Transgene Expression Facilitated by the *v-src* Splice Acceptor Can Impair Replication Kinetics and Lead to Genomic Instability of Rous Sarcoma Virus-Based Vectors^{∇†}

Daniel Portsmouth,^{1,2} Daria Deitermann,¹ Brian Salmons,³
Walter H. Günzburg,^{1,2} and Matthias Renner^{3*}

Research Institute of Virology and Biomedicine, University of Veterinary Medicine,¹ Christian-Doppler Laboratory for Gene Therapeutic Vector Development,² and Austrianova Biotechnology GmbH,³ Vienna, Austria

Received 9 August 2007/Accepted 19 November 2007

Rous sarcoma virus (RSV) can be used for the simple generation of high-titer replication-competent retroviral (RCR) vectors. Retroviruses undergo frequent genomic recombination, however, and vectors with reduced replication kinetics are rapidly overgrown by mutant forms. Vector design is hence critical to vector efficacy. In this study, two different designs of RSV-based RCR vectors were evaluated. Vectors in which transgene expression was facilitated by the *v-src* splice acceptor were revealed to have greatly reduced replication kinetics and genomic stability in comparison to vectors in which transgene expression was mediated by an internal ribosome entry site in the 3' untranslated region.

Rous sarcoma virus (RSV) is the only known replication-competent retrovirus to have naturally acquired an oncogene (*v-src*) (5, 12–14). Heterologous sequences can be expressed instead of the *v-src* gene in RSV-based replication-competent retroviral (RCR) vectors (3, 7, 8, 10, 11), whereby high-titer vector stocks can be generated simply by transfecting cells and allowing the vector to spread, without the need for the multiple transfection and/or selection protocols required for the production of replication-defective retroviral vectors. In addition to the use of RSV-based RCR vectors as tools in avian systems, RSV-based RCR vectors can be genetically pseudotyped with the amphotropic 4070A *env* gene from murine leukemia virus (MLV) to allow transgene delivery to mammalian cells (1, 2). Since RSV does not replicate in mammalian cells, genetic pseudotyping of RSV-based RCR vectors also provides an interesting alternative to replication-deficient MLV and human immunodeficiency virus vectors for human gene therapy (6).

The modification of retroviral genomes by the addition of heterologous sequences, however, can lead to impaired replication kinetics, and due to the inherent propensity of retroviral genomes to undergo recombinations during reverse transcription, viral mutants lacking heterologous sequences will emerge as predominant subpopulations if they replicate with kinetics superior to those of the parental genotype. Such mutants may arise from RSV-based RCR vectors during vector production in avian cell culture or during replication in avian species *in vivo*. Hence, vector designs which allow rapid replication are desirable, since these should lead to more reliable transgene expression in target cells. We have therefore investigated the

replication and transgene propagation kinetics of RSV-based RCR vectors utilizing different strategies to facilitate marker gene expression and have characterized the nucleotide sequences and/or structural motifs associated with the generation of predominant mutants arising during vector replication.

RCR vectors expressing fluorescent marker proteins were generated by modifying the genetically pseudotyped amphotropic vector RCASBPM2C(4070A) (1), whereby the expression of enhanced green fluorescent protein (eGFP) and *Drosophila* sp. red fluorescent protein (DsRed) from vectors RCASBPM2C(4070A)eGFP and RCASBPM2C(4070A)DsRed or RCANBPM2C(4070A)IRESeGFP and RCANBPM2C(4070A)IRESDsRed was mediated either by an additional spliced message facilitated by the *v-src* splice acceptor (SA) or by an encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) (9) in the 3' untranslated region (UTR), respectively (Fig. 1A).

To evaluate transgene propagation kinetics, vectors were used to inoculate DF-1 cells at a relative multiplicity of infection (MOI) of 0.001, calculated by comparing the reverse transcriptase (RT) activity of each vector preparation (determined using a product-enhanced RT [PERT] assay) to that of a stock of vector RCASBPM2C(4070A)eGFP of known titer. Following passaging and fluorescence-activated cell sorter (FACS) analyses every 2 or 3 days postinfection, it became apparent that vectors RCANBPM2C(4070A)IRESeGFP and RCANBPM2C(4070A)IRESDsRed propagated the eGFP and DsRed genes in infected cells much more efficiently than vectors RCASBPM2C(4070A)eGFP and RCASBPM2C(4070A)DsRed (Fig. 2A). The relatively slow transgene propagation kinetics of vectors RCASBPM2C(4070A)eGFP and RCASBPM2C(4070A)DsRed suggested that the utilization of the *v-src* SA to facilitate transgene expression may interfere with virus replication. To investigate this hypothesis, PERT analyses were performed every 2 or 3 days following the infection of DF-1 cells at an MOI of 0.001. Additionally, to determine the impact of the *v-src* SA *per se*, vector RCASBPM2C(4070A)

* Corresponding author. Mailing address: MR, Austrianova Biomanufacturing AG, Veterinärplatz 1, A-1210 Vienna, Austria. Phone: 43-1-25077-2301. Fax: 43-1-25077-2390. E-mail: renner@austrianova.com.

† Supplemental material for this article may be found at <http://jvi.asm.org/>.

[∇] Published ahead of print on 5 December 2007.

FIG. 1. RSV-based RCR vector constructs and experimental setup. (A) Vectors are based on the Schmidt-Ruppin A strain of RSV (white vector backbone). In vector RCASBPM2C(4070A), the RSV *env* gene is replaced by the MLV 4070A amphotropic *env* gene, and the RSV upstream direct repeat (UDR) and *v-src* gene are deleted. A unique ClaI restriction enzyme site is present downstream of the *v-src* SA. In vectors RCASBPM2C(4070A)eGFP and RCASBPM2C(4070A)DsRed, the eGFP and DsRed genes, respectively, are inserted into the unique ClaI site and expressed from a spliced message utilizing the *v-src* SA. In vectors RCANBPM2C(4070A)IRESeGFP and RCANBPM2C(4070A)IRESDsRed, the eGFP and DsRed genes, respectively, are inserted downstream of an EMCV IRES element fused to the 3' end of the 4070A *env* gene. SD, splice donor; DDR, downstream direct repeat; R, repeat region. (B) DF-1 cells were inoculated with RSV-based RCR vector stocks at a predetermined MOI. For transgene propagation analysis, infected cells were passaged every 2 or 3 days until the percentage of fluorescent protein-expressing cells, as measured by FACS analysis, ceased to increase from one passage to the next, at which point genomic DNA and virion RNA were isolated. For vector stability analysis via serial infection cycles, diluted cell-free supernatant was transferred from infected cells 4 or 5 days after infection and used to inoculate fresh DF-1 cells.

was compared with vector RCASBPM2C(4070A)delSA, in which the *v-src* SA is deleted (Fig. 1A). PERT assay data revealed that vectors harboring either the eGFP or DsRed gene indeed replicated much more efficiently when the transgene was expressed via the EMCV IRES in the 3' UTR than when it was expressed from a spliced message facilitated by the

v-src SA (Fig. 2B). Interestingly, however, there was no difference in the replication kinetics of vectors RCASBPM2C(4070A) and RCASBPM2C(4070A)delSA, indicating that the presence of the *v-src* SA per se is not detrimental to virus replication (Fig. 2B).

To evaluate the genomic stability of transgene-expressing

FIG. 2. Replication and transgene propagation kinetics and genomic stability of RSV-based RCR vectors. (A) DF-1 cells were inoculated with vectors RCANBPM2C(4070A)IRESeGFP (closed squares), RCANBPM2C(4070A)IRESDsRed (open squares), RCASBPM2C(4070A)eGFP (closed triangles), and RCASBPM2C(4070A)DsRed (open triangles) at a relative MOI of 0.001. Cells were subsequently passaged and subjected to FACS analysis every 2 or 3 days. +ve, positive. (B) Replication kinetics were determined by PERT assay of cell-free supernatants harvested at each passage from cells infected with vectors RCANBPM2C(4070A)IRESeGFP (closed squares), RCANBPM2C(4070A)IRESDsRed (open squares), RCASBPM2C(4070A)eGFP (closed triangles), RCASBPM2C(4070A)DsRed (open triangles), RCASBPM2C(4070A) (open circles), and RCASBPM2C(4070A)delSA (open diamonds) at a relative MOI of 0.001. Shown are the RT activities at each passage, relative to those of the inoculates used to initiate the infections. (C) DF-1 cells were inoculated with vectors RCANBPM2C(4070A)IRESeGFP (black bars) and RCANBPM2C(4070A)IRESDsRed (white bars) at an MOI of 0.001 and RCASBPM2C(4070A)eGFP (gray bars) at an MOI of 0.1 and passaged and subjected to FACS analysis every 2 or 3 days. New infection cycles were initiated by inoculation with cell-free virus-containing supernatant harvested 4 or 5 days after the inoculation in the preceding infection cycle; supernatants were diluted 1,000-fold for vectors RCANBPM2C(4070A)IRESeGFP and RCANBPM2C(4070A)IRESDsRed and 10-fold for vector RCASBPM2C(4070A)eGFP. Shown are the maximum percentages of eGFP- and DsRed-expressing DF-1 cells detected following the passaging of cells at each infection cycle.

vectors, multiple serial infection cycles were performed with DF-1 cells inoculated with vectors RCANBPM2C(4070A) IRESeGFP (MOI, 0.001), RCANBPM2C(4070A) IRESdsRed (MOI, 0.001), and RCASBPM2C(4070A)eGFP (MOI, 0.1) and passaged 2 days postinfection. Four days postinfection, filtered, diluted supernatant from infected cells was used to inoculate fresh DF-1 cells in a second infection cycle (Fig. 1B). This process was repeated for a total of 10 serial infection cycles. Even with the greatly reduced initial MOI and greatly increased dilution factor used in infection cycles with vectors RCANBPM2C(4070A) IRESeGFP and RCANBPM2C(4070A) IRESdsRed (MOI, 0.001; dilution, 1000-fold) compared to those used with vector RCASBPM2C(4070A)eGFP (MOI, 0.1; dilution, 10-fold), the former were able to propagate transgene expression over more infection cycles than the latter (Fig. 2C).

To characterize predominant deletion mutants arising during vector replication, PCR was carried out using genomic DNA extracted from DF-1 cells infected with both transgene-expressing and non-transgene-expressing vectors at an initial MOI of 0.001. PCR was also performed with DNA extracted from cells infected with vectors RCANBPM2C(4070A) IRESeGFP and RCANBPM2C(4070A) IRESdsRed following multiple serial 1,000-fold-diluting infection cycles. Analyses of PCR products from DNA extracted from cells infected with vectors RCANBPM2C(4070A) IRESeGFP and RCANBPM2C(4070A) IRESdsRed from infection cycles 2, 4, 6, 8, and 10 revealed that major deletion mutants did not emerge within the first two infection cycles (Fig. 3A and B, lanes 1), becoming evident only during subsequent infection cycles (Fig. 3A and B, lanes 2 through 5). Analyses of PCR products from DNA extracted from four independent infection experiments with vectors RCASBPM2C(4070A)eGFP and RCASBPM2C(4070A)DsRed, on the other hand, revealed that major deletion mutants became dominant in the vector population within a single infection cycle (Fig. 3C, lanes 1 to 4 and 5 to 8, respectively). Thus, PCR results indicate that vectors from which transgene expression is mediated by an EMCV IRES in the 3' UTR are genomically much more stable than vectors from which transgene expression is mediated by a spliced message facilitated by the *v-src* SA. PCR analysis of DNA extracted from DF-1 cells infected with vectors RCASBPM2C(4070A) (Fig. 3D, lanes 1 and 2) and RCASBPM2C(4070A)delSA (Fig. 3D, lanes 3 and 4) gave rise to a single band of a size expected for the respective parental vectors, indicating that mutants did not emerge during the replication of these vectors.

Sequencing analyses of PCR products revealed that several common recombination junctions were involved in the formation of deletion mutants, both in different mutants of the same vector and in deletion mutants of different vectors, indicating that, at least to a certain extent, recombinogenic hot spots exist in the vectors evaluated in this study (see the supplemental material). Eight of the sequenced recombination junctions were associated with direct repeats and two were associated with inverted repeats, but over half of the sequenced recombination junctions revealed no role for nucleotide homology (see the supplemental material). No differences in the nucleotide sequences associated with IRES- or *v-src*-containing vectors could be ascertained. Strikingly, however, the *v-src* SA was deleted from almost every sequenced recombinant of vectors RCASBPM2C(4070A)eGFP and RCASBPM2C(4070A)DsRed,

FIG. 3. PCR analysis of integrated proviral vectors in infected cells. Primers binding at the 3' end of the 4070A *env* gene and in the U3 region of the RSV 3' long terminal repeat were used to perform PCR analysis of genomic DNA extracted from infected cells initially inoculated with RSV-based RCR vectors at a relative MOI of 0.001. The sizes of the products expected for the parental vectors are indicated with large arrows. M, DNA marker. (A) Vector RCANBPM2C(4070A)IRESeGFP; DNA extracted at the ends of infection cycles 2, 4, 6, 8, and 10 (lanes 1 to 5, respectively). (B) Vector RCANBPM2C(4070A)IRESdsRed; DNA extracted at the ends of infection cycles 2, 4, 6, 8, and 10 (lanes 1 to 5, respectively). (C and D) Vectors RCASBPM2C(4070A)eGFP (lanes 1 to 4) and RCASBPM2C(4070A)DsRed (lanes 5 to 8) (C) and vectors RCASBPM2C(4070A) (lanes 1 and 2) and RCASBPM2C(4070A)delSA (lanes 3 and 4) (D); DNA extracted following five passages of cells from infection cycle 1. Each lane represents DNA extracted from an independent infection experiment. The nomenclature for each sequenced product indicates the parental vector (first three letters), the infection experiment in which the product was detected (first number; each number indicates inoculation with a different vector stock), and the relative size of each PCR product (second number; the lowest number indicates the smallest sequenced PCR product). RIE, RCANBPM2C(4070A)IRESeGFP; RID, RCANBPM2C(4070A)IRESdsRed; RAE, RCASBPM2C(4070A)eGFP; RAD, RCASBPM2C(4070A)DsRed.

but not from the parental vector RCASBPM2C(4070A) (see the supplemental material), supporting the hypothesis that the presence of the *v-src* SA is not detrimental to RSV replication per se but only in combination with the presence of downstream transgenic sequences.

FIG. 4. Northern analysis was performed with total cellular RNA extracted from DF-1 cells either infected with vector RCANBPM2C(4070A)IRES-eGFP or RCASBPM2C(4070A)eGFP at infection cycle 1 (lanes 1 and 2, respectively) or infected with vector RCASBPM2C(4070A)eGFP at infection cycles 1 and 10 (lanes 3 and 4, respectively). Probes specific for the long terminal repeat-*gag* region and the eGFP gene were both used in the same hybridization reaction.

It is possible that the use of an additional SA to facilitate the production of an additional subgenomic message for transgene expression leads to a concomitant decrease in the genomic length and the amount of *env*-spliced RNA in infected cells, which may, in turn, lead to a decrease in the amount of infectious virus particles being released, thus conferring the observed replicative disadvantage on vectors containing an additional SA. To investigate this hypothesis, Northern analyses were performed using RNA extracted from DF-1 cells infected with vectors (i) RCANBPM2C(4070A)IRES-eGFP and RCASBPM2C(4070A)eGFP at infection cycle 1 and (ii) RCASBPM2C(4070A)eGFP at infection cycles 1 and 10. At infection cycle 1, DF-1 cells infected with vector RCANBPM2C(4070A)IRES-eGFP contained more *env*-spliced and full-length RNA than DF-1 cells infected with vector RCASBPM2C(4070A)eGFP, which contained large amounts of eGFP gene-spliced RNA and only poorly detectable full-length or *env*-spliced messages (Fig. 4, lanes 1 and 2, respectively). In DF-1 cells infected with vector RCASBPM2C(4070A)eGFP at cycle 10, however, full-length and *env*-spliced messages were more abundant than those in DF-1 cells infected with vector RCASBPM2C(4070A)eGFP at cycle 1 (Fig. 4, lane 4 and 3, respectively). Hence, the results of the Northern analyses of RNA extracted from infected cells support the hypothesis that the use of an additional SA to facilitate the production of an additional subgenomic message for transgene expression leads to a decrease in the genomic length and the amount of *env*-spliced RNA in infected cells and that vector mutants which contain deletions spanning the *v-src* SA no longer generate this subgenomic message, such that more full-length and *env*-spliced messages are available for virus production, in turn leading to an increase in vector replication kinetics. In this respect, the observation that vector RCASBPM2C(4070A), which contains the *v-src* SA but does not replicate more slowly than vector RCASBPM2C(4070A), from which the SA is deleted, may be explained by the absence of additional downstream transgene sequences containing putative exonic splicing enhancers, which are known to play an important role in determining the level of splicing to upstream SAs (4).

Taken together, our results indicate that genetically pseudotyped, amphotropic RSV-based RCR vectors which utilize an EMCV IRES in the 3' UTR for transgene expression replicate more efficiently and are genomically more stable than equivalent vectors from which transgenes are expressed via a subgenomic message facilitated by the *v-src* SA and that, hence, predominant deletion mutants are less likely to arise during the generation of vector stocks by passaging in avian cells. Moreover, the findings we present here may be of relevance for RSV vectors in general and, thus, may also have implications for the use of replicating RSV vectors as research tools in avian systems, in which improved replication kinetics and vector stability represent a potentially significant advantage.

We thank Stephen H. Hughes (Human Immunodeficiency Virus Drug Resistance Program, National Cancer Institute at Frederick, MD) for providing plasmid pRCASBPM2C(4070A) and Noriyuki Kasahara (University of California—Los Angeles, Los Angeles, CA) for providing plasmid pACE-GFP. Many thanks also to Elzbieta Knapp, Magdalena Pusch, Sonja Sabitzer, and Reinhard Ertl for help and advice with FACS analysis and PERT assays.

This work was financed by the Christian-Doppler Forschungsgesellschaft, Austria.

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. Bell, E. J., and P. M. Brickell. 1997. Replication-competent retroviral vectors for expressing genes in avian cells in vitro and in vivo. *Mol. Biotechnol.* **7**:289–298.
4. Blencowe, B. J. 2000. Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends Biochem. Sci.* **25**:106–110.
5. Hanafusa, H., C. C. Halpern, D. L. Buchhagen, and S. Kawai. 1977. Recovery of avian sarcoma virus from tumors induced by transformation-defective mutants. *J. Exp. Med.* **146**:1735–1747.
6. 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 virus vectors. *Hum. Gene Ther.* **18**:691–700.
7. Hughes, S., and E. Kosik. 1984. Mutagenesis of the region between *env* and *src* of the SR-A strain of Rous sarcoma virus for the purpose of constructing helper-independent vectors. *Virology* **136**:89–99.
8. Hughes, S. H. 2004. The RCAS vector system. *Folia Biol. (Praha)* **50**:107–119.
9. Jang, S. K., and E. Wimmer. 1990. Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein. *Genes Dev.* **4**:1560–1572.
10. Murakami, M., H. Watanabe, Y. Niikura, T. Kameda, K. Saitoh, M. Yamamoto, Y. Yokouchi, A. Kuroiwa, K. Mizumoto, and H. Iba. 1997. High-level expression of exogenous genes by replication-competent retrovirus vectors with an internal ribosomal entry site. *Gene* **202**:23–29.
11. Petropoulos, C. J., and S. H. Hughes. 1991. Replication-competent retrovirus vectors for the transfer and expression of gene cassettes in avian cells. *J. Virol.* **65**:3728–3737.
12. Rohrschneider, L. R., R. N. Eisenman, and C. R. Leitch. 1979. Identification of a Rous sarcoma virus transformation-related protein in normal avian and mammalian cells. *Proc. Natl. Acad. Sci. USA* **76**:4479–4483.
13. Wang, L. H., C. C. Halpern, M. Nadel, and H. Hanafusa. 1978. Recombination between viral and cellular sequences generates transforming sarcoma virus. *Proc. Natl. Acad. Sci. USA* **75**:5812–5816.
14. Wang, L. H., C. Moscovici, R. E. Karess, and H. Hanafusa. 1979. Analysis of the *src* gene of sarcoma viruses generated by recombination between transformation-defective mutants and quail cellular sequences. *J. Virol.* **32**:546–556.