Transgenic mouse produced by retroviral transduction of male germ-line stem cells. *PNAS* **99** 13090–13095.

Nagano M, Ryu BY, Brinster CJ, Avarbock MR & Brinster RL 2003 Maintenance of mouse male germ line stem cells *in vitro*. *Biology of Reproduction* **68** 2207–2214.

Naito M, Matsubara Y, Harumi T, Tagami T, Kagami H, Sakurai M & Kuwana T 1999 Differentiation of donor primordial germ cells into functional gamets in the gonads of mixed-sex germline chimeric chickensproduced by transfer of primordial germ cells isolated from embryonic blood. *Journal of Reproduction and Fertility* **117** 291–298.

Ogawa T, Dobrinski I, Avarbock MR & Brinster RL 2000 Transplantation of male germ line stem cells restores fertility in infertile mice. *Nature Medicine* **6** 29–34.

Pain B, Clark ME, Shen M, Nakazawa H, Sakurai M, Samarut J & Etches RJ 1996 Long-term *in vitro* culture and characterization of avian embryonic stem cells with multiple morphogenetic potentialities. *Development* **122** 2339–2348.

Petitte JN, Liu G & Yang Z 2004 Avian pluripotent stem cells. *Mechanisms of Development* **121** 1159–1168.

Salter DW, Smith EJ, Hughes SH, Wright SE & Crittenden LB 1987 Transgenic chickens: insertion of retroviral genes into the chicken germ line. *Virology* **157** 236–240.

Sang H 2004 Prospects for transgenesis in the chick. *Mechanisms of Development* **121** 1179–1186.

Thurston RJ & Korn N 2000 Spermiogenesis in comercial poultry species: anatomy and control. *Poultry Science* **79** 1650–1668.

Trefil P, Polak J, Poplstein M, Mikus M, Kotrbova A & Rozinek J 2003 Preparation of fowl tetstes as recipient organs to germ-line chimeras by means of γ-radiation. *British Poultry Science* **44** 643–650.

Trefil P, Mičáková A, Mucksová J, Hejnar J, Bakst MR, Kalina J, Poplštejn M & Brillard J-P 2006 Restoration of spermatogenesis and male fertility by transplantation of dispersed testicular cells in the chicken. *Biology of Reproduction* **75** 575–581.

Vick L, Li Z & Simkiss K 1993 Transgenic birds from transformed primordial germ cells. *Proceedings of the Royal Society of London, Series B: Biological Sciences* **251** 179–182.

Received 27 September 2006
First decision 23 October 2006
Revised manuscript received 23 April 2007
Accepted 1 June 2007

JOURNAL OF VIROLOGY, Mar. 2008, p. 2097–2105 Vol. 82, No. 5
0022-538X/08/$08.00+0 doi:10.1128/JVI.02206-07

A Single-Amino-Acid Substitution in the Tvb^{S1} Receptor Results in Decreased Susceptibility to Infection by Avian Sarcoma and Leukosis Virus Subgroups B and D and Resistance to Infection by Subgroup E In Vitro and In Vivo$^{\nabla}$

Markéta Reinišová,[1]§ Filip Šenigl,[1]§ Xueqian Yin,[2] Jiří Plachý,[1] Josef Geryk,[1] Daniel Elleder,[1]† Jan Svoboda,[1] Mark J. Federspiel,[2]* and Jiří Hejnar[1]*

Department of Cellular and Viral Genetics, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Prague, Czech Republic,[1] and Department of Molecular Medicine, Mayo Clinic, Rochester, Minnesota[2]

Received 9 October 2007/Accepted 10 December 2007

The avian sarcoma and leukosis virus (ASLV) family of retroviruses contains five highly related envelope subgroups (A to E) thought to have evolved from a common viral ancestor in the chicken population. Three genetic loci in chickens determine the susceptibility or resistance of cells to infection by the subgroup A to E ASLVs. Some inbred lines of chickens display phenotypes that are somewhere in between either efficiently susceptible or resistant to infection by specific subgroups of ASLV. The *tvb* gene encodes the receptor for subgroups B, D, and E ASLVs. The wild-type Tvb^{S1} receptor confers susceptibility to subgroups B, D, and E ASLVs. In this study, the genetic defect that accounts for the altered susceptibility of an inbred chicken line, line M, to infection by ASLV(B), ASLV(D), and ASLV(E) was identified. The *tvb* gene in line M, *tvb^{r2}*, encodes a mutant Tvb^{S1} receptor protein with a substitution of a serine for a cysteine at position 125 (C125S). Here, we show that the C125S substitution in Tvb^{S1} significantly reduces the susceptibility of line M cells to infection by ASLV(B) and ASLV(D) and virtually eliminates susceptibility to ASLV(E) infection both in cultured cells and in the incidence and growth of avian sarcoma virus-induced sarcomas in chickens. The C125S substitution significantly reduces the binding affinity of the Tvb^{S1} receptor for the subgroup B, D, and E ASLV envelope glycoproteins. These are the first results that demonstrate a possible role of the cysteine-rich domain 3 in the function of the Tvb receptors.

Retroviruses cause serious diseases in animals and humans. The disease process begins with the virus infecting a cell(s), a process mediated by the interaction of the retroviral envelope glycoproteins with specific cell surface proteins that act as receptors (14, 24). A proper viral-glycoprotein–receptor interaction initiates conformational changes in the viral glycoproteins that ultimately result in the fusion of the viral and cellular membranes and entry of viral components (9). Despite the complexity and specificity of the viral-glycoprotein–receptor interaction required for virus entry, families of closely related retroviruses have evolved their glycoproteins to use different cellular proteins as receptors. Presumably, the presence of multiple viral subgroups that utilize different receptors is an advantage for viruses in overcoming host resistance. Resistance to retroviral infection occurs when the specific receptor protein is not available. Genetic alteration(s) can account for resistance, resulting in the complete lack of receptor protein expression or the expression of an aberrant protein not suitable as a viral receptor. In addition, receptors can be saturated with viral glycoproteins expressed by the cell, physically blocking receptor accessibility, a phenomenon known as receptor interference (14, 24).

The avian sarcoma and leukosis virus (ASLV) family of retroviruses contains five highly related envelope subgroups (A to E) thought to have evolved in the chicken population from a common viral ancestor (4, 5, 24). Three genetic loci in chickens determine the susceptibility or resistance of cells to infection by the subgroup A to E ASLVs. Susceptibility to subgroup A ASLVs is determined by the *tva* locus; susceptibility to subgroup C ASLVs by the *tvc* locus; and susceptibility to the subgroup B, D, and E ASLVs by the *tvb* locus. The Tva proteins are related to the low-density lipoprotein receptor family (6, 25). The Tvb proteins are related to the tumor necrosis factor receptor (TNFR) family (2, 3, 7). The Tvc proteins are most closely related to mammalian butyrophilins of the immunoglobulin superfamily (11). The normal functions and ligands of the Tva, Tvb, and Tvc proteins in birds are unknown.

The genetic defects that account for resistance to infection by specific ASLVs, *tva^r*, *tvb^r*, and *tvc^r* alleles, have been identified in some lines of inbred chickens (10, 11, 16). The mutations found in the resistant alleles result in premature termination codons, or frameshifts, leading to severely truncated proteins or no receptor expression. In addition, single-amino-acid substitutions, often changing a cysteine residue, that result in a dramatic decrease in the affinity of the ASLV envelope glycoprotein interaction for the mutant receptor protein have been identified. An understanding of the variation in receptor

* Corresponding author. Mailing address for J. Hejnar: Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Vídeňská 1083, CZ-14220 Prague 4, Czech Republic. Phone: 420 296 443 443. Fax: 420 223 410 955. E-mail: hejnar@img.cas.cz. Mailing address for Mark J. Federspiel: Department of Molecular Medicine, Mayo Clinic, 200 First St., SW, Rochester, MN 55905. Phone: (507) 284-8895. Fax: (507) 266-2122. E-mail: federspiel.mark@mayo.edu.

† Present address: The Infectious Disease Laboratory, The Salk Institute, La Jolla, San Diego, CA 92037-1099.

§ M.R. and F.Š. contributed equally to this work.

$^{\nabla}$ Published ahead of print on 19 December 2007.

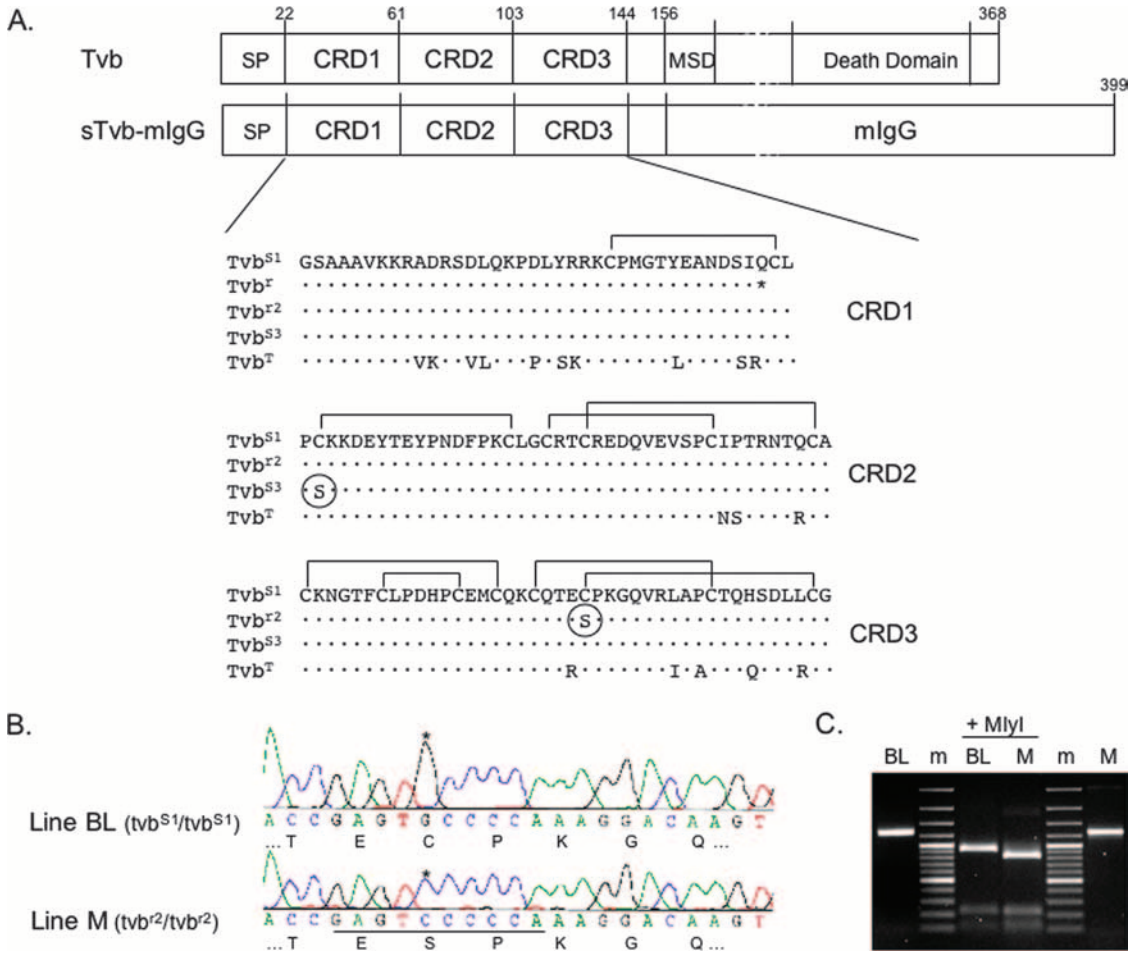


FIG. 1. The *tvb*^{r2} allele in chicken line M contains a C125S substitution in CRD3. (A) Schematic representations of the Tvb proteins and a soluble form of Tvb fused to an mIgG are shown (not drawn to scale). The numbering of the Tvb proteins includes the signal peptide (SP) for historical reasons. MSD, membrane-spanning domain. The deduced amino acid sequences of the CRDs of the known Tvb receptors, including the turkey ortholog, Tvb^{T}, are compared, with residues different from the Tvb^{S1} receptor indicated. The C125S substitution in Tvb^{R2} and the C62S substitution in Tvb^{S3} are circled. The putative disulfide bonds in the Tvb CRDs are indicated based on comparisons with the related TNFR proteins and the X-ray structure of the human DR5 TRAIL receptor protein. (B) Chromatograms of partial nucleotide and deduced amino acid sequences of CRD3s from chicken line M and outbred chicken line BL cDNAs. The G-to-C transversion is highlighted with an asterisk; the new MlyI site in line M is underlined. (C) A 1,259-bp *tvb* fragment was produced by reverse transcription-PCR using RNA from line BL and line M. After MlyI digestion, a 920-bp fragment was produced from the wild-type *tvb*^{s1} cDNA from line BL, while the additional MlyI site in the line M *tvb*^{r2} cDNA produced an 820-bp fragment. m, DNA marker 100-bp ladder.

proteins that can be encountered by ASLVs in chickens, including related receptor orthologs in other birds, provides valuable information on the evolutionary pressures on these viruses and enables the identification of domains of the receptor and envelope glycoprotein critical for an efficient virus-receptor interaction and subsequent virus entry.

The structures of TNFR-related proteins have shown that homologous proteins (e.g., the Tvb proteins) likely contain three cysteine-rich domains (CRDs) in the extracellular domain (Fig. 1A) (15). Two naturally occurring *tvb* susceptibility alleles have been identified in chickens; the *tvb*^{s1} allele confers susceptibility to subgroups B, D, and E ASLVs, while the *tvb*^{s3} allele confers susceptibility to only subgroup B and D ASLVs (3, 7). The substitution of a serine for a cysteine at residue 62 in the CRD2 domain of Tvb^{S3} distinguishes the Tvb^{S1} and Tvb^{S3} proteins. Presumably, this mutation alters the structure of CRD2, resulting in the loss of ASLV(E) binding and entry. The major determinants of the Tvb receptor important for entry of ASLV(B), ASLV(D), and ASLV(E) were identified previously using a variety of mutational approaches (1, 17, 18). A 15-amino-acid peptide in the Tvb CRD1 domain appears to be sufficient for ASLV(B) and ASLV(D) binding and entry. In contrast, the main determinants of Tvb important for ASLV(E) binding and entry reside in the CRD2 domain. These studies did not demonstrate any significant role of the CRD3 domain in the binding and entry of ASLV(B), ASLV(D), or ASLV(E). A *tvb* allele, *tvb*^{r}, which contains a premature stop codon at residue 57 explaining the resistance to infection by all three subgroups (16), was identified and characterized from inbred line 7_2 chickens. Finally, a *tvb* homolog, *tvb*^{t}, that confers susceptibility only to ASLV(E) (2), has been identified in turkeys.

Some inbred lines of chickens display phenotypes that are somewhere in between completely susceptible and resistant to

infection by specific subgroups of ASLV. This altered susceptibility phenotype could be the result of mutations in the receptor that alter but do not eliminate the ability of ASLV to use the protein as a receptor. We hypothesized that such mutations, which reduce but do not eliminate the binding affinity of the mutant receptor for the viral glycoproteins, could explain this phenotype. In this study, the genetic defect that accounts for the altered susceptibility of an inbred chicken line, line M, to infection by ASLV(B), ASLV(D), and ASLV(E) was identified. The *tvb* gene in line M, *tvb*^{r2}, encodes a mutant Tvb^{S1} receptor protein with a substitution of a serine for a cysteine at position 125 (C125S). These are the first results that demonstrate a possible role of CRD3 in the efficient functioning of the Tvb receptors.

MATERIALS AND METHODS

Chicken lines. The inbred chicken lines M (Black Minorca), L15, CC.R1, and WA and the outbred chicken population Brown Leghorn (BL) have been maintained at the Institute of Molecular Genetics, Prague (20). Hens and cockerels were kept in individual cages (4,200 cm^2) under standard conditions. Feed and water were provided ad libitum, and a light-dark photoperiod of 12 h-12 h was applied. Fertilized eggs were incubated at 38°C and 60% relative humidity in a forced-air incubator with a tilting motion every 2 h through a 90° angle. All experiments were performed in accordance with Czech legal requirements for animal handling.

Cell culture and virus propagation. Chicken embryo fibroblasts (CEFs) were prepared from 10-day-old embryos from line M, WA, and BL chickens as described previously (12). The DF-1 chicken fibroblast cell line (12a), the Rous sarcoma virus (RSV)-transformed 16Q quail cell line (19), and CEFs were grown in a mixture of two parts Dulbecco's modified Eagle's medium and one part F-12 medium supplemented with 5% calf serum, 5% fetal calf serum, 1% chicken serum, and penicillin/streptomycin (100 mg/ml each) in a 5% CO_2 atmosphere at 37°C. RCASBP(A)GFP, RCASBP(B)GFP, RCASBP(C)GFP, and RCASBP(D)GFP viruses (12, 13) were propagated by transfection of plasmid DNA containing the reporter vector into DF-1 cells, which are free of closely related endogenous retrovirus loci. The RCASBP(E)GFP virus was propagated in line L15 CEFs. Transfection was performed by applying a mixture of 10 μg plasmid DNA and 100 μl linear polyethylenimine (Polysciences; molecular weight, 25,000; 1 mg/ml; pH 7.5) in 12 ml serum-free culture medium on subconfluent DF-1 cells grown in 100 mm plates. Virus spread was observed as an increasing proportion of green fluorescent protein (GFP)-positive cells, and viral stocks were established from cell supernatants at day 4 or 5 after transfection. Viral stocks were cleared of debris by centrifugation at 2,000 × *g* for 10 min at 10°C and stored at −80°C.

The transforming subgroup A, B, C, D, and E ASLV specificities for in vivo sarcoma induction were produced by rescuing the replication-defective Bryan high-titer-RSV present in the 16Q cell line. DF-1 cells were infected with RCASBP(A)GFP, RCASBP(B)GFP, RCASBP(C)GFP, RCASBP(D)GFP, or RCASBP(E)GFP, and virus spread was monitored by fluorescence. After 4 days, the infected GFP-positive DF-1 cells were mixed with 16Q cells (19) and cultivated for a further 5 days. Viral stocks containing GFP reporter viruses, as well as transforming viruses of the same subgroup, were centrifuged at 2,000 × *g* for 10 min at 10°C and stored at −80°C. The titers of the transforming viruses were quantitated by an *src* focus assay on BL CEFs and reached titers of 10^2 to 10^5 focus-forming unit (FFU) per ml.

Reverse transcription-PCR of the *tvb*^{r} allele from chicken line M. We prepared total RNA from blood collected from line M chickens using the RNeasy total RNA isolation system (Qiagen Inc.). Reverse transcription was carried out with 1 μg total RNA, Moloney murine leukemia virus reverse transcriptase (Promega), and oligo(dT)$_{15}$ primers (Promega). The cDNA resulting from this reaction was PCR amplified with primers TVB3 (5′-CAGACCTCCAGAAGCCA GAC-3′) and TVB6 (5′-CGAGAGCACTGTCACAGAGAT-3′) with *Taq* polymerase (TaKaRa). The conditions for the amplification were as follows: 2 min at 94°C; 34 cycles of 15 s at 94°C, annealing for 40 s at 57°C, and 3 min at 72°C; and final extension for 10 min at 72°C. The final PCR product was treated with ExoSAP-IT (USB) and then sequenced using the BigDye Terminator v3.1 cycle-sequencing kit (PE Applied Biosystems). A 1,259-bp fragment of *tvb* was amplified from the cDNA using primers TVB5 (5′-TCCTAACTCGGTCCGAA TCC-3′) and TVB6 and the Expand Long Template Polymerase (Roche). The conditions for amplification were 2 min at 94°C and 35 cycles of 15 s at 94°C, annealing for 40 s at 56°C, and 90 s at 68°C; final extension was for 7 min at 68°C. The product was digested with MlyI, and the products were separated by agarose electrophoresis.

Virus entry analyzed by FACS. Line BL or M CEFs were seeded at 5 × 10^5 per 60-mm plate and infected with RCASBP viruses the next day: 10^5 infectious units were applied in 0.5 ml medium, and after 1 h, the volume of the medium was increased to 4 ml. The percentage of GFP-positive cells was quantitated by fluorescence-activated cell sorting (FACS) using an LSR II analyzer (Becton Dickinson) on days 1, 2, 3, and 6 postinfection. Each day, one half of the cell culture was used for FACS and the other half was passaged on a new dish. For FACS analysis, trypsinized cells were first washed in phosphate-buffered saline (PBS) and then analyzed.

DNA constructs. A gene encoding a soluble form of the chicken Tvb^{S1C125S} receptor (*stvb*^{s1C125S}*-mIgG*) was constructed as described previously for soluble forms of the chicken Tvb^{S3} receptor (*stvb*^{s3}*-mIgG*) and Tvb^{S1} receptor (*stvb*^{s1}*-mIgG*) (11). These genes encode the extracellular domain of the ASLV receptor fused to the constant region of a mouse immunoglobulin G (mIgG) heavy chain and are in the CLA12NCO adaptor plasmid (12). The *stvb*^{s1C125S}*-mIgG*, *stvb*^{s1}*-mIgG*, and *stvb*^{S3}*-mIgG* gene cassettes were isolated as ClaI fragments and subcloned into the ClaI site of the RCASBP(A) vector. DF-1 cells were infected with each virus, and infected cell supernatants that contained either the Tvb^{S1C125S}-mIgG, Tvb^{S3}-mIgG, or Tvb^{S1}-mIgG receptor proteins were collected. The supernatants were cleared by centrifugation at 2,000 × *g* for 10 min at 4°C and stored in aliquots at −80°C.

The recombinant, replication-competent ASLV vectors RCASBP(A)GFP, RCASBP(B)GFP, RCASBP(C)GFP, RCASBP(D)GFP, and RCASBP(E)GFP viruses containing the GFP gene have been described previously (12). The receptor subgroup of the viral envelope glycoprotein is denoted in parentheses [e.g., subgroup B by (B)].

Immunoprecipitations and Western immunoblot analysis. The Tvb^{S1C125S}-mIgG, Tvb^{S3}-mIgG, and Tvb^{S1}-mIgG receptor proteins were immunoprecipitated separately from cell culture supernatants with anti-mIgG-agarose beads (Sigma) and analyzed by Western immunoblotting as previously described (11).

Binding affinity was analyzed by FACS. Line L15 CEFs and line L15 CEFs infected with either RCASBP(B), RCASBP(D), or RCASBP(E) were removed from culture with trypsin de Larco (Quality Biological, Inc.) and washed with PBS. The cells were fixed with 4% paraformaldehyde in PBS at room temperature for 15 min and then washed with PBS. Approximately 1 × 10^6 cells in PBS supplemented with 1% calf serum (PBS-CS) were incubated with supernatant containing one of the receptor-mIgG fusion proteins on ice for 30 min. The DF-1 cells were then washed with PBS-CS and incubated with either 5 μl of goat anti-mIgG (heavy plus light chains) linked to phycoerythrin or 5 μl of goat anti-rabbit IgG (rIgG) (heavy plus light chains) linked to phycoerythrin (Kirkegaard & Perry Laboratories, Gaithersburg, MD) in PBS-CS (1 ml total volume) on ice for 30 min. The cell–soluble receptor-IgG–Ig-phycoerythrin complexes were washed with PBS-CS, resuspended in 0.5 ml PBS-CS, and analyzed with a Becton Dickinson FACSCalibur using CELLQuest 3.1 software.

***K*$_D$ calculations.** The maximum possible fluorescence and apparent dissociation constant (K_D) value for each data set obtained from the FACS binding assays were estimated by fitting the data via nonlinear least squares to a log logistic growth curve function: $f(y) = M/[1 + e^{-r(\log x - \log K_d)}]$, where y is the mean fluorescence, M is the maximum fluorescence, r is the rate, x is the concentration of the receptor-mIgG or SU-rIgG fusion protein, and K_d is the dissociation constant, defined as the concentration of the receptor-mIgG or SU-rIgG fusion protein at half-maximal binding.

In vivo sarcoma induction and monitoring. Ten-day-old chickens from lines M, CC.R1, and L15 were inoculated with 10^2 or 10^3 FFU of transforming virus rescued from 16Q cells in 0.1 ml of Iscove's Dulbecco's modified Eagle's medium subcutaneously into the pectoral muscle. The growth of sarcomas at the site of inoculation was monitored by calculating the area of prominent tumor. Transparent foil was placed on the tumor, the contours of the tumor were traced, and the area of the tumor was calculated in mm^2 (21). Birds bearing vast nonregressing tumors were sacrificed before they reached the terminal stage.

RESULTS

The *tvb* allele from the inbred chicken line M (*tvb*^{r2}) encodes a mutant Tvb^{S1} receptor with the C125S substitution in CRD3. cDNAs were synthesized from mRNAs isolated from inbred chicken line M and line WA CEFs and used to PCR amplify

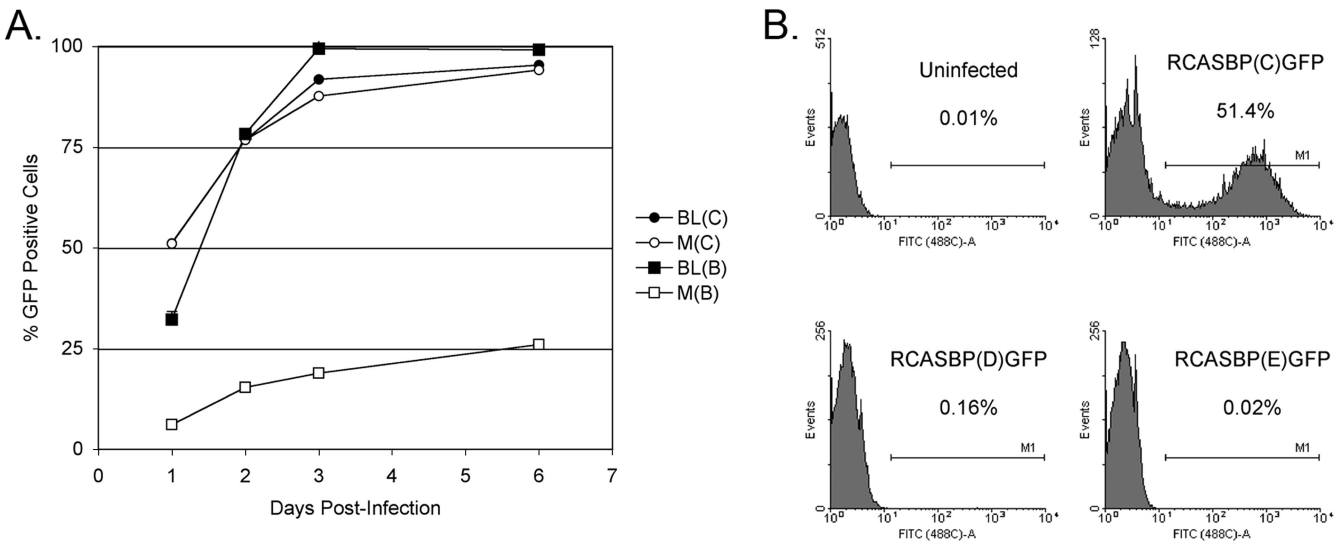


FIG. 2. The time course of infection of BL and line M (M) CEFs with ASLVs. CEFs were infected with replication-competent ASLVs encoding the GFP reporter proteins RCASBP(B)GFP, RCASBP(C)GFP, RCASBP(D)GFP, and RCASBP(E)GFP. (A) The proportion of GFP-positive cells was determined by FACS on the indicated day postinfection; the percentage of GFP-positive cells is indicated as a mean of three parallel dishes. (B) Line M CEFs were infected at an MOI of 0.1, and the percentage of GFP-positive cells was quantitated by FACS 1 day postinfection. In the histograms, the relative GFP fluorescence is plotted against the cell count and the percentage of GFP-positive cells is indicated as a mean of two parallel dishes.

tvb sequences and to clone the *tvb* cDNAs. We determined the nucleotide and deduced amino acid sequences of the coding regions of *tvb* receptor cDNA clones and compared the nucleotide sequences with the sequences of tvb^{s1} and tvb^{s3} genes (3, 7). The tvb^{s1} allele of line WA encodes the same single-nucleotide mutation in codon 58 (CAG to UAG) that results in a premature termination codon described previously for the chicken line 7_2 tvb^r allele (16). However, the *tvb* gene in line M contains a different single-nucleotide mutation that changes the cysteine at position 125 to serine (UGC to UCC) in the Tvb^{S1} receptor protein (Fig. 1A) and is designated tvb^{r2}. This single-nucleotide change creates an additional MlyI recognition site in tvb^{r2} compared to tvb^{s1} (Fig. 1B). The presence of this polymorphic marker was further demonstrated by digestion of a 1,259-bp fragment of *tvb* cDNA amplified from line M RNA. The diagnostic 920-bp MlyI fragment was obtained using the control tvb^{s1} cDNA from line BL; a shortened 820-bp fragment was obtained after MlyI cleavage of tvb^{r2} cDNA (Fig. 1C).

The C125S substitution in the Tvb^{S1} receptor reduces the susceptibility of line M cells to subgroup B, D, and E ASLV infection. To determine the effect of the C125S substitution in the chicken Tvb^{S1} receptor on ASLV susceptibility, line M and BL CEFs were infected at a multiplicity of infection (MOI) of 0.1 with RCASBPGFP reporter virus, a replication-competent ASLV vector encoding GFP, and the time course of infection was followed by quantitating the percentage of green fluorescent cells by flow cytometry. Line BL CEFs are susceptible to subgroup A, B, C, and D ASLVs and were used as a positive control. As expected, both line M and line BL CEFs were efficiently infected by RCASBP(C)GFP, as shown in Fig. 2A, with almost one-half of the cells infected at 1 day postinfection and virtually all cells infected by day 3. A very different result was obtained when the CEFs were infected with RCASBP(B)-GFP. As expected, line BL CEFs were efficiently infected with RCASBP(B)GFP at a rate similar to that of RCASBP(C)GFP infection (Fig. 2A). However, RCASBP(B)-GFP infected line M CEFs much less efficiently, with only 6.6% of the cells infected at day 1 (5- to 10-fold lower than line BL), and virus spread through the cells very slowly, with only ~25% of the cell population infected by day 6. In a separate experiment, line M CEFs were infected with RCASBP(C)GFP, RCASBP(D)GFP, or RCASBP(E)GFP at 0.1 MOI, and the infected cells were quantitated the next day (Fig. 2B). As in the first experiment, ~50% of the line M CEFs were infected with the control RCASBP(C)GFP virus, but only 0.16% of the line M CEFs were infected with RCASBP(D)GFP and virtually no line M CEFs were infected with RCASBP(E)GFP. These data clearly demonstrate, at least in cultured cells, that the C125S substitution in the Tvb^{S1} receptor resulted in lower susceptibility of line M CEFs to ASLV(B) infection, including a significant decrease in the rate of virus spread, perhaps an even lower level of susceptibility to ASLV(D) infection and virus spread, and almost complete resistance to infection by ASLV(E).

The C125S substitution in Tvb^{S1} lowers the binding affinity for the ASLV envelope glycoproteins. Soluble forms of the Tvb receptors fused to an mIgG domain (Fig. 1A) were used to estimate by FACS the binding affinities of the Tvb^{S1}, $Tvb^{S1C125S}$, and Tvb^{S3} receptors for ASLV envelope glycoproteins expressed on the surfaces of infected line L15 CEFs, as described previously (11). The integrity of the soluble Tvb (sTvb)-mIgG proteins was verified by immunoprecipitation and Western immunoblot analysis (Fig. 3A). The concentration of each protein stock was determined by an enzyme-linked immunosorbent assay for mIgG as described previously (10). The $sTvb^{S1}$-mIgG protein bound the ASLV(B)-infected cells (Fig. 3B), the ASLV(D)-infected cells (Fig. 3C), and the ASLV(E)-infected cells (Fig. 3D) with subnanomolar affinities

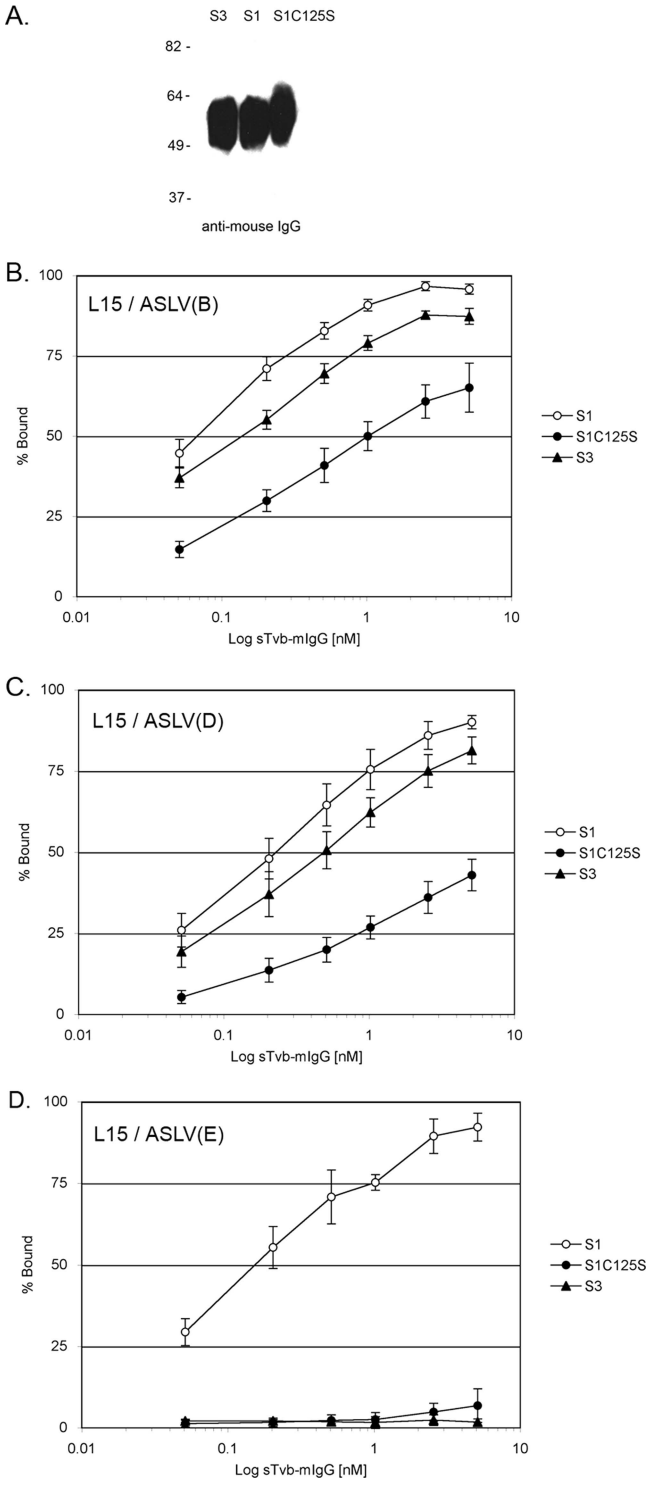


FIG. 3. Binding affinities of ASLV envelope glycoproteins for Tvb receptors. (A) Western immunoblot analysis of the soluble forms of the Tvb proteins, sTvb^{S3}-mIgG (S3), sTvb^{S1}-mIgG (S1), and sTvb^{S1C125S} (S1C125S), immunoprecipitated with anti-mIgG agarose beads, denatured and separated by sodium dodecyl sulfate-12% polyacrylamide gel electrophoresis, and transferred to nitrocellulose. The mIgG-tagged proteins were probed with anti-mIgG-conjugated to horseradish peroxidase and visualized by chemiluminescence. Molecular masses (in kilodaltons) are given on the left. (B to D) Line L15 CEFs chronically infected with ASLV(B) (B), ASLV(D) (C), or ASLV(E) (D) were fixed with paraformaldehyde and incubated with different amounts of soluble receptor. The viral glycoprotein–soluble-receptor complexes were bound to goat-anti-mIgG linked to phycoerythrin. The amount of phycoerythrin bound to the cells was quantitated by FACS, and the maximum fluorescence was estimated (see Materials and Methods). The data were plotted as percent maximum fluorescence. The values shown are the averages and standard deviations from four experiments.

TABLE 1. Estimated binding affinities of soluble forms of the Tvb receptors for ASLV envelope glycoproteins expressed on the surfaces of L15 CEFs

Cells	Apparent K_D (nM)[a]		
	sTvb^{S1}-mIgG	sTvb^{S1C125S}-mIgG	sTvb^{S3}-mIgG
L15/ASLV(B)	0.07 ± 0.01	1.27 ± 0.65	0.15 ± 0.03
L15/ASLV(D)	0.23 ± 0.08	8.77 ± 3.17	0.50 ± 0.14
L15/ASLV(E)	0.17 ± 0.05	≫1,000[b]	NDB[c]

[a] Apparent K_D values were estimated by fitting data via nonlinear least squares to log logistic growth curve function as described in Materials and Methods. Each result is the average and standard deviation from three experiments.
[b] Binding of the sTvb^{S1C125S}-mIgG receptor was detectable but not sufficient to calculate a reliable apparent K_D (Fig. 3).
[c] NDB, no detectable binding.

(Table 1). As expected, the sTvb^{S3}-mIgG protein bound subgroup B and D glycoproteins with affinities similar to that of sTvb^{S1}-mIgG but did not bind subgroup E glycoproteins at a detectable level. Binding of the sTvb^{S1C125S}-mIgG protein to all three ASLV glycoproteins could be detected, but with significantly lower affinities than for sTvb^{S1}-mIgG: ASLV(B) at 10- to 25-fold lower affinity, ASLV(D) at 25- to 50-fold lower affinity, and ASLV(E) at barely detectable levels. The reduction in the binding affinity of sTvb^{S1C125S}-mIgG correlates with the reduced infectivity of subgroups B, D, and E ASLVs in line M CEFs (Fig. 2).

Compared to chickens with a wild-type Tvb^{S1} receptor, the incidence and growth of sarcomas induced by subgroups B, D, and E ASVs were significantly altered in line M chickens. In order to see the in vivo effects of the C125S substitution in the Tvb^{S1} receptor, line M chickens were infected with transforming viruses and the formation and growth rates of the induced sarcomas were measured. ASLVs containing the *src* oncogene (avian sarcoma virus [ASV]) were produced by the rescue of *env*-defective Bryan high-titer-RSV from 16Q cells (19) by infection with the appropriate envelope subgroup RCASB-PGFP virus. The titer of the rescued transforming virus stocks was quantitated by in vitro focus assay. Ten-day-old line M chicks and age-matched controls of line CC.R1 and line L15 were challenged with 1,000 FFU or 100 FFU of the ASV stock injected into the pectoral muscle. Line CC.R1 is susceptible to subgroup B, C, and D ASLV infection. Line L15 is susceptible to subgroup A, B, D, and E ASLVs—a rare inbred chicken line susceptible to subgroup E viruses (8). The incidence and growth rates of sarcomas induced at the site of virus inoculation were quantitated.

As expected, line CC.R1 was susceptible to sarcoma formation induced by ASV(C), ASV(B), and ASV(D) infection while line M was susceptible to sarcoma formation induced by ASV(A) and ASV(C) infection. In these cases, the tumor in-

cidence was 100% in both lines, with similar tumor growth rates, depending on the initial dose of ASV (Fig. 4A, B, and C). Animals that received the 1,000-FFU dose were terminated due to tumor burden by day 18 and those that received the 100-FFU dose by days 23 to 28. Very different tumor incidence and progression were observed with ASV(B) and ASV(D) infections of line M chicks. The 1,000-FFU ASV(B) dose induced sarcomas in 100% of the infected line M animals, but the tumors grew at significantly lower rates than in line CC.R1, with all animals requiring termination by day 28 (Fig. 4B). The 100-FFU ASV(B) dose induced sarcomas in only 45% of the infected line M animals (four of nine chicks), and the tumors that did form grew very slowly. A similar pattern was observed with ASV(D) infection of line M. The 1,000-FFU ASV(D) dose induced tumors in 100% of the infected animals, but the tumors grew at an even lower rate (Fig. 4C) than ASV(B)-induced tumors (Fig. 4B). The 100-FFU ASV(D) dose induced sarcomas in only 45% of the infected animals (four of nine chicks), and the tumors that did form grew even more slowly. Finally, ASV(E) induced sarcomas in 100% of the infected line L15 animals, and the tumor growth rate was dependent on the virus dose (Fig. 4D). However, even at the 1,000-FFU ASV(E) dose, no infected line M animals (zero of eight chicks) produced detectable tumors even after 100 days postinfection. Assuming that the continual cell-to-cell spread of a transforming virus by reinfection is the major factor facilitating the progression of ASV-induced sarcomas (22), we conclude that the C125S substitution in the Tvb^{S1} receptor is responsible for the reduced susceptibility of line M to ASLV(B) and ASLV(D) infection and the complete reisistance to ASLV(E) infection in vivo.

DISCUSSION

In this study, we describe the identification of the molecular defect in the *tvb* gene of inbred line M chickens, the C125S substitution in the Tvb^{S1} receptor, that significantly decreases the sensitivity of these birds to infection by ASLV(B) and ASLV(D) and completely abrogates their sensitivity to ASLV(E) infection. This single-amino-acid substitution in the Tvb^{R2} receptor in the inbred line M chickens significantly reduces the binding affinity of the Tvb^{S1} receptor for the ASLV glycoproteins and explains the decreased susceptibility to the ASLV(B) and ASLV(D) infection and resistance to ASLV(E) infection. This is the first reported example of a mutation in the ASLV receptors that results in a quantitative effect on ASLV susceptibility and pathogenesis. Various levels of resistance to human immunodeficiency virus type 1 infection have been reported that are explained by polymorphisms in the CD4 receptor and/or the viral coreceptors that result in a significant reduction in the efficiency of the human immunodeficiency virus type 1 glycoprotein interaction with the mutant receptor (23). Therefore, one mechanism whereby this initial interaction between the receptor and virus required for entry can cause selective pressure on the virus population is the interaction of the viral glycoproteins with variant receptors, resulting in quantitative differences in host sensitivity. The altered susceptibility of line M for infection by ASLV(B), ASLV(D), and ASLV(E) was observed both in cultured CEFs and in chicks challenged with ASV.

Originally, the major determinants of the interaction of the ASLV(B), ASLV(D), and ASLV(E) glycoproteins with the Tvb receptors were thought to be independent of CRD3. Indeed, mutant receptors with CRD3 completely deleted were constructed in previous studies and appeared to bind the viral glycoproteins and to function as receptors for these three ASLV subgroups as efficiently as wild-type Tvb (Fig. 5). A 15-amino-acid domain in Tvb CRD1, residues 32 to 46, has been reported to contain the interaction determinants for subgroup B and D ASLVs: viral entry can be mediated using only this domain as a synthesized peptide (18). Residues L36, Q37, L41, and Y42 in this domain are critical for ASLV(B) and ASLV(D) binding and entry (18). The disulfide bond pattern of CRD1 did not appear to be important for Tvb to function as a receptor for ASLV(B) and ASLV(D) in these studies. From the sequence alignment of the Tvb receptor with several TNFR family members, and based on the X-ray structure of the DR5 TRAIL receptor, the $Tvb_{32\text{-}46}$ domain is predicted to be in the nonstructured N terminus of the CRD1 domain, perhaps explaining the retention of receptor function without a particular pattern of disulfide bonds in CRD1.

In contrast to ASLV(B) and ASLV(D), both CRD1 and CRD2 of Tvb are required for efficient ASLV(E) binding and entry (17). The structural integrity of the protein, including the disulfide bonds and residues Y67, N72, and D73 in CRD2, are critical for proper receptor function. Modeling Tvb on the X-ray structure of DR5 TRAIL receptor predicts that these critical residues reside in a highly structured region of the protein. The Tvb^{S3} protein contains the C62S substitution, which presumably alters the structure of the critical CRD2, eliminating binding of the ASLV(E) glycoproteins and virus entry but having no detectable effect on ASLV(B) and ASLV(D) binding and entry (Fig. 5). It has also been suggested that the Tvb^{S1} receptor protein may exist in at least two forms that result from alternate patterns of protein folding and disulfide bonds: type 1 confers susceptibility to subgroup B, D, and E ASLVs; type 2 confers susceptibility to only subgroup B and D ASLVs (1).

The C125S substitution likely alters the folding and final structure of at least CRD3 in the Tvb^{R2} protein, since the C125 residue in Tvb is predicted to exist as a disulfide bond with C143 in CRD3 (Fig. 5, Tvb^{R2} A). Published studies have demonstrated that Tvb receptors with the CRD3 domain deleted (TvbΔCRD3) appear to function as efficiently as a wild-type Tvb^{S1} receptor (17). Therefore, a simple model requiring the interaction of the viral glycoproteins with two domains for efficient virus binding and entry, ASLV(B) and ASLV(D) with the $Tvb_{32\text{-}46}$ domain and CRD3 and ASLV(E) with CRD2 and CRD3, does not seem likely. The $Tvb_{32\text{-}46}$ domain expressed in a heterologous protein can also act as an ASLV(B) and ASLV(D) receptor (18). Invariably, experiments that study the functions of different domains and/or mutations in receptor function express the proteins at very high levels compared to the levels found naturally in chicken cells. In an earlier study, we reported that the *tva*^{r} allele found in line C chickens contains a cysteine-to-tryptophan substitution that significantly reduces the binding affinity of ASLV(A) glycoproteins for the Tva receptor (10). However, the ectopic expression of the Tva^R receptor at high levels in resistant cells conferred susceptibility to ASLV(A) infection at levels similar to wild-type

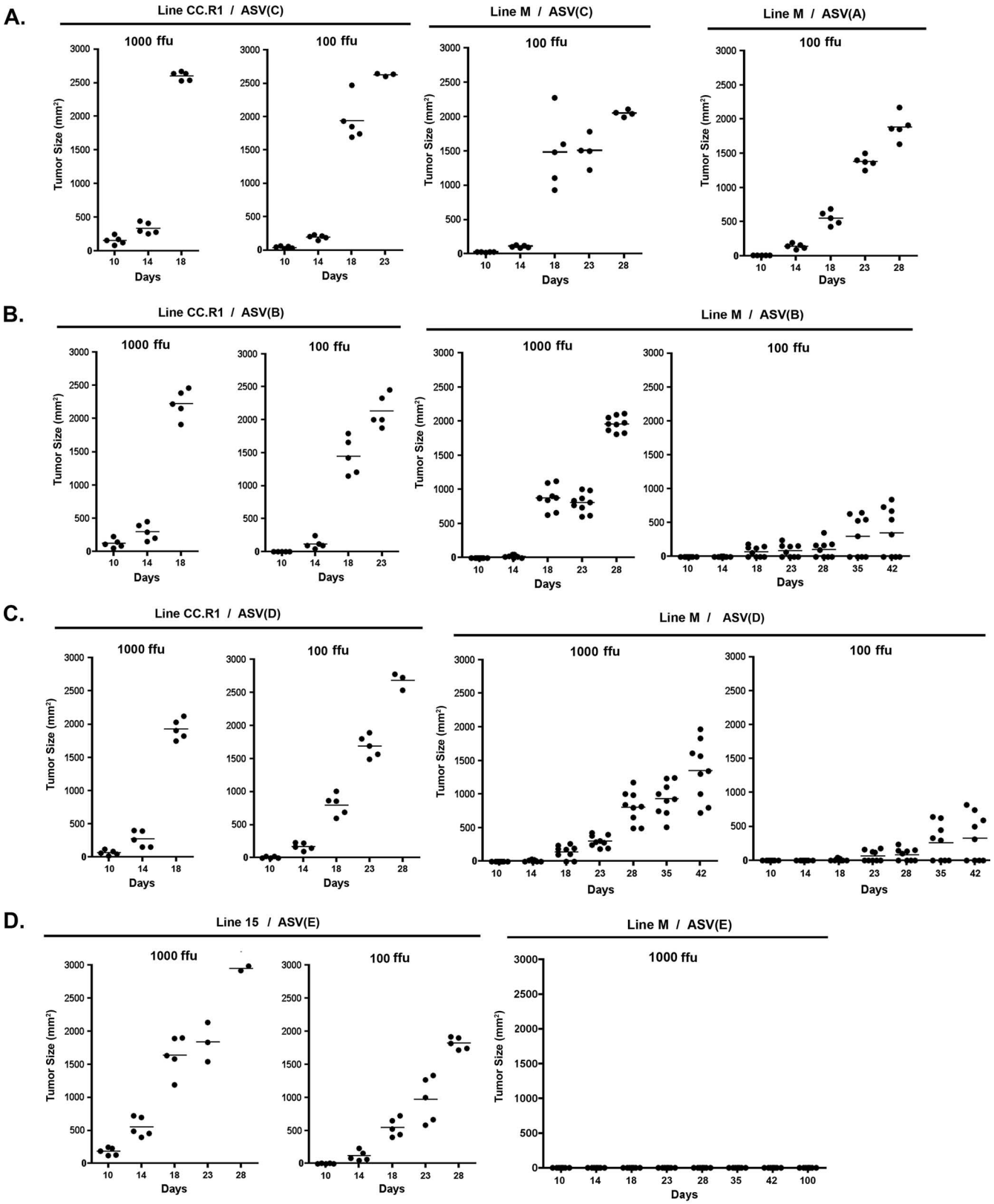


FIG. 4. The growth of sarcomas induced after infection with 1,000 FFU or 100 FFU of ASV in line CC.R1, line L15, or line M 10-day-old chicks. The sizes of tumors induced in chicks infected with ASV(A) or ASV(C) (A), ASV(B) (B), ASV(D) (C), or ASV(E) (D) were measured on the days indicated. The tumor size in each bird is indicated by an individual dot, with the average tumor size for the group shown with a horizontal bar.

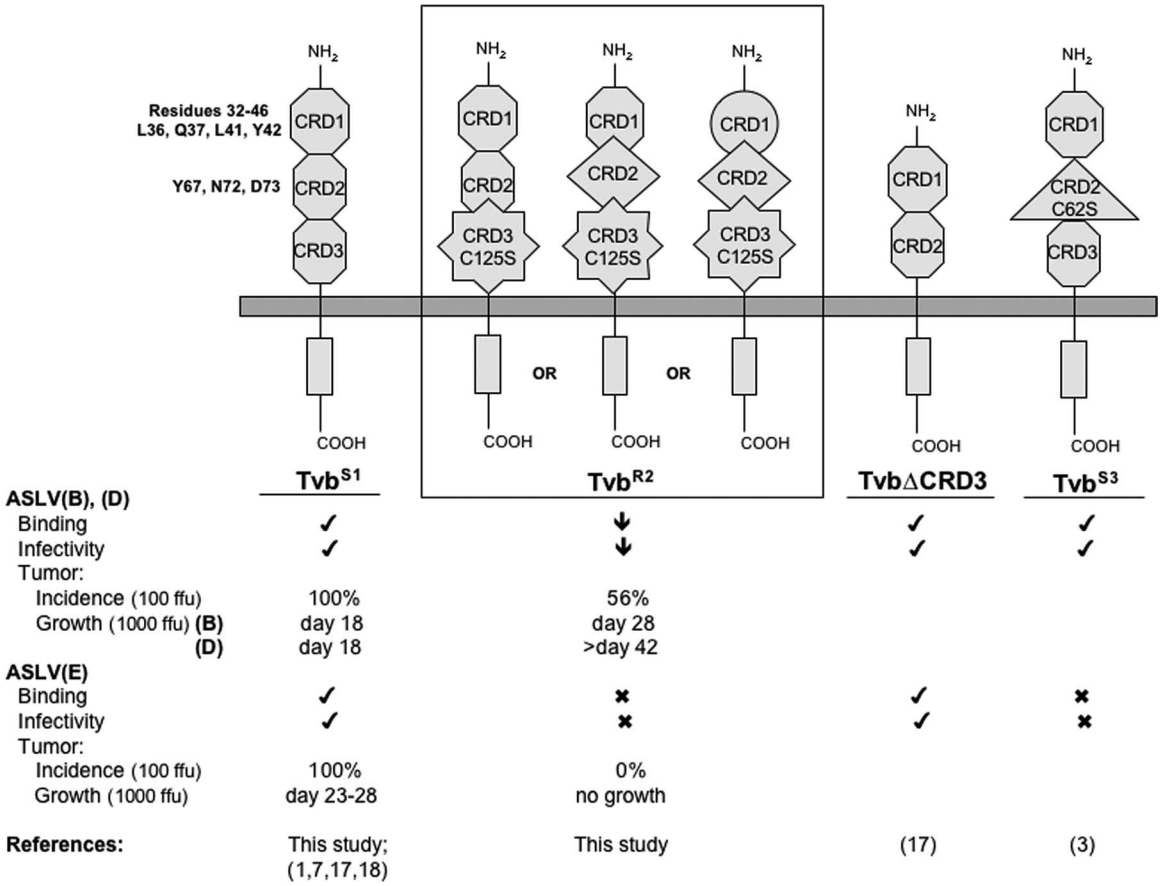


FIG. 5. Schematic representations of membrane-bound forms of the known and functional chicken Tvb proteins. Hypothetical models are shown for Tvb^{S1}, Tvb^{R2} (A, B, and C), the published Tvb^{S1} with CRD3 deleted (TvbΔCDR3), and Tvb^{S3}. Tvb^{S1} depicts the wild-type structure of the CRDs (hexagons). The different shapes of the CRDs denote altered structures induced by either the C125S substitution in Tvb^{R2} or the C62S substitution in Tvb^{S3}. The abilities of the proteins to function as ASLV(B), ASLV(D), and ASLV(E) receptors are indicated. Check mark, wild-type activity; ↓, significantly reduced activity; X, little or no detectable activity. Tumor incidence (the percentages of chickens infected with 100 FFU of ASVs that formed tumors) and growth (the day when all chickens were sacrificed due to tumor burden from infection with 1,000 FFU ASVs) were summarized from Fig. 4.

Tva. The true resistant phenotype of Tva^{R} became apparent experimentally only when very low levels of the receptors were expressed: wild-type Tva still conferred a high level of susceptibility, while the Tva^{R} protein was >1,000-fold less efficient as an ASLV(A) receptor. It may be that when the Tvb proteins are expressed at wild-type levels, mutant proteins with only the $Tvb_{32\text{-}46}$ domain or just CRD1 and CRD2 may still function but at a lower efficiency than receptors that also contain CRD3. The Tvb receptors expressed at levels found on chicken cells may require the interaction of the ASLV glycoproteins with the known domains and CRD3 on the receptor for optimal function.

Alternatively, since the Tvb proteins may naturally exist in several forms, the C125S mutation may result in structural alterations in other regions of the protein. Given the existence of two types of Tvb^{S1} receptors with altered disulfide bonds and the Tvb^{S3} receptor, both with putative alterations in the CRD2 domain that eliminate ASLV(E) binding and entry, it seems obvious to propose that the C125S substitution may also cause alterations in the CRD2 domain structure. Another possible model of the Tvb^{R2} protein to explain the altered ASLV susceptibility phenotype would have an altered CRD3 structure interfering with and/or altering the structure of CRD2, thereby blocking the use of the protein as an ASLV(E) receptor. However, in contrast to the phenotype of the Tvb^{S3} receptor, the binding and entry of ASLV(B) and ASLV(D) are also affected in the Tvb^{R2} receptor. Therefore, if C125S alters both CRD3 and CRD2, the structure of CRD2 in Tvb^{R2} would likely be different than Tvb^{S3} and might inhibit the interaction of ASLV(B) and ASLV(D) with CRD1 (Fig. 5, Tvb^{R2} B). Finally, the C125S substitution could result in an alteration in the structures of all three CRDs (Fig. 5, Tvb^{R2} C), eliminating the use of the protein as an ASLV(E) receptor and significantly reducing, but not eliminating, the binding and entry of ASLV(B) and ASLV(D). Other combinations of these and other models could account for the phenotype of the Tvb^{R2} receptor, including the altered structure of CRD3 resulting from the C125S substitution altering the presentation of intact CRD1 and CRD2 on the membrane so that the virus is sterically blocked from effective interactions.

The normal functions of the Tvb proteins in chickens are unknown and can be estimated only from their similarity to the

TNFR family of proteins. Until we identify these functions, the effect of the C125S substitution on the normal activity of this protein cannot be evaluated. The ASLV glycoproteins may interact with regions of the receptor that do not compete with the interaction of normal ligand(s) of the protein, providing possible scenarios in which the protein can still act as a receptor but loses normal function, and vice versa. Subgroup E ASLVs are endogenous to the germ lines of almost all chickens. To control the spread and accumulation of endogenous ASLV proviruses and possible pathogenesis, polymorphisms in the ASLV receptors that altered the binding of the ASLV glycoproteins while retaining normal function might have offered a positive selective advantage in the chicken population. Selection of mutant Tvb proteins that have lost the ability to function as an ASLV(E) receptor while possibly retaining the "normal" Tvb function may have provided the greatest selective advantage, resulting in fixing of the tvb^{s3} and tvb^{r2} loci in the germ lines of certain lines of inbred chickens.

ACKNOWLEDGMENTS

We thank Viěra Hoserová, Helena Burešová, and Lenka Mikušová for technical assistance with molecular cloning and cell cultures.

This work was supported by the Grant Agency of the Czech Republic (grants no. 523/07/1171 to J.H. and 523/07/1282 to J.G.), by the Academy of Sciences of the Czech Republic (grant no. AV0Z50529514), and by National Institutes of Health grant AI48682 (M.J.F.).

REFERENCES

1. **Adkins, H. B., S. C. Blacklow, and J. A. T. Young.** 2001. Two functionally distinct forms of a retroviral receptor explain the nonreciprical receptor interference among subgroups B, D, and E avian leukosis viruses. J. Virol. **75:**3520–3526.
2. **Adkins, H. B., J. Brojatsch, J. Naughton, M. M. Rolls, J. M. Pesola, and J. A. T. Young.** 1997. Identification of a cellular receptor for subgroup E avian leukosis virus. Proc. Natl. Acad. Sci. USA **94:**11617–11622.
3. **Adkins, H. B., J. Brojatsch, and J. A. T. Young.** 2000. Identification and characterization of a shared TNFR-related receptor for subgroup B, D and E avian leukosis viruses reveal cysteine residues required specifically for subgroup E viral entry. J. Virol. **74:**3572–3578.
4. **Barnard, R. J., D. Elleder, and J. A. T. Young.** 2006. Avian sarcoma and leukosis virus-receptor interactions: from classical genetics to novel insights into virus-cell membrane fusion. Virology **344:**25–29.
5. **Barnard, R. J., and J. A. T. Young.** 2003. Alpharetrovirus envelope-receptor interactions. Curr. Top. Microbiol. Immunol. **281:**107–136.
6. **Bates, P., J. A. T. Young, and H. E. Varmus.** 1993. A receptor for subgroup A Rous sarcoma virus is related to the low density lipoprotein receptor. Cell **74:**1043–1051.
7. **Brojatsch, J., J. Naughton, M. M. Rolls, K. Zingler, and J. A. T. Young.** 1996. CAR1, a TNFR-related protein, is a cellular receptor for cytopathic avian leukosis-sarcoma viruses and mediates apoptosis. Cell **87:**845–855.
8. **Crittenden, L. B., S. McMahon, M. S. Halpern, and A. M. Fadley.** 1987. Embryonic infection with the endogenous avian leukosis virus Rous-associated virus-0 alters resposes to exogenous avian leukosis virus infection. J. Virol. **61:**722–725.
9. **Earp, L. J., S. E. Delos, H. E. Park, and J. M. White.** 2005. The many mechanisms of viral membrane fusion proteins. Curr. Top. Microbiol. Immunol. **285:**25–66.
10. **Elleder, D., D. C. Melder, K. Trejbalová, J. Svoboda, and M. J. Federspiel.** 2004. Two different molecular defects in the Tva receptor gene explain the resistance of two tva^r lines of chickens to infection by subgroup A avian sarcoma and leukosis viruses. J. Virol. **78:**13489–13500.
11. **Elleder, D., V. Stepanets, D. C. Melder, F. Šenigl, J. Geryk, P. Pajer, J. Plachý, J. Hejnar, J. Svoboda, and M. J. Federspiel.** 2005. The receptor for the subgroup C avian sarcoma and leukosis viruses, Tvc, is related to mammalian butrophilins, members of the immunoglobulin superfamily. J. Virol. **79:**10408–10419.
12. **Federspiel, M. J., and S. H. Hughes.** 1997. Retroviral gene delivery. Methods Cell Biol. **52:**179–214.

12a. **Himly, M., D. N. Foster, I. Bottoli, J. S. Iacovoni, and P. K. Vogt.** 1998. The DF-1 chicken fibroblast cell line: transformation induced by diverse oncogenes and cell death resulting from infection by avian leukosis viruses. Virology **248:**295–304.

13. **Hughes, S. H.** 2004. The RCAS vector system. Folia Biol. **50:**107–119.
14. **Hunter, E.** 1997. Viral entry and receptors., p. 71–120. *In* J. M. Coffin, S. H. Hughes, and H. E. Varmus (ed.), Retroviruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
15. **Hymowitz, S. G., H. W. Christinger, G. Fuh, M. Ultsch, M. O'Connell, R. F. Kelley, A. Ashkenazi, and A. de Vos.** 1999. Triggering cell death: the crystal structure of Apo2L/TRAIL in a complex with death receptor 5. Mol. Cell **4:**563–571.
16. **Klucking, S., H. B. Adkins, and J. A. T. Young.** 2002. Resistance of infection by subgroups B, D and E avian sarcoma and leukosis viruses is explained by a premature stop codon within a resistance allele of the *tvb* receptor gene. J. Virol. **76:**7918–7921.
17. **Klucking, S., and J. A. T. Young.** 2004. Amino acid residues Tyr-67, Asn-72, and Asp-73 of the Tvb receptor are important for subgroup E avian sarcoma and leukosis virus interaction. Virology **318:**371–380.
18. **Knauss, D. J., and J. A. T. Young.** 2002. A Fifteen-amino-acid Tvb peptide serves as a minimal soluble receptor for subgroup B avian leukosis and sarcoma viruses. J. Virol. **76:**5404–5410.
19. **Murphy, H. M.** 1977. A new, replication-defective variant of the Bryan high-titer strain Rous sarcoma virus. Virology **77:**705–721.
20. **Plachý, J.** 2000. The chicken—a laboratory animal of the class *Aves*. Folia Biol. **46:**17–24.
21. **Plachý, J., J. Hejnar, K. Trtková, K. Trejbalová, J. Svoboda, and K. Hála.** 2001. DNA vaccination against *v-src* oncogene-induced tumors in congenic chickens. Vaccine **19:**4526–4535.
22. **Ponten, J.** 1964. The in vivo growth mechanism of avian Rous sarcoma. Natl. Cancer Inst. Monogr. **17:**131–145.
23. **Reiche, E. M., A. M. Bonametti, J. C. Voltarelli, H. K. Morimoto, and M. A. Watanabe.** 2007. Genetic polymorphisms in the chemokine and chemokine receptors: impact on clinical course and therapy of the human immunodeficiency virus type 1 infection. Curr. Med. Chem. **14:**1325–1334.
24. **Weiss, R. A.** 1992. Cellular receptors and viral glycoproteins involved in retrovirus entry, p. 1–108. *In* J. A. Levy (ed.), The retroviruses, vol. 2. Plenum Press, New York, NY.
25. **Young, J. A. T., P. Bates, and H. E. Varmus.** 1993. Isolation of a chicken gene that confers susceptibility to infection by subgroup A avian leukosis and sarcoma viruses. J. Virol. **67:**1811–1816.

JVI Accepts, published online ahead of print on 11 June 2008
J. Virol. doi:10.1128/JVI.00419-08

The Core Element of a CpG Island Protects Avian Sarcoma and Leukosis Virus-Derived Vectors from Transcriptional Silencing

Filip Šenigl, Jiří Plachý, Jiří Hejnar*

Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, CZ-14220 Prague 4, Czech Republic

*__Corresponding Author:__ *Jiří Hejnar, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Vídeňská 1083, CZ-14220 Prague 4, Czech Republic. Tel.: (420) 296 443 443, Fax: (420) 224 310 955, E-mail: hejnar@img.cas.cz*

Running Title: Anti-silencing protection of retroviral vectors

Manuscript information: 30 text pages, 5 figures

Abstract Word Count: 160

Text Word Count: 5161

Abbreviations: ASLV, avian sarcoma and leukosis virus; VSV-G, vesicular stomatitis virus glycoprotein; RCAS, replication-competent ASLV vector with a splice acceptor; MLV, murine leukemia virus; HIV-1, human immunodeficiency virus type 1; LTR, long terminal repeat; GFP, green fluorescent protein; IRES, internal ribosomal entry site; IE, island element

ABSTRACT

Unmethylated CpG islands are known to maintain adjacent promoters transcriptionally active. In the adenosine phosphoribosyltransferase gene-adjacent CpG island, the protection from transcriptional silencing can be attributed to the short CpG-rich core element containing Sp1 binding sites. We report here the insertion of this CpG island core element, IE, into the long terminal repeat of a retroviral vector derived from Rous sarcoma virus, which normally suffers from progressive transcriptional silencing in mammalian cells. IE insertion into specific position between enhancer and promoter sequences led to efficient protection of the integrated vector from silencing and gradual CpG methylation in rodent and human cells. Individual cell clones with IE-modified reporter vectors display high levels of reporter expression for a sustained period and without substantial variegation in the cell culture. The presence of Sp1 binding sites is important for the protective effect of IE but at least some part of the entire anti-silencig capacity is maintained in IE with mutated Sp1 sites. We suggest that this strategy of anti-silencing protection by CpG island core element may prove generally useful in retroviral vectors.

INTRODUCTION

Among retroviral vectors used for gene transfer, transgenesis, or even for gene therapy, the avian sarcoma and leukosis viruses (ASLV)-based vectors are of special importance. They have been used for infection of both avian and mammalian cells expressing specific surface receptors. Alternatively, these vectors were pantropized by vesicular stomatitis virus glycoprotein (VSV-G). The replication-competent ASLV vector with a splice acceptor (RCAS) was developed to facilitate the production of high-titer virus in avian cells (11) and efficient infection of mammalian cells when the avian envelope gene is replaced with murine envelope genes, amphotropic or ecotropic (1,2). A transgenic mouse expressing the chicken receptor for the ASLV-A subgroup was established as a platform for *in vivo* infection of mammalian host with ASLV (13) and similar system with ASLV-C subgroup receptor (8) is under preparation. Successful transduction of hematopoietic progenitor cells in Rhesus monkey (26) brings the RCAS vector system and other ASLV-based vectors closer to use as a tool for gene therapy.

Replication of retroviruses requires cellular cofactors for proper entry, uncoating, transcription, splicing, and assembly. ASLVs cannot replicate in mammalian cells because of the lack of specific avian cofactors required for correct splicing, polyprotein cleavage, virus assembly etc (48). Therefore, ASLV-derived vectors such as RCAS do not produce any infectious progeny in mammalian cells *in vitro* (12) or *in vivo* (26,39). In addition, mammalian cells do not contain any endogenous retroviruses, which are capable of recombination with ASLVs, but a high-titer virus can be produced in *ev*-loci-free DF-1 chicken cell line (23). The ASLV-based vectors are thus very stable and safe for gene therapy. Another advantage of ASLV-based vectors is their integration pattern, which differs from murine leukemia virus (MLV) or

human immunodeficiency virus type 1 (HIV-1). Genome-wide analyses of retrovirus integration showed that HIV-1-based vectors integrate preferentially into gene-rich regions (9) and particularly into open chromatin of highly transcribed genes (44). MLV also integrates with a high preference for expressed genes, particularly into the transcriptional start sites of genes (51), and such integration might transcriptionally activate adjacent proto-oncogenes as shown in the SCID-X1 gene therapy trial (17). In comparison with MLV and HIV-1, ASLVs display the weakest preference for integration into genes and has no bias for promoter regions (35,38,42). Hence, in this way ASLV-based vectors might be safer than the widely used lentiviral or MLV-based vectors.

Lentiviral vectors have recently been used predominantly due to their ability to transduce differentiated non-dividing cells. However, the presence of at least weak nuclear localization signal was shown in ASLV (30). Active nuclear import of the preintegration complex and efficient transduction of nondividing cells and terminally differentiated neurons was shown *in vitro* (14,18,28). Increased transduction times improved the efficiency of gene transfer into various types of hematopoietic cells in vivo (26). Optimization of transduction protocols might further increase this efficiency and broaden the use of ASLV-derived vectors in mammalian cells, particularly in the hematopoietic cell gene therapy.

Another obstacle in the use of ASLV-derived vectors in mammalian cells is the transcriptional silencing of integrated proviruses. In general, the expression of retrovirus-driven gene reporters is not stable in long-term *in vitro* cultures and gradual silencing of transduced vectors correlates with epigenetic changes of retroviral long terminal repeats (LTRs). CpG methylation of DNA and/or modifications of histones in nucleosomes linked to the promoter region were found in silenced proviruses *in vitro*. Particularly, the silencing of murine leukemia virus (MLV) and human immunodeficiency virus type 1 (HIV-1) has been characterized in detail (4,25,31,36,40).

It was found that the chromatin environment at the site of retrovirus integration determines the transcriptional activity of the integrated provirus (27). Rous sarcoma virus (RSV) and ASLV-derived vectors are not progressively silenced in chicken cells however, in the cells of heterologous mammalian hosts, they are efficiently methylated and transcriptionally silenced (45, 22). Protection from silencing and position-dependent suppressive effects of surrounding chromatin is, therefore, highly needed for the ASLV-based vectors. Various anti-methylation and insulation strategies have been applied to increase stability and position-independence of expression of lentiviral and MLV-based vectors (6,43,49,52,53). We have observed previously that RSV proviruses are methylation- and silencing-resistant when flanked with the CpG island from the mouse adenine phosphoribosyl-transferase (*aprt*) gene (20). In the present study, we have analysed the effect of the short island element (IE), the core sequence of the Syrian hamster *aprt*-CpG island (46), inserted within the RSV LTR. We show here the stable and position-independent expression of the reporter vector, which suggests that this type of LTR modification might be used in construction of vectors with protracted transcriptional activity and for the improvement of ALV-based vectors toward the therapeutic application.

MATERIALS AND METHODS

Construction of retroviral vectors. We used the plasmids pLPCX, pLXRN, pIRES2-EGFP (all Clontech, Mountain View, CA) and pH19KE (20) to generate the series of constructs used in this study. We inserted the neomycin resistance gene (*neo*^{r}) as 1,419 bp *Bgl*II - *Eco*RV fragment from the pLXRN plasmid into the multiple cloning site of the pIRES2-EGFP cut with *Bgl*II and

*Sma*I to form pN-IRES-G. The pRMR construct was made by ligation of the 3,445 bp *Bst*EII-*Bal*I RSV LTR-bearing fragment of the pH19KE and 2,548 bp *Psh*AI-*Pvu*II inner fragment of the pLPCX. The cassette containing *neo*^{r}, IRES, and EGFP, the 2,864 bp *Eco*47III – *Hpa*I fragment of the pN-IRES-G, was then ligated into the 4,579 bp pRMR backbone fragment cut with *Sma*I and *Xho*I. The resulting retrovirus vector pRNIG (Fig. 1a) was used for the construction of all insertion variants of the RSV LTR. The pMNIG vector was constructed by ligation of 3,509 bp *Bst*E II – *Cla* I cassette of *neo*^{r}, IRES, and EGFP from pRNIG and 4,458 bp *Bst*EII – *Cla*I LTR fragment of pLPCX. The 5´ and 3´ LTR sequences in pLPCX come from Moloney murine leukemia virus and its replication-defective derivative Moloney murine sarcoma virus, respectively. Minor sequence differences between these LTRs enabled discrimination of the cloned plasmid retrovirus DNA from the proviral DNA that went through reverse transcription.

Three unique restriction sites were introduced into the LTRs of the pRNIG by *in vitro* mutagenesis (see below). The *Bsi*W I site at the position –226 upstream from the transcription start site and the *Bst*Z17 I site at the position –89 were introduced into the 3´LTR and *Hpa*I site at the position +27 in the 5´LTR. The unique restriction sites were used for the insertion of the IE element. The pRNIG1+IE and pRNIG1-IE were constructed by insertion of the IE element into the *Hpa*I restriction site in sense and anti-sense orientation, respectively. The pRNIG2 or pRNIG3 variants were formed analogically by insertion of the IE element into the *Bst*Z17I or *Bsi*WI restriction site, respectively. The pRNIG2-2IE variant was created analogically by inserting two copies of the IE element into the *Bst*Z17I restriction site in anti-sense orientation (Fig. 1c). Vectors with insertion of the IE with mutated Sp1 binding sites (see below) were designated RNIG3+MIE, RNIG3-MIE and RNIG2-MIE. The correct insertion and orientation of IEs was confirmed by DNA sequencing.

Assembly of the IE. The core element of the CpG island of the *aprt* gene (46) was constructed by gene assembly PCR using eight oligonucleotides: 1F 5´-agtcgtatactccagcaaatgcgttacttcctgcc-3´, 2F 5´-aaaagccagcctccccgcaacccac-3´, 3F 5´-tctcccagaggccccgccccgtccc-3´, 4F 5´-gccccctcccggcctctcctcgtgctgg-3´, 1R 5´-tgcggggaggctggcttttggcagg-3´, 2R 5´-gggcggggcctctgggagagtgggt-3´, 3R 5´-cgaggagaggccgggaggggcgggacg-3´ and 4R 5´-atgcgtatactccttagggagcgatccagca-3´. These oligonucleotides were combined in pairs, 1F+1R, 2F+2R, 3F+3R, and 4F+4R in four assembly reactions, 1 – 4, respectively. The cycling conditions were as follows: 96°C for 2 min, 4 cycles of 95°C for 40 s, 70°C for 10 s, 0.3°C/s ramp to 25°C and 72°C for 1 min and 35 cycles of 95°C for 30 s, 54°C for 30 s and 72°C for 30 s, the final extension was at 72°C for 3 min. Products of reactions 1 and 2 were mixed and PCR was set-up with primers 1F and 2R. Reactions 3 and 4 were processed analogically with primers 3F and 4R. The PCR conditions were as follows: 96°C for 2 min and 35 cycles of 95°C for 30 s, 50°C for 30 s and 72°C for 40 s, the final extension was at 72°C for 5 min. The products of these two reactions were mixed into one reaction and the PCR, under the conditions used previously, was performed using primers 1F and 4R, which provide a 142 bp product with *Bst*Z17I restriction sites on both ends. To prepare the IE with *Bsi*WI restriction sites, we performed a PCR reaction of the final product with primers 1Fb 5´-agtccgtacgtccagcaaatgcgttacttcctgc-3´ and 4Rb 5´-agtccgtacgtccttagggagcgatccagc-3´.

Site-directed mutagenesis. All site-directed mutagenesis experiments were performed with the Transformer site-directed mutagenesis kit (Clontech, Mountain View, CA) according to the manufacturer's protocol. The assembled IE was cloned into the pGEM-T Easy vector (Promega, Madison, WI). For mutation of Sp1 binding sites, we used a mutagenic primer 5´-ctcccagaggcccattcccgtcctttcccctcccggc-3´, which also served as a selection primer. The selection

was performed with *Dra*II restriction enzyme. To introduce unique cloning sites into the pRNIG construct we used mutagenic primer 5´-ggaaatgtagtcgtacgcaatactcctg-3´ for introduction of the *Bsi*WI cloning site, mutagenic primer 5´-cttattaggaaggtatacagacgggtc-3´ for introduction of the *Bst*Z17I site, and mutagenic primer 5`- acattggtgttaacctgggttg-3` for introduction of the *Hpa*I site. The selection of the vectors with the new sites was performed with alternate use of two selection primers, select *Sca*I/*Bgl*II primer 5´-gtgactggtgagatctcaaccaag-3´ and reselect *Bgl*II/*Sca*I primer 5´-gtgactggtgagtactcaaccaag-3´.

Cell Culture. The packaging GP293 cell line (Clontech, Mountain View, CA) was maintained in D-MEM/F12 Eagle's modified medium (Sigma, St Louis, MO) supplemented with 5% of newborn calf serum, 5% of foetal calf serum (both Gibco-BRL, Gaithersburg, MD), and penicillin/streptomycin (100 mg/ml each). HEK293 human embryo kidney cell line, NIL-2 hamster fibroblastoid cell line (7), QT6 quail methylcholanthrene-transformed cell line (37), and DF1 chicken fibroblastic cell line (23) were maintained in D-MEM/F12 Eagle's modified medium supplemented with 5% of newborn calf serum, 2% of foetal calf serum, 1% of chicken serum (Gibco BRL, Gaithersburg, MD), and penicillin/streptomycin (100 mg/ml each). The tissue cultures were cultivated at 37°C in 3% CO_2 atmosphere. In reactivation experiments, the 5-AzaC and TSA treatment was performed with 4 μM 5-azaC (Sigma) and with 0.5 μM or 1 μM TSA (Sigma) for 4 days.

Vector propagation. The MNIG and RNIG vectors and their modifications were propagated by transfection of plasmid DNA containing the proviral forms of reporter vectors together with the plasmid pVSV-G (Clontech, Mountain View, CA) into GP293 cells. The amount of 1.5 x 10^7

GP293 cells was seeded on 140 mm Petri dishes. After 24 hours, the cells were cotransfected with 10 μg of pVSV-G and 50 μg of the vector plasmid. Cotransfection was performed by calcium phosphate precipitation. The culture medium containing the vector particles was collected 1, 2, and 3 days post-transfection. The collected viral stocks were clarified by centrifugation at 200 x g for 10 min at 4°C. The supernatant was collected and centrifuged at 24 000 rpm for 2 hours 30 min at 4°C in rotor SW28, Beckman Optima100 (Beckman, Fullerton, CA). The pellet was resuspended in a culture medium with 10% of newborn calf serum, frozen and stored in -80°C. The titration of infectious virus particles was performed by serial dilution of the virus stock and subsequent infection of DF1 cells. In repeated experiments, vectors with modified and unmodified LTRs reached similar titers within the range of 2 x 10^4 – 10^5 IU/ml. Twenty-four hours post infection, 400 μg/ml of G418 (Sigma, St Louis, MO) was introduced and the cells were selected for 15 days. The number of G418-resistant colonies was counted after the selection.

Transduction of cells and FACS analysis. The amount of 5 x 10^5 cells per 60 mm Petri dish was seeded and after 5 hours, 200 μl of viral suspension with 15 μg/ml polybrene was applied to the cell culture and allowed to adsorb for 40 min at room temperature. After the adsorption, the fresh medium was added up to 4 ml and the cells were placed at 37°C and 3% CO_2. Twenty-four hours post infection, selection with 400 μg/ml G418 (Sigma) was introduced and changed every two or three days. After 15 days of selection (6 days in reactivation experiment), G418 was withdrawn. In one-week intervals, the cell cultures were analysed with the LSR II cytometer (Becton-Dickinson, San Jose, CA) and the frequencies of the GFP^+ cells were assessed. At specific intervals, the cultures were sorted with a FACSVantage SE (Becton-

Dickinson) device according to the presence or absence of GFP expression. In the case of clonal FACS analysis, the cell clones were obtained by limit dilution of transduced cell cultures.

Methylation analysis. The genomic DNA isolated from infected cells was treated with sodium bisulfite using the EpiTect bisulfite kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The semi-nested PCR of the upper strand was performed with primers complementary to the U3 region of the RSV LTR and the leader region (Fig. 5b) comprising all but one CpGs within the LTR. Sequences of the primers were as follows: 5´-gttttataaggaaagaaaag-3´ (upper), 5´-aacccccaaataaaaaacccc-3´ (lower-inner) and 5´-aaacaaaaatctccaaatcc-3´ (lower-outer). PCR reactions were carried out with 200 ng of DNA at 25 cycles of 95°C for 1 min, 58°C for 2 min and 72°C for 90 s. The PCR products were cloned into pGEM-T Easy vector (Promega) and sequenced by using universal pUC/M13 forward primer.

RESULTS

The core element of the CpG island stabilizes long-term expression of the RSV LTR-driven retroviral vector.

We constructed retroviral vectors RNIG and MNIG with neomycin resistance (neo^r) and enhanced green fluorescence protein (GFP) genes under the control of the RSV and MLV LTRs, respectively. Both genes are expressed from a single bicistronic mRNA by virtue of the internal ribosomal entry site (IRES) from the encephalomyocarditis virus (Fig. 1a). The RSV LTR contains a highly efficient promoter, but is prone to CpG methylation and transcriptional silencing upon infection of mammalian cells (21,22). For the anti-silencing protection of RSV

LTR, we have chosen the IE of the hamster *aprt* CpG island characterized previously as an efficient protective motif (46). The IE comprises a 120 bp sequence that includes eight CpG dinucleotides and two high affinity Sp1 binding sites (Fig. 1b) and is responsible for the majority of the CpG island properties. Single IE was inserted in both orientations into three positions within the RSV LTR, i. in the U5 region downstream to the promoter and transcription start, ii. in the middle of the U3 region between the enhancer and the promoter, and iii. at the beginning of the U3 region, upstream to the enhancer. Furthermore, two IEs were inserted in an anti-sense orientation between the enhancer and the promoter. The U5 region was targeted in the 5´LTR and U3 at the 3´LTR, so that after the reverse transcription of the corresponding RNA all modifications appear within both LTRs of the resulting provirus. The complete set of modified RSV LTRs with the names of corresponding retroviral vectors is shown in Fig. 1c.

To study the stability of the long-term expression, we packaged the wild-type and modified vectors in GP293 cells and transduced them into two avian cell lines, chicken DF1 and quail QT6, and two mammalian cell lines, hamster NIL-2 and human HEK 293. The cells that contained transcriptionally active transduced genes were selected with G418 and the expression of GFP was followed at regular intervals after the G418 withdrawal by microscopic inspection of GFP variegation and fluorescence-activated cell sorting (FACS). All vectors with either wild-type or modified LTRs exhibited very stable expression in the DF1 chicken cell line during *in vitro* cultivation. Only few GFP-negative cells appeared after two months of cultivation without selective pressure (data not shown). A similar lack of variegation was also observed in QT6 cells transduced with RNIG, RNIG2-2IE, and MNIG vectors. We isolated six monocellular clones of the G418-resistant cells after transduction with each of these three vectors and cultivated them separately for more than two months. All clones exhibited very stable expression often without any sign of silencing (data not shown). There was only one clone from the QT6 cell line

transduced with the vector with unmodified RSV LTR which was progressively silenced, and after 70 days of cultivation (without selection) 22 % of the cells were GFP-negative.

Different behavior of integrated vectors was observed after the transduction of mammalian cells. The vectors RNIG and MNIG were progressively silenced in NIL-2 cells whereas the vectors with inserted IE element exhibited varying degrees of protection from transcriptional silencing (Fig. 2,3a-c). The insertion of a single or duplicated IE between the promoter and the enhancer nearly completely stabilized the long-term expression of RNIG vectors, irrespective of the orientation (Fig. 2, 3b). Because of the significant variability among cell cultures transduced with the same vector modification, we isolated several G418-resistant NIL-2 clones after infection with individual vectors and observed GFP expression over the course of time. Various numbers of clones were inspected, from six clones with the RNIG2+IE vector up to 19 clones with the RNIG2-2IE vector. The data are shown individually for both the wild-type RNIG and RNIG2-2IE vectors and cumulatively for the remaining vectors. The RNIG vector was progressively silenced in most cell clones. Out of nine clones isolated, only two were without variegation and exhibited stable GFP expression (Fig. 3d), probably due to insertion into a particular site in the host cell genome, which supported efficient transcription of the integrated retrovirus. In contrast, almost no silencing was observed in clones bearing the vector RNIG2-2IE modified by insertion of double IE in the –89 position. We analysed 19 clones of NIL-2 cells transduced with this vector and observed weak silencing of GFP expression only in two of them. They exhibited 2 % and 8 % of GFP negative cells after 91 days of cultivation. The remaining 17 clones did not exhibit any sign of silencing (Fig. 3e).

The variability in the rate of GFP silencing among individual cell clones corresponded with the protective effect of the IE insertion. There was vast variability in the case of poorly protective LTR modifications, but even here we also observed rare, unsilenced clones with stable GFP

expression. On the other hand, the clones bearing very stable vectors were rather uniform, with weak silencing in only few clones as in the case of the RNIG2-2 vector (Fig. 3e). All experiments mentioned were confirmed in the human HEK 293 cell line. Data concerning silencing and anti-silencing protection correspond to the data obtained in NIL-2 cells. The only difference was generally slower progression of silencing in the HEK 293 cell line. The intensity of GFP fluorescence was similar in NIL-2 and HEK 293 cells and strikingly higher in QT6 cells. This corresponds with higher level of transcription driven by the RSV LTR in avian cells (21). Insertion of IE, even in tandem between the enhancer and the promoter, did not significantly increase the mean intensity of GFP fluorescence, which means that the transcription and titers of modified vectors will not be affected. Much more striking differences in mean fluorescence intensity (MFI) were found among individual clones with vectors integrated into various genomic positions (data not shown).

The Sp1 binding sites are important for the protective role of the CpG island core element.

The IE used in this study comprises two Sp1 binding sites. We introduced point mutations into the Sp1 binding sites that abolished the protein binding capacity of the DNA sequence. RNIG2-IE, RNIG3+IE, and RNIG3-IE vectors were mutagenized in this manner and named RNIG2-MIE, RNIG3+MIE, and RNIG3-MIE, respectively. All Sp1-mutated vectors exhibited significantly decreased protective effect on GFP expression in NIL-2 cells compared to their nonmutated counterparts. The mutated vectors RNIG3+MIE and RNIG3-MIE were silenced more rapidly than the original vector without any element inserted (Fig. 3f). In avian DF-1 and QT6 cells, the GFP expression from the RNIG3-MIE vector was stable and no silencing was observed (data not shown).

Reactivation of the silenced vectors with 5-azacytidine and trichostatine A.

We have assessed the possibility to revert the transcriptional suppression of integrated proviruses by DNA methyltransferase inhibitor 5-azacytidine (5azaC) and/or histone deacetylase inhibitor trichostatine A (TSA). We sorted the GFP-negative cells from silencing-prone NIL-2 clones transduced with RNIG and RNIG3-MIE vectors 11 days after neomycin withdrawal and treated the cell cultures at regular time intervals with the drugs, either alone or in combination. An increase in the percentage of GFP-positive cells was assessed by FACS analysis. Both 5-azaC and TSA applied separately reactivated GFP expression, but the combination of both inhibitors provided the strongest effect (Fig. 4). The majority of silenced vectors were reactivated by 5-azaC and TSA 21 days after G418 withdrawal. However, the effect of these drugs gradually decreased and after 112 days of cultivation without selection, only 2 % of the vectors were reactivated by the treatment (Fig. 4).

DNA methylation analysis of the integrated vectors.

Because transcriptional silencing of genes or proviruses is usually caused by DNA methylation of promoters, we analysed and compared the CpG methylation patterns of the vector LTRs that were either silenced or transcriptionally active. For this analysis, we chose two NIL-2 clonal cell cultures transduced by the RNIG vector, one with a high and the other with very low percentage of GFP-positive cells and one clonal cell culture transduced by the RNIG3-IE vector with 30% of GFP-positive cells. These clones were sampled approximately nine weeks after G418 selection of transduced cells. Bisulfite sequencing of the vectors displayed almost fully methylated 5´LTR sequences in the cell clone with silent RNIG proviruses. In contrast, we found almost no CpG methylation in the cell clone without RNIG vector silencing. The cell clone with 30% GFP-

positive cells was intermedite as to the methylation of 5´LTR of the RNIG3-IE vector (Fig. 5a). There was no obvious methylation pattern and the methylated CpG dinucleotides were observed throughout the entire LTR. To show that the tandem of two IEs inserted in the LTR of RNIG2-2IE vector protects itself and the adjacent DNA sequence from CpG methylation, we analysed the methylation in two clones from the experiment shown in Fig. 3e. The bisulfite sequencing confirmed the presence of two IEs and the non-methylated status of whole LTR (Fig. 5b). The level of CpG methylation was extremely low here, 1.6 and 4.0%.

In order to see the early phase of vector silencing, we examined the CpG methylation in the rare GFP-negative cells that appeared soon after the selection of vector-transduced cells. We selected these cells from two NIL-2 cell cultures transduced with RNIG and RNIG3-MIE vectors by FACS after 21 days of cultivation without selective pressure. As it is difficult to separate the weakly expressing cells, the resulting cell cultures were enriched for GFP-negative cells but also contained, respectively, 15% and 28% of positive cells, respectively (Fig. 5c). The density of CpG methylation within LTR sequences was 28% and 19%, respectively, much lower than that of the nine-week-old cell culture transduced with RNIG3IE, which contained a comparable percentage of silenced cels (Fig. 5c). We therefore conclude that the silencing of integrated proviruses is associated with CpG methylation of promoter sequence, however there are other mechanisms of silencing that precede the onset of heavy methylation of proviral LTRs.

DISCUSSION

In this study, we report the capacity of the CpG island core element from the hamster *aprt* gene, the IE, to stabilize long-term expression of a retroviral vector and protect it from transcriptional silencing. Our data document that insertion of the 120 bp IE into a specific site of the LTR

ensures efficient transcription of ASLV-derived vectors and overcomes the repressive chromosomal position effects in non-permissive mammalian cells. We have found the best protective effect in LTRs with two IEs in tandem inserted between the promoter and enhancer sequences. None of the 19 analyzed cell clones transduced with this modification of the reporter vector exhibited substantial vector silencing after nine weeks of cultivation. We suggest our optimized vector design as a new strategy in improving retroviral vectors toward efficient gene transfer and therapy.

There are two basic strategies of how to counteract the silencing and variegation of retroviral vectors, based on both gammaretroviruses and lentiviruses (10). The first is the elimination of silencers defined *e.g*. in the LTR and primer binding site of MLV (5). Multiple-point mutation of all CpGs in the LTR region also stabilizes the expression of MLV vectors in embryonic stem cells (49). In a complementary approach, matrix attachment regions or insulators have been employed to prevent position effects of adjacent cellular silencers (6,24,43,52). Our previous experiments with the mouse *aprt* CpG island (20) adjacent to the RSV reporter provirus indicated that the anti-silencing and anti-methylation protection might have been caused also by some insulatory effect. We have found that insertion of the CpG island protects the downstream enhancer/promoter in an orientation-dependent manner without any increase of the promoter strength. The range of the anti-methylation effect was relatively narrow covering just the RSV LTR.

Here, we have employed the short core element of the hamster *aprt* CpG island as defined by Siegfried *et al*. (46), which can be inserted into the retroviral LTR. This IE comprises the protective effect of the whole CpG island, but differs from it in several characteristics. The observed stabilization of reporter expression was strongly dependent on the position of the IE within the LTR. The IE inserted at the beginning of the U3 region had only a minor effect, but

insertion between viral enhancers and promoter ensured efficient anti-silencing protection. We do not possess exact data on the promoter strength of IE-modified LTRs, but an estimate of the mean fluorescence intensity (MFI) of non-cloned cell cultures infected with individual vectors demonstrates that there are only mild differences in comparison with the wild-type RSV LTR (data not shown). Furthermore, these differences do not correlate with the anti-silencing effect. For example, the silencing-prone vector RNIG3-MIE exerted slightly higher MFI than the RNIG vector whereas the well protected vector RNIG2+IE displayed slightly decreased MFI. Substantial differences in MFI were found among individual clones irrespective the transduced vector (data not shown), which points to the strong influence of the integration site. In our previous study (32), we showed the increased expression driven by the RSV LTR enriched in Sp1 sites without the context of CpG island. Together, these data are in contradiction with the effect of insulator elements defined as sequences capable of obstructing outside enhancers to influence promoters. The range of anti-methylation effect is limited to approximately 150 bp from the IE (20,46). This might explain the weak or no protection effect of the IE inserted upstream of the enhancer and the best protection from the IE inserted between the enhancer and the promoter. Both transcriptional regulators are in this way within the range of IE anti-silencing influence. The position, 89 bp upstream to the transcription start site, best corresponds to the position of the element within the hamster *aprt* gene, 107 bp upstream from the transcription start site. This distance is probably favourable for keeping the promoter transcriptionally active. In contrast to the full-length mouse CpG island, the effect of the IE orientation is only weak.

Transcriptional silencing of retroviruses has been described in two experimental settings. First, complete transcriptional silencing occurs shortly after infection, most probably during the process of integration. Second, the selected reporter-positive cells variegate during the time course of cultivation. Even cell clones that inherited initially active and uniformly integrated

reporter proviruses variegate and eventually loose expression of the transduced reporter. In our case, we describe just the second type of provirus silencing. To assess the early peri-integration silencing, we performe the colony-forming assay that, in a preliminary experiment, displays a lower number of reporter-positive cells or selected cell colonies that did not correspond to the virus dose determined independently (data not shown). Furthermore, the rate of silencing observed in our experiments might be underestimated. We start our GFP monitoring with completely GFP-positive clonal cell cultures after two weaks of G418 selection. The cells, in which the silencing occurred most rapidly, could have been already selected against at that time. If we monitored the loss of GFP-expressing cells a few days after the infection, we would not obtain reliable data due to the background from non-infected cells as well as from cells expressing transiently the unintegrated forms of retrovirus.

Several studies have described the important role of Sp1 binding sites in the anti-methylation capacity of CpG islands (3,20,33). We have tested the effect of Sp1 mutations in three of our modified LTRs. In all three vectors, mutation of the Sp1 sites significantly increased the provirus silencing during prolonged cultivation, abrogating the protective effect of the inserted IE. In particular, the vector RNIG3-MIE was silenced even more efficiently than the vector with an unmodified LTR. This is probably due to the accumulation of CpG dinucleotides that, without the context of Sp1 sites, become a target of DNA methyltransferases (3). The most stable vector, RNIG2-IE, retains a part of this anti-silencing capacity even after mutation of Sp1 sites. It indicates that Sp1 binding sites are involved in the anti-silencing capacity of IE, but it is not the only *cis*-acting element involved. We do not know whether the binding of Sp1 factor to the IE is necessary for its anti-silencing activity. Nevertheless, the CpG islands in cells derived from Sp1 knock-out mouse embryos are maintained methylation-free (34) suggesting that Sp1 binding sites act in *cis*. Finally, it turns out that the methylation status of CpG islands establishes in early

embryonic cells and than maintains in differentiated cells. Based on aberrant methylation of CpG island in the modified *β-actin* allele, Strathdee et al (47) propose a two-step process to defining a CpG island and the role of Sp1 binding sites might be different in pluripotent and differentiated cells. It will be very interesting to check the anti-methylation and anti-silencing capacity of our IE-modified LTRs also *in vivo* in transgenes passing the germ line and early embryo.

The reactivation experiments showed additive effects of 5-azac and TSA and revealed that vector silencing was associated with DNA methylation and/or histone deacetylation. The inter-dependence of both mechanisms of provirus silencing was described in detail in experiments with lentiviral vectors (19). Katz et al. (29) also found silenced ASLV-based vectors to be reactivable by TSA alone. We observed that this capability strongly depends on the time elapsed from transduction, with a gradual loss of reactivation. This can be explained by the slow and multistep process of heterochromatinization, which requires further histone and chromatin modifiers, finally locking the provirus in a constitutively inactive form (16).

The efficient silencing of ASLV-based vectors has been described in rodent cells (22, 45, 48 for the review). To be sure that this phenomenon and our approach to anti-silencing protection of vectors can be applied to other system, we have performed these experiments in parallel with the human HEK 293 cell line. The results obtained are basically the same as those results seen in NIL-2 cells for all vectors with modified 5´LTR, but the silencing of the wild-type RNIG vector was slower and, correspondingly, the anti-silencing effects of IE insertions were less pronounced. Because the silencing can be associated with the distributive DNA methylation by Dnmt1 (50) during the S-phase of the cell cycle, it is probable that the slower silencing in HEK 293 cells correlates with slower proliferation and longer doubling time in comparison to NIL-2. Also, other cell-specific factors may play a role in the different rate of silencing (15,41).

Transcriptional silencing is a principal obstacle for the use of retroviral and lentiviral vectors in gene transfer and gene therapy. Our strategy of anti-methylation protection of ASLV-based vectors by CpG island core element indicates that retroviral vector completely resistant to silencing can be constructed. This modification may prove immediately useful whenever *e.g* RCAS vectors are used in mammalian cells and it could be part of a general optimized vector design if also proven to be effective in gammaretroviral and lentiviral vectors.

ACKNOWLEDGEMENTS

We thank Jan Svoboda and Jasper Manning (both at the Institute of Molecular Genetics, Prague) for their helpful comments, encouragement and carefull reading of the manuscript; Dana Kučerová, Věra Hoserová, and Helena Burešová for excellent technical assistance. This research was supported by the Grant Agency of the Czech Republic (Grants No. 204/05/0939 and 523/07/1171 to JH) and by the Academy of Sciences of the Czech Republic (grant No. AV0Z50529514).

REFERENCES

1. **Barsov, E.V., and S.H. Hughes.** 1996. Gene transfer into mammalian cells by a Rous sarcoma virus-based retroviral vector with the host range of the amphotropic murine leukemia virus. J. Virol. **70:**3922-3929.

2. **Barsov, E.V., W.S. Payne, and S.H. Hughes.** 2001. Adaptation of chimeric retroviruses in vitro and in vivo: isolation of avian retroviral vectors with extended host range. J. Virol. **75:**4973-4983.

3. **Brandeis M., D. Frank, I. Keshet, Z. Siegfried, M. Mendelsohn, A. Nemes, V. Temper, A. Razin, H. Cedar.** 1994. Sp1 elements protect a CpG island from de novo methylation. Nature **371**:435-8.

4. **Challita, P.M., and D.B. Kohn.** 1994. Lack of expression from a retroviral vector after transduction of murine hematopoietic stem cells is associated with methylation in vivo. Proc. Natl. Acad. Sci. USA **91:**2567-2571.

5. **Challita, P.M., D. Skelton, A. El-Khoueiry, X.J. Yu, K. Weinberg, and D.B. Kohn.** 1995. Multiple modifications in *cis* elements of the long terminal repeat of retroviral vectors lead to increased expression and decreased DNA methylation in embryonic carcinoma cells. J. Virol. **69:**748-755.

6. **Dang, Q., J. Auten, and I. Plavec.** 2000. Human beta interferon scaffold attachment region inhibits de novo methylation and confers long-term, copy number-dependent expression to a retroviral vector. J. Virol. **74:**2671-2678.

7. **Diamond, L.** 1967. Two spontaneously transformed cell lines derived from the same hamster embryo culture. Int. J. Cancer **2**:143-152.

8. **Elleder, D., V. Stepanets, D.C. Melder, F. Šenigl, J. Geryk, P. Pajer, J. Plachý, J. Hejnar, J. Svoboda, and M.J. Federspiel.** 2005. The receptor for the subgroup C avian sarcoma and leukosis viruses, Tvc, is related to mammalian butyrophilins, members of the immunoglobulin superfamily. J. Virol. **79:**10408-10419.

9. **Elleder, D., A. Pavlíček, J. Pačes, and J. Hejnar.** 2002. Preferential integration of human immunodeficiency virus type 1 into genes, cytogenetic R bands and GC-rich DNA regions: insight from the human genome sequence. FEBS Lett. **517:**285-286.

10. **Ellis, J.** 2005. Silencing and variegation of gammaretrovirus and lentivirus vectors. Human Gene Ther. **16:**1241-1246.

11. **Federspiel, M.J., and S.H. Hughes.** 1994. Effects of the gag region on genome stability: avian retroviral vectors that contain sequences from the Bryan strain of Rous sarcoma virus. Virology **203:**211-220.

12. **Federspiel, M.J., D.A. Swing, B. Eagleson, S.W. Reid, and S.H. Hughes.** 1996. Expression of transduced genes in mice generated by infecting blastocysts with avian leukosis virus-based retroviral vectors. Proc. Natl. Acad. Sci. USA **93:**4931-4936.

13. **Federspiel, M.J., P. Bates, J.A. Young, H.E. Varmus, and S.H. Hughes.** 1994. A system for tissue-specific gene targeting: transgenic mice susceptible to subgroup A avian leukosis virus-based retroviral vectors. Proc. Natl. Acad. Sci. USA **91:**11241-11245.

14. **Greger, J.G., R.A. Katz, K. Taganov, G.F. Rall, and A.M. Skalka.** 2004. Transduction of terminally differentiated neurons by avian sarcoma virus. J. Virol. **78:**4902-4906.

15. **Greger, J.G., R.A. Katz, A.M. Ishov, G.G. Maul, and A.M. Skalka.** 2005. The cellular protein Daxx interacts with avian sarcoma virus integrase and viral DNA to repress viral transcription. J. Virol. **79:**4610-4618.

16. **Groth, A., W. Rocha, A. Verreault, and G. Almouzni.** 2007. Chromatin challenges during DNA replication and repair. Cell **128:**21-733.

17. **Hacein-Bey-Abina, S., C. Von Kalle, M. Schmidt, M.P. McCormack, N. Wulffraat, P. Leboulch, A. Lim, C.S. Osborne, R. Pawliuk, E. Morillon, R. Sorensen, A. Forster, P. Fraser, J.I. Cohen, G. de Saint Basile, I. Alexander, U. Wintergerst, T. Frebourg, A. Aurias, D. Stoppa-Lyonnet, S. Romana, I. Radford-Weiss, F. Gross, F. Valensi, E. Delabesse, E. MacIntyre, F. Sigaux, J. Soulier, L.E. Leiva, M. Wissler, C. Prinz, T.H. Rabbitts, F. Le Deist, A. Fischer, and M. Cavazzana-Calvo.** 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science **302:**415-419. Erratum in: Science 302: 568, 2003.

18. **Hatziioannou, T. and S.P. Goff.** 2001. Infection of nondividing cells by Rous sarcoma virus. J Virol. **75**:9526-31.

19. **He J., Q. Yang, and L.J. Chang.** 2005. Dynamic DNA methylation and histone modifications contribute to lentiviral transgene silencing in murine embryonic carcinoma cells. J Virol. **79:**13497-508.

20. **Hejnar, J., P. Hájková, J. Plachý, D. Elleder, V. Stepanets, and J. Svoboda.** 2001. CpG island protects Rous sarcoma virus-derived vectors integrated into nonpermissive cells from DNA methylation and transcriptional suppression. Proc. Natl. Acad. Sci. USA **98:**565-569.

21. **Hejnar, J., J. Plachý, J. Geryk, O. Machoň, K. Trejbalová, R.V. Gunataka, and J. Svoboda.** 1999. Inhibition of the rous sarcoma virus long terminal repeat-driven transcription by in vitro methylation: different sensitivity in permissive chicken cells versus mammalian cells. Virology **255:**171-181.

22. **Hejnar, J., J. Svoboda, J. Geryk, V.J. Fincham, and R. Hák.** 1994. High rate of morphological reversion in tumor cell line H-19 associated with permanent transcriptional suppression of the LTR, v-src, LTR provirus. Cell Growth Differ. **5:**277-285.

23. **Himly, M., D.N. Foster, I. Bottoli, J.S. Iacovoni, and P.K. Vogt.** 1998. The DF-1 chicken fibroblast cell line: transformation induced by diverse oncogenes and cell death resulting from infection by avian leukosis viruses. Virology **248:**295-304.

24. **Hino, S, J. Fan, S. Taguwa, K. Akasaka, and M. Matsuoka.** 2004 Sea urchin insulator protects lentiviral vector from silencing by maintaining active chromatin structure. Gene Ther. **11**:819-28.

25. **Hoeben, R.C., A.A. Migchielsen, R.C. van der Jagt, H. van Ormondt, and A.J. van der Eb.** 1991. Inactivation of the Moloney murine leukemia virus long terminal repeat in murine fibroblast cell lines is associated with methylation and dependent on its chromosomal position. J. Virol. **65:**904-912.

26. **Hu, J., A. Ferris, A. Larochelle, A.E. Krouse, M.E. Metzger, R.E. Donahue, S.H. Hughes, and C.E. Dunbar.** 2007. Transduction of Rhesus Macaque Hematopoietic Stem and Progenitor Cells with Avian Sarcoma and Leukosis Viral Vectors. Hum. Gene Ther. **18:**691-700.

27. **Jordan, A., P. Defechereux, and E. Verdin.** 2001. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. EMBO J. **20:**1726-1738.

28. **Katz, R.A., J.G. Greger, K. Darby, P. Boimel, G.F. Rall, and A.M. Skalka.** 2002. Transduction of interphase cells by avian sarcoma virus. J. Virol. **76:**5422-5434.

29. **Katz, RA, E. Jack-Scott, A. Narezkina, I. Palagin, P. Boimel, J. Kulkosky, E. Nicolas, J.G. Greger, and A.M. Skalka.** 2007. High-frequency epigenetic repression and silencing of retroviruses can be antagonized by histone deacetylase inhibitors and transcriptional activators, but uniform reactivation in cell clones is restricted by additional mechanisms. J Virol. **81:**2592-604.

30. **Kukolj, G., R.A. Katz, and A.M. Skalka.** 1998. Characterization of the nuclear localization signal in the avian sarcoma virus integrase. Gene. **26**:157-63

31. **Lorincz, M.C., D. Schuebeler, S.C. Goeke, M. Walters, M. Groudine, and D.I.K. Martin.** 2000. Dynamic analysis of proviral induction and De Novo methylation: implications for a histone deacetylase-independent, methylation density-dependent mechanism of transcriptional repression. Mol. Cell. Biol. **20:**842-850.

32. **Machoň, O., V. Strmen, J. Hejnar, J. Geryk, and J. Svoboda.** 1998. Sp1 binding sites inserted into the rous sarcoma virus long terminal repeat enhance LTR-driven gene expression. Gene **208:**73-82.

33. **Macleod, D, J. Charlton, J. Mullins, A.P. Bird.** 1994. Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. Genes Dev. **8:**2282-92.

34. **Marin, M., Karis, A., Visser, P., Grosveld, F., and S. Philipsen.** 1997. Transcription factor Sp1 is essential for early embryonic development but dispensable for cell growth and differentiation. Cell **89:**619-628.

35. **Mitchell, R.S., B.F. Beitzel, A.F. Schroeder, P. Shinn, H. Chen, C.C. Berry, J.R. Ecker, and F.D. Bushman.** 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol. **2:**E234.

36. **Mok, H.P., S. Javed, and A. Lever.** 2007. Stable gene expression occurs from a minority of integrated HIV-1-based vectors: transcriptional silencing is present in the majority. Gene Ther. **14:**741-751.

37. **Moscovici, C., M.G. Moscovici, and H. Jimenez.** 1977. Continous tissue culture cell lines derived from chemically induced tumors of japanese quail. Cell **11**:95-103.

38. **Narezkina, A., K.D. Taganov, S. Litwin, R. Stoyanova, J. Hayashi, C. Seeger, A.M. Skalka, and R.A. Katz.** 2004. Genome-wide analyses of avian sarcoma virus integration sites. J Virol. **78:**11656-11663.

39. **Pao, W., D.S. Klimstra, G.H. Fisher, and H.E. Varmus.** 2003. Use of avian retroviral vectors to introduce transcriptional regulators into mammalian cells for analyses of tumor maintenance. Proc. Natl. Acad. Sci. USA **100:**8764-8769.

40. **Pawliuk, R., K.A. Westerman, M.E. Fabry, E. Payen, R. Tighe, E.E. Bouhassira, S.A. Acharya, J. Ellis, I.M. London, C.J. Eaves, R.K. Humphries, Y. Beuzard, R.L. Nagel, and P. Leboulch.** 2001. Correction of sickle cell disease in transgenic mouse models by gene therapy. Science **294:**2368-2371.

41. **Poleshko, A., I. Pelagin, R. Zhang, P. Boimel, C. Castagna, P.D. Adams, A.M. Skalka, and R.A. Katz.** 2008. Identification of cellular proteins that maintain retroviral epigenetic silencing: evidence for an antiviral response. J. Virol. **82:**2313-2323.

42. **Reinišová, M., A. Pavlíček, P. Divina, J. Geryk, J. Plachý, and J. Hejnar.** 2008. Target site preferences of subgroup C Rous sarcoma virus integration into the chicken DNA. The Open Genomics Journal, *In press*.

43. **Rivella, S., J.A. Callegari, C. May, C.W. Tan, and M. Sadelain.** 2000. The cHS4 insulator increases the probability of retroviral expression at random chromosomal integration sites. J. Virol. **74:**4679-4687.

44. **Schroeder, A.R., P. Shinn, H. Chen, C.C. Berry, J.R. Ecker, and F. Bushman.** 2002. HIV-1 integration in the human genome favors active genes and local hotspots. Cell **110:**521-529.

45. **Searle, S., D.A. Gillespie, D.J. Chiswell, and J.A. Wyke.** 1984. Analysis of the variations in proviral cytosine methylation that accompany transformation and morphological reversion in a line of Rous sarcoma virus-infected Rat-1 cells. Nucleic Acids Res. **12:**5193-5210.

46. **Siegfried, Z., S. Eden, M. Mendelsohn, X. Feng, B.Z. Tsuberi, and H. Cedar.** 1999. DNA methylation represses transcription in vivo. Nat. Genet. **22:**203-206.

47. **Strathdee, D., Whitelaw, C.B.A., and A.J. Clark.** 2008. Distal transgene insertion affects CpG island maintenance during differentiation. J. Biol. Chem. **283:**11509-11515.

48. **Svoboda, J., J. Hejnar, J. Geryk, D. Elleder, and Z. Vernerová.** 2000. Retroviruses in foreign species and the problem of provirus silencing. Gene **261:**181-188.

49. **Swindle, C.S., H.G. Kim, and C.A. Klug.** 2004. Mutation of CpGs in the murine stem cell virus retroviral vector long terminal repeat represses silencing in embryonic stem cells. J. Biol. Chem. **279:**34-41.

50. **Vilkaitis, G., I. Suetake, S. Klimašauskas, and S. Tajima.** 2005. Processive methylation of hemimethylated CpG sites by mouse Dnmt1 methyltransferase. J. Biol. Chem. **280:**64-72.

51. **Wu, X., Y. Li, B. Crise, and S.M. Burgess.** 2003. Transcription start regions in the human genome are favored targets for MLV integration. Science **300:**1749-1751.

52. **Yannaki, E., J. Tubb, M. Aker, G. Stamatoyannopoulos, and D.W. Emery.** 2002. Topological constraints governing the use of the chicken HS4 chromatin insulator in oncoretrovirus vectors. Mol. Ther. **5:**589-598.

53. **Zhang, F., S.I. Thornhill, S.J. Howe, M. Ulaganathan, A. Schambach, J. Sinclair, C. Kinnon, H.B. Gaspar, M. Antoniou, and A.J. Thrasher.** 2007. Lentiviral vectors containing an enhancer-less ubiquitously-acting chromatin opening element (UCOE) provide highly reproducible and stable transgene expression in haematopoietic cells. Blood **110:**14481457.

FIGURE LEGENDS

Figure 1 Schematic representation of retroviral vectors used throughout this study. **(a)** The basic vectors RNIG and MNIG containing the *neo*^{r} and EGFP genes

driven from the 5´ LTR of Rous sarcoma virus (RSV) or murine leukemia virus (MLV), respectively. PBS, primer-binding site; ψ^+, encapsidation signal from MLV; IRES, internal ribosome entry site; PPT, polypurine tract. **(b)** The nucleotide sequence of the island element (IE) of the hamster adenosin phosphoribosyltransferase (*aprt*) CpG island. The Sp1 binding sites are in frames, the CpG dinucleotides are in bold cases. The substitution of three bases in each Sp1 binding site that abolished the Sp1 binding capacity is depicted. **(c)** Modifications of the RNIG vector LTRs. Single or duplicated IEs with intact or mutated Sp1 binding sites were inserted in the depicted unique restriction sites in sense and/or anti-sense orientations. The orientation of IE shown by arrow, the bicistronic coding regions and 3´LTR not shown.

Figure 2 FACS analysis of GFP expression from modified and unmodified vector LTRs. NIL-2, HEK 293, and QT6 cell lines transduced with RNIG and RNIG2-2IE vectors were selected by G418 and FACS-analyzed 57 days after selection withdrawal. One clone from each cell line with each of the two vectors are shown. In histograms, the relative GFP fluorescence is plotted against the cell count and the percentage of GFP-positive cells is indicated. The range of GFP positivity is shown as a horizontal bar. As a negative control, the autofluorescence of mock-infected NIL-2, HEK 293, and QT6 cells is shown.

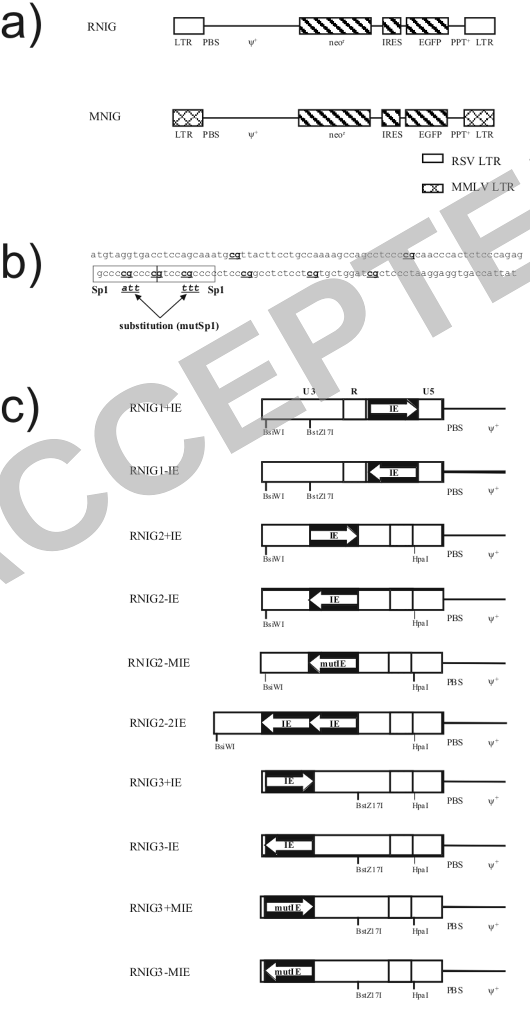
Figure 3 Stability of the GFP expression from retroviral vectors with modified and unmodified LTRs during long-term cultivation of transduced cells. NIL-2 cells were infected with retroviral vectors modified by IE inserted into the U5 region of LTR (**a**),

between the LTR promoter and enhancer (**b**), and at the beginning of the U3 region (**c**). Transduced cells were selected with G418 for 15 days, the percentage of GFP-positive cells was quantified by FACS in regular intervals after the selection was withdrawn, and the cells were cultivated without the selection. Each point represents the mean of three culture dishes. (**d,e**) NIL-2 cells were infected with the modified and unmodified vectors, individual clones of G418-resistant cells were isolated after 7 days of selection and cultivated separately without selection. The time course of vector silencing was followed by FACS at regular intervals and the percentage of GFP-positive cells is given as the mean of all cell clones transduced with the same vector construct. (**e**) Stability of the GFP expression after selection withdrawal of 19 individual cell clones transduced with the RNIG2-2IE vector. The diagrams of 17 non-silenced cell clones merge into one. (**d**) Stability of the GFP expression of nine individual cell clones transduced with the control RNIG vector. (**f**) Stability of expression of vectors modified by insertion of IE with mutated Sp1 binding sites. The population of NIL-2 cells was transduced with the vectors, selected withG418, and individual clones were isolated and cultivated separately. Each curve represents the mean percentage of GFP-positive cells of all cell clones transduced with the same vector construct.

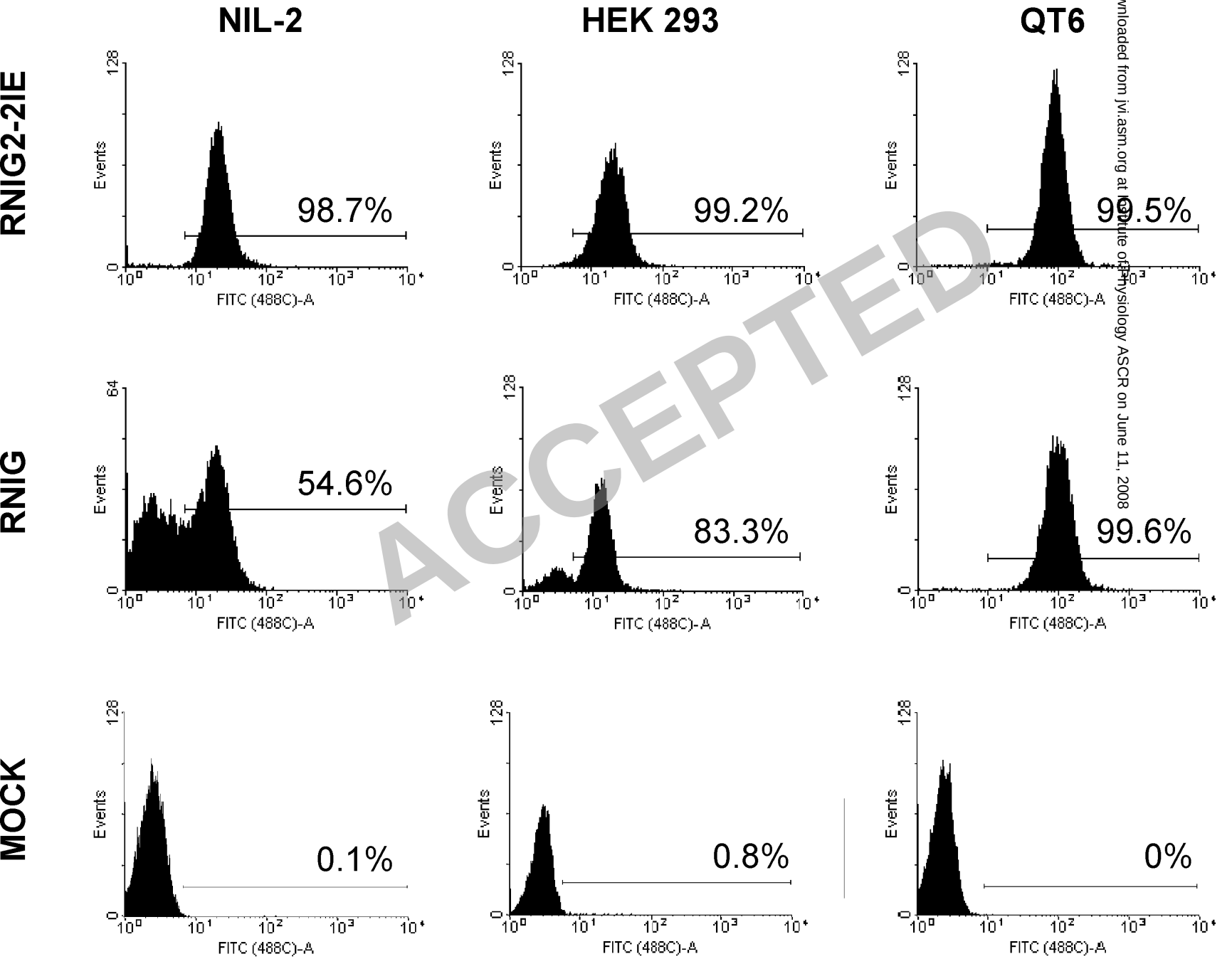
Figure 4 Reactivation of silenced vectors. NIL-2 cells were infected with silencing-prone vectors RNIG and RNIG3-MIE with mutated Sp1 binding sites. The cells were selected with G418 and after six days of selection, the selection was withdrawn. After 18-day-cultivation, the GFP-negative cells were sorted by FACS. Each GFP-negative population was then divided into two subpopulations six days after the sorting: one was treated with 5-azaC and/or TSA for four days and the proportion of GFP-expressing cells

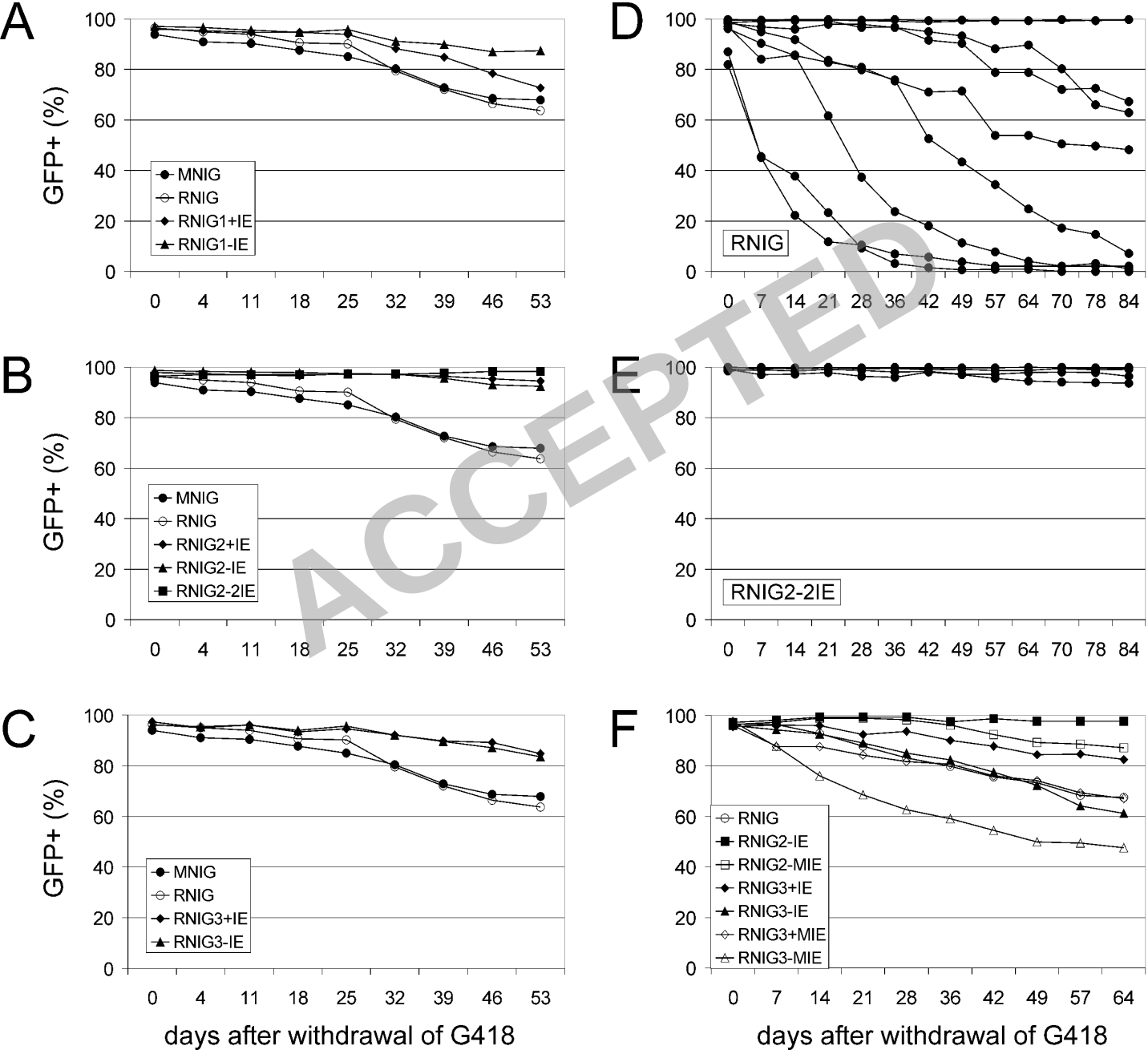
was assessed. The nontreated GFP-negative cell population was cultivated further and the treatment was repeated in regular intervals. Each experiment was done in triplicate.

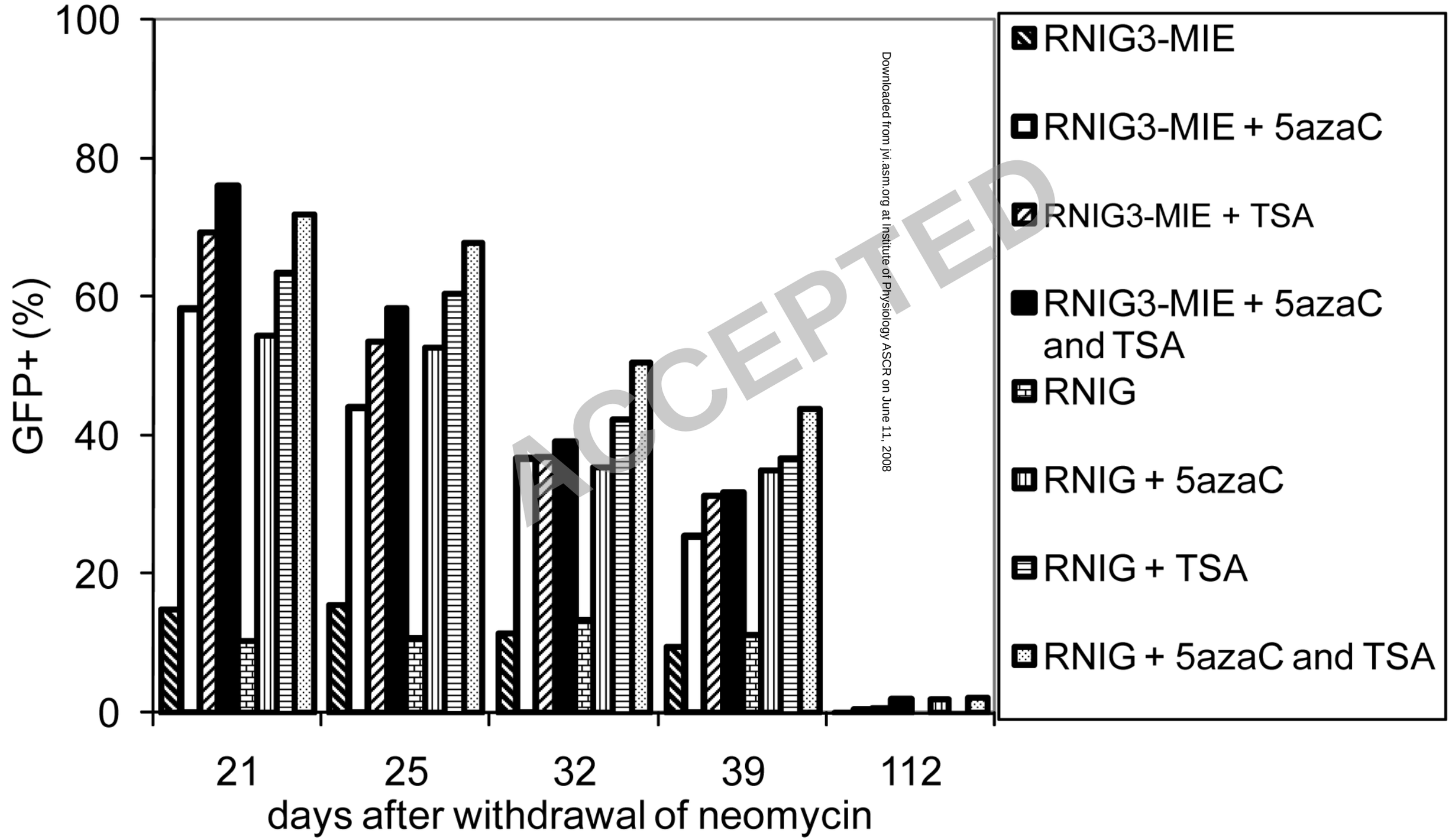
Figure 5 DNA methylation of integrated vectors. The methylation status of CpGs within the LTR of integrated vectors was assayed by the bisulfite technique. (**a**) CpG methylation status of the 5´LTR sequence of silencing-prone vectors RNIG and RNIG3-IE in three cell clones with different level of GFP silencing. NIL-2 cells transduced with vectors were sampled nine weeks after selection withdrawal and subjected to the bisulfite sequencing. (**b**) CpG methylation status of the 5´LTR sequence of the RNIG2-2IE vector protected from CpG methylation in two cell clones. NIL-2 cells transduced with RNIG2-2IE vector were sampled nine weeks after selection withdrawal and subjected to the bisulfite sequencing. (**c**) CpG methylation status of the 5´LTR sequence of silencing-prone vectors RNIG and RNIG3-MIE in two cell cultures enriched with GFP-negative cells early after vector infection. NIL-2 cells transduced with vectors and enriched with GFP-negative cells by FACS three weeks after selection withdrawal were sampled and subjected to the bisulfite sequencing. Each line with circles represents one independent LTR sequence of the PCR product obtained from the bisulfite-treated DNA. Percentages of GFP-positive cells and methylated CpGs are indicated. Methylated CpGs are depicted by solid circles, non-methylated CpGs are indicated by open circles. Localization of primers used for semi-nested PCR after the sodium bisulfite treatment and the site of insertion of the IE in the RNIG3-IE vector are depicted as the horizontal and vertical arrows, respectively.

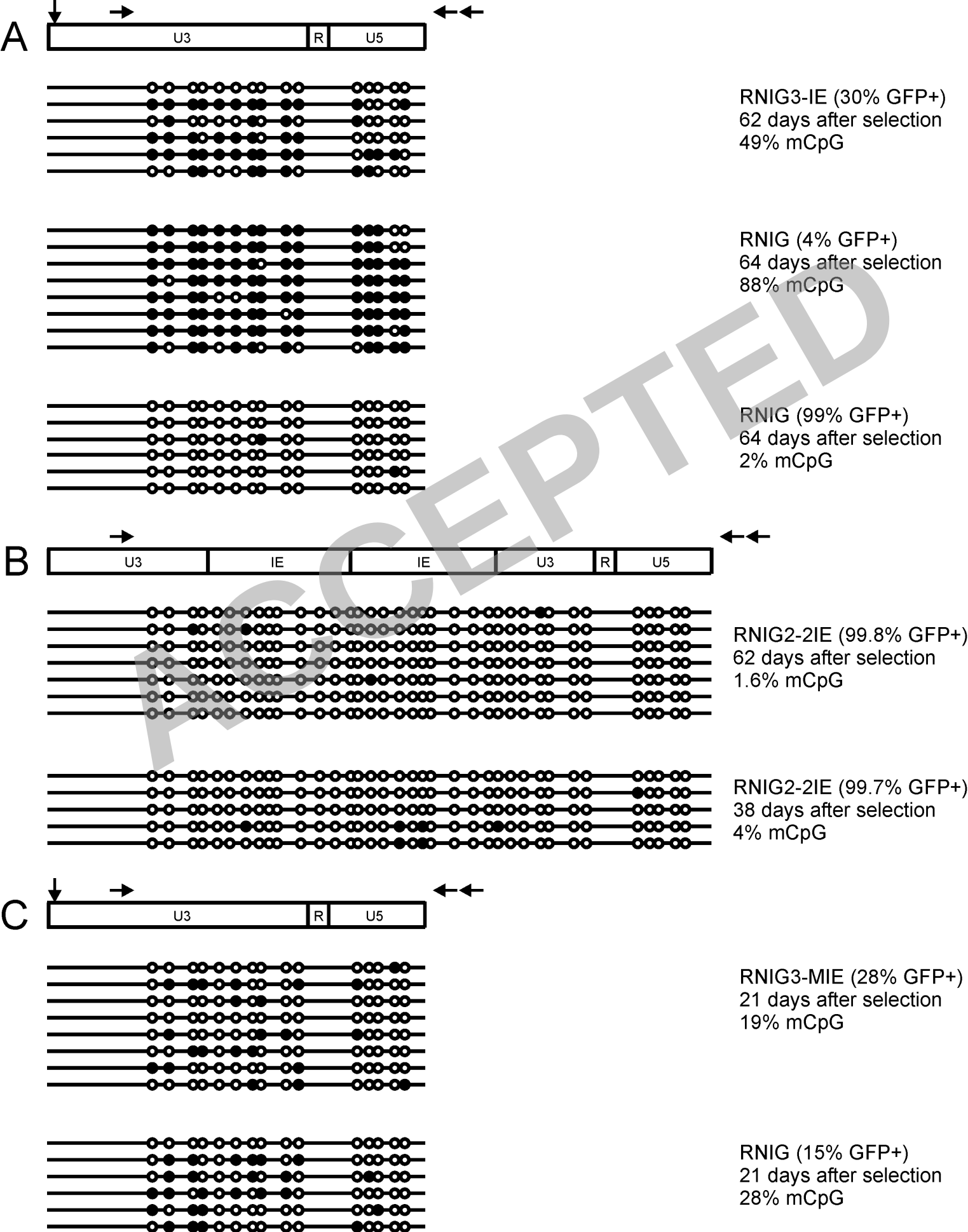


a)
RNIG
LTR PBS ψ+ neor IRES EGFP PPT+ LTR
MNIG
LTR PBS ψ+ neor IRES EGFP PPT+ LTR
RSV LTR
MMLV LTR
b)
atgtaggtgacctccagcaaatgcgttacttcctgccaaaagccagcctccccgcaacccactctcccagag
gccccgccccgtccccgcccctccccgcctctcctcgtgctggatcgctccctaaggaggtgaccattat
Sp1 att ttt Sp1
substitution (mutSp1)
c)
U3 R U5
RNIG1+IE
IE
BsiWI BstZ17I PBS ψ+
RNIG1-IE
IE
BsiWI BstZ17I PBS ψ+
RNIG2+IE
IE
BsiWI HpaI PBS ψ+
RNIG2-IE
IE
BsiWI HpaI PBS ψ+
RNIG2-MIE
mutIE
BsiWI HpaI PBS ψ+
RNIG2-2IE
IE IE
BsiWI HpaI PBS ψ+
RNIG3+IE
IE
BstZ17I HpaI PBS ψ+
RNIG3-IE
IE
BstZ17I HpaI PBS ψ+
RNIG3+MIE
mutIE
BstZ17I HpaI PBS ψ+
RNIG3-MIE
mutIE
BstZ17I HpaI PBS ψ+




A
GFP+ (%)
MNIG
RNIG
RNIG1+IE
RNIG1-IE
B
GFP+ (%)
MNIG
RNIG
RNIG2+IE
RNIG2-IE
RNIG2-2IE
C
GFP+ (%)
MNIG
RNIG
RNIG3+IE
RNIG3-IE
days after withdrawal of G418
D
RNIG
E
RNIG2-2IE
F
RNIG
RNIG2-IE
RNIG2-MIE
RNIG3+IE
RNIG3-IE
RNIG3+MIE
RNIG3-MIE
days after withdrawal of G418




A
U3
R
U5
RNIG3-IE (30% GFP+)
62 days after selection
49% mCpG
RNIG (4% GFP+)
64 days after selection
88% mCpG
RNIG (99% GFP+)
64 days after selection
2% mCpG
B
U3
IE
IE
U3
R
U5
RNIG2-2IE (99.8% GFP+)
62 days after selection
1.6% mCpG
RNIG2-2IE (99.7% GFP+)
38 days after selection
4% mCpG
C
U3
R
U5
RNIG3-MIE (28% GFP+)
21 days after selection
19% mCpG
RNIG (15% GFP+)
21 days after selection
28% mCpG

The Core Element of a CpG Island Protects Avian Sarcoma and Leukosis Virus-Derived Vectors from Transcriptional Silencing

Filip Šenigl, Jiří Plachý, Jiří Hejnar*

Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, CZ-14220 Prague 4, Czech Republic

*__Corresponding Author:__ *Jiří Hejnar, Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Vídeňská 1083, CZ-14220 Prague 4, Czech Republic. Tel.: (420) 296 443 443, Fax: (420) 224 310 955, E-mail: hejnar@img.cas.cz*

Running Title: Anti-silencing protection of retroviral vectors

Manuscript information: 30 text pages, 5 figures

Abstract Word Count: 160

Text Word Count: 5161

Abbreviations: ASLV, avian sarcoma and leukosis virus; VSV-G, vesicular stomatitis virus glycoprotein; RCAS, replication-competent ASLV vector with a splice acceptor; MLV, murine leukemia virus; HIV-1, human immunodeficiency virus type 1; LTR, long terminal repeat; GFP, green fluorescent protein; IRES, internal ribosomal entry site; IE, island element

ABSTRACT

Unmethylated CpG islands are known to maintain adjacent promoters transcriptionally active. In the adenosine phosphoribosyltransferase gene-adjacent CpG island, the protection from transcriptional silencing can be attributed to the short CpG-rich core element containing Sp1 binding sites. We report here the insertion of this CpG island core element, IE, into the long terminal repeat of a retroviral vector derived from Rous sarcoma virus, which normally suffers from progressive transcriptional silencing in mammalian cells. IE insertion into specific position between enhancer and promoter sequences led to efficient protection of the integrated vector from silencing and gradual CpG methylation in rodent and human cells. Individual cell clones with IE-modified reporter vectors display high levels of reporter expression for a sustained period and without substantial variegation in the cell culture. The presence of Sp1 binding sites is important for the protective effect of IE but at least some part of the entire anti-silencig capacity is maintained in IE with mutated Sp1 sites. We suggest that this strategy of anti-silencing protection by CpG island core element may prove generally useful in retroviral vectors.

INTRODUCTION

Among retroviral vectors used for gene transfer, transgenesis, or even for gene therapy, the avian sarcoma and leukosis viruses (ASLV)-based vectors are of special importance. They have been used for infection of both avian and mammalian cells expressing specific surface receptors. Alternatively, these vectors were pantropized by vesicular stomatitis virus glycoprotein (VSV-G). The replication-competent ASLV vector with a splice acceptor (RCAS) was developed to facilitate the production of high-titer virus in avian cells (11) and efficient infection of mammalian cells when the avian envelope gene is replaced with murine envelope genes, amphotropic or ecotropic (1,2). A transgenic mouse expressing the chicken receptor for the ASLV-A subgroup was established as a platform for *in vivo* infection of mammalian host with ASLV (13) and similar system with ASLV-C subgroup receptor (8) is under preparation. Successful transduction of hematopoietic progenitor cells in Rhesus monkey (26) brings the RCAS vector system and other ASLV-based vectors closer to use as a tool for gene therapy.

Replication of retroviruses requires cellular cofactors for proper entry, uncoating, transcription, splicing, and assembly. ASLVs cannot replicate in mammalian cells because of the lack of specific avian cofactors required for correct splicing, polyprotein cleavage, virus assembly etc (48). Therefore, ASLV-derived vectors such as RCAS do not produce any infectious progeny in mammalian cells *in vitro* (12) or *in vivo* (26,39). In addition, mammalian cells do not contain any endogenous retroviruses, which are capable of recombination with ASLVs, but a high-titer virus can be produced in *ev*-loci-free DF-1 chicken cell line (23). The ASLV-based vectors are thus very stable and safe for gene therapy. Another advantage of ASLV-based vectors is their integration pattern, which differs from murine leukemia virus (MLV) or

human immunodeficiency virus type 1 (HIV-1). Genome-wide analyses of retrovirus integration showed that HIV-1-based vectors integrate preferentially into gene-rich regions (9) and particularly into open chromatin of highly transcribed genes (44). MLV also integrates with a high preference for expressed genes, particularly into the transcriptional start sites of genes (51), and such integration might transcriptionally activate adjacent proto-oncogenes as shown in the SCID-X1 gene therapy trial (17). In comparison with MLV and HIV-1, ASLVs display the weakest preference for integration into genes and has no bias for promoter regions (35,38,42). Hence, in this way ASLV-based vectors might be safer than the widely used lentiviral or MLV-based vectors.

Lentiviral vectors have recently been used predominantly due to their ability to transduce differentiated non-dividing cells. However, the presence of at least weak nuclear localization signal was shown in ASLV (30). Active nuclear import of the preintegration complex and efficient transduction of nondividing cells and terminally differentiated neurons was shown *in vitro* (14,18,28). Increased transduction times improved the efficiency of gene transfer into various types of hematopoietic cells in vivo (26). Optimization of transduction protocols might further increase this efficiency and broaden the use of ASLV-derived vectors in mammalian cells, particularly in the hematopoietic cell gene therapy.

Another obstacle in the use of ASLV-derived vectors in mammalian cells is the transcriptional silencing of integrated proviruses. In general, the expression of retrovirus-driven gene reporters is not stable in long-term *in vitro* cultures and gradual silencing of transduced vectors correlates with epigenetic changes of retroviral long terminal repeats (LTRs). CpG methylation of DNA and/or modifications of histones in nucleosomes linked to the promoter region were found in silenced proviruses *in vitro*. Particularly, the silencing of murine leukemia virus (MLV) and human immunodeficiency virus type 1 (HIV-1) has been characterized in detail (4,25,31,36,40).

It was found that the chromatin environment at the site of retrovirus integration determines the transcriptional activity of the integrated provirus (27). Rous sarcoma virus (RSV) and ASLV-derived vectors are not progressively silenced in chicken cells however, in the cells of heterologous mammalian hosts, they are efficiently methylated and transcriptionally silenced (45, 22). Protection from silencing and position-dependent suppressive effects of surrounding chromatin is, therefore, highly needed for the ASLV-based vectors. Various anti-methylation and insulation strategies have been applied to increase stability and position-independence of expression of lentiviral and MLV-based vectors (6,43,49,52,53). We have observed previously that RSV proviruses are methylation- and silencing-resistant when flanked with the CpG island from the mouse adenine phosphoribosyl-transferase (*aprt*) gene (20). In the present study, we have analysed the effect of the short island element (IE), the core sequence of the Syrian hamster *aprt*-CpG island (46), inserted within the RSV LTR. We show here the stable and position-independent expression of the reporter vector, which suggests that this type of LTR modification might be used in construction of vectors with protracted transcriptional activity and for the improvement of ALV-based vectors toward the therapeutic application.

MATERIALS AND METHODS

Construction of retroviral vectors. We used the plasmids pLPCX, pLXRN, pIRES2-EGFP (all Clontech, Mountain View, CA) and pH19KE (20) to generate the series of constructs used in this study. We inserted the neomycin resistance gene (*neo*^{r}) as 1,419 bp *Bgl*II - *Eco*RV fragment from the pLXRN plasmid into the multiple cloning site of the pIRES2-EGFP cut with *Bgl*II and

*Sma*I to form pN-IRES-G. The pRMR construct was made by ligation of the 3,445 bp *Bst*EII-*Bal*I RSV LTR-bearing fragment of the pH19KE and 2,548 bp *Psh*AI-*Pvu*II inner fragment of the pLPCX. The cassette containing *neo*^{r}, IRES, and EGFP, the 2,864 bp *Eco*47III – *Hpa*I fragment of the pN-IRES-G, was then ligated into the 4,579 bp pRMR backbone fragment cut with *Sma*I and *Xho*I. The resulting retrovirus vector pRNIG (Fig. 1a) was used for the construction of all insertion variants of the RSV LTR. The pMNIG vector was constructed by ligation of 3,509 bp *Bst*E II – *Cla* I cassette of *neo*^{r}, IRES, and EGFP from pRNIG and 4,458 bp *Bst*EII – *Cla*I LTR fragment of pLPCX. The 5´ and 3´ LTR sequences in pLPCX come from Moloney murine leukemia virus and its replication-defective derivative Moloney murine sarcoma virus, respectively. Minor sequence differences between these LTRs enabled discrimination of the cloned plasmid retrovirus DNA from the proviral DNA that went through reverse transcription.

Three unique restriction sites were introduced into the LTRs of the pRNIG by *in vitro* mutagenesis (see below). The *Bsi*W I site at the position –226 upstream from the transcription start site and the *Bst*Z17 I site at the position –89 were introduced into the 3´LTR and *Hpa*I site at the position +27 in the 5´LTR. The unique restriction sites were used for the insertion of the IE element. The pRNIG1+IE and pRNIG1-IE were constructed by insertion of the IE element into the *Hpa*I restriction site in sense and anti-sense orientation, respectively. The pRNIG2 or pRNIG3 variants were formed analogically by insertion of the IE element into the *Bst*Z17I or *Bsi*WI restriction site, respectively. The pRNIG2-2IE variant was created analogically by inserting two copies of the IE element into the *Bst*Z17I restriction site in anti-sense orientation (Fig. 1c). Vectors with insertion of the IE with mutated Sp1 binding sites (see below) were designated RNIG3+MIE, RNIG3-MIE and RNIG2-MIE. The correct insertion and orientation of IEs was confirmed by DNA sequencing.

Assembly of the IE. The core element of the CpG island of the *aprt* gene (46) was constructed by gene assembly PCR using eight oligonucleotides: 1F 5´-agtcgtatactccagcaaatgcgttacttcctgcc-3´, 2F 5´-aaaagccagcctccccgcaacccac-3´, 3F 5´-tctcccagaggccccgccccgtccc-3´, 4F 5´-gccccctcccggcctctcctcgtgctgg-3´, 1R 5´-tgcggggaggctggcttttggcagg-3´, 2R 5´-gggcggggcctctgggagagtgggt-3´, 3R 5´-cgaggagaggccgggaggggcgggacg-3´ and 4R 5´-atgcgtatactccttagggagcgatccagca-3´. These oligonucleotides were combined in pairs, 1F+1R, 2F+2R, 3F+3R, and 4F+4R in four assembly reactions, 1 – 4, respectively. The cycling conditions were as follows: 96°C for 2 min, 4 cycles of 95°C for 40 s, 70°C for 10 s, 0.3°C/s ramp to 25°C and 72°C for 1 min and 35 cycles of 95°C for 30 s, 54°C for 30 s and 72°C for 30 s, the final extension was at 72°C for 3 min. Products of reactions 1 and 2 were mixed and PCR was set-up with primers 1F and 2R. Reactions 3 and 4 were processed analogically with primers 3F and 4R. The PCR conditions were as follows: 96°C for 2 min and 35 cycles of 95°C for 30 s, 50°C for 30 s and 72°C for 40 s, the final extension was at 72°C for 5 min. The products of these two reactions were mixed into one reaction and the PCR, under the conditions used previously, was performed using primers 1F and 4R, which provide a 142 bp product with *Bst*Z17I restriction sites on both ends. To prepare the IE with *Bsi*WI restriction sites, we performed a PCR reaction of the final product with primers 1Fb 5´-agtccgtacgtccagcaaatgcgttacttcctgc-3´ and 4Rb 5´-agtccgtacgtccttagggagcgatccagc-3´.

Site-directed mutagenesis. All site-directed mutagenesis experiments were performed with the Transformer site-directed mutagenesis kit (Clontech, Mountain View, CA) according to the manufacturer's protocol. The assembled IE was cloned into the pGEM-T Easy vector (Promega, Madison, WI). For mutation of Sp1 binding sites, we used a mutagenic primer 5´-ctcccagaggcccattcccgtcctttcccctcccggc-3´, which also served as a selection primer. The selection

was performed with *Dra*II restriction enzyme. To introduce unique cloning sites into the pRNIG construct we used mutagenic primer 5´-ggaaatgtagtcgtacgcaatactcctg-3´ for introduction of the *Bsi*WI cloning site, mutagenic primer 5´-cttattaggaaggtatacagacgggtc-3´ for introduction of the *Bst*Z17I site, and mutagenic primer 5`- acattggtgttaacctgggttg-3` for introduction of the *Hpa*I site. The selection of the vectors with the new sites was performed with alternate use of two selection primers, select *Sca*I/*Bgl*II primer 5´-gtgactggtgagatctcaaccaag-3´ and reselect *Bgl*II/*Sca*I primer 5´-gtgactggtgagtactcaaccaag-3´.

Cell Culture. The packaging GP293 cell line (Clontech, Mountain View, CA) was maintained in D-MEM/F12 Eagle's modified medium (Sigma, St Louis, MO) supplemented with 5% of newborn calf serum, 5% of foetal calf serum (both Gibco-BRL, Gaithersburg, MD), and penicillin/streptomycin (100 mg/ml each). HEK293 human embryo kidney cell line, NIL-2 hamster fibroblastoid cell line (7), QT6 quail methylcholanthrene-transformed cell line (37), and DF1 chicken fibroblastic cell line (23) were maintained in D-MEM/F12 Eagle's modified medium supplemented with 5% of newborn calf serum, 2% of foetal calf serum, 1% of chicken serum (Gibco BRL, Gaithersburg, MD), and penicillin/streptomycin (100 mg/ml each). The tissue cultures were cultivated at 37°C in 3% CO_2 atmosphere. In reactivation experiments, the 5-AzaC and TSA treatment was performed with 4 μM 5-azaC (Sigma) and with 0.5 μM or 1 μM TSA (Sigma) for 4 days.

Vector propagation. The MNIG and RNIG vectors and their modifications were propagated by transfection of plasmid DNA containing the proviral forms of reporter vectors together with the plasmid pVSV-G (Clontech, Mountain View, CA) into GP293 cells. The amount of 1.5 x 10^7

GP293 cells was seeded on 140 mm Petri dishes. After 24 hours, the cells were cotransfected with 10 μg of pVSV-G and 50 μg of the vector plasmid. Cotransfection was performed by calcium phosphate precipitation. The culture medium containing the vector particles was collected 1, 2, and 3 days post-transfection. The collected viral stocks were clarified by centrifugation at 200 x g for 10 min at 4°C. The supernatant was collected and centrifuged at 24 000 rpm for 2 hours 30 min at 4°C in rotor SW28, Beckman Optima100 (Beckman, Fullerton, CA). The pellet was resuspended in a culture medium with 10% of newborn calf serum, frozen and stored in -80°C. The titration of infectious virus particles was performed by serial dilution of the virus stock and subsequent infection of DF1 cells. In repeated experiments, vectors with modified and unmodified LTRs reached similar titers within the range of $2 \times 10^4 - 10^5$ IU/ml. Twenty-four hours post infection, 400 μg/ml of G418 (Sigma, St Louis, MO) was introduced and the cells were selected for 15 days. The number of G418-resistant colonies was counted after the selection.

Transduction of cells and FACS analysis. The amount of 5×10^5 cells per 60 mm Petri dish was seeded and after 5 hours, 200 μl of viral suspension with 15 μg/ml polybrene was applied to the cell culture and allowed to adsorb for 40 min at room temperature. After the adsorption, the fresh medium was added up to 4 ml and the cells were placed at 37°C and 3% CO_2. Twenty-four hours post infection, selection with 400 μg/ml G418 (Sigma) was introduced and changed every two or three days. After 15 days of selection (6 days in reactivation experiment), G418 was withdrawn. In one-week intervals, the cell cultures were analysed with the LSR II cytometer (Becton-Dickinson, San Jose, CA) and the frequencies of the GFP^+ cells were assessed. At specific intervals, the cultures were sorted with a FACSVantage SE (Becton-

Dickinson) device according to the presence or absence of GFP expression. In the case of clonal FACS analysis, the cell clones were obtained by limit dilution of transduced cell cultures.

Methylation analysis. The genomic DNA isolated from infected cells was treated with sodium bisulfite using the EpiTect bisulfite kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The semi-nested PCR of the upper strand was performed with primers complementary to the U3 region of the RSV LTR and the leader region (Fig. 5b) comprising all but one CpGs within the LTR. Sequences of the primers were as follows: 5´-gttttataaggaaagaaaag-3´ (upper), 5´-aacccccaaataaaaaacccc-3´ (lower-inner) and 5´-aaacaaaaatctccaaatcc-3´ (lower-outer). PCR reactions were carried out with 200 ng of DNA at 25 cycles of 95°C for 1 min, 58°C for 2 min and 72°C for 90 s. The PCR products were cloned into pGEM-T Easy vector (Promega) and sequenced by using universal pUC/M13 forward primer.

RESULTS

The core element of the CpG island stabilizes long-term expression of the RSV LTR-driven retroviral vector.

We constructed retroviral vectors RNIG and MNIG with neomycin resistance (neo^r) and enhanced green fluorescence protein (GFP) genes under the control of the RSV and MLV LTRs, respectively. Both genes are expressed from a single bicistronic mRNA by virtue of the internal ribosomal entry site (IRES) from the encephalomyocarditis virus (Fig. 1a). The RSV LTR contains a highly efficient promoter, but is prone to CpG methylation and transcriptional silencing upon infection of mammalian cells (21,22). For the anti-silencing protection of RSV

LTR, we have chosen the IE of the hamster *aprt* CpG island characterized previously as an efficient protective motif (46). The IE comprises a 120 bp sequence that includes eight CpG dinucleotides and two high affinity Sp1 binding sites (Fig. 1b) and is responsible for the majority of the CpG island properties. Single IE was inserted in both orientations into three positions within the RSV LTR, i. in the U5 region downstream to the promoter and transcription start, ii. in the middle of the U3 region between the enhancer and the promoter, and iii. at the beginning of the U3 region, upstream to the enhancer. Furthermore, two IEs were inserted in an anti-sense orientation between the enhancer and the promoter. The U5 region was targeted in the 5´LTR and U3 at the 3´LTR, so that after the reverse transcription of the corresponding RNA all modifications appear within both LTRs of the resulting provirus. The complete set of modified RSV LTRs with the names of corresponding retroviral vectors is shown in Fig. 1c.

To study the stability of the long-term expression, we packaged the wild-type and modified vectors in GP293 cells and transduced them into two avian cell lines, chicken DF1 and quail QT6, and two mammalian cell lines, hamster NIL-2 and human HEK 293. The cells that contained transcriptionally active transduced genes were selected with G418 and the expression of GFP was followed at regular intervals after the G418 withdrawal by microscopic inspection of GFP variegation and fluorescence-activated cell sorting (FACS). All vectors with either wild-type or modified LTRs exhibited very stable expression in the DF1 chicken cell line during *in vitro* cultivation. Only few GFP-negative cells appeared after two months of cultivation without selective pressure (data not shown). A similar lack of variegation was also observed in QT6 cells transduced with RNIG, RNIG2-2IE, and MNIG vectors. We isolated six monocellular clones of the G418-resistant cells after transduction with each of these three vectors and cultivated them separately for more than two months. All clones exhibited very stable expression often without any sign of silencing (data not shown). There was only one clone from the QT6 cell line

transduced with the vector with unmodified RSV LTR which was progressively silenced, and after 70 days of cultivation (without selection) 22 % of the cells were GFP-negative.

Different behavior of integrated vectors was observed after the transduction of mammalian cells. The vectors RNIG and MNIG were progressively silenced in NIL-2 cells whereas the vectors with inserted IE element exhibited varying degrees of protection from transcriptional silencing (Fig. 2,3a-c). The insertion of a single or duplicated IE between the promoter and the enhancer nearly completely stabilized the long-term expression of RNIG vectors, irrespective of the orientation (Fig. 2, 3b). Because of the significant variability among cell cultures transduced with the same vector modification, we isolated several G418-resistant NIL-2 clones after infection with individual vectors and observed GFP expression over the course of time. Various numbers of clones were inspected, from six clones with the RNIG2+IE vector up to 19 clones with the RNIG2-2IE vector. The data are shown individually for both the wild-type RNIG and RNIG2-2IE vectors and cumulatively for the remaining vectors. The RNIG vector was progressively silenced in most cell clones. Out of nine clones isolated, only two were without variegation and exhibited stable GFP expression (Fig. 3d), probably due to insertion into a particular site in the host cell genome, which supported efficient transcription of the integrated retrovirus. In contrast, almost no silencing was observed in clones bearing the vector RNIG2-2IE modified by insertion of double IE in the –89 position. We analysed 19 clones of NIL-2 cells transduced with this vector and observed weak silencing of GFP expression only in two of them. They exhibited 2 % and 8 % of GFP negative cells after 91 days of cultivation. The remaining 17 clones did not exhibit any sign of silencing (Fig. 3e).

The variability in the rate of GFP silencing among individual cell clones corresponded with the protective effect of the IE insertion. There was vast variability in the case of poorly protective LTR modifications, but even here we also observed rare, unsilenced clones with stable GFP

expression. On the other hand, the clones bearing very stable vectors were rather uniform, with weak silencing in only few clones as in the case of the RNIG2-2 vector (Fig. 3e). All experiments mentioned were confirmed in the human HEK 293 cell line. Data concerning silencing and anti-silencing protection correspond to the data obtained in NIL-2 cells. The only difference was generally slower progression of silencing in the HEK 293 cell line. The intensity of GFP fluorescence was similar in NIL-2 and HEK 293 cells and strikingly higher in QT6 cells. This corresponds with higher level of transcription driven by the RSV LTR in avian cells (21). Insertion of IE, even in tandem between the enhancer and the promoter, did not significantly increase the mean intensity of GFP fluorescence, which means that the transcription and titers of modified vectors will not be affected. Much more striking differences in mean fluorescence intensity (MFI) were found among individual clones with vectors integrated into various genomic positions (data not shown).

The Sp1 binding sites are important for the protective role of the CpG island core element.

The IE used in this study comprises two Sp1 binding sites. We introduced point mutations into the Sp1 binding sites that abolished the protein binding capacity of the DNA sequence. RNIG2-IE, RNIG3+IE, and RNIG3-IE vectors were mutagenized in this manner and named RNIG2-MIE, RNIG3+MIE, and RNIG3-MIE, respectively. All Sp1-mutated vectors exhibited significantly decreased protective effect on GFP expression in NIL-2 cells compared to their nonmutated counterparts. The mutated vectors RNIG3+MIE and RNIG3-MIE were silenced more rapidly than the original vector without any element inserted (Fig. 3f). In avian DF-1 and QT6 cells, the GFP expression from the RNIG3-MIE vector was stable and no silencing was observed (data not shown).

Reactivation of the silenced vectors with 5-azacytidine and trichostatine A.

We have assessed the possibility to revert the transcriptional suppression of integrated proviruses by DNA methyltransferase inhibitor 5-azacytidine (5azaC) and/or histone deacetylase inhibitor trichostatine A (TSA). We sorted the GFP-negative cells from silencing-prone NIL-2 clones transduced with RNIG and RNIG3-MIE vectors 11 days after neomycin withdrawal and treated the cell cultures at regular time intervals with the drugs, either alone or in combination. An increase in the percentage of GFP-positive cells was assessed by FACS analysis. Both 5-azaC and TSA applied separately reactivated GFP expression, but the combination of both inhibitors provided the strongest effect (Fig. 4). The majority of silenced vectors were reactivated by 5-azaC and TSA 21 days after G418 withdrawal. However, the effect of these drugs gradually decreased and after 112 days of cultivation without selection, only 2 % of the vectors were reactivated by the treatment (Fig. 4).

DNA methylation analysis of the integrated vectors.

Because transcriptional silencing of genes or proviruses is usually caused by DNA methylation of promoters, we analysed and compared the CpG methylation patterns of the vector LTRs that were either silenced or transcriptionally active. For this analysis, we chose two NIL-2 clonal cell cultures transduced by the RNIG vector, one with a high and the other with very low percentage of GFP-positive cells and one clonal cell culture transduced by the RNIG3-IE vector with 30% of GFP-positive cells. These clones were sampled approximately nine weeks after G418 selection of transduced cells. Bisulfite sequencing of the vectors displayed almost fully methylated 5´LTR sequences in the cell clone with silent RNIG proviruses. In contrast, we found almost no CpG methylation in the cell clone without RNIG vector silencing. The cell clone with 30% GFP-

positive cells was intermedite as to the methylation of 5´LTR of the RNIG3-IE vector (Fig. 5a). There was no obvious methylation pattern and the methylated CpG dinucleotides were observed throughout the entire LTR. To show that the tandem of two IEs inserted in the LTR of RNIG2-2IE vector protects itself and the adjacent DNA sequence from CpG methylation, we analysed the methylation in two clones from the experiment shown in Fig. 3e. The bisulfite sequencing confirmed the presence of two IEs and the non-methylated status of whole LTR (Fig. 5b). The level of CpG methylation was extremely low here, 1.6 and 4.0%.

In order to see the early phase of vector silencing, we examined the CpG methylation in the rare GFP-negative cells that appeared soon after the selection of vector-transduced cells. We selected these cells from two NIL-2 cell cultures transduced with RNIG and RNIG3-MIE vectors by FACS after 21 days of cultivation without selective pressure. As it is difficult to separate the weakly expressing cells, the resulting cell cultures were enriched for GFP-negative cells but also contained, respectively, 15% and 28% of positive cells, respectively (Fig. 5c). The density of CpG methylation within LTR sequences was 28% and 19%, respectively, much lower than that of the nine-week-old cell culture transduced with RNIG3IE, which contained a comparable percentage of silenced cels (Fig. 5c). We therefore conclude that the silencing of integrated proviruses is associated with CpG methylation of promoter sequence, however there are other mechanisms of silencing that precede the onset of heavy methylation of proviral LTRs.

DISCUSSION

In this study, we report the capacity of the CpG island core element from the hamster *aprt* gene, the IE, to stabilize long-term expression of a retroviral vector and protect it from transcriptional silencing. Our data document that insertion of the 120 bp IE into a specific site of the LTR

ensures efficient transcription of ASLV-derived vectors and overcomes the repressive chromosomal position effects in non-permissive mammalian cells. We have found the best protective effect in LTRs with two IEs in tandem inserted between the promoter and enhancer sequences. None of the 19 analyzed cell clones transduced with this modification of the reporter vector exhibited substantial vector silencing after nine weeks of cultivation. We suggest our optimized vector design as a new strategy in improving retroviral vectors toward efficient gene transfer and therapy.

There are two basic strategies of how to counteract the silencing and variegation of retroviral vectors, based on both gammaretroviruses and lentiviruses (10). The first is the elimination of silencers defined *e.g*. in the LTR and primer binding site of MLV (5). Multiple-point mutation of all CpGs in the LTR region also stabilizes the expression of MLV vectors in embryonic stem cells (49). In a complementary approach, matrix attachment regions or insulators have been employed to prevent position effects of adjacent cellular silencers (6,24,43,52). Our previous experiments with the mouse *aprt* CpG island (20) adjacent to the RSV reporter provirus indicated that the anti-silencing and anti-methylation protection might have been caused also by some insulatory effect. We have found that insertion of the CpG island protects the downstream enhancer/promoter in an orientation-dependent manner without any increase of the promoter strength. The range of the anti-methylation effect was relatively narrow covering just the RSV LTR.

Here, we have employed the short core element of the hamster *aprt* CpG island as defined by Siegfried *et al*. (46), which can be inserted into the retroviral LTR. This IE comprises the protective effect of the whole CpG island, but differs from it in several characteristics. The observed stabilization of reporter expression was strongly dependent on the position of the IE within the LTR. The IE inserted at the beginning of the U3 region had only a minor effect, but

insertion between viral enhancers and promoter ensured efficient anti-silencing protection. We do not possess exact data on the promoter strength of IE-modified LTRs, but an estimate of the mean fluorescence intensity (MFI) of non-cloned cell cultures infected with individual vectors demonstrates that there are only mild differences in comparison with the wild-type RSV LTR (data not shown). Furthermore, these differences do not correlate with the anti-silencing effect. For example, the silencing-prone vector RNIG3-MIE exerted slightly higher MFI than the RNIG vector whereas the well protected vector RNIG2+IE displayed slightly decreased MFI. Substantial differences in MFI were found among individual clones irrespective the transduced vector (data not shown), which points to the strong influence of the integration site. In our previous study (32), we showed the increased expression driven by the RSV LTR enriched in Sp1 sites without the context of CpG island. Together, these data are in contradiction with the effect of insulator elements defined as sequences capable of obstructing outside enhancers to influence promoters. The range of anti-methylation effect is limited to approximately 150 bp from the IE (20,46). This might explain the weak or no protection effect of the IE inserted upstream of the enhancer and the best protection from the IE inserted between the enhancer and the promoter. Both transcriptional regulators are in this way within the range of IE anti-silencing influence. The position, 89 bp upstream to the transcription start site, best corresponds to the position of the element within the hamster *aprt* gene, 107 bp upstream from the transcription start site. This distance is probably favourable for keeping the promoter transcriptionally active. In contrast to the full-length mouse CpG island, the effect of the IE orientation is only weak.

Transcriptional silencing of retroviruses has been described in two experimental settings. First, complete transcriptional silencing occurs shortly after infection, most probably during the process of integration. Second, the selected reporter-positive cells variegate during the time course of cultivation. Even cell clones that inherited initially active and uniformly integrated

reporter proviruses variegate and eventually loose expression of the transduced reporter. In our case, we describe just the second type of provirus silencing. To assess the early peri-integration silencing, we performe the colony-forming assay that, in a preliminary experiment, displays a lower number of reporter-positive cells or selected cell colonies that did not correspond to the virus dose determined independently (data not shown). Furthermore, the rate of silencing observed in our experiments might be underestimated. We start our GFP monitoring with completely GFP-positive clonal cell cultures after two weaks of G418 selection. The cells, in which the silencing occurred most rapidly, could have been already selected against at that time. If we monitored the loss of GFP-expressing cells a few days after the infection, we would not obtain reliable data due to the background from non-infected cells as well as from cells expressing transiently the unintegrated forms of retrovirus.

Several studies have described the important role of Sp1 binding sites in the anti-methylation capacity of CpG islands (3,20,33). We have tested the effect of Sp1 mutations in three of our modified LTRs. In all three vectors, mutation of the Sp1 sites significantly increased the provirus silencing during prolonged cultivation, abrogating the protective effect of the inserted IE. In particular, the vector RNIG3-MIE was silenced even more efficiently than the vector with an unmodified LTR. This is probably due to the accumulation of CpG dinucleotides that, without the context of Sp1 sites, become a target of DNA methyltransferases (3). The most stable vector, RNIG2-IE, retains a part of this anti-silencing capacity even after mutation of Sp1 sites. It indicates that Sp1 binding sites are involved in the anti-silencing capacity of IE, but it is not the only *cis*-acting element involved. We do not know whether the binding of Sp1 factor to the IE is necessary for its anti-silencing activity. Nevertheless, the CpG islands in cells derived from Sp1 knock-out mouse embryos are maintained methylation-free (34) suggesting that Sp1 binding sites act in *cis*. Finally, it turns out that the methylation status of CpG islands establishes in early

embryonic cells and than maintains in differentiated cells. Based on aberrant methylation of CpG island in the modified *β-actin* allele, Strathdee et al (47) propose a two-step process to defining a CpG island and the role of Sp1 binding sites might be different in pluripotent and differentiated cells. It will be very interesting to check the anti-methylation and anti-silencing capacity of our IE-modified LTRs also *in vivo* in transgenes passing the germ line and early embryo.

The reactivation experiments showed additive effects of 5-azac and TSA and revealed that vector silencing was associated with DNA methylation and/or histone deacetylation. The inter-dependence of both mechanisms of provirus silencing was described in detail in experiments with lentiviral vectors (19). Katz et al. (29) also found silenced ASLV-based vectors to be reactivable by TSA alone. We observed that this capability strongly depends on the time elapsed from transduction, with a gradual loss of reactivation. This can be explained by the slow and multistep process of heterochromatinization, which requires further histone and chromatin modifiers, finally locking the provirus in a constitutively inactive form (16).

The efficient silencing of ASLV-based vectors has been described in rodent cells (22, 45, 48 for the review). To be sure that this phenomenon and our approach to anti-silencing protection of vectors can be applied to other system, we have performed these experiments in parallel with the human HEK 293 cell line. The results obtained are basically the same as those results seen in NIL-2 cells for all vectors with modified 5´LTR, but the silencing of the wild-type RNIG vector was slower and, correspondingly, the anti-silencing effects of IE insertions were less pronounced. Because the silencing can be associated with the distributive DNA methylation by Dnmt1 (50) during the S-phase of the cell cycle, it is probable that the slower silencing in HEK 293 cells correlates with slower proliferation and longer doubling time in comparison to NIL-2. Also, other cell-specific factors may play a role in the different rate of silencing (15,41).

Transcriptional silencing is a principal obstacle for the use of retroviral and lentiviral vectors in gene transfer and gene therapy. Our strategy of anti-methylation protection of ASLV-based vectors by CpG island core element indicates that retroviral vector completely resistant to silencing can be constructed. This modification may prove immediately useful whenever *e.g* RCAS vectors are used in mammalian cells and it could be part of a general optimized vector design if also proven to be effective in gammaretroviral and lentiviral vectors.

ACKNOWLEDGEMENTS

We thank Jan Svoboda and Jasper Manning (both at the Institute of Molecular Genetics, Prague) for their helpful comments, encouragement and carefull reading of the manuscript; Dana Kučerová, Věra Hoserová, and Helena Burešová for excellent technical assistance. This research was supported by the Grant Agency of the Czech Republic (Grants No. 204/05/0939 and 523/07/1171 to JH) and by the Academy of Sciences of the Czech Republic (grant No. AV0Z50529514).

REFERENCES

1. **Barsov, E.V., and S.H. Hughes.** 1996. Gene transfer into mammalian cells by a Rous sarcoma virus-based retroviral vector with the host range of the amphotropic murine leukemia virus. J. Virol. **70:**3922-3929.

2. **Barsov, E.V., W.S. Payne, and S.H. Hughes.** 2001. Adaptation of chimeric retroviruses in vitro and in vivo: isolation of avian retroviral vectors with extended host range. J. Virol. **75:**4973-4983.

3. **Brandeis M., D. Frank, I. Keshet, Z. Siegfried, M. Mendelsohn, A. Nemes, V. Temper, A. Razin, H. Cedar.** 1994. Sp1 elements protect a CpG island from de novo methylation. Nature **371**:435-8.

4. **Challita, P.M., and D.B. Kohn.** 1994. Lack of expression from a retroviral vector after transduction of murine hematopoietic stem cells is associated with methylation in vivo. Proc. Natl. Acad. Sci. USA **91:**2567-2571.

5. **Challita, P.M., D. Skelton, A. El-Khoueiry, X.J. Yu, K. Weinberg, and D.B. Kohn.** 1995. Multiple modifications in *cis* elements of the long terminal repeat of retroviral vectors lead to increased expression and decreased DNA methylation in embryonic carcinoma cells. J. Virol. **69:**748-755.

6. **Dang, Q., J. Auten, and I. Plavec.** 2000. Human beta interferon scaffold attachment region inhibits de novo methylation and confers long-term, copy number-dependent expression to a retroviral vector. J. Virol. **74:**2671-2678.

7. **Diamond, L.** 1967. Two spontaneously transformed cell lines derived from the same hamster embryo culture. Int. J. Cancer **2**:143-152.

8. **Elleder, D., V. Stepanets, D.C. Melder, F. Šenigl, J. Geryk, P. Pajer, J. Plachý, J. Hejnar, J. Svoboda, and M.J. Federspiel.** 2005. The receptor for the subgroup C avian sarcoma and leukosis viruses, Tvc, is related to mammalian butyrophilins, members of the immunoglobulin superfamily. J. Virol. **79:**10408-10419.

9. **Elleder, D., A. Pavlíček, J. Pačes, and J. Hejnar.** 2002. Preferential integration of human immunodeficiency virus type 1 into genes, cytogenetic R bands and GC-rich DNA regions: insight from the human genome sequence. FEBS Lett. **517:**285-286.

10. **Ellis, J.** 2005. Silencing and variegation of gammaretrovirus and lentivirus vectors. Human Gene Ther. **16:**1241-1246.

11. **Federspiel, M.J., and S.H. Hughes.** 1994. Effects of the gag region on genome stability: avian retroviral vectors that contain sequences from the Bryan strain of Rous sarcoma virus. Virology **203:**211-220.

12. **Federspiel, M.J., D.A. Swing, B. Eagleson, S.W. Reid, and S.H. Hughes.** 1996. Expression of transduced genes in mice generated by infecting blastocysts with avian leukosis virus-based retroviral vectors. Proc. Natl. Acad. Sci. USA **93:**4931-4936.

13. **Federspiel, M.J., P. Bates, J.A. Young, H.E. Varmus, and S.H. Hughes.** 1994. A system for tissue-specific gene targeting: transgenic mice susceptible to subgroup A avian leukosis virus-based retroviral vectors. Proc. Natl. Acad. Sci. USA **91:**11241-11245.

14. **Greger, J.G., R.A. Katz, K. Taganov, G.F. Rall, and A.M. Skalka.** 2004. Transduction of terminally differentiated neurons by avian sarcoma virus. J. Virol. **78:**4902-4906.

15. **Greger, J.G., R.A. Katz, A.M. Ishov, G.G. Maul, and A.M. Skalka.** 2005. The cellular protein Daxx interacts with avian sarcoma virus integrase and viral DNA to repress viral transcription. J. Virol. **79:**4610-4618.

16. **Groth, A., W. Rocha, A. Verreault, and G. Almouzni.** 2007. Chromatin challenges during DNA replication and repair. Cell **128:**21-733.

17. **Hacein-Bey-Abina, S., C. Von Kalle, M. Schmidt, M.P. McCormack, N. Wulffraat, P. Leboulch, A. Lim, C.S. Osborne, R. Pawliuk, E. Morillon, R. Sorensen, A. Forster, P. Fraser, J.I. Cohen, G. de Saint Basile, I. Alexander, U. Wintergerst, T. Frebourg, A. Aurias, D. Stoppa-Lyonnet, S. Romana, I. Radford-Weiss, F. Gross, F. Valensi, E. Delabesse, E. MacIntyre, F. Sigaux, J. Soulier, L.E. Leiva, M. Wissler, C. Prinz, T.H. Rabbitts, F. Le Deist, A. Fischer, and M. Cavazzana-Calvo.** 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science **302:**415-419. Erratum in: Science 302: 568, 2003.

18. **Hatziioannou, T. and S.P. Goff.** 2001. Infection of nondividing cells by Rous sarcoma virus. J Virol. **75**:9526-31.

19. **He J., Q. Yang, and L.J. Chang.** 2005. Dynamic DNA methylation and histone modifications contribute to lentiviral transgene silencing in murine embryonic carcinoma cells. J Virol. **79:**13497-508.

20. **Hejnar, J., P. Hájková, J. Plachý, D. Elleder, V. Stepanets, and J. Svoboda.** 2001. CpG island protects Rous sarcoma virus-derived vectors integrated into nonpermissive cells from DNA methylation and transcriptional suppression. Proc. Natl. Acad. Sci. USA **98:**565-569.

21. **Hejnar, J., J. Plachý, J. Geryk, O. Machoň, K. Trejbalová, R.V. Gunataka, and J. Svoboda.** 1999. Inhibition of the rous sarcoma virus long terminal repeat-driven transcription by in vitro methylation: different sensitivity in permissive chicken cells versus mammalian cells. Virology **255:**171-181.

22. **Hejnar, J., J. Svoboda, J. Geryk, V.J. Fincham, and R. Hák.** 1994. High rate of morphological reversion in tumor cell line H-19 associated with permanent transcriptional suppression of the LTR, v-src, LTR provirus. Cell Growth Differ. **5:**277-285.

23. **Himly, M., D.N. Foster, I. Bottoli, J.S. Iacovoni, and P.K. Vogt.** 1998. The DF-1 chicken fibroblast cell line: transformation induced by diverse oncogenes and cell death resulting from infection by avian leukosis viruses. Virology **248:**295-304.

24. **Hino, S, J. Fan, S. Taguwa, K. Akasaka, and M. Matsuoka.** 2004 Sea urchin insulator protects lentiviral vector from silencing by maintaining active chromatin structure. Gene Ther. **11**:819-28.

25. **Hoeben, R.C., A.A. Migchielsen, R.C. van der Jagt, H. van Ormondt, and A.J. van der Eb.** 1991. Inactivation of the Moloney murine leukemia virus long terminal repeat in murine fibroblast cell lines is associated with methylation and dependent on its chromosomal position. J. Virol. **65:**904-912.

26. **Hu, J., A. Ferris, A. Larochelle, A.E. Krouse, M.E. Metzger, R.E. Donahue, S.H. Hughes, and C.E. Dunbar.** 2007. Transduction of Rhesus Macaque Hematopoietic Stem and Progenitor Cells with Avian Sarcoma and Leukosis Viral Vectors. Hum. Gene Ther. **18:**691-700.

27. **Jordan, A., P. Defechereux, and E. Verdin.** 2001. The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. EMBO J. **20:**1726-1738.

28. **Katz, R.A., J.G. Greger, K. Darby, P. Boimel, G.F. Rall, and A.M. Skalka.** 2002. Transduction of interphase cells by avian sarcoma virus. J. Virol. **76:**5422-5434.

29. **Katz, RA, E. Jack-Scott, A. Narezkina, I. Palagin, P. Boimel, J. Kulkosky, E. Nicolas, J.G. Greger, and A.M. Skalka.** 2007. High-frequency epigenetic repression and silencing of retroviruses can be antagonized by histone deacetylase inhibitors and transcriptional activators, but uniform reactivation in cell clones is restricted by additional mechanisms. J Virol. **81:**2592-604.

30. **Kukolj, G., R.A. Katz, and A.M. Skalka.** 1998. Characterization of the nuclear localization signal in the avian sarcoma virus integrase. Gene. **26**:157-63

31. **Lorincz, M.C., D. Schuebeler, S.C. Goeke, M. Walters, M. Groudine, and D.I.K. Martin.** 2000. Dynamic analysis of proviral induction and De Novo methylation: implications for a histone deacetylase-independent, methylation density-dependent mechanism of transcriptional repression. Mol. Cell. Biol. **20:**842-850.

32. **Machoň, O., V. Strmen, J. Hejnar, J. Geryk, and J. Svoboda.** 1998. Sp1 binding sites inserted into the rous sarcoma virus long terminal repeat enhance LTR-driven gene expression. Gene **208:**73-82.

33. **Macleod, D, J. Charlton, J. Mullins, A.P. Bird.** 1994. Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. Genes Dev. **8:**2282-92.

34. **Marin, M., Karis, A., Visser, P., Grosveld, F., and S. Philipsen.** 1997. Transcription factor Sp1 is essential for early embryonic development but dispensable for cell growth and differentiation. Cell **89:**619-628.

35. **Mitchell, R.S., B.F. Beitzel, A.F. Schroeder, P. Shinn, H. Chen, C.C. Berry, J.R. Ecker, and F.D. Bushman.** 2004. Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol. **2:**E234.

36. **Mok, H.P., S. Javed, and A. Lever.** 2007. Stable gene expression occurs from a minority of integrated HIV-1-based vectors: transcriptional silencing is present in the majority. Gene Ther. **14:**741-751.

37. **Moscovici, C., M.G. Moscovici, and H. Jimenez.** 1977. Continous tissue culture cell lines derived from chemically induced tumors of japanese quail. Cell **11**:95-103.

38. **Narezkina, A., K.D. Taganov, S. Litwin, R. Stoyanova, J. Hayashi, C. Seeger, A.M. Skalka, and R.A. Katz.** 2004. Genome-wide analyses of avian sarcoma virus integration sites. J Virol. **78:**11656-11663.

39. **Pao, W., D.S. Klimstra, G.H. Fisher, and H.E. Varmus.** 2003. Use of avian retroviral vectors to introduce transcriptional regulators into mammalian cells for analyses of tumor maintenance. Proc. Natl. Acad. Sci. USA **100:**8764-8769.

40. **Pawliuk, R., K.A. Westerman, M.E. Fabry, E. Payen, R. Tighe, E.E. Bouhassira, S.A. Acharya, J. Ellis, I.M. London, C.J. Eaves, R.K. Humphries, Y. Beuzard, R.L. Nagel, and P. Leboulch.** 2001. Correction of sickle cell disease in transgenic mouse models by gene therapy. Science **294:**2368-2371.

41. **Poleshko, A., I. Pelagin, R. Zhang, P. Boimel, C. Castagna, P.D. Adams, A.M. Skalka, and R.A. Katz.** 2008. Identification of cellular proteins that maintain retroviral epigenetic silencing: evidence for an antiviral response. J. Virol. **82:**2313-2323.

42. **Reinišová, M., A. Pavlíček, P. Divina, J. Geryk, J. Plachý, and J. Hejnar.** 2008. Target site preferences of subgroup C Rous sarcoma virus integration into the chicken DNA. The Open Genomics Journal, *In press*.

43. **Rivella, S., J.A. Callegari, C. May, C.W. Tan, and M. Sadelain.** 2000. The cHS4 insulator increases the probability of retroviral expression at random chromosomal integration sites. J. Virol. **74:**4679-4687.

44. **Schroeder, A.R., P. Shinn, H. Chen, C.C. Berry, J.R. Ecker, and F. Bushman.** 2002. HIV-1 integration in the human genome favors active genes and local hotspots. Cell **110:**521-529.

45. **Searle, S., D.A. Gillespie, D.J. Chiswell, and J.A. Wyke.** 1984. Analysis of the variations in proviral cytosine methylation that accompany transformation and morphological reversion in a line of Rous sarcoma virus-infected Rat-1 cells. Nucleic Acids Res. **12:**5193-5210.

46. **Siegfried, Z., S. Eden, M. Mendelsohn, X. Feng, B.Z. Tsuberi, and H. Cedar.** 1999. DNA methylation represses transcription in vivo. Nat. Genet. **22:**203-206.

47. **Strathdee, D., Whitelaw, C.B.A., and A.J. Clark.** 2008. Distal transgene insertion affects CpG island maintenance during differentiation. J. Biol. Chem. **283:**11509-11515.

48. **Svoboda, J., J. Hejnar, J. Geryk, D. Elleder, and Z. Vernerová.** 2000. Retroviruses in foreign species and the problem of provirus silencing. Gene **261:**181-188.

49. **Swindle, C.S., H.G. Kim, and C.A. Klug.** 2004. Mutation of CpGs in the murine stem cell virus retroviral vector long terminal repeat represses silencing in embryonic stem cells. J. Biol. Chem. **279:**34-41.

50. **Vilkaitis, G., I. Suetake, S. Klimašauskas, and S. Tajima.** 2005. Processive methylation of hemimethylated CpG sites by mouse Dnmt1 methyltransferase. J. Biol. Chem. **280:**64-72.

51. **Wu, X., Y. Li, B. Crise, and S.M. Burgess.** 2003. Transcription start regions in the human genome are favored targets for MLV integration. Science **300:**1749-1751.

52. **Yannaki, E., J. Tubb, M. Aker, G. Stamatoyannopoulos, and D.W. Emery.** 2002. Topological constraints governing the use of the chicken HS4 chromatin insulator in oncoretrovirus vectors. Mol. Ther. **5:**589-598.

53. **Zhang, F., S.I. Thornhill, S.J. Howe, M. Ulaganathan, A. Schambach, J. Sinclair, C. Kinnon, H.B. Gaspar, M. Antoniou, and A.J. Thrasher.** 2007. Lentiviral vectors containing an enhancer-less ubiquitously-acting chromatin opening element (UCOE) provide highly reproducible and stable transgene expression in haematopoietic cells. Blood **110:**14481457.

FIGURE LEGENDS

Figure 1 Schematic representation of retroviral vectors used throughout this study. **(a)** The basic vectors RNIG and MNIG containing the *neo*^{r} and EGFP genes

driven from the 5´ LTR of Rous sarcoma virus (RSV) or murine leukemia virus (MLV), respectively. PBS, primer-binding site; ψ^+, encapsidation signal from MLV; IRES, internal ribosome entry site; PPT, polypurine tract. **(b)** The nucleotide sequence of the island element (IE) of the hamster adenosin phosphoribosyltransferase (*aprt*) CpG island. The Sp1 binding sites are in frames, the CpG dinucleotides are in bold cases. The substitution of three bases in each Sp1 binding site that abolished the Sp1 binding capacity is depicted. **(c)** Modifications of the RNIG vector LTRs. Single or duplicated IEs with intact or mutated Sp1 binding sites were inserted in the depicted unique restriction sites in sense and/or anti-sense orientations. The orientation of IE shown by arrow, the bicistronic coding regions and 3´LTR not shown.

Figure 2 FACS analysis of GFP expression from modified and unmodified vector LTRs. NIL-2, HEK 293, and QT6 cell lines transduced with RNIG and RNIG2-2IE vectors were selected by G418 and FACS-analyzed 57 days after selection withdrawal. One clone from each cell line with each of the two vectors are shown. In histograms, the relative GFP fluorescence is plotted against the cell count and the percentage of GFP-positive cells is indicated. The range of GFP positivity is shown as a horizontal bar. As a negative control, the autofluorescence of mock-infected NIL-2, HEK 293, and QT6 cells is shown.

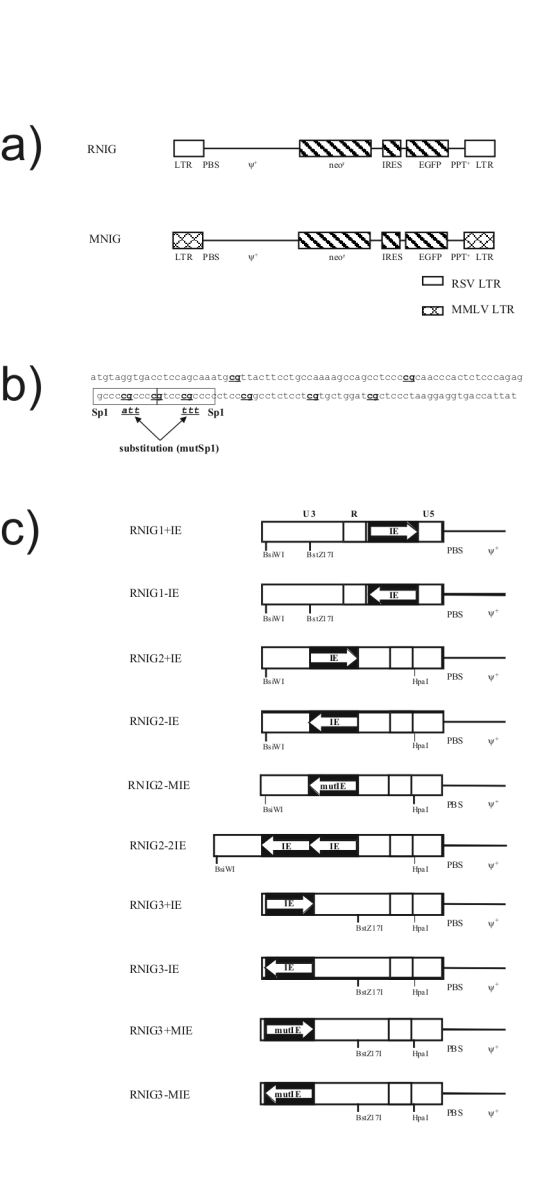
Figure 3 Stability of the GFP expression from retroviral vectors with modified and unmodified LTRs during long-term cultivation of transduced cells. NIL-2 cells were infected with retroviral vectors modified by IE inserted into the U5 region of LTR (**a**),

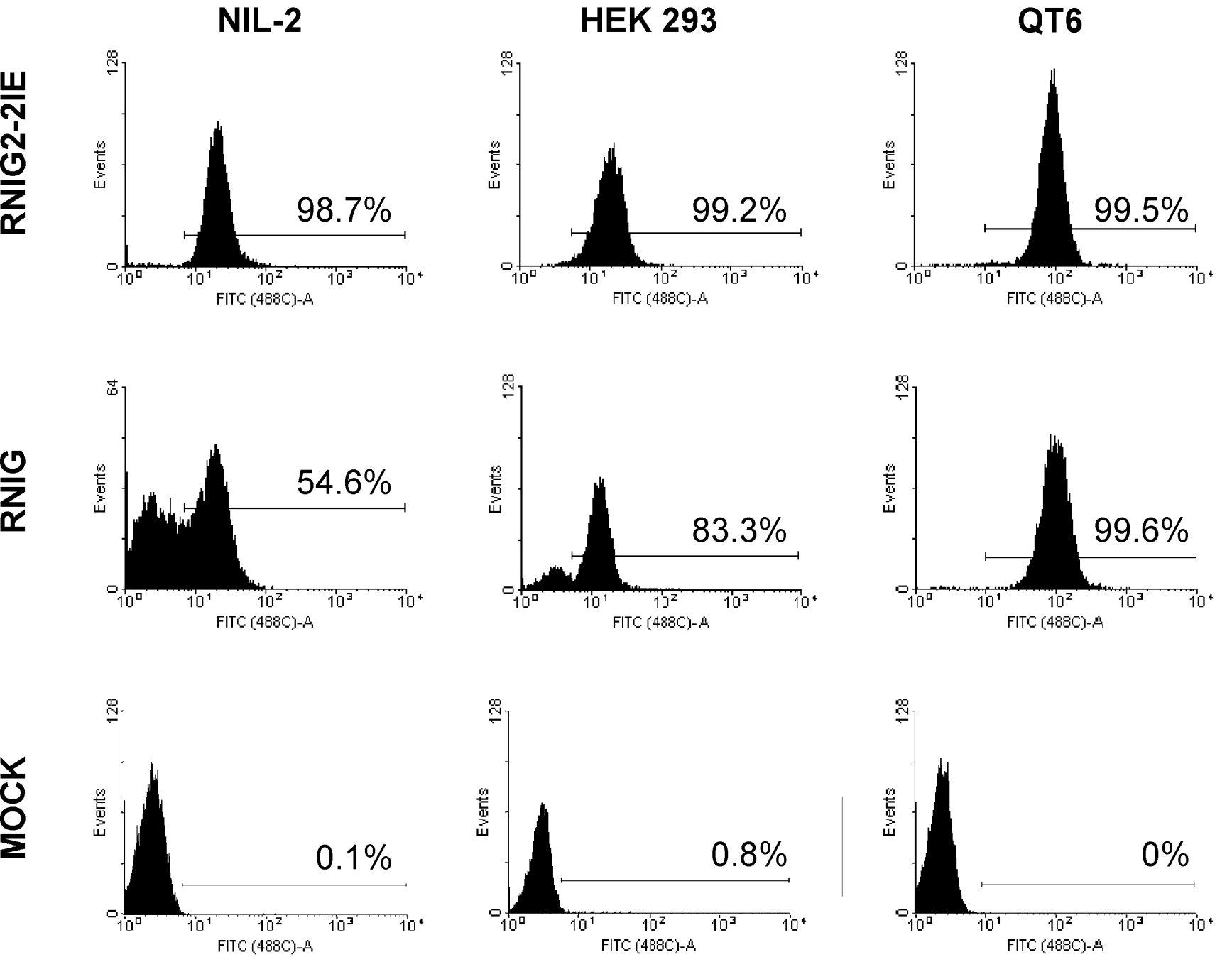
between the LTR promoter and enhancer (**b**), and at the beginning of the U3 region (**c**). Transduced cells were selected with G418 for 15 days, the percentage of GFP-positive cells was quantified by FACS in regular intervals after the selection was withdrawn, and the cells were cultivated without the selection. Each point represents the mean of three culture dishes. (**d,e**) NIL-2 cells were infected with the modified and unmodified vectors, individual clones of G418-resistant cells were isolated after 7 days of selection and cultivated separately without selection. The time course of vector silencing was followed by FACS at regular intervals and the percentage of GFP-positive cells is given as the mean of all cell clones transduced with the same vector construct. (**e**) Stability of the GFP expression after selection withdrawal of 19 individual cell clones transduced with the RNIG2-2IE vector. The diagrams of 17 non-silenced cell clones merge into one. (**d**) Stability of the GFP expression of nine individual cell clones transduced with the control RNIG vector. (**f**) Stability of expression of vectors modified by insertion of IE with mutated Sp1 binding sites. The population of NIL-2 cells was transduced with the vectors, selected withG418, and individual clones were isolated and cultivated separately. Each curve represents the mean percentage of GFP-positive cells of all cell clones transduced with the same vector construct.

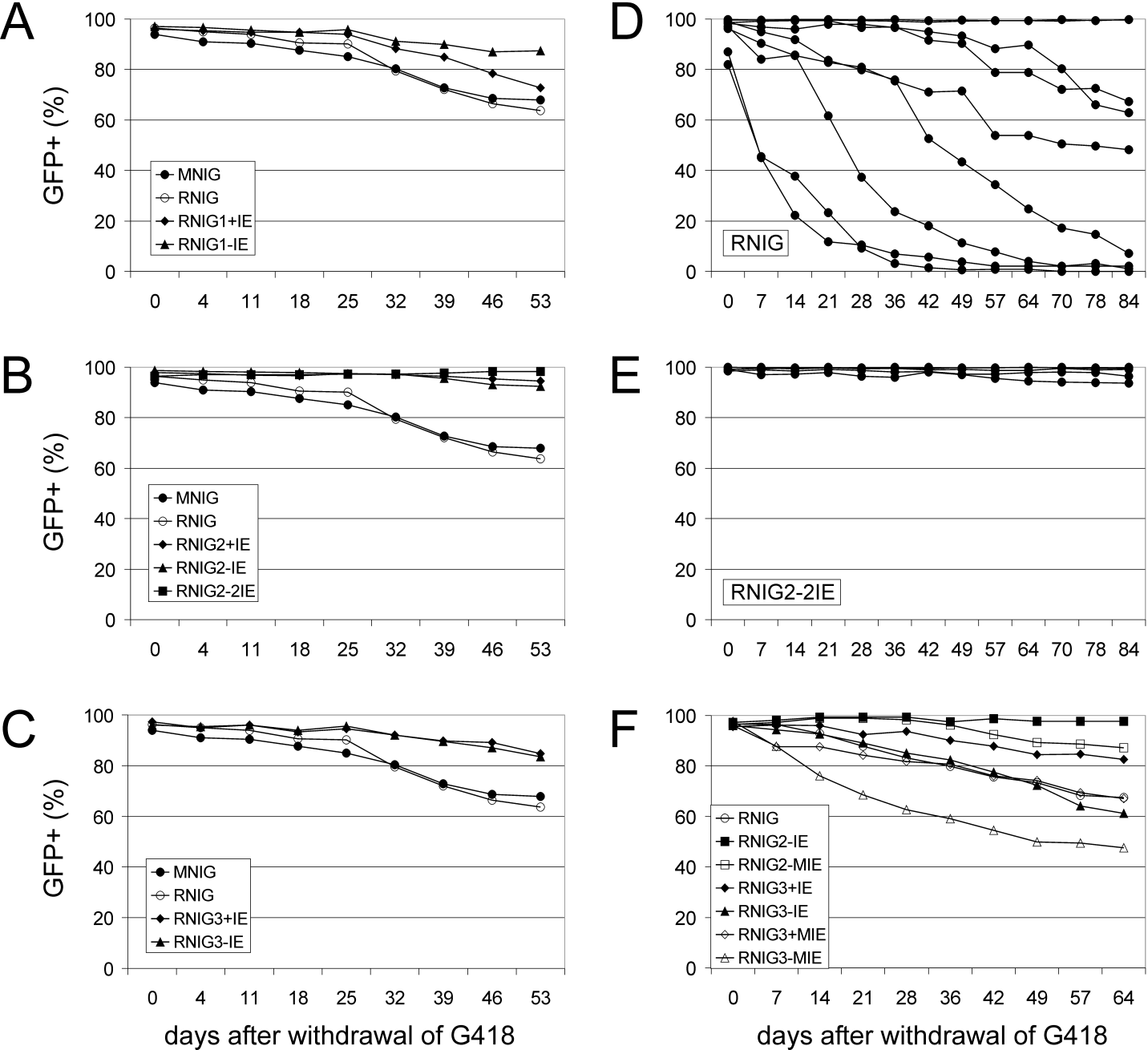
Figure 4 Reactivation of silenced vectors. NIL-2 cells were infected with silencing-prone vectors RNIG and RNIG3-MIE with mutated Sp1 binding sites. The cells were selected with G418 and after six days of selection, the selection was withdrawn. After 18-day-cultivation, the GFP-negative cells were sorted by FACS. Each GFP-negative population was then divided into two subpopulations six days after the sorting: one was treated with 5-azaC and/or TSA for four days and the proportion of GFP-expressing cells

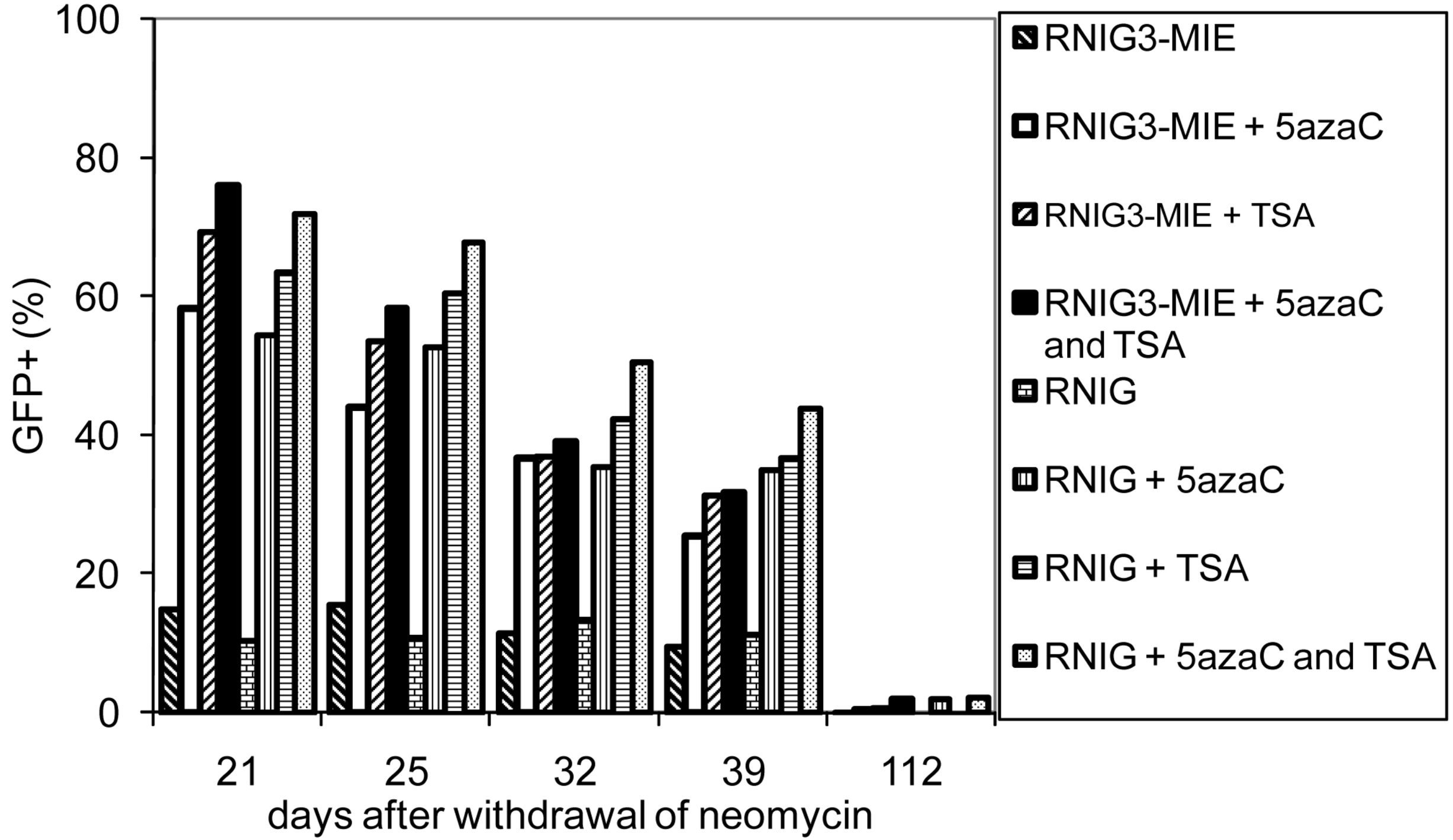
was assessed. The nontreated GFP-negative cell population was cultivated further and the treatment was repeated in regular intervals. Each experiment was done in triplicate.

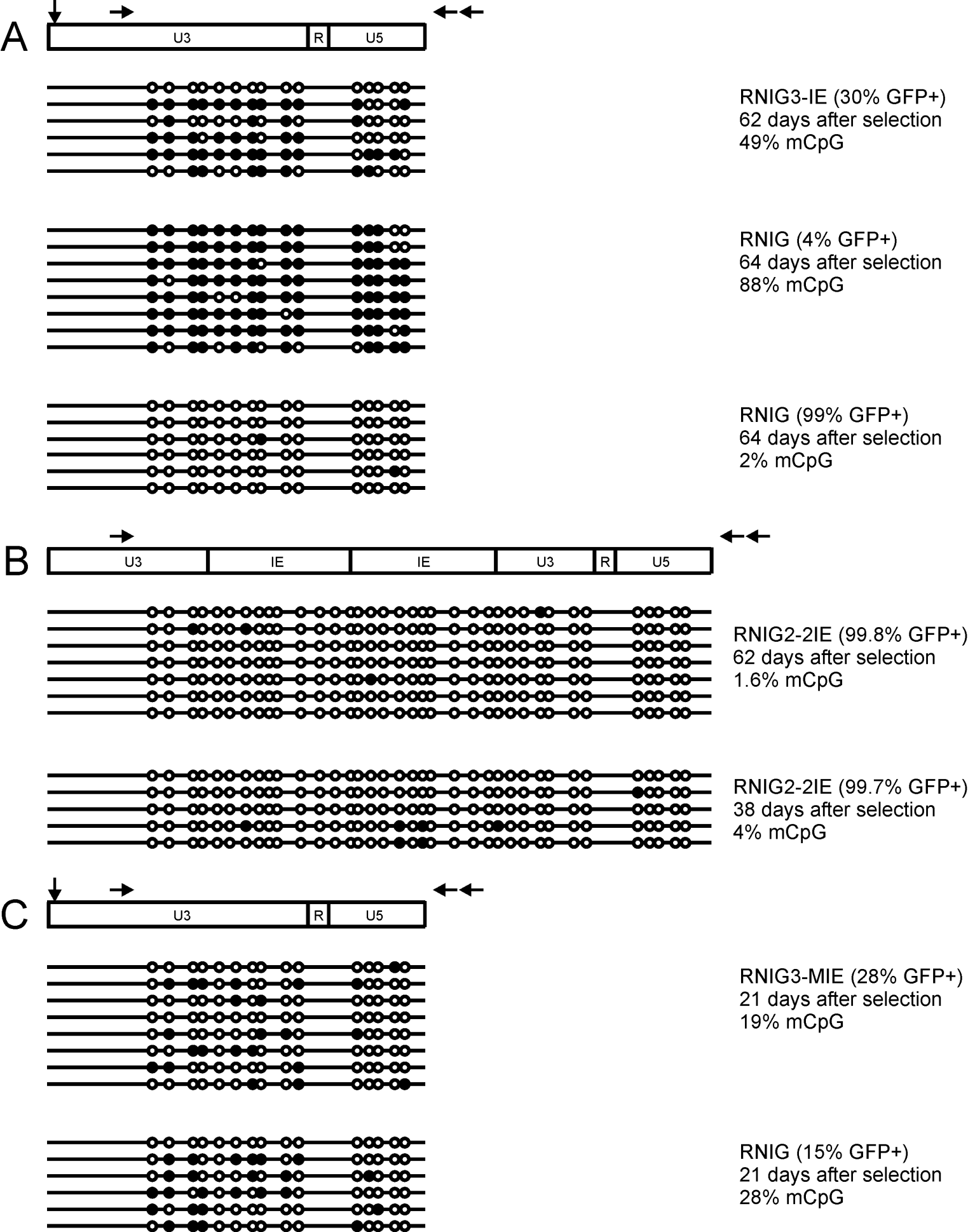
Figure 5 DNA methylation of integrated vectors. The methylation status of CpGs within the LTR of integrated vectors was assayed by the bisulfite technique. (**a**) CpG methylation status of the 5´LTR sequence of silencing-prone vectors RNIG and RNIG3-IE in three cell clones with different level of GFP silencing. NIL-2 cells transduced with vectors were sampled nine weeks after selection withdrawal and subjected to the bisulfite sequencing. (**b**) CpG methylation status of the 5´LTR sequence of the RNIG2-2IE vector protected from CpG methylation in two cell clones. NIL-2 cells transduced with RNIG2-2IE vector were sampled nine weeks after selection withdrawal and subjected to the bisulfite sequencing. (**c**) CpG methylation status of the 5´LTR sequence of silencing-prone vectors RNIG and RNIG3-MIE in two cell cultures enriched with GFP-negative cells early after vector infection. NIL-2 cells transduced with vectors and enriched with GFP-negative cells by FACS three weeks after selection withdrawal were sampled and subjected to the bisulfite sequencing. Each line with circles represents one independent LTR sequence of the PCR product obtained from the bisulfite-treated DNA. Percentages of GFP-positive cells and methylated CpGs are indicated. Methylated CpGs are depicted by solid circles, non-methylated CpGs are indicated by open circles. Localization of primers used for semi-nested PCR after the sodium bisulfite treatment and the site of insertion of the IE in the RNIG3-IE vector are depicted as the horizontal and vertical arrows, respectively.


a)
RNIG
LTR PBS ψ+ neo^r IRES EGFP PPT+ LTR
MNIG
LTR PBS ψ+ neo^r IRES EGFP PPT+ LTR
RSV LTR
MMLV LTR
b)
atgtaggtgacctccagcaaatggttacttcctgccaaaagccagcctcccgcaacccactctcccagag
gcccgcgccccgtccgcccctcccggcctctcctcgtgctggatcgctccctaaggaggtgaccattat
Sp1 att ttt Sp1
substitution (mutSp1)
c)
U3 R U5
RNIG1+IE
IE
BsiWI BstZI 71 PBS ψ+
RNIG1-IE
IE
BsiWI BstZI 71 PBS ψ+
RNIG2+IE
IE
BsiWI HpaI PBS ψ+
RNIG2-IE
IE
BsiWI HpaI PBS ψ+
RNIG2-MIE
mutIE
BsiWI HpaI PBS ψ+
RNIG2-2IE
IE IE
BsiWI HpaI PBS ψ+
RNIG3+IE
IE
BstZI 71 HpaI PBS ψ+
RNIG3-IE
IE
BstZI 71 HpaI PBS ψ+
RNIG3+MIE
mutIE
BstZI 71 HpaI PBS ψ+
RNIG3-MIE
mutIE
BstZI 71 HpaI PBS ψ+


NIL-2
HEK 293
QT6
RNIG2-2IE
RNIG
MOCK
Events
FITC (488C)-A
98.7%
99.2%
99.5%
54.6%
83.3%
99.6%
0.1%
0.8%
0%


A
GFP+ (%)
MNIG
RNIG
RNIG1+IE
RNIG1-IE
0 4 11 18 25 32 39 46 53
B
GFP+ (%)
MNIG
RNIG
RNIG2+IE
RNIG2-IE
RNIG2-2IE
0 4 11 18 25 32 39 46 53
C
GFP+ (%)
MNIG
RNIG
RNIG3+IE
RNIG3-IE
0 4 11 18 25 32 39 46 53
days after withdrawal of G418
D
RNIG
0 7 14 21 28 36 42 49 57 64 70 78 84
E
RNIG2-2IE
0 7 14 21 28 36 42 49 57 64 70 78 84
F
RNIG
RNIG2-IE
RNIG2-MIE
RNIG3+IE
RNIG3-IE
RNIG3+MIE
RNIG3-MIE
0 7 14 21 28 36 42 49 57 64
days after withdrawal of G418


100
80
60
40
20
0
GFP+ (%)
21
25
32
39
112
days after withdrawal of neomycin
RNIG3-MIE
RNIG3-MIE + 5azaC
RNIG3-MIE + TSA
RNIG3-MIE + 5azaC and TSA
RNIG
RNIG + 5azaC
RNIG + TSA
RNIG + 5azaC and TSA


A
U3
R
U5
RNIG3-IE (30% GFP+)
62 days after selection
49% mCpG
RNIG (4% GFP+)
64 days after selection
88% mCpG
RNIG (99% GFP+)
64 days after selection
2% mCpG
B
U3
IE
IE
U3
R
U5
RNIG2-2IE (99.8% GFP+)
62 days after selection
1.6% mCpG
RNIG2-2IE (99.7% GFP+)
38 days after selection
4% mCpG
C
U3
R
U5
RNIG3-MIE (28% GFP+)
21 days after selection
19% mCpG
RNIG (15% GFP+)
21 days after selection
28% mCpG

Available online at www.sciencedirect.com
ScienceDirect
Archives of Biochemistry and Biophysics 455 (2006) 136–143
ABB
www.elsevier.com/locate/yabbi

Regulation of c-Src activity by the expression of wild-type v-Src and its kinase-dead double Y416F-K295N mutant

Martina Vojtěchová *, Filip Šenigl, Eva Šloncová, Zdena Tuháčková

Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Flemingovo n. 2, 166 37 Praha 6, Czech Republic

Received 30 June 2006, and in revised form 13 September 2006
Available online 4 October 2006

Abstract

Active, wild-type v-Src and its kinase-dead double Y416F–K295N mutant were expressed in hamster fibroblasts. Expression of the active v-Src induced activation of endogenous c-Src and increased general protein-tyrosine phosphorylation in the infected cells. Expression of the kinase-dead mutant induced hypophosphorylation of Tyr416 of the endogenous c-Src. The inactivation of c-Src was reversible, as confirmed by *in vitro* kinase activity of c-Src immunoprecipitated from the kinase-dead v-Src-expressing cells. Both activation and inactivation of c-Src may be explained by direct interaction of the v-Src and c-Src that may either facilitate transphosphorylation of the regulatory Tyr416 in the activation loop, or prevent it by formation of transient dead-end complexes of the Y416F–K295N mutant with c-Src. The interaction was also indicated by co-localization of v- and c-Src proteins in immunofluorescent images of the infected cells. These results suggest that dimerization of Src plays an important role in the regulation of Src tyrosine kinase activity.

Keywords: c-Src; v-Src; Kinase-dead double Y416F–K295N v-Src mutant; Regulation of Src kinase activity; Protein-tyrosine phosphorylation

Tyrosine phosphorylation of protein kinases that in turn phosphorylate other substrate proteins, thus transferring the impulse through the cell, plays an important role in the signal transduction. Two main groups of protein tyrosine kinases are known, according to their function and localization in the cell: the transmembrane kinases that act as receptors for various ligands and the non-receptor tyrosine kinases located in the cytosol and acting as intermediate intracellular signaling proteins [1].

The non-receptor tyrosine kinases of the Src family contain an amino-terminal acylation site, a unique domain, Src homology 3 and 2 domains that are necessary for protein–protein interactions, a catalytic domain, and the carboxy-terminal (C-terminal) tail with negative regulatory tyrosine. The catalytic domain consists of two parts (N- and C-lobes), with the active site of the enzyme located in the cleft between the lobes. The ATP-binding site is located in the N-kinase lobe and part of the C-kinase lobe forms an activation loop with another regulatory tyrosine (Tyr416 in the c-Src). The primary structures of the v-Src oncoprotein [2] and of c-Src proteins from man, mouse [3] and several other sources have been determined and only minor differences have been found (96% homology between v-Src and c-Src proteins). The largest difference between v-Src and c-Src was found in the C-terminal tail. The C-terminal tail of c-Src contains the regulatory Tyr527 that is phosphorylated by C-terminal Src kinase, Csk, a member of another family of non-receptor tyrosine kinases [4]. In the molecule of v-Src, this acid C-terminal tail is replaced by a shorter sequence without regulatory Tyr527 and therefore v-Src is not downregulated by Csk-catalyzed phosphorylation. Apart from the C-terminus, the largest difference in the amino-acid sequence was found within the SH2 domain. The catalytic activity of Src family tyrosine kinases is regulated by several types of intramolecular interactions stabilizing the inactive, closed form of the Src protein. The SH3 domain binds directly to the N-kinase lobe of the catalytic domain and to the linker region between the SH2 domain and the catalytic domain, and the SH2 domain binds to the phosphorylated regulatory Tyr527.

* Corresponding author. Fax: +4202 33320702.
E-mail address: vojtech@img.cas.cz (M. Vojtěchová).

0003-9861/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.abb.2006.09.011

Phosphorylation of Tyr416 in the activation loop of the catalytic domain causes disarrangement of the closed inactive form and stimulates complete activation of the Src kinase. This phosphorylation is also able to prevent the inactivation of Src by Csk [5]. Phosphorylation of the Tyr416 is known to be an autophosphorylation event; however, there are some doubts whether this autophosphorylation is an intramolecular reaction or whether it is in fact a transphosphorylation between two Src molecules. Several authors [6,7] reported that the Tyr416 autophosphorylation is an entirely intramolecular reaction. Osusky et al. [8] later proved that the reaction follows, at least in part, a pseudomonomolecular mechanism. On the other hand, there is evidence for the transphosphorylation character of Tyr416 autophosphorylation in the kinetic experiments of Barker et al. [9] and in the work of Cooper and McAuley [10], who observed trans-autophosphorylation of a catalytically inactive Src mutant. However, no Src dimers have been detected so far by the chromatographic methods usually employed in this type of studies.

We have previously found that the transformation of hamster fibroblasts by Rous sarcoma virus resulted in high expression and activity of the v-Src protein, which was accompanied by an increased activity of endogenous c-Src and by deregulation of several signal transduction pathways [11–13]. Somewhat paradoxically, transformation of hamster fibroblasts by the v-*src* oncogene also induced an increased activity of Csk, which is expected to downregulate endogenous c-Src. However, the effect of Csk seemed to be offset by the presence of v-Src, and thus a possibility of v-Src generating the active double phosphorylated form of c-Src had to be considered. Therefore, we examined the effects of the presence of active and inactive v-Src protein on the activity and phosphorylation status of the endogenous c-Src kinase in hamster fibroblasts.

Materials and methods

All general reagents were from Sigma (USA), unless otherwise stated. Protein G–Sepharose 4B and [γ-^{32}P]-ATP were obtained from GE Healthcare Bio-Sciences. Molecular weight markers, chemiluminescence detection reagent Lumi-Glo, anti-phospho-Src (Tyr416), anti-phospho-Src (Tyr527) and anti-nonphospho-Src (Tyr527) antibodies and HP-conjugated anti-rabbit and anti-mouse antibodies were obtained from Cell Signaling Technology, Danvers, MA, USA. Monoclonal antibody LA 074 against the N-terminus of p60 c-Src (IgG2a) was from LP-016 mouse hybridoma cells [14], and the anti-avian Src antibody clone EC 10 (IgG2b) was from Upstate Biotechnologies, Lake Placid, NY, USA. The anti-α-tubulin antibody was provided by Dr. P. Dráber from our Institute [11]. The Alexa 594-conjugated anti-mouse antibodies and Alexa 488-conjugated anti-rabbit antibody were purchased from Molecular Probes (Eugene, OR, USA). DAPI dye (4′-6-diamidino-2-phenylindole)[1] for nuclei staining was from Roche.

v-src mutagenesis

The RSV long terminal repeat (LTR)-driven v-*src* in the pH19KE construct containing the LTR, v-*src*, LTR proviral structure [15] and the same construct with Y416F mutation in the v-*src* gene [16] were used. The sequence AAGACTCTG of the v-*src* Y416F gene coding for K295, T296 and L297 was replaced by AATACTCTT coding for N295, T296 and L297 using the Transformer Site-Directed Mutagenesis Kit (Clontech Laboratories, Inc.). All mutagenesis and selection procedures were done according to the manufacturer's instructions with the switch *Hind*III/*Sca*I selection primer and mutation primer 5′ TGCCGGGCTTAAGAGTATTTATG GCCA 3′. The mutated sequence creates a new diagnostic *Afl*II cleavage site (underlined). We obtained the construct pH19FN with mutations Y416F and K295N in the v-*src* gene.

Retroviral vector construction

The pH19KE and pH19FN were used as a source of the v-*src* gene. The constructs were cleaved with *Sac*I and *Cla*I. The 2-kb *Sac*I-*Cla*I fragment containing the v-*src* gene was treated with mung-bean nuclease to remove single-strand overhangs. The pLXRN construct (Clontech Laboratories, Inc.) was cleaved with *Hpa*I and treated with calf intestinal phosphatase. Plasmid pLSRN was created by blunt-end ligation of the 2-kb *Sac*I-*Cla*I fragment of the pH19KE into the pLXRN. Analogically, the plasmid pLmSRN was created by ligation of the 2-kb fragment of the pH19FN.

Virus production and cell infection

The packaging cell line GP293 (Clontech) was grown in Dulbecco's modified Eagle's medium-nutrient mixture F-12 Ham (Sigma) supplemented with 5% calf serum and antibiotics (100 U/ml penicillin and 100 μg/ml streptomycin) in a 5% CO_2 atmosphere at 37 °C. Subconfluent cultures of GP293 cells were cotransfected with plasmid pVSV-G (Clontech Laboratories, Inc.) and pLSRN or pLmSRN by the calcium phosphate precipitation method. The medium was collected 24 and 48 h after transfection and filtered with a 0.45 μm pore syringe filter. The collected medium (virus stock) was used for the infection of hamster fibroblast line NIL-2 [17]. Subconfluent cultures of NIL2 cells were infected with the virus stock with 10 μg/ml polybrene. After 24 h of incubation in growth medium, successfully infected cells were selected by cultivation in the presence of Geneticin (800 mg/ml of medium) for 7 days, then frozen and kept at −80 °C for further experiments.

Cell culture and preparation of cell extracts

Cells were grown in Petri dishes (6 cm) in Dulbecco's modified Eagle's medium with Ham's nutrient mixture F-12 (1:1) containing 5% (v/v) fetal calf serum (FCS) for 2 or 3 days to form monolayers. Cells were then washed twice with cold phosphate buffered saline (PBS pH 7.4) and extracted in 0.5 ml ice-cold extraction buffer (50 mM HEPES, pH 7.4, 1% Triton X-100, 50 mM NaF, 40 mM β-glycerophosphate, 5 mM $Na_2H_2P_2O_7$, 5 mM EGTA, 1 mM Na_3VO_4, 1 mM benzamidine, 0.5 mM PMSF, and 0.4 μM of each of the proteinase inhibitors pepstatin, antipain and leupeptin). Cell debris was pelleted by centrifugation at 20,000*g* for 10 min at 4 °C and supernatants were measured for protein concentration.

Western blotting and immunodetection

Cell extracts (40 μg of protein) were subjected to SDS–PAGE followed by electrotransfer of proteins to PVDF membranes (Immobilon P, Millipore) and immunodetected with specific antibodies, as described previously [12,13].

Immunoprecipitation of proteins

For immunoprecipitation of Src proteins, 200 μl of cell extract (400 μg of protein) was incubated overnight at 4 °C with 1 μl LA 074 monoclonal antibody or 1 μl anti-avian Src EC-10 and 10 μl of protein G–Sepharose 4B, as described previously [12,13]. The resulting immunoprecipitates were used either for the SDS–PAGE/immunoblot analysis or for the *in vitro* kinase assays.

[1] *Abbreviations used:* DAPI, 4′-6-diamidino-2-phenylindole; LTR, long terminal repeat; FCS, fetal calf serum; PBS, phosphate buffered saline.

In vitro kinase activity

The Src tyrosine kinase reaction was performed in 20 μl of 50 mM Tris–HCl, pH 7.5, 5 mM $MgCl_2$, 1 mM $MnCl_2$, 5 mM DTT, 10 μM Na_3VO_4, 50 μM ATP, 1 μCi [γ-^{32}P]ATP and 100 μM Src optimal peptide substrate AEEEIYGEFEAKKKK, as described previously [18]. The assays were run in triplicates and the data obtained in the absence of the peptide substrate were subtracted from the results.

Immunocytochemistry

Cells cultured on slides in a 24-well plate were washed with PBS and fixed for 20 min with 4% paraformaldehyde in PBS and subsequently permeabilized with methanol:acetone (1:1 v/v) for 1 min. After several washes with PBS, the cells were blocked with 4% BSA in 0.2% Triton X-100 for 2 h at room temperature and incubated overnight at 4 °C with appropriate primary antibodies (1:200 in the blocking solution). The cells were stained in the dark with secondary anti-mouse antibodies conjugated with Alexa 594 or anti-rabbit antibody conjugated with Alexa 488 (1:1000). The nuclei were visualized by DAPI staining. Fluorescent images were captured with Nikon Eclipse E600W microscope using Lucia software (Laboratory Imaging, Ltd.).

Results

The kinase-dead double Y416F–K295N mutant was prepared by introducing the K295N point mutation into the Y416F H19 v-*src* mutant [16]. The kinase-dead double Y416F–K295N v-*src* mutant and the active, wild-type H19 v-*src* genes were introduced into embryonic hamster fibroblasts NIL-2 using retroviral vectors. Two transgenic cell lines expressing v-Src protein were created; NIL-A expressing the active v-Src protein and NIL-B expressing kinase-dead v-Src protein. NIL-C cells infected with an empty retroviral vector and the non-infected control NIL-D cells do not express v-Src protein.

Src protein expression and phosphorylation

The presence of v-Src protein in the infected cell lines was examined by immunocytochemistry using the specific anti-avian Src antibody followed by fluorescent microscopy (Fig. 1B, F, J and N). The c-Src protein was visualized with nonphospho-Src (Tyr527) antibody that recognizes the amino-acid sequence surrounding Tyr527 of c-Src only when the Tyr527 is dephosphorylated (Fig. 1A, E, I and M). The merged image (Fig. 1C and G) revealed co-localization of v-Src and c-Src proteins in both the NIL-A and NIL-B cells. No presence of v-Src (and consequently, no co-localization of v-Src with c-Src) was detected in the NIL-C cells infected with the pLXRN plasmid alone or in the non-infected NIL-D cells, as shown in Fig. 1K and O.

In NIL-A cells expressing the active wild-type v-Src, cytoplasmic localization of the v-Src was observed (Fig. 1B). The cytoplasmic localization was also displayed by the nonphospho-Src (Tyr527) in the NIL-A cells (Fig. 1A), as well as in the NIL-C cells infected with an empty expression vector (Fig. 1I) and non-infected NIL-D cells (Fig. 1M). On the contrary, NIL-B cells expressing the kinase-dead double mutant v-Src exhibited prevalent nuclear localization of both the v-Src and the c-Src dephosphorylated at Tyr 527 (Figs. 1E and F, respectively).

Immunocytochemistry of the infected cells performed with the phospho-Src (Tyr416) antibody recognizing the active form of both Src proteins revealed their predominant cytoplasmic localization in all the infected and control cells (Fig. 1R, S, T and U). These results suggest that the Src proteins localized into the nucleus are inactive and that the c-Src co-localizing there with the kinase-dead v-Src is a double-dephosphorylated (Tyr416 and Tyr5527) form.

Immunoblotting of the cell extracts confirmed that in both v-*src* infected cell lines NIL-A and NIL-B, the expression of Src protein is elevated compared to NIL-C cells carrying only the infection vector or to the non-infected control cells NIL-D (Fig. 2D). As expected, v-Src protein was found only in NIL-A and NIL-B cells, not in NIL-C or NIL-D cells (Fig. 2A). In NIL-A cells infected with the active v-*src*, highly increased phosphorylation of the Src regulatory Tyr416 was observed (Fig. 2B), implying increased Src activity *in vivo*. This was confirmed by elevated protein-tyrosine phosphorylation similar to the tyrosine phosphorylation of proteins extracted from the RSV-transformed cells H19, as it is demonstrated by immunoblot analysis of the cell extracts using a general anti-phospho-tyrosine antibody (Fig. 3).

In NIL-B cells, the expression of the kinase-dead double Y416F–K295N mutant induced a significant decrease in the phosphorylation of regulatory Tyr416 of Src protein (about 50%) when compared to NIL-C cells, infected with the vector alone, or to the non-infected NIL-D cells (Fig. 2B). However, the decrease in the phosphorylation of Tyr416 was not accompanied by an increase in Tyr527 phosphorylation in the kinase-dead v-Src-expressing NIL-B cells. On the other hand, in NIL-A cells expressing active v-Src, an increase in the phosphorylation of both Tyr416 and Tyr527 was found (Fig. 2B and C). The Csk expression followed the same pattern as Tyr527 phosphorylation of Src (Supplementary data). These observations may indicate the Csk preference for c-Src phosphorylated on Tyr416, as it has been reported by Wang et al. [19].

In vivo activity of v-Src and c-Src

To analyze v-Src and c-Src separately, the individual Src proteins were immunoprecipitated from the infected cells. Total Src including both v- and c-Src proteins was immunoprecipitated from the cell extracts with LA-074 antibody that recognizes both the viral Src protein and its cellular counterpart. The v-Src was precipitated separately using specific anti-avian Src antibody, and the c-Src was then immunoprecipitated from the residual supernatants with anti-Src antibody LA-074.

Regulatory Tyr416 of the v-Src protein immunoprecipitated from NIL-A cells was highly phosphorylated, as seen in Fig. 4A. The increase in Tyr416 phosphorylation of the endogenous c-Src immunoprecipitated from the NIL-A cells was also observed, when compared to the c-Src immu-

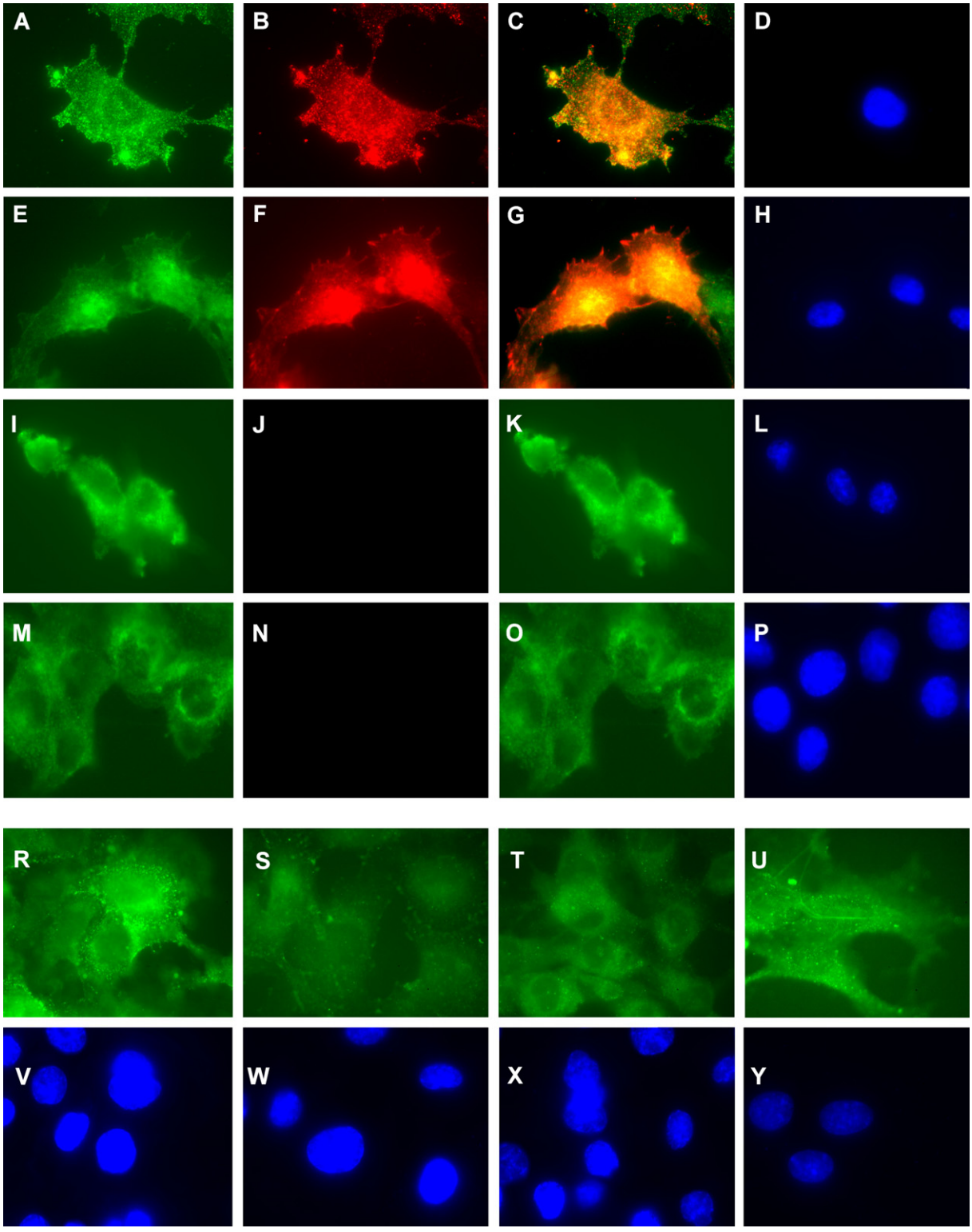


Fig. 1. Immunofluorescent analysis of the Src proteins expressed in cell lines derived from embryonic hamster fibroblasts NIL-2 infected with wild-type- and kinase-dead double mutant v-*src*. Upper panel: localization of the c-Src and v-Src proteins in the cell lines derived from v-*src* infected cells NIL-2 was evaluated by fluorescent microscopy following immunocytochemistry using nonphospho-Src (Tyr527) (A, E, I, M) and anti-avian Src (B, F, J, N) antibodies. C, G, K, O are the merged images and D, H, L, P are the DAPI-stained nuclei. (A–D) NIL-A cells infected with pH19KE (wild-type v-Src). (E–H) NIL-B cells infected with the kinase-inactive mutant pH19FN (Y416F–K295N v-Src). (I–L) NIL-C cells infected with the pLXRN plasmid alone. (M–P) NIL-D, non-infected cells. Lower panel: the Tyr416 phosphorylation status of src proteins was analyzed with phospho-Src (Tyr416) antibody (R–U). V–Y are the DAPI stained nuclei. (R, V) NIL-A cells infected with pH19KE (wild-type v-Src). (S, W) NIL-B cells infected with the kinase-inactive mutant pH19FN (Y416F–K295N v-Src). (T, X) NIL-C cells infected with the pLXRN plasmid alone. (U, Y) NIL-D, non-infected cells. Fluorescent images were captured with Nikon Eclipse E600W microscope. Magnification, 100×.

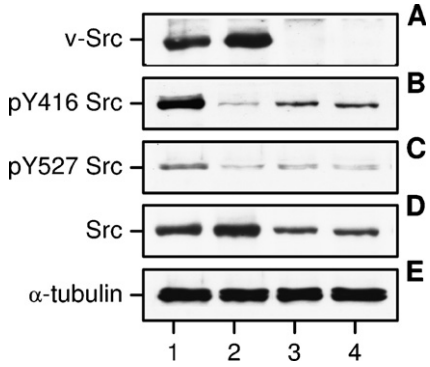


Fig. 2. Expression and phosphorylation of Src protein in the cell lines derived from embryonic hamster fibroblasts NIL-2 infected with wild-type- and kinase-dead double mutant v-*src*. Crude cell extracts (40 μg) of the cells were analyzed by SDS–PAGE followed by immunoblotting using appropriate antibody as indicated for individual experiments. Data are typical of those obtained from at least three separate experiments. (A) Expression of the v-Src protein in the infected cells was analyzed using the monoclonal anti-avian Src antibody (clone EC10). (B) Phosphorylation of tyrosine 416 of Src in the infected cells was analyzed using the polyclonal anti-phospho-Src (Tyr416) antibody. (C) Phosphorylation of tyrosine 527 in the infected cells was analyzed using the polyclonal anti-phospho-Src (Tyr527) antibody. (D) Src protein expression in the infected cells was analyzed using the monoclonal anti-Src LA-074 antibody. (E) Loading control was obtained by using the anti-α-tubulin antibody. Lane 1, NIL-A cells infected with pH19KE (wild-type v-Src). Lane 2, NIL-B cells infected with the kinase-inactive mutant pH19FN (Y416F–K295N v-Src). Lane 3, NIL-C cells infected with the pLXRN plasmid alone. Lane 4, NIL-D, non-infected cells.

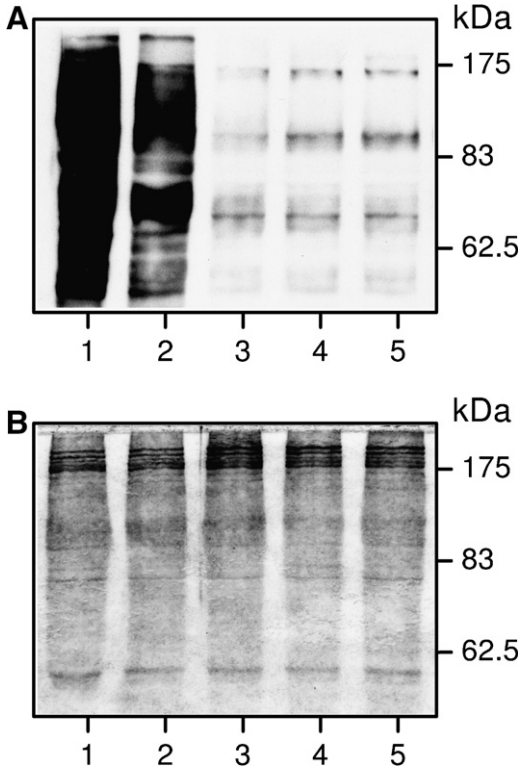


Fig. 3. Protein-tyrosine phosphorylation in the presence of active and kinase-dead v-Src protein in the infected cells and in the RSV-transformed cells H19. Crude cell extracts (40 μg) of the cells were analyzed by SDS–PAGE and immunoblotting. (A) Anti-phospho-tyrosine (clone PY99). (B) Coomassie brilliant blue staining of the membrane (Loading control). Lane 1, H19 cells, RSV-transformed hamster fibroblasts. Lane 2, NIL-A cells infected with pH19KE (wild-type v-Src). Lane 3, NIL-B cells infected with the kinase-inactive mutant pH19FN (Y416F–K295N v-Src). Lane 4, NIL-C cells infected with the pLXRN plasmid alone. Lane 5, NIL-D, non-infected cells.

noprecipitated from NIL-C cells (infected with the vector alone) or to the non-infected cells NIL-D (Fig. 4B).

As expected, no phosphorylation of regulatory tyrosine 416 of v-Src immunoprecipitated from the NIL-B cells which express kinase-dead v-Src was found (Fig. 4A). Also, very low phosphorylation of the regulatory Tyr416 was found in c-Src immunoprecipitated from NIL-B cells (Fig. 4B), in accordance with results of the analysis of Tyr416 phosphorylation in whole cell extracts (Fig. 2B), indicating inactivation of endogenous c-Src in the presence of kinase-dead v-Src protein.

In vitro interaction of v-Src and c-Src proteins

Co-immunoprecipitation of v-Src and c-Src proteins *in vitro* could support the possibility that v-Src and c-Src proteins interact and, possibly, form dimers *in vivo*. To that aim, c-Src was immunoprecipitated from crude cell extracts of the infected cells using the anti-nonphospho-Src (Tyr527) antibody and the presence of co-precipitated v-Src protein was analyzed with the anti-avian Src antibody. v-Src protein was found in the nonphospho-Src (Tyr527) immunoprecipitates from both NIL-A cells infected with the active, wild-type v-Src and in the NIL-B cells expressing the kinase-dead double mutant protein (Fig. 4C). However, it was not possible to detect the presence of c-Src in the v-Src immunoprecipitates of the infected cells using the anti-nonphospho-Src (Tyr527) antibody, most likely because the anti-avian Src antibody, which binds to the N-terminal part of v-Src, prevents the interaction with the c-Src protein.

In vitro kinase activity of v-Src and c-Src

The *in vitro* kinase activity of the total Src protein (a combination of both v- and c-Src proteins immunoprecipitated by LA-074 antibody from the infected cells) and also with the v-Src and c-Src immunoprecipitated separately (Fig. 5) was assayed using the Src optimal peptide substrate. We were primarily interested in the changes in c-Src activity induced by the expression of v-Src; and the Src optimal substrate enabled us to read even small changes in the *in vitro* activity of c-Src. However, because the Src optimal peptide is more specific substrate for c-Src than for the v-Src [18], the v-Src activities may be seemingly lower when assayed with this peptide.

The only detectable activity of v-Src appeared in NIL-A cells. No activity of v-Src immunoprecipitated from NIL-B cells was observed, evidently because mutation of the conserved lysine 295 in the ATP-binding site prevents v-Src from binding the nucleotide substrate, and in combination with the Y416F mutation in the activation loop renders the

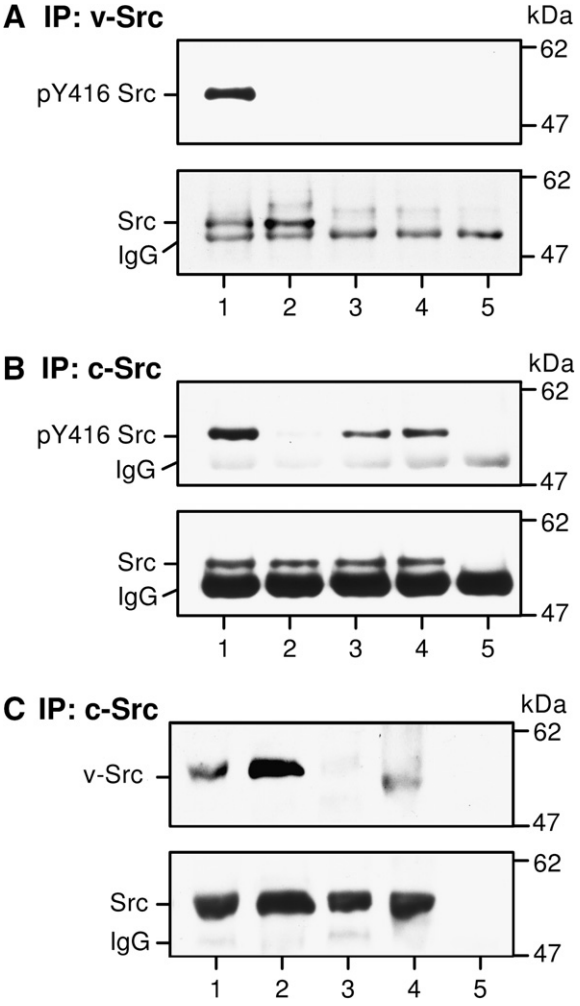


Fig. 4. Phosphorylation of Tyr 416 in v-Src and c-Src proteins in the cell lines derived from embryonic hamster fibroblasts NIL-2 infected with wild-type- and kinase-dead double mutant v-*src*. (A) v-Src protein immunoprecipitated from the cell extracts (400 μg) with monoclonal anti-v-Src antibody as described in Material and methods and examined by SDS–PAGE followed by immunoblotting using polyclonal anti-phospho-Src (Tyr416) antibody (upper panel) or monoclonal anti-Src antibody LA-074 (lower panel). (B) The c-Src protein was immunoprecipitated from supernatants (obtained after v-Src immunoprecipitation of cell extracts) with monoclonal anti-Src antibody LA-074 and examined by SDS–PAGE followed by immunoblotting using polyclonal anti-phospho-Src (Tyr416) antibody (upper panel) or monoclonal anti-Src antibody LA-074 (lower panel). (C) The c-Src protein was immunoprecipitated from the cell extracts of the infected cells with the nonphospho-Src (Tyr527) antibody as described in Material and methods, examined by SDS–PAGE and the presence of co-precipitated v-Src protein was detected using the monoclonal anti-avian Src antibody (clone EC10) (upper panel). The presence of Src protein was detected with monoclonal anti-Src antibody LA-074 (lower panel). Lane 1, NIL-A cells infected with pH19KE (wild-type v-Src). Lane 2, NIL-B cells infected with the kinase-inactive mutant pH19FN (Y416F–K295N v-Src). Lane 3, NIL-C cells infected with the pLXRN plasmid alone. Lane 4, NIL-D, non-infected cells. Lane 5, negative control (extraction buffer). Position of IgGs is indicated.

enzyme completely inactive. The absence of kinase activity in the v-Src immunoprecipitates from NIL-B cells also confirms the absence of c-Src protein in the anti-avian Src immunoprecipitates.

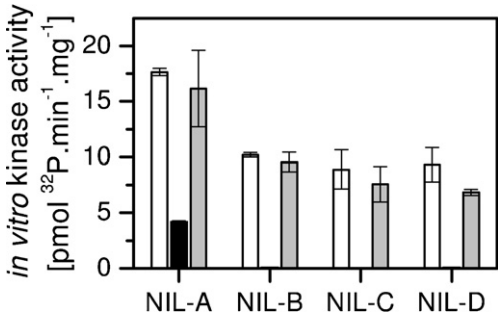


Fig. 5. *In vitro* kinase activity of Src immunoprecipitated from the cell lines derived from embryonic hamster fibroblasts NIL-2 infected with wild-type- and kinase-dead double mutant v-*src*. Src kinase activity was assayed in the immunoprecipitates as described in Material and methods. Open bars, activity of total Src protein immunoprecipitated with the LA-074 antibody. Black bars, activity of v-Src immunoprecipitated with the anti-avian Src antibody. Grey bars, activity of c-Src immunoprecipitated with the LA-074 antibody from the residual supernatant after the v-Src immunoprecipitation. NIL-A, cells infected with pH19KE (wild-type v-Src). NIL-B, cells infected with the kinase-inactive mutant pH19FN (Y416F–K295N v-Src). NIL-C, cells infected with the pLXRN plasmid alone. NIL-D, non-infected cells.

The *in vitro* kinase activity of both total Src and c-Src immunoprecipitated from NIL-A cells was increased compared to protein kinase activity of c-Src in NIL-C and the non-infected NIL-D cells. Similarly, increased protein kinase activity of the endogenous c-Src was previously also observed in RSV-transformed cells that expressed highly active v-Src [12].

Unexpectedly, no decrease of *in vitro* kinase activity of the c-Src immunoprecipitated from NIL-B cells that expressed the kinase-dead v-Src mutant was observed, compared to the activity of c-Src in NIL-C and NIL-D cells (Fig. 5). This suggests complete reversibility of the inactivation of c-Src, probably by the auto-activation of Src protein by MgATP during the course of the kinase assay. The autophosphorylation is very fast, and under the conditions of the end-point radioactive assay it cannot be detected. It can be observed as an initial lag of reaction course of the continuous spectrophotometric assay coupling protein phosphorylation by ATP to NADPH oxidation via pyruvate kinase, phosphoenolpyruvate and lactate dehydrogenase [9].

The decrease of *in vivo* c-Src kinase activity observed in the NIL-B cells expressing kinase-dead double mutant v-Src therefore seems to be caused rather by double dephosphorylation of c-Src protein (suggested by the immunochemical data) than by its downregulation by the Csk.

Discussion

The activation of Src is an autophosphorylation process, the most likely mechanism being the autophosphorylation *in trans*, i.e., two Src protein molecules binding together and phosphorylating one another in their activation loops. Phosphorylation of regulatory Tyr416 in the activation loop of the catalytic domain of Src kinases is supposed to

induce conformational changes in the activation loop that facilitate the transfer of phosphate from ATP to protein substrates, thus enabling the enzyme to reach full kinase activity [20]. The open, active conformation of c-Src is stabilized by the interaction of its C-terminal tail with a hydrophobic pocket in the C-terminal lobe of the catalytic domain [21].

This Tyr416-phosphorylated form of Src is the preferred target for downregulation by C-terminal Src kinase (Csk) [19,22]. The first step of the downregulation is the phosphorylation of Tyr527 in the C-terminal tail by Csk. However, as long as the regulatory Tyr416 in the activation loop remains phosphorylated, the activity of Src is not affected by the phosphorylation of Tyr527 [23,24]. Following dephosphorylation of Tyr416, a re-arrangement takes place in the protein molecule, phosphorylated Tyr527 binds to the SH2 domain, and the resulting deformation of the catalytic domain causes the inactivation of the enzyme. Amino-acid sequences of the v-Src and of the c-Src are almost identical (96% of homology). In spite of the great overall similarity of these two proteins, there are several point mutations that affect the binding properties of both the SH3 and SH2 domains, disabling several intramolecular bonds between the SH3 and catalytic domains, thus rendering the v-Src protein less compact and destabilizing it [25].

Formation of hybrid dimers of v- and c-Src proteins may explain both types of cellular Src response to the presence of the overexpressed v-Src protein described in this paper, i.e., activation of the c-Src in the presence of active v-Src and its inactivation induced by the kinase-dead Src mutant. Higher flexibility of the v-Src molecule may facilitate its interaction with the c-Src protein, leading to increased transphosphorylation of the regulatory Tyr416 in the activation loop and stabilization of the c-Src molecule.

There are two possible causes of the decrease in phosphorylation of the c-Src regulatory Tyr416 in the presence of the kinase-dead v-Src mutant. First, the kinase-dead v-Src protein can act directly as a competitive inhibitor of Tyr416 phosphorylation catalyzed by the c-Src protein. However, this is not likely the only mechanism of inactivation of c-Src induced by kinase-dead Src mutants, because a similar decrease of c-Src activity has been observed in human colon cancer cells expressing a double K295M Y527F kinase-dead c-*src* mutant with the intact regulatory Tyr416 [26], which cannot act as a competitive inhibitor of the endogenous c-Src. The decrease in Tyr416 phosphorylation of c-Src can therefore be explained by considering the enzymatic properties of the Src tyrosine kinase. The kinetic mechanism of Src tyrosine kinase is sequential Bi Bi rapid equilibrium random [27], i.e., either of the two substrates, protein or ATP, is able to bind as first. The v-Src protein is inactivated by mutation of the conserved Lys295 (essential for ATP-binding); however, it is able to bind its protein substrates, and thus can bind the c-Src protein. Formation of transient dead-end complexes of Y416F–K295N v-Src mutant with c-Src in which neither of the two proteins can phosphorylate the other may lead to a decrease of availability of c-Src molecules for homodimer formation with other c-Src molecules necessary for the transphosphorylation, and, eventually, to the inactivation of endogenous c-Src.

Therefore, the inactivation of endogenous c-Src in the presence of ectopic kinase-dead v-Src may confirm the existence of Src protein dimers. Together with the activation of endogenous c-Src observed in the presence of active v-Src and the presence of v-Src protein in the anti-nonphosphoSrc (Tyr527) immunoprecipitates, our results suggest that dimerization of Src protein can play an important role in the regulation of Src tyrosine kinase activity at the protein level.

Acknowledgments

Dr. Jiří Hejnar from the Institute of Molecular Genetics very kindly provided the Y416F v-*src* mutant that was used to prepare the double Y416F–K295N mutant. Excellent technical assistance of Mrs. M. Čechová and Mrs. I. Dosoudilová is gratefully acknowledged. This work was supported by Grant 301/04/0550 from the Grant Agency of the Czech Republic and the Research Project AVOZ 50520514 from the Academy of Sciences of the Czech Republic.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.abb.2006.09.011.

References

[1] T. Erpel, S.A. Courtneidge, Curr. Opin. Cell Biol. 7 (1995) 176–182.
[2] J. Bodor, E. Poliak, J. Pichrtová, J. Geryk, J. Svoboda, Nucleic Acids Res. 17 (1989) 8869.
[3] R.L. Strausberg, E.A. Feingold, L.H. Grouse, J.G. Derge, R.D. Klausner, F.S. Collins, L. Wagner, C.M. Shenmen, G.D. Schuler, S.F. Altschul, B. Zeeberg, K.H. Buetow, et al., Proc. Natl. Acad. Sci. USA 99 (2002) 16899–16903.
[4] J.T. Parsons, S.J. Parsons, Curr. Opin. Cell Biol. 9 (1997) 187–192.
[5] G. Sun, A.K. Sharma, R.J.A. Budde, Oncogene 17 (1998) 1587–1595.
[6] Y. Sugimoto, E. Erikson, Y. Graziani, R.L. Erikson, J. Biol. Chem. 260 (1985) 13838–13843.
[7] D. Feder, J.M. Bishop, J. Biol. Chem. 265 (1990) 8205–8211.
[8] M. Osusky, S.J. Taylor, D. Shalloway, J. Biol. Chem. 270 (1995) 25729–25732.
[9] S.C. Barker, D.B. Kassel, D. Weigl, X. Huang, M.A. Luther, W.B. Knight, Biochemistry 34 (1995) 14843–14851.
[10] J.A. Cooper, A. MacAuley, Proc. Natl. Acad. Sci. USA 85 (1988) 4232–4236.
[11] Z. Tuháčková, V. Sovová, E. Šloncová, C.G. Proud, Int. J. Cancer 81 (1999) 963–969.
[12] Z. Tuháčková, M. Vojtěchová, J. Hlaváček, M. Ruzzene, V. Sovová, L.A. Pinna, Biochem. Biophys. Res. Commun. 290 (2002) 790–795.
[13] M. Vojtěchová, E. Šloncová, D. Kučerová, J. Jiřička, V. Sovová, Z. Tuháčková, FEBS Lett. 543 (2003) 81–86.
[14] A. Mukherjee, L. Arnaud, J.A. Cooper, J. Biol. Chem. 278 (2003) 40806–40814.
[15] O. Machoň, J. Hejnar, P. Hájková, J. Geryk, J. Svoboda, Gene 174 (1996) 9–17.
[16] J.V. Plachý, J. Hejnar, K. Trtková, K. Trejbalová, J. Svoboda, K. Hála, Vaccine 19 (2001) 4526–4535.
[17] L. Diamond, Int. J. Cancer 2 (1967) 143–152.

[18] M. Vojtěchová, Z. Tuháčková, J. Hlaváček, J. Velek, V. Sovová, Arch. Biochem. Biophys. 421 (2003) 277–282.
[19] D. Wang, X.-Y. Huang, P.A. Cole, Biochemistry 40 (2001) 2004–2010.
[20] M.T. Brown, J.A. Cooper, Biochim. Biophys. Acta 1287 (1996) 121–149.
[21] S.W. Cowan-Jacob, G. Fendrich, P.W. Manley, W. Jahnke, D. Fabbro, J. Liebetanz, T. Meyer, Structure 13 (2005) 861–871.
[22] Z. Songyang, S.E. Shoelson, J. McGlade, P. Olivier, T. Pawson, X.R. Bustelo, M. Barbacid, H. Sabe, H. Hanafusa, T. Yi, R. Ren, D. Baltimore, S. Ratnofsky, R.A. Feldman, L.C. Cantley, Mol. Cell. Biol. 14 (1994) 2777–2785.
[23] R.J. Boerner, D.B. Kassel, S.C. Barker, B. Ellis, P. DeLacy, W.B. Knight, Biochemistry 35 (1996) 9519–9525.
[24] G. Sun, A.K. Sharma, R.J. Budde, Oncogene 17 (1998) 1587–1595.
[25] S.F. Falsone, S. Leptihn, A. Osterauer, M. Haslbeck, J. Buchner, J. Mol. Biol. 344 (2004) 281–291.
[26] G.J. Griffiths, M.Y. Koh, V.G. Brunton, C. Cawthorne, N.A. Reeves, M. Greaves, M.J. Tilby, D.G. Pearson, C.J. Ottley, P. Workman, M.C. Frame, C. Dive, J. Biol. Chem. 279 (2004) 46113–46121.
[27] R.J. Boerner, S.C. Barker, W.B. Knight, Biochemistry 34 (1995) 16419–16423.