Transcriptional activation of *cKi-ras* proto-oncogene resulting from retroviral promoter insertion

Stephen P.Trusko, Eric K.Hoffman and Donna L.George*

Howard Hughes Medical Institute and Department of Human Genetics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104-6072, USA

Received July 20, 1989; Revised and Accepted October 2, 1989

EMBL accession nos. X15887, X15888

ABSTRACT

Enhanced expression of the *cKi-ras* proto-oncogene in a bone marrow-derived mouse cell line, 416B, has been shown to be associated with the integration of Friend viral DNA into the cellular gene. Here we report the results of experiments designed to clarify the molecular mechanism responsible for the *cKi-ras* overexpression. Based on primer extension analyses and DNA sequencing of *cKi-ras* cDNA clones, we have obtained evidence that the 416B cells contain viral-host chimaeric transcripts that initiate within the 3' long terminal repeat (LTR) of the integrated provirus. Processing of the transcripts from the rearranged *cKi-ras* gene includes an unexpected splicing event associated with the fortuitous creation of a cryptic donor splice site at the junction between the proviral and cellular DNA sequences. These data demonstrate that enhanced *cKi-ras* expression in the 416B cells results from a retroviral promoter insertion mechanism of transcriptional activation.

INTRODUCTION

Slowly oncogenic retroviruses that lack obvious transforming sequences can nonetheless alter the expression and/or structure of host-cell proto-oncogenes by insertional activation mechanisms (1–3). Such activation of cellular genes by integrated proviruses can occur via transcription from strong promoter elements within the viral LTRs and/or through proviral-mediated enhanced transcription from normal or cryptic cellular promoter sequences. The protein-encoding domains of the affected genes may remain intact or may be altered by such integration events.

The Friend murine leukemia virus (Fr-MLV) does not contain an oncogene but may contribute to transformation, at least in part, by insertional activation of particular cellular genes. In the 416B mouse cell line, there is a 25 to 30-fold overexpression of the *cKi-ras* gene that does not result from gene amplification but is associated with the integration of a segment of the Friend proviral DNA into that gene (4–6). However, the exact nature of the mechanism responsible for *cKi-ras* overexpression in these cells remained to be elucidated. In previous studies using RNA blot analyses, we obtained evidence that transcripts from the rearranged *cKi-ras* gene (*rcKi-ras*) do not contain sequences from the 5'-most non-coding exon ϕ (6); thus, it was proposed that transcription may initiate within the viral LTR. However, based on Northern blot analyses, the *cKi-ras* transcripts in the 416B cells are indistinguishable in size from those seen in control cells; novel-sized transcripts, previously noted for some viral-host chimaeric mRNAs, have not been detected (5, 6). Data presented in a previous report did not provide evidence that Fr-MLV LTR sequences are present in *cKi-ras* mRNAs in the 416B cells (5); however, attempts to address this question using RNA blotting methods are complicated by the very high level of Fr-MLV RNAs in these cells.

Here we describe the results of experiments designed to clarify the molecular mechanism responsible for elevated *cKi-ras* expression in the 416B cells. This information could contribute to our understanding of the ways in which the transforming potential of this proto-oncogene can be activated. Previous studies have shown that the *cKi-ras* transforming potential can be activated by the presence of point mutations affecting critical amino acids, and probably by high levels of expression of an otherwise normal gene product (2, 3, 7, 8). Except for such point mutations, no structural alteration affecting the *cKi-ras* product has yet been reported as an activating event for this gene. In this regard, it would be important to determine if proviral insertion has altered the *cKi-ras* protein coding domain.

Based on primer extension studies and DNA sequence analysis of *cKi-ras* cDNAs, we have obtained data demonstrating that transcription initiates within the 3' LTR of the integrated provirus. Processing of the transcripts from the rearranged gene includes an unexpected splicing event apparently associated with the fortuitous creation of a donor splice site at the junction between the proviral and cellular DNA sequences, thus coincidentally generating mRNAs that are similar in size to those from the unrearranged gene.

MATERIALS AND METHODS

Cell Lines

The 416B cell line was isolated from a female BDF₁ (C57BL/6×DBA/2) bone marrow culture that had been infected with Friend virus complex (5, 9). The properties of the Y1 cells, which exhibit amplification and overexpression of a normal *cKi-ras* gene, have been described (7, 10).

RNA Isolation

Total cellular RNA was prepared with guanidine hydrochloride as described (11), and poly A⁺ RNA isolated using Hybond-mAP affinity paper (Amersham) as recommended by the manufacturer.

DNA Sequencing

DNA fragments were subcloned into M13mp18 and M13mp19 vectors or into plasmid pGEM4 (Promega Biotec) and sequenced according to published protocols (12).

Primer Extension Analysis and cDNA Cloning

A synthetic oligonucleotide corresponding to 20 bp located within the first coding exon of *cKi-ras* (5'-ACAAGTTTATACTCAGTCAT-3') was prepared by the DNA Synthesis Center of the Cancer Center of the Hospital of the University of Pennsylvania. Approximately 10 ng (2×10⁶cpm) of the ³²P-labeled primer was annealed to poly A⁺ RNA and primer extension reactions were carried out as previously detailed (13, 14). Samples were ethanol precipitated, recovered, and analyzed on 6% polyacrylamide-8M urea gels. The same synthetic oligonucleotide was used to prime synthesis of cDNAs using 416B poly A⁺ RNAs as a template. First strand synthesis was carried out as described above. Second strand synthesis was performed using published protocols (15). The primer extended cDNAs were blunt-ended with the Klenow fragment of DNA polymerase I and ligated to blunt-ended pGEM4, linearized at the EcoRI site. Recombinant plasmids were transfected into E. coli HB101, and approximately 1–2×10³ ampicillin-resistant colonies were screened by hybridization with the oligonucleotide described above.

Figure 1. Physical map of the 5' end of the mouse *cKi-ras* gene illustrating the position of Fr-MLV DNA integration in the 416B cell line. The region of normal *cKi-ras* transcription initiation upstream of the untranslated exon ϕ is marked by the small arrow and the site of translation initiation within exon 1 is indicated by ATG. The large arrow indicates the site of insertion of Fr-MLV DNA 332 bp upstream of exon 1.

RESULTS

Figure 1 shows the organization of the rearranged *rcKi-ras* gene in the 416B cells and summarizes data presented elsewhere (6, 14). In the normal *cKi-ras* gene, the 5' noncoding exon ϕ is located about 3 kilobase pairs (kb) upstream of the first coding exon (exon 1). Approximately 3 kb of Fr-MLV DNA, encompassing the *env* gene and 3' LTR have integrated within the intron separating exon ϕ and exon 1, in the same 5'–3' transcriptional orientation. The integrated provirus is not a full-length representation of the viral genome: a segment of approximately 5 kb, containing the 5'LTR-*gag-pol* region, has been deleted. The 3' provirus-cell junction is situated 332 bp upstream of exon 1.

The normal *cKi-ras* gene initiates transcription from multiple sites which have been mapped to the 5' end of exon ϕ (see reference 14 and Figure 2 below). Based on RNA blot analysis, the majority of the *cKi-ras* transcripts in the 416B cells, presumably encoded

Figure 2. Primer extension analysis of *cKi-ras* transcripts. Reactions were carried out as described in the text using samples of poly A⁺ RNAs from the 416B cells (2 μ g), Y1 cells (2 μ g) or 3T3 cells (10 μ g), or using yeast tRNA (20 μ g) as a control, and the extension products were analyzed on 6% acrylamide-8M urea gels. Lanes: (1) primer-extended 416B RNA; (2) primer-extended Y1 RNA; (3) primer-extended 3T3 RNA; (4) primer-extended reactions carried out with yeast tRNA; (5) ³²P end-labelled pBR322 MspI fragments as molecular weight markers. The numbers on the right indicate the sizes of some of the marker bands in bases.

Figure 3. Comparison of the nucleotide sequence of the *rcKi-ras* genomic DNA (capital letters) to the sequence of the cDNA clone pGR3 (lower case letters). Putative transcriptional control elements in the viral LTR (CAAT and TATAA boxes) are underlined. The region of transcription initiation mapped by primer extension analysis is designated by an asterisk. The junction between the Fr-MLV DNA and *cKi-ras* intron DNA is indicated by a large arrow, and the normal *cKi-ras* translation start codon (ATG) in exon 1 is boxed.

by the *rcKi-ras* gene, do not include sequences derived from exon ϕ (6). To characterize the transcription initiation site(s) and the nature of the 5' ends of the *rcKi-ras* transcripts, we used two approaches: (1) primer extension analysis; (2) DNA sequencing of *rcKi-ras* cDNAs.

Primer Extension Analysis

For the primer extension experiments, we used a synthetic oligonucleotide primer (5'-ACAAGTTTATACTCAGTCAT-3') corresponding to the first 20 nucleotides of coding material in exon 1 (bases 11–31). When the experiments were carried out using 416B poly A⁺ RNA, one major extension product was seen, having a calculated size of about 177 bp (Figure 2, lane 1). Based on the primer used, the extension products should contain 31 bp of DNA from exon 1 and 146 bp of more 5' material. The predominant extension products are probably homologous to *rcKi-ras* transcripts, given that any transcripts encoded by the normal *cKi-ras* gene in these cells would be expected to be produced at a 25–30 fold lower level. Data supporting this conclusion are presented below. No extension products were seen in control reactions carried out using yeast tRNA (Figure 2, lane 4).

In contrast to the results obtained with the 416B RNA, multiple extension products, ranging in size from 170–220 nucleotides, were seen if the primer extension studies were carried out using control poly A⁺ RNA isolated from Y1 mouse adrenal tumor cells, in which there is a 25 to 30-fold amplification and overexpression of a normal *cKi-ras* gene (Figure 2, lane 2 and reference 14), or if the primer extension studies were carried out using control RNA isolated from mouse 3T3 cells in which there is no amplification or overexpression of the *cKi-ras* gene (Figure 2, lane 3 and reference 14). The reduced intensity of the 3T3 extension products, relative to those in lanes 1 and 2, are consistent with the 25–30 fold lower abundance of *cKi-ras* transcripts in the 3T3 cells versus the 416B or Y1 cells. It should be noted that, after longer exposures of the autoradiogram, minor extension products were seen in the 416B lane, and these have sizes comparable to some of the primer extension products seen in the Y1 lane (data not shown). These minor primer extension products could well be derived from the 25 fold less abundant mRNAs encoded

by the normal *cKi-ras* gene, or by transcripts initiating normally from the *rcKi-ras* gene.

Construction and Sequencing of cDNAs

Transcriptional control elements are present within the LTR of the integrated provirus (16, 17). If transcription from the *rcKi-ras* gene initiates 20–35 bp downstream of the TATAA box (18) within the LTR and reads through the viral sequences as well as the *rcKi-ras* intron sequences to exon 1 then, using the oligonucleotide primer described above, we would have expected to see primer extension products that are 460–480 bp in size, rather than the 177 bp products actually observed. Two possible explanations for this discrepancy are suggested: (1) a promoter insertion mechanism is not utilized in the *rcKi-ras* gene, but rather transcriptional enhancement associated with a cryptic cellular promoter is involved; (2) transcription from the *rcKi-ras* gene does utilize the promoter elements within the viral LTR, but these transcripts are subject to post-transcriptional processing involving one or more splicing events. To distinguish between these possibilities, we further characterized the sequences present at the 5' ends of mRNAs. The same oligonucleotide primer described above, containing sequences complementary to those in exon 1, was used to prime synthesis of partial cDNAs using 416B poly A⁺ RNA as a template. From the primer-extended cDNA mini-library, we isolated two clones that hybridized to a *cKi-ras* exon 1 probe. The inserts in both clones are 159 bp in size and have the same nucleotide sequence. The size of the cDNA inserts (160 bp) is in good agreement with that of the major products (177 bp) generated in the primer-extension analysis using 416B RNA (Figure 2). The cDNA sequence was compared to that of the *rcKi-ras* genomic DNA surrounding the provirus-cell junction (Figure 3). One end of the cDNA (5') matches the sequence within the viral LTR just downstream of the TATAA box. The cDNA sequence is homologous to that of the viral LTR DNA up to the position of the provirus-cell junction, where it then diverges from that of the *cKi-ras* intron sequence. At that point, the cDNA sequence begins an exact match to the DNA contained within *cKi-ras* exon 1 (located 332 bp downstream) up to and including DNA sequences complementary to the primer used for the cDNA construction. The point of nucleotide sequence divergence between the genomic DNA and the primer-extended cDNA that occurs at the provirus-cell junction (CAG/GTA) corresponds reasonably well to a consensus splice donor sequence ${}^{\text{C}}_{\text{A}}\text{AG/GT}^{\text{A}}_{\text{G}}$ (19). These data support the conclusion that most, if not all, of the transcripts from the *rcKi-ras* gene initiate within the viral LTR and are processed using a cryptic splice donor site, which was apparently created during the viral integration event, and the splice acceptor site (at exon 1) that is normally used in the generation of *cKi-ras* mRNAs.

DISCUSSION

The results obtained in these experiments explain a number of features associated with the *cKi-ras* overexpression in the 416B cells: (1) utilization of strong promoter/enhancer elements within the viral LTR is consistent with the 25–30 fold overexpression of the mRNAs; (2) the majority of transcripts in the 416B cells do not include sequences from exon ϕ (6); (3) processing of the transcripts via the splicing event at the provirus-cell DNA junction accounts for the fact that the major products produced in the primer extension reaction (Figure 2) are smaller than might otherwise have been expected (177 bp vs. 480 bp); (4) the number of nucleotides derived from the viral LTR (~146 bp) is very similar to the number normally derived from *cKi-ras* exon ϕ (140–190 bp) thus accounting for the observation that, by Northern blot analysis, the overexpressed *rcKi-ras* hybrid mRNAs in the 416B cells are indistinguishable in size from the transcripts encoded by a normal

cKi-ras gene. The normal *cKi-ras* mRNAs, like those derived from a number of other proto-oncogenes, have fairly long 5' untranslated sequences (20) which are G+C rich (~82% G+C). The possible cellular function of the 5' non-translated exon ϕ and the potential consequences, if any, of replacing it with viral LTR sequences (~51% G+C) remain to be determined.

A number of mechanisms have been documented whereby gene activation results from proviral integration events (1–3). In the *rcKi-ras* gene, the use of a 3' LTR as a transcriptional promoter of downstream *cKi-ras* sequences is associated with a viral integration event that resulted in a deletion of the upstream (5') LTR. Such a situation has been described in a number of other tumors in which enhanced expression of a potential cellular oncogene follows retroviral insertion (1, 21–23). Previous studies have provided a possible explanation for such an association, in that deletion events which encompass the 5' LTR could allow more efficient use of the 3' LTR as a strong promoter element by eliminating interference from the upstream elements (21–24).

Our DNA sequence analysis has shown that nucleotides derived from exon 1 in the *rcKi-ras* mRNAs match those present in normal *cKi-ras* transcripts and do not contain activating point mutations (Figure 3 and reference 6). We have also determined that the normal ATG translation initiation codon within *cKi-ras* exon 1 remains the first ATG in the transcripts from the *rcKi-ras* gene (see Figure 3). We assume, therefore, that the proviral integration event has not disturbed the normal coding region of the gene. This conclusion is supported by the report that immunoprecipitation reactions carried out with a monoclonal antibody against the p21 *ras* proteins detect molecules in the 416B cells that are 25-fold more abundant than those in control mouse cells, but are otherwise indistinguishable in size (5, 7). To date, therefore, specific point mutations or elevated levels of expression of a normal or mutated *cKi-ras* protein remain the only documented mechanisms for activating the transforming potential of this gene. With regard to the latter, two molecular mechanisms have now been described that result in overexpression of the *cKi-ras* product: gene amplification (7, 10) and, as shown here, retroviral promoter insertion. It is likely that the enhanced production of *cKi-ras* protein molecules, resulting from a promoter insertion mechanism of gene activation, directly contributes to the transformed properties of the 416B cells.

ACKNOWLEDGEMENTS

This work was supported in part by NIH grants CA34462 and GM32592. D.L.G. is an Assistant Investigator of the Howard Hughes Medical Institute.

*To whom correspondence should be addressed at University of Pennsylvania School of Medicine, Department of Human Genetics, Clinical Research Building, 422 Curie Boulevard, Philadelphia, PA 19104-6145, USA

REFERENCES

1. Nusse, R. (1986) *Trends Genet.*, **2**, 244–247.
2. Bishop, J.M. (1987) *Science*, **235**, 305–311.
3. Varmus, H. (1984) *Ann. Rev. Genet.*, **18**, 553–612.
4. Scolnick, E.M., Weeks, M.O., Shih, T.Y., Ruscetti, S.K. and Dexter, T.M. (1981) *Mol. Cell. Biol.*, **1**, 66–74.
5. Ellis, R.W., DeFeo, D., Furth, M. and Scolnick, E.M. (1982) *Mol. Cell. Biol.*, **2**, 1339–1345.
6. George, D.L., Glick, B., Trusko, S. and Freeman, N. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 1651–1655.
7. Schwab, M., Alitalo, K., Varmus, H.E., Bishop, J.M. and George, D.L. (1983) *Nature*, **303**, 497–501.

8. Pulciani, S., Santos, E., Long, L.K., Sorrentino, V. and Barbacid, M. (1985) *Mol. Cell. Biol.*, **5**, 2836–2841.
9. Dexter, T.M., Allen, T.D., Scott, D. and Teich, N.M. (1979) *Nature*, **277**, 471–474.
10. George, D.L., Scott, A.F., Trusko, S., Glick, B., Ford, E. and Dorney, D. J. (1985) *EMBO J.*, **4**, 1199–1203.
11. Liu, C., Slate, D.L., Gravel, R. and Ruddle, F.H. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4503–4506.
12. Sanger, F., Nicklin, S., and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463–5467.
13. Friedman, E.Y. and Rosbash, M. (1977) *Nucleic Acids Res.*, **4**, 3455–3471.
14. Hoffman, E.K., Trusko, S.P., Freeman, N. and George, D.L. (1987) *Mol. Cell. Biol.*, **7**, 2592–2596.
15. Gubler, U. and Hoffman, P.J. (1983) *Gene*, **25**, 263–269.
16. Koch, w., Zimmermann, W., Oliff, A. and Friedrich, R. (1984) *J. Virol.*, **49**, 828–840.
17. Clark, S.P., Kaufhold, R., Chan, A. and Mak, T.W. (1985) *Virology*, **144**, 481–484.
18. Breathnach, R. and Chambon, P. (1981) *Ann. Rev. Biochem.*, **50**, 349–383.
19. Mount, S.M. (1982) *Nucleic Acids Res.*, **10**, 459–472.
20. Kozak, M. (1987) *Nucleic Acids Res.*, **15**, 8125–8148.
21. Payne, G.S., Courtneidge, S.A., Crittenden, L.B., Fadly, A.M., Bishop, J.M. and Varmus, H.E. (1981), *Cell*, **23**, 311–322.
22. Neel, B.G., Hayward, W.S., Robinson, H.L., Fang, J. and Astrin, S.M. (1981) *Cell*, **23**, 323–334.
23. Fung, Y.-K.T., Fadly, A.M., Crittenden, L.B. and Wng, H.-J. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 3418–3422.
24. Cullen, B.R., Lomedico, P.T. and Ju, G. (1984) *Nature*, **307**, 241–245.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.