Review

Mechanisms of Retrovirus-Induced Oncogenesis

(retroviral integration / cancer / oncogenes / retroviruses)

G. SOURVINOS[1], C. TSATSANIS[2], D. A. SPANDIDOS[1]

[1]Laboratory of Virology, and [2]Laboratory of Clinical Chemistry, Medical School, University of Crete, Heraklion, Crete, Greece

ABSTRACT. Retroviruses are implicated in a series of human and animal tumours such as leukaemias, mammary tumours or skin cancer. The mechanism that they use to induce tumour formation varies. Insertional mutagenesis is a common mechanism in rodent, feline and avian retroviruses, where the retrovirus integrates into the host genome and affects the transcription of the neighbouring genes. Cloning of these affected genes led to identification of a series of oncogenes that play a significant role in the induction of human neoplasms. Retrovirus insertion also serves as a model to identify collaborating oncogenes. Human retroviruses use different, more complex mechanisms contributing to oncogenesis. Studies of the propagation and induction mechanisms used by retroviruses have given insight to the understanding of oncogenesis.

Retroviral life cycle

Retroviruses are plus RNA viruses with a mode of replication in which the RNA genome is reverse-transcribed into DNA and integrated into the host's genome as a provirus (Löwer, 1999). Two identical elements are reconstituted at the 5' and 3' ends of the provirus: the long terminal repeats (LTRs). They contain the promoter, enhancer and repressor sequences for tissue-specific expression, as well as the polyadenylation signal. All genera of retroviruses encode three basic genes, which are expressed and processed as polyproteins: *gag* encoding the core proteins, *ptr/pol* encoding the protease, polymerase and integrase (also designated reverse transcriptase, RT), and *env* encoding the envelope proteins. To express the *env* gene, the *gag* and *prt/pol* genes are spliced out as intronic sequences. The amount of spliced versus unspliced transcripts is controlled in the simplest form by *cis*-acting signals: the secondary structures in the transcripts mediate binding of regulating cellular factors. Genera with a complex genome encode accessory proteins as adapters for cellular factors to modulate expression and infectivity, inducing transactivators of transcription, e.g. *tat* in human immunodeficiency virus (HIV) (Cullen, 1998), regulators of expression of virus proteins, e.g. *rev* in HIV (Cullen, 1998), and infectivity factors, e.g. the superantigen *sag* of the mouse mammary tumour virus (Golovkina et al., 1998).

Retrovirus infection is initiated by binding of the viral envelope glycoprotein to a cell surface receptor (Jonkers and Berns, 1996). For several retroviruses the relevant receptors have been described (Weiss and Tailor, 1995). After entry of the retrovirus into the host cell the viral RNA is copied into double-stranded, blunt-ended, linear DNA molecule, still contained in a core particle. This sequence of priming, RNA hydrolysis and template switching events is catalyzed by the viral reverse transcriptase, resulting in a DNA molecule which is slightly longer than the RNA template and contains identical LTRs. The next step in the infectious process involves import of the viral DNA into the nucleus and insertion into the host genome. This process shows a strong dependence on cell division, presumably requiring the breakdown of the nuclear membrane during mitosis (Lewis and Emerman, 1994). The actual integration reaction is mediated by the viral integrase protein and results in an integrated provirus, which is stably retained in the infected cell and is transmitted upon cell division (Goff, 1992). The integration of the proviral DNA is very precise with respect to the viral nucleotide sequences; however, it appears to be relatively random with respect to cellular DNA sequences, showing a mild preference for actively transcribed regions of chromatin (Mooslehner et al., 1990; Mansour et al., 1993) or for certain specific sites within the host DNA of unknown structure (Shih et al., 1988). By using *in vitro* integration systems it has been shown that target-site selection is highly non-random in naked DNA and is determined by local DNA structures such as methylated CpG dinucleotide repeats, rather than by the sequence or overall structure (Kitamura et al., 1992). Furthermore, it has been shown that nucleosome-associated regions of minichromosomes are efficiently used as integration targets (Pryciak

Received September 9, 2000.

Corresponding author: D. A. Spandidos, Medical School, University of Crete, Heraklion, Crete, Greece. Tel.: +3 (01) 7227809; Fax: +3 (01) 7252922; e-mail: spandido@hol.gr.

Abbreviations: ALV – avian leucosis virus, ATL – adult T-cell leukaemia, BLV – bovine leucosis virus, FeLV – feline leukaemia virus, HIV – human immunodeficiency virus, HTLV – human T-cell lymphotropic virus, LTR – long terminal repeat, MMTV – mouse mammary tumour virus, MuLV - murine leukaemia virus, Sag - superantigen.

et al., 1992). The integrated provirus serves as a template for production of viral RNAs and proteins encoded by their corresponding genes. The ability of a retrovirus to propagate in a specific cell type is mainly determined by the tissue specificity of the enhancer sequences in the LTR and by the viral envelope gene.

Mechanisms of gene activation by insertional mutagenesis

The ability of an integrated provirus to activate the transforming potential of a flanking gene is in all cases mediated by the transcriptional control sequences present in the viral LTR (Athas et al., 1994). Depending on the integration site and the transcriptional orientation of the provirus with respect to the cellular gene, these sequences are capable of initiating, enhancing and/or terminating transcription of host sequences, resulting in high levels of mRNA encoding the complete amino acid sequence of a protein or in production of aberrant transcripts encoding mutant proteins.

Gene activation by the promoter insertion mechanism requires integration of a provirus in the same transcriptional orientation as the target gene (Fig. 1). Transcription initiates from the viral promoter in either the 5' or the 3' LTR, replacing the function of the normal promoter. Activation by promoter insertion is frequently associated with viral deletions when the 3' LTR promoter is used. This often results in deletion of the 5' LTR, suggesting that transcription driven by the 5' promoter and proceeding into the 3' LTR may negatively influence transcription from the 3' LTR (Cullen et al., 1984). When the 5' LTR promoter is used, transcription results in the formation of fusion transcripts containing both viral and cellular sequences, due to frequent read-through at the 3' LTR polyadenylation site and subsequent splicing using the subgenomic mRNA splice donor or cryptic splice donor sites (Herman and Coffin, 1986; Coffin and Moore, 1990; Swain and Coffin, 1993).

Transcriptional enhancement is probably the most frequent mechanism of gene activation by insertional mutagenesis. The high frequency may be explained by the fact that activation by enhancement allows more flexibility with respect to proviral orientation and distance between the provirus and the target gene (Krimpenfort et al., 1988). Furthermore, a much larger DNA domain may be an effective target for proviral insertions. Transcripts

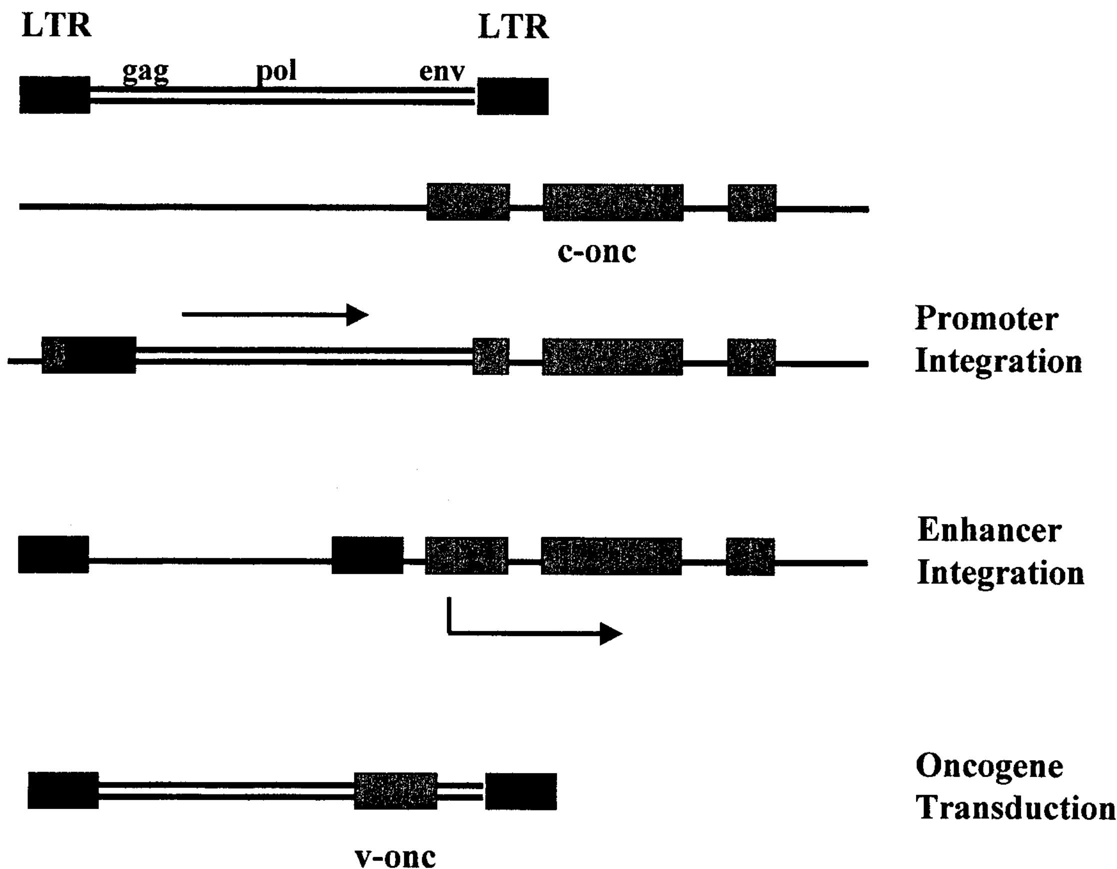

Fig. 1. Modes of oncogene activation by proviral insertion. A cellular oncogene can be activated when the provirus integrates at the 5' end and in the same transcriptional orientation, where the viral promoter drives transcription of the oncogene. When the provirus integrates at the 5' or the 3' end outside the coding sequence, the proviral LTR acts as transcriptional enhancer to the cellular oncogene, inducing its overexpression. Alternatively, a virus can transduce an oncogene by integrating it into the viral genome.