JOURNAL OF VIROLOGY, Feb. 1989, p. 1015–1018
0022-538X/89/021015-04$02.00/0

Vol. 63, No. 2

# The LTR, v-*src*, LTR Provirus Generated in the Mammalian Genome by *src* mRNA Reverse Transcription and Integration

JOSEF BODOR AND JAN SVOBODA*

*Department of Cellular and Viral Genetics, Institute of Molecular Genetics, Czechoslovak Academy of Sciences, 166 37 Prague, Czechoslovakia*

Received 20 June 1988/Accepted 14 October 1988

**Different types of altered proviruses of Rous sarcoma virus (RSV) have been detected in mammalian tumor cell lines. We cloned and sequenced one of these altered proviruses with the structure LTR, v-*src*, LTR. The presence of an intact viral splice junction, as well as duplications of the chromosomal sequence GCGGGG flanking the two 2-base-pair-deleted LTRs, demonstrated reverse transcription and normal retroviral integration of *src* mRNA in mammalian cells. In addition, a 1-nucleotide deletion 2 bases upstream from the AAUAAA polyadenylation signal is suspected to be responsible for the absence of a poly(A) track in the *src* mRNA present in virions of rescued viruses.**

The simplified transcriptionally and translationally active LTR, v-*src*, LTR provirus has been detected in the H-19 hamster tumor cell line transformed by the Prague strain of Rous sarcoma virus (RSV) subgroup C (PR-RSV-C) rescued from XC cells (provirus II) as reported previously (10, 20, 35, 39). It was demonstrated that this cryptic provirus is functional, because if properly complemented (by fusion with chicken fibroblasts preinfected with Rous-associated virus 1), transforming viruses can be rescued from H-19 cells which encapsidate RNA having the character of nonpolyadenylated *src* mRNA (37). According to restriction enzyme analysis and S1 mapping, the proviral structures of most of the rescued viruses, as well as of the original H-19 provirus, corresponded to that of reverse-transcribed *src* mRNA (8, 37). The aim of this work was to elucidate the structure of the LTR, v-*src*, LTR provirus integrated in the hamster genome to understand its genesis.

For cloning, we used the finding that the whole LTR, v-*src*, LTR provirus, together with flanking cellular sequences, is contained within a 5.9-kilobase-pair *Hin*dIII fragment and is excised by *Eco*RI as a 2.6-kilobase-pair fragment (37, 38).

*Hin*dIII fragments 5.4 to 6.0 kilobase pairs long were cut out of agarose and electroeluted, and the purified DNA was ligated with λNM1149 DNA opened at the *Hin*dIII site (18, 21). After molecular cloning, both of the expected *Hin*dIII and *Eco*RI fragments were observed (data not shown). The molecular clone obtained was designated λH19clC2.1.

The integrity of λH19clC2.1 was verified in cotransfection experiments in which 10 μg of each of λH19clC2.1 and λ311411 DNAs containing a complete MAV-1 insert (25) was used to treat chicken fibroblasts (9). Transforming virus was detected 12 days after transfection ($7 \times 10^1$ focus-forming units/ml) and reached a titer of $6.5 \times 10^2$ focus-forming units/ml after 4 weeks of culture. The relatively low titer of this transforming virus has already been noticed in virus rescue experiments (8). This indicates that in spite of the presence of the main structures required for packaging in *src* mRNA (see below), some additional *cis*-acting and packaging signals of the *gag* region are absent (1, 2, 14, 27, 33). Moreover, a lack of appropriate structure and the small size of *src* mRNA might be responsible for decreased binding to the encapsidation proteins (5), and the efficient packaging of viral genomic RNA might compete with that of *src* mRNA (7).

λH19clC2.1 DNA was digested with *Hin*dIII or *Eco*RI and restriction enzyme *Acc*I, *Apa*I, *Bal*I, *Bam*HI, *Bgl*I, *Bgl*II, *Bst*EII, *Eco*RV, *Eco*47III, *Hin*cII, *Hpa*I, *Kpn*I, *Nar*I, *Nco*I, *Nde*I, *Pst*I, *Pvu*I, *Pvu*II, *Sac*I, *Sma*I, *Sph*I, *Sst*II, *Stu*I, *Xba*I, or *Xho*I and hybridized to the $^{32}$P-labeled *src*11 probe (the *Pvu*II fragment spanning base pairs [bp] 8098 to 8671 of PR-RSV-C) or a long terminal repeat (LTR) 2 probe (the *Hae*III fragment spanning bp 8992 to 9335 of PR-RSV-C) or probe pATV-8, which detects all RSV genes (13). In this way, a restriction enzyme cleavage site map of λH19clC2.1 was obtained (Fig. 1). This map corresponds to reverse-transcribed *src* mRNA after expected RSV genomic RNA splicing.

To determine the precise structure of λH19clC2.1 relative to PR-RSV-C, we examined the nucleotide sequence of the provirus restriction fragments that seemed likely to contain important features. The viral leader, including the primer-binding site for Trp tRNA from which synthesis of the minus-strand DNA initiates during reverse transcription, was found to be intact (data not shown). As expected, a deletion was found to have occurred precisely between the splice donor site, which is located 18 bases downstream from the beginning of the *gag* gene sequence, and the acceptor site for the subgenomic *src* mRNA, which is located 75 bases upstream from the v-*src* initiation codon (Fig. 2).

This arrangement demonstrates that *src* mRNA was used as the template for genesis of the LTR, v-*src*, LTR provirus. With a packaging mutant deleted in the leader sequence, high-frequency (50%) generation of proviral structures which contain the splice junction of *src* mRNA was recently reported in chicken fibroblasts (15, 22). The published size (2.3 kilobase pairs) indicates that the proviruses retained the deleted leader. In contrast, our provirus was found in only 1 of 24 mammalian tumors tested (J. Pichrtová, Ph.D. thesis, Institute of Molecular Genetics, Czechoslovak Academy of Sciences, Prague, Czechoslovakia, 1988) but retained all structures of *src* mRNA (see below). Therefore, it most likely represents an infrequent event of unaltered *src* mRNA packaging and reverse transcription. In fact, the *src* mRNA is equipped with the leader sequence containing the main packaging signals (14–16, 27, 32) and with one copy of the

* Corresponding author.

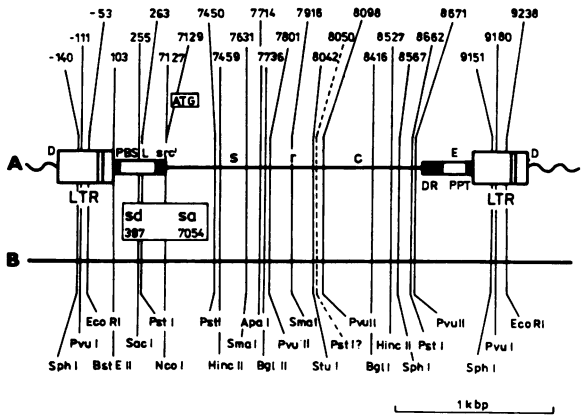


FIG. 1. Restriction enzyme cleavage site map of the proviral part of the molecular clone λH19clC2.1. (A) Proposed structure of the LTR, v-*src*, LTR genome corresponding to reverse-transcribed *src* mRNA after regular retroviral splicing in which the splice donor site (sd) at position 397 would be exactly joined to the splice acceptor site (sa) for subgenomic *src* mRNA at position 7,054. Restriction enzyme cleavage sites are numbered according to Schwartz et al. (26), and each number designates the first nucleotide of the recognition sequence. (B) Actual restriction map obtained by Southern blot restriction analysis which satisfactorily fits with the proposed structure (vertical lines), except for one *Pst*I site at position 8,050 (indicated by a broken line). Wavy lines represent flanking hamster chromosomal DNA. Abbreviations: D, expected duplications of hamster chromosomal DNA after integration; PBS, primer-binding site; L, viral leader; src', noncoding region of viral *src* gene; DR, direct repeat; E, inverted repeat element; PPT, polypurine tract. The LTRs are boxed.

direct repeat to the right of v-*src* (see below) which is implicated also in RNA encapsidation (30). The *src* mRNA probably was transmitted to the hamster cell by the chicken-generated RSV particles responsible for H-19 cell transformation (38, 39). Because of general nonpermissiveness of mammalian cells to RSV (36), it is difficult to imagine that a

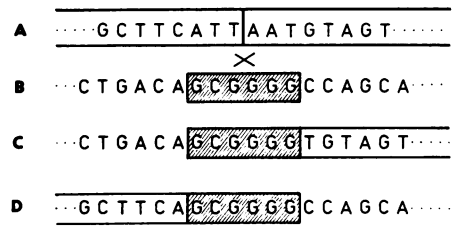


FIG. 3. Model of integration of *src* mRNA into the hamster genome via normal retroviral integration based on actual sequence data (see the text). (A) Nucleotide sequence of the deduced integration site of the circular double-LTR integration precursor of reverse-transcribed *src* mRNA. LTRs are in open boxes. (B) Deduced target site of hamster chromosomal DNA (hatched box). (C) Sequence of the left edge of the LTR, v-*src*, LTR genome. (D) Sequence of the right edge of the LTR, v-*src*, LTR genome.

*src* mRNA produced in RSV-infected hamster cells can be transmitted to other cells. Packaging and reverse transcription of subgenomic viral mRNAs have been reported for the *env* mRNA (17, 31) and repeatedly observed in the *myc* mRNA of the MH2 virus, which harbors two oncogenes (3, 19, 24).

Furthermore, the rest of the noncoding sequences at the 3′ end of the cloned LTR, v-*src*, LTR structure, e.g., the inverted repeat element, one copy of the direct repeat sequence localized downstream from v-*src*, and especially the polypurine tract from which plus-strand DNA synthesis initiates during reverse transcription (29, 40), were also found without changes (Fig. 2). The same is true for both LTRs of the LTR, v-*src*, LTR provirus, which were fully conserved after reverse transcription and integration, including the imperfect inverted repeat at both outer ends of the LTRs, with one exception discussed below (data not shown). Duplications of the chromosomal sequence GCGGGG are found flanking the two 2-bp-deleted LTRs (Fig. 3).

Retroviruses are supposed to produce equally long duplications of host cell DNA regardless of the host cell (42).

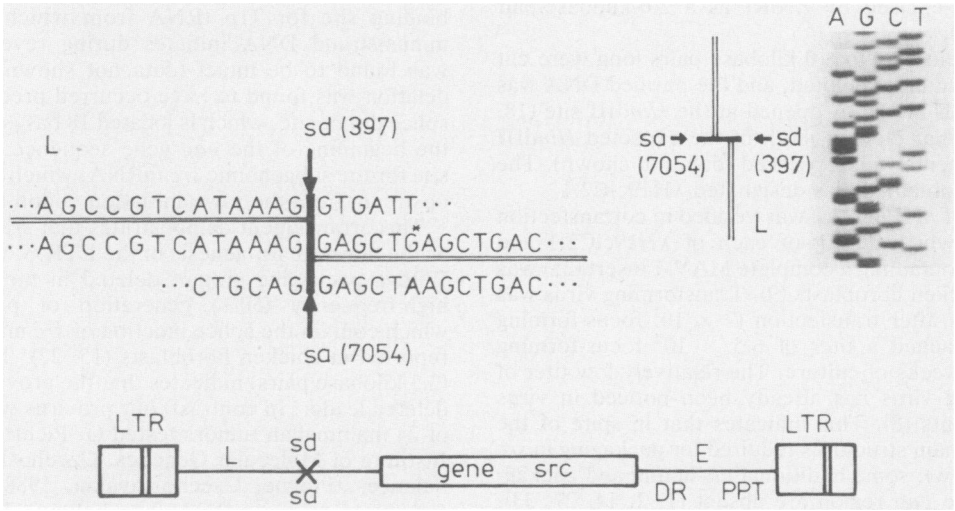


FIG. 2. Nucleotide sequence analysis of the splice junction site of the LTR, v-*src*, LTR genome. The deletion was found to have occurred between the splice donor site (sd) at position 397 and the acceptor site (sa) for subgenomic *src* mRNA at position 7,054. The only change near this splice junction is at position 7,059, where G (asterisk) substitutes for A according to the predicted sequence (26). Sequencing was performed with Sequenase in the M13 mp18, M13 mp19, pUC18, and pUC19 cloning system by standard methods (18). Abbreviations: L, viral leader; DR, direct repeat; E, inverted repeat element; PPT, polypurine tract.

RSV-related viruses were shown to produce hexanucleotide duplications during integration (11, 12). In accord with this, we have observed a hexanucleotide repeat, GCGGGG. This leads us to suggest a model of integration of *src* mRNA into the hamster genome via normal avian retroviral reverse transcription and integration (for reviews of retroviral DNA integration, see references 4, 23, and 28).

All of the sequences compared with the nucleotide sequence of PR-RSV-C (26) were highly conserved. Only 1-bp changes were noted (in 1,595 bases examined, two 1-bp changes and one 1-bp deletion were detected). One of the two 1-bp changes found was an A-to-G transition at position 7,059, 6 bases downstream from the *src* mRNA splice acceptor site (Fig. 2). The same change has been correlated with the increase of *src* mRNA splicing efficiency (34). Such elevated *src* mRNA synthesis might have created a situation favorable for its packaging in virions required for generation of the LTR, v-*src*, LTR provirus in H-19 cells. A deletion of 1 bp, AT, 2 bp upstream from the polyadenylation signal, AATAAA, was found in the U3 region of the 3′ LTR but not in the 5′ LTR. Because of this location, the deletion should be maintained in virions rescued from H-19 cells and might influence the pattern of the viral RNA polyadenylation signal (6). In agreement with this assumption are our previous findings that the *src* mRNA present in virions of viruses rescued from H-19 cells is nonpolyadenylated, in contrast to the polyadenylated *src* mRNA isolated from transformed cells (37).

Here we document for the first time that unaltered *src* mRNA may be reverse transcribed and integrated in the mammalian genome in a way characteristic of the whole provirus. These results confirm and extend previous conclusions based on restriction analysis and S1 mapping (8, 38). The data presented thus provide additional support, at the molecular level, for Temin's original idea suggesting transmission of an oncogene via an RNA transcript and its integration into the cell genome (41).

We are grateful to Howard M. Temin for valuable comments. We are indebted to Jana Lepasant for helpful advice and to Noreen E. Murray for providing cloning vector λNM1149 and *Escherichia coli* NM553 and NM514. We thank J. Geryk for providing H-19 tissue cultures and his expertise and J. Pichrtová for unpublished data. We also thank V. Hoserová, J. Nohavová, and L. Mikušová for excellent technical assistance.

## LITERATURE CITED

1. **Arrigo, S., M. Yun, and K. Beemon.** 1987. *cis*-Acting regulatory elements within *gag* genes of avian retroviruses. Mol. Cell. Biol. **7**:388–397.
2. **Bender, M. A., T. D. Palmer, R. E. Gelinas, and A. D. Miller.** 1987. Evidence that the packaging signal of Moloney murine leukemia virus extends into the *gag* region. J. Virol. **61**: 1639–1646.
3. **Biegalke, B., and M. Linial.** 1987. Retention or loss of v-*mil* sequences after propagation of MH2 virus in vivo or in vitro. J. Virol. **61**:1949–1956.
4. **Colicelli, J., and S. P. Goff.** 1988. Sequence and spacing requirements of a retrovirus integration site. J. Mol. Biol. **199**:47–59.
5. **Darlix, J. L.** 1986. Control of Rous sarcoma virus RNA translation and packaging by the 5′ and 3′ untranslated sequences. J. Mol. Biol. **189**:421–434.
6. **Dougherty, J. P., and H. M. Temin.** 1987. A promoterless retroviral vector indicates that there are sequences in U3 required for 3′ RNA processing. Proc. Natl. Acad. Sci. USA **84**:1197–1201.
7. **Embretson, J. E., and H. M. Temin.** 1987. Lack of competition results in efficient packaging of heterologous murine retroviral RNAs and reticuloendotheliosis virus encapsidation-minus RNAs by the reticuloendotheliosis virus helper cell line. J. Virol. **61**:2675–2683.
8. **Geryk, J., J. Pichrtová, R. V. Guntaka, S. Gowda, and J. Svoboda.** 1986. Characterization of transforming viruses rescued from a hamster tumour cell line harbouring the v-*src* gene flanked by long terminal repeats. J. Gen. Virol. **67**:2395–2404.
9. **Graham, F. L., and J. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–461.
10. **Grófová, M., J. Bízik, S. Bláhová, J. Geryk, and J. Svoboda.** 1985. Identification of transformation-specific proteins synthesized in cryptovirogenic mammalian cells. Folia Biol. (Prague) **31**:152–160.
11. **Hishinuma, F., P. J. DeBona, S. Astrin, and A. M. Skalka.** 1981. Nucleotide sequence of acceptor site and termini of integrated avian endogenous provirus *ev*1: integration creates a 6 bp repeat of host DNA. Cell **23**:155–164.
12. **Hughes, S. H., A. Mutschler, J. M. Bishop, and H. E. Varmus.** 1981. A Rous sarcoma virus provirus is flanked by short direct repeats of a cellular DNA sequence present in only one copy prior to integration. Proc. Natl. Acad. Sci. USA **78**:4299–4303.
13. **Katz, R. A., C. A. Omer, J. H. Weis, S. A. Mitsialis, A. J. Faras, and R. V. Guntaka.** 1982. Restriction endonuclease and nucleotide sequence analyses of molecularly cloned unintegrated avian tumor virus DNA: structure of large terminal repeats in circle junctions. J. Virol. **42**:346–351.
14. **Katz, R. A., R. W. Terry, and A. M. Skalka.** 1986. A conserved *cis*-acting sequence in the 5′ leader of avian sarcoma virus is required for packaging. J. Virol. **59**:163–167.
15. **Kawai, S., and T. Koyama.** 1984. Characterization of a Rous sarcoma virus mutant defective in packaging its own genomic RNA: biological properties of mutant TK15 and mutant-induced transformants. J. Virol. **51**:147–153.
16. **Linial, M., E. Medeiros, and W. S. Hayward.** 1978. An avian oncovirus mutant (SE21Q1B) deficient in genomic RNA: biological and biochemical characterization. Cell **15**:1371–1381.
17. **Majors, J. E., and H. E. Varmus.** 1983. Nucleotide sequencing of an apparent proviral copy of *env* mRNA defines determinants of expression of the mouse mammary tumor virus *env* gene. J. Virol. **47**:495–504.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 390–401. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Martin, P., C. Henry, F. Ferre, C. Bechade, A. Begue, C. Calothy, B. Debuire, D. Stehelin, and S. Saule.** 1986. Characterization of a *myc*-containing retrovirus generated by propagation of an MH2 viral subgenomic RNA. J. Virol. **57**:1191–1194.
20. **Mitsialis, S. A., R. A. Katz, J. Svoboda, and R. V. Guntaka.** 1983. Studies on the structure and organization of avian sarcoma proviruses in the rat XC cell line. J. Gen. Virol. **64**: 1885–1893.
21. **Murray, N. E.** 1983. Phage λ and molecular cloning, p. 395–431. *In* R. W. Hendrix, J. W. Roberts, F. Stahl, and R. A. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Nishizawa, M., T. Koyama, and S. Kawai.** 1987. Frequent segregation of more-defective variants from a Rous sarcoma virus packaging mutant, TK15. J. Virol. **61**:3208–3213.
23. **Panganiban, A. T.** 1985. Retroviral DNA integration. Cell **42**:5–6.
24. **Patchinsky, T., H. W. Jansen, H. Blöcker, R. Frank, and K. Bister.** 1986. Structure and transforming function of transduced mutant alleles of the chicken c-*myc* gene. J. Virol. **59**:341–353.
25. **Perbal, B., J. S. Lipsick, J. Svoboda, R. F. Silva, and M. A. Baluda.** 1985. Biologically active proviral clone of myeloblastosis-associated virus type 1: implications for genesis of avian myeloblastosis virus. J. Virol. **56**:240–244.
26. **Schwartz, D. E., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32**:853–869.
27. **Shank, P. R., and M. Linial.** 1980. Avian oncovirus mutant (SE21Qlb) deficient in genomic RNA: characterization of a deletion in the provirus. J. Virol. **36**:450–456.
28. **Shoemaker, C., S. Goff, E. Gilboa, M. Paskind, V. Mitra, and D.**

**Baltimore.** 1980. Structure of a cloned circular Moloney murine leukemia virus molecule containing an inverted segment: implications for retrovirus integration. Proc. Natl. Acad. Sci. USA **77**:3932–3936.
29. **Sorge, J., and S. H. Hughes.** 1982. Polypurine tract adjacent to the U3 region of the Rous sarcoma virus genome provides a *cis*-acting function. J. Virol. **43**:482–488.
30. **Sorge, J., W. Ricci, and S. H. Hughes.** 1983. *cis*-Acting RNA packaging locus in the 115-nucleotide direct repeat of Rous sarcoma virus. J. Virol. **48**:667–675.
31. **Stacey, D. W.** 1980. Expression of a subgenomic retroviral messenger RNA. Cell **21**:811–820.
32. **Stoker, A. W., and M. J. Bissel.** 1988. Development of avian sarcoma and leukosis virus-based vector-packaging cell lines. J. Virol. **62**:1008–1015.
33. **Stoltzfus, C. M., L. J. Chang, T. P. Cripe, and L. P. Turek.** 1987. Efficient transformation by Prague A Rous sarcoma virus plasmid DNA requires the presence of *cis*-acting regions within the *gag* gene. J. Virol. **61**:3401–3409.
34. **Stoltzfus, C. M., S. K. Lorenzen, and S. L. Berberich.** 1987. Noncoding region between the *env* and *src* genes of Rous sarcoma virus influences splicing efficiency at the *src* gene 3′ splice site. J. Virol. **61**:177–184.
35. **Svoboda, J.** 1960. Presence of chicken tumour virus in the sarcoma of the adult rat inoculated after birth with Rous sarcoma tissue. Nature (London) **186**:980–981.
36. **Svoboda, J.** 1986. Rous sarcoma virus. Intervirology **26**:1–60.
37. **Svoboda, J., M. Dvořák, R. V. Guntaka, and J. Geryk.** 1986. Transmission of (LTR, v-*src*, LTR) without recombination with a helper virus. Virology **153**:314–317.
38. **Svoboda, J., J. Geryk, V. Lhoták, and J. Pichrtová.** 1985. Altered and cryptic proviral sequences in mammalian tumor cell lines and characterization of rescued ASVs, p. 353–359. *In* Proceedings of the 16th FEBS Congress, Part B. VNU Science Press BV, Utrecht, The Netherlands.
39. **Svoboda, J., V. Lhoták, J. Geryk, S. Saule, M. B. Raes, and D. Stehelin.** 1983. Characterization of exogenous proviral sequences in hamster tumor cell lines transformed by Rous sarcoma virus rescued from XC cells. Virology **128**:195–209.
40. **Taylor, J., and L. Sharmeen.** 1987. Retrovirus genome replication: priming specificities of plus-strand DNA synthesis. J. Cell Sci. **7**:189–195.
41. **Temin, H. M.** 1971. Mechanism of cell transformation by RNA tumor viruses. Annu. Rev. Microbiol. **25**:609–649.
42. **Varmus, H. E.** 1983. Retroviruses, p. 411–503. *In* J. A. Shapiro (ed.), Mobile genetic elements. Academic Press, Inc., New York.