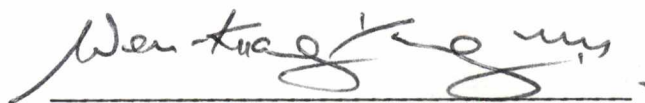


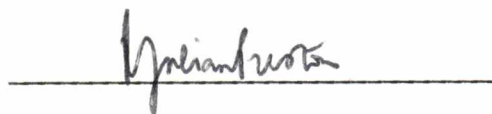
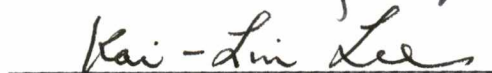
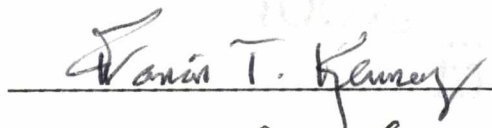
To the Graduate Council:

I am submitting herewith a dissertation written by Kave Noghrei Nikbakht entitled "Characterization of Endogenous Murine Leukemia Virus-Related Sequences From the RFM/Un Strain Mouse." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biomedical Sciences.



Dr. Wen K. Yang, Major Professor

We have read this dissertation
and recommend its acceptance:



Accepted for the Council:



Vice Provost
and Dean of The Graduate School

CHARACTERIZATION OF ENDOGENOUS MURINE LEUKEMIA VIRUS-RELATED SEQUENCES
FROM THE RFM/Un STRAIN MOUSE

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Kave Noghrei Nikbakht

August 1986

DEDICATION

I dedicate this dissertation to my parents for all their support, encouragement, and love they have given me throughout the years.

ACKNOWLEDGMENTS

The author thanks Dr. Wen K. Yang, his principal supervisor, for his advice and encouragement throughout this project. He also thanks Drs. Francis T. Kenney, Julian Preston, Kai-Lin Lee, and Lawrence R. Boone who agreed to be his committee members in preparation of this dissertation and who helped him with their discussions and suggestions. He thanks Dr. Chin-Yih Ou for teaching him the techniques used in recombinant DNA work, Den-Mei Yang for performing the DNA transfection experiments and preparing various monolayer cell line cultures, Kenneth R. Isham for advice in RNA analysis techniques, and Fred E. Myer for helping in recombinant plasmid DNA preparation. His appreciation also goes to other faculty and staff of the ORNL Biology Division, Oak Ridge Graduate School of Biomedical Sciences, and laboratory colleagues who helped him in various ways throughout his graduate studies. He thanks Neva Hair and Willena Carter for typing this dissertation and Nette Crowe for illustrations. Finally, he expresses his deepest thanks to his wife, Patsy Ann, who believed in him and supported him in the pursuit of his education.

Research supported jointly by the National Institute of Environmental Health Sciences (Y01-ES-4-0118), and National Cancer Institute (R01-CA30308 and CA09104) and the Office of Health and Environmental Research, U.S. Department of Energy, under contract DE-AC05-84OR21400 with the Martin Marietta Energy Systems, Inc.

ABSTRACT

The structures of specific regions of the murine leukemia virus-related proviral DNA important in regulation of transcription, replication and host-range type specificity were determined by nucleotide sequence analysis in several endogenous MuLV-related sequences isolated from the RFM/Un mouse. Transcriptional activities and infectious properties of selected molecular clones of these sequences were also examined after transfection into a heterologous host cell line. Restriction endonuclease map analyses of 5 recombinant clones showed that only 5' portion of proviruses were present in pRFM#1 and pRFM#9; that pRFM#16 and pRFM#17 were defective at least in the pol gene and that pRFM#6 harbored a full-length MuLV-related provirus. Analysis of the region which corresponds to the site of tRNA primer binding for initiation of (-) strand DNA synthesis in functional retroviruses demonstrated that 4 out of 5 clones contained binding sites with perfect complementary matches (18/18) with the 3'-terminal nucleotide sequence of a major rat glutamine tRNA isoacceptor. Furthermore, these sites showed dispersed 5 base mismatches with the 3'-terminal 18 nucleotide sequence of proline tRNA which serves as a primer in all ecotropic MuLV strains examined thus far. Structural analysis of the integration site of pRFM#6 provirus showed that each inverted repeat at the terminus of the provirus was linked to a short direct repeat of cellular origin, a structure similar to those of integrated infectious proviruses or transposable gene elements. The 5' and 3' LTRs of pRFM#6 provirus were

identical except for 2 base transitions and 1 nucleotide deletion/addition in the U3 region. When compared with the infectious RFV provirus LTR, most of the transcriptional regulatory elements were identical and occurred at nearly the same distances from the mRNA cap site. While strong homologies were observed in the U5 and R regions, the U3 of pRFM#6 LTR contained a novel ≈ 170 bp sequence segment which was absent in RFV LTR. Sequence analysis of the endonuclease gene, predicted a polypeptide sequence of 390 amino acids for pRFM#6 provirus which was 15 residues shorter than the predicted endonuclease peptide sequence of AKV, an endogenous type C retrovirus of AKR mouse. Two polypeptides showed 86% sequence homology. Both by DNA gel blot analysis and nucleotide sequence analyses, the 5'-env gene of pRFM#6 provirus was found to be similar to that of the mink cell focus forming (MCF)-type murine leukemia virus but significantly different from those of ecotropic and xenotropic MuLVs. The cellular DNA site harboring the pRFM#6 provirus in RFM/Un mice, also harbors the provirus in genomes of BALB/c, AKR, C3H and CBA but not in those of C57BL/6 and NFS/N mice. The human HT1080 fibrosarcoma cells transfected with pRFM#6 and pRFM#16 DNAs, contained low amounts of MuLV-specific transcripts. A similar splicing mechanism was found in human and murine cells as evidenced by the presence of a 24S transcript encoding the env proteins in RNAs from both systems.

TABLE OF CONTENTS

CHAPTER	PAGE
I. MURINE LEUKEMIA VIRUSES	1
General Introduction	1
The Taxonomy and Morphology of Murine Leukemia Viruses	1
Genomic Structure and Replication Cycle of the	
Murine Leukemia Viruses	1
Proteins of MuLVs	8
Host Range of MuLVs	10
Endogenous MuLV and <u>MuLV-Related</u> Proviral Sequences .	11
Objectives	15
II. CHARACTERIZATION OF THE PRIMER BINDING SITES FOR MINUS	
STRAND DNA SYNTHESIS IN ENDOGENOUS <u>MuLV-RELATED</u> PROVIRAL	
DNA SEQUENCES	16
Introduction	16
Materials and Methods	20
Molecular cloning of DNA fragments containing	
the endogenous <u>MuLV-related</u> sequences	20
Preparation of phage DNA	20
Large scale preparation of plasmid DNA by	
hydroxylapatite chromatography	22
Digestion of DNA with restriction enzymes,	
electrophoresis and gel blotting	25
Hybridization of DNA fragments to specific probes . .	25
Preparation of molecular probes	26

CHAPTER	PAGE
5'-End labeling of cloned DNA fragments	27
Restriction endonuclease mapping	27
Nucleotide sequence determination	28
Results	28
Cloning of endogenous <u>MuLV-related</u> proviral DNA	
sequences	28
Restriction endonuclease maps	31
Size heterogeneity of the LTRs in pRFM recombinant	
clones	37
Identification of DNA fragments carrying the tRNA	
primer binding site	40
Nucleotide sequence analyses of the tRNA primer	
binding sites and adjacent regions in pRFM	
recombinant clones	40
Discussion	46
III. CHARACTERIZATION OF A MOLECULAR CLONE OF A FULL-LENGTH	
ENDOGENOUS <u>MuLV-RELATED</u> PROVIRAL DNA	49
Introduction	49
Materials and Methods	51
Molecular cloning and restriction enzyme mapping	
of pRFM#6	51
Preparation of DNA fragments containing the LTRs	
or specific structural genes from pRFM#6	51

Preparation of transformation-competent <u>E. coli</u>	
cells by CaCl_2 treatment	51
Transformation of competent <u>E. coli</u> cells	52
DNA gel blot analysis	53
Extraction of 5'-end labeled fragments from	
acrylamide gels	53
Preparation of molecular probes	54
Results	55
Restriction endonuclease map of pRFM#6	55
Identification of DNA fragments carrying intact	
5' and 3' LTRs	58
Nucleotide sequences of the 5' and 3' LTRs	67
Sequences of viral origin adjacent to 5' and	
3' LTRs	67
Sequences of cellular origin adjacent to 5' and	
3' LTRs (cell-provirus junctions)	72
Comparison of LTR structures in pRFM#6 provirus	
and RFV, the endogenous ecotropic MuLV of the	
RFM/Un mouse	76
Endonuclease gene structure of pRFM#6	79
Characterization of the env gene - predicted	
host-range for pRFM#6 provirus	85
Strain distribution of the pRFM#6 provirus	91
Discussion	94

CHAPTER	PAGE
IV. EXPRESSION AND INFECTIVITY OF THE ENDOGENOUS <u>MuLV-RELATED</u>	
PROVIRAL DNA SEQUENCES	100
Introduction	100
Materials and Methods	102
Cell lines and recombinant clones	102
DNA transfection and clone selection	103
Reverse transcriptase assay	103
DNA extraction from monolayer culture cells	104
DNA gel blot analyses	107
Molecular probes	107
Total RNA extraction from neomycin-resistant cell	
cultures	107
RNA gel blot analysis	109
Results	110
Transient expression of p0.28-9.5 and pRFM#6 plasmids	
in cell lines from heterologous species	110
Stable transfection of human fibrosarcoma HT 1080	
cell line with p0.28-9.5, pRFM#6 and pRFM#16	
recombinant plasmids	112
Reverse transcriptase assay of the culture media	
from HT 1080 transfectants	115
Transcriptional activities of integrated pRFM	
plasmids in HT 1080 transfectants	117

CHAPTER	PAGE
Effect of 5-azacytidine on transcriptional activities of pRFM plasmids: methylation state of pRFM LTRs in heterologous cells	121
Splicing of <u>MuLV-related</u> transcripts in HT 1080 transfectants	126
Discussion	129
V. SUMMARY	133
LIST OF REFERENCES	138
VITA	153

LIST OF TABLES

TABLE	PAGE
1. Approximate sizes of restriction fragments generated upon digestion of pRFM#6 recombinant plasmid DNA with various enzymes	59
2. Reverse transcriptase assay of the culture media from various cell lines transfected with p0.28-9.5 and pRFM#6 plasmids	111
3. Reverse transcriptase assay of the culture media from human HT 1080 stable transfectants	116

LIST OF FIGURES

FIGURE		PAGE
1.	A simplified view of the genomic structure and replication cycle of retroviruses including regeneration of the viral genomic RNA via a double-stranded DNA intermediate. .	3
2.	A proposed model for reverse transcription	17
3.	Hybridization of recombinant phage plaques to [IS]-specific probe	29
4.	Restriction endonuclease maps of four endogenous <u>MuLV-related</u> proviral DNA clones isolated from the RFM/Un mouse DNA . .	32
5.	Demonstration of the presence of two LTRs in pRFM#16 by gel blot analysis (1% agarose).	35
6.	Gel blot analysis of the internal LTR PstI-KpnI fragments in endogenous <u>MuLV-related</u> proviral DNA clones.	38
7.	Gel blot analysis of the BglII restriction fragments from cloned endogenous <u>MuLV-related</u> proviral DNAs of the RFM/Un mouse.	41
8.	Plus-strand nucleotide sequences of the LTR U5, inverted repeat (IR), tRNA primer binding site and the immediate downstream variable (V) and conserved regions in four endogenous <u>MuLV-related</u> proviral DNA clones from the RFM/Un mouse.	43
9.	Restriction endonuclease map of the mouse DNA cloned in pRFM#6	56

FIGURE	PAGE
10. Digestion of pRFM#6 recombinant clone DNA with various restriction enzymes	60
11. Identification of DNA fragments containing the 5'-LTR of the proviral genome in pRFM#6	63
12. Identification of DNA fragments containing the 3'-LTR of the proviral genome in pRFM#6	65
13. Strategy used in the sequencing of the LTRs in pRFM#6	68
14. Nucleotide sequences of the 5' and 3' LTRs and adjacent cellular and viral junctions in pRFM#6.	70
15. The autoradiograms of sequencing gels showing cell-provirus junctions of pRFM#6	74
16. Comparison of the LTR structures and primer binding sites for (-) strand DNA synthesis in pRFM#6 and RFV proviruses .	77
17. Strategy used in sequencing of 1.5 kbp SmaI/SmaI restriction fragment of pRFM#6 (Figure 9, map positions 5.1 to 6.6 kbp) that includes the 3' portion of the pol gene (the endonuclease) and the 5' portion of the env gene . . .	80
18. Comparison of nucleotide and predicted amino acid sequences of the corresponding 3' portion of pol genes in pRFM#6 and AKV	82
19. Hybridization of a MCF-specific probe, to the env gene of pRFM#6	87

FIGURE	PAGE
20. Comparison of nucleotide and predicted amino acid sequences of the corresponding 5' portion of env genes in pRFM#6, AKR MCF247 MuLV and NFS xenotropic MuLV	89
21. Detection of the endogenous <u>MuLV-related</u> proviral genome of pRFM#6 in various inbred mice	92
22. Gel blot demonstrating the absence of pRFM#6 LTR in the genomic DNAs of C57BL/6 and NFS/N mice.	95
23. Gel blot analysis of genomic DNAs isolated from the human HT 1080 transfectants.	113
24. RNA blot analysis of three human HT 1080 transfectants and producer NIH3T3 cells (transfected with p0.28-9.5 ecotropic provirus)	118
25. Effect of 5-azacytidine on transcriptional activities of the integrated pRFM plasmids in human HT 1080 transfectants	122
26. The methylation state at SmaI site in the LTRs of integrated pRFM#16 provirus in HT 1080 transfectants	124
27. Detection of a MuLV-specific 24S transcript in human HT 1080 transfectants	127

CHAPTER I

MURINE LEUKEMIA VIRUSES

GENERAL INTRODUCTION

The Taxonomy and Morphology of Murine Leukemia Viruses

Murine leukemia viruses (MuLVs) belong to the subfamily Oncovirinae which together with two other subfamilies, Lentivirinae and Spumavirinae constitute the family Retroviridae (1). The members of the Retroviridae family include all viruses containing a RNA genome and a reverse transcriptase (RNA-dependent DNA polymerase) activity (2). The formation of a double-stranded DNA intermediate during the replication cycle is a unique characteristic of the Retroviridae family. The Oncovirinae subfamily encompasses all oncogenic as well as a large number of nononcogenic but closely related retroviruses.

With respect to their morphology as revealed in the electron microscope, MuLVs and the majority of other retroviruses are classified as C-type particles (3,4,5). These particles are roughly spherical in shape and have a diameter of approximately 100 nm. Each mature viral particle contains a spherical core or nucleoid enclosed in an outer envelope. The outer envelope, usually containing projecting spikes, is assembled on the cell membrane as the virion is budded and released into the extracellular space. The key features of C-type particles are the absence of any discernible intracytoplasmic particle forms before the budding has started, and a centrally located core in mature virions.

Genomic Structure and Replication Cycle of the Murine Leukemia Viruses

Among the mammalian leukemia viruses, the genomic structure of murine

leukemia viruses (in particular Moloney MuLV) has been studied most extensively. As for other retroviruses, the genome in the native state is a complex of two identical molecules of single-stranded RNA joined at or near their 5' termini (6,7,8,9). This diploid complex has a sedimentation coefficient of 70S and can be dissociated into 35S subunits by gentle heating or treatment with denaturing agents. The significance of the diploid genome, which is unique among the genomes of known animal viruses is not well understood but may play an important role in the complex mechanism of the viral replication. Shown in Figure 1 is the genomic structure and replication cycle of a haploid subunit of the MuLV using the genome of Moloney leukemia virus as a representative (10,11, 12). The 8.3 kilobase RNA genome of MoMuLV contains three coding regions as well as noncoding sequences with more or less defined functions. The coding regions include the gag (group specific antigens) gene (≈ 1.7 Kb) which is the most 5'-proximal structural gene and codes for internal virion capsid proteins; the pol (polymerase) gene (≈ 3.6 Kb) which encodes the virion reverse transcriptase and other enzymatic activities (RNase H, protease, endonuclease) involved in viral replication; the env (envelope) gene (≈ 2.0 Kb) which encodes proteins involved in the recognition of cellular receptors and probably virion penetration into the host cell. The noncoding sequences are present as discrete units at both ends of the genomic RNA. The R region is a short sequence (68 nucleotides) repeated directly at both ends of the genome. These direct repeats presumably play an important role in transfer of the newly synthesized DNA chain during the process of reverse transcription

Figure 1: A simplified view of the genomic structure and replication cycle of retroviruses including regeneration of the viral genomic RNA via a double-stranded DNA intermediate. The 8.3 kb RNA genome of Moloney MuLV is shown here as a prototype for other retroviruses.

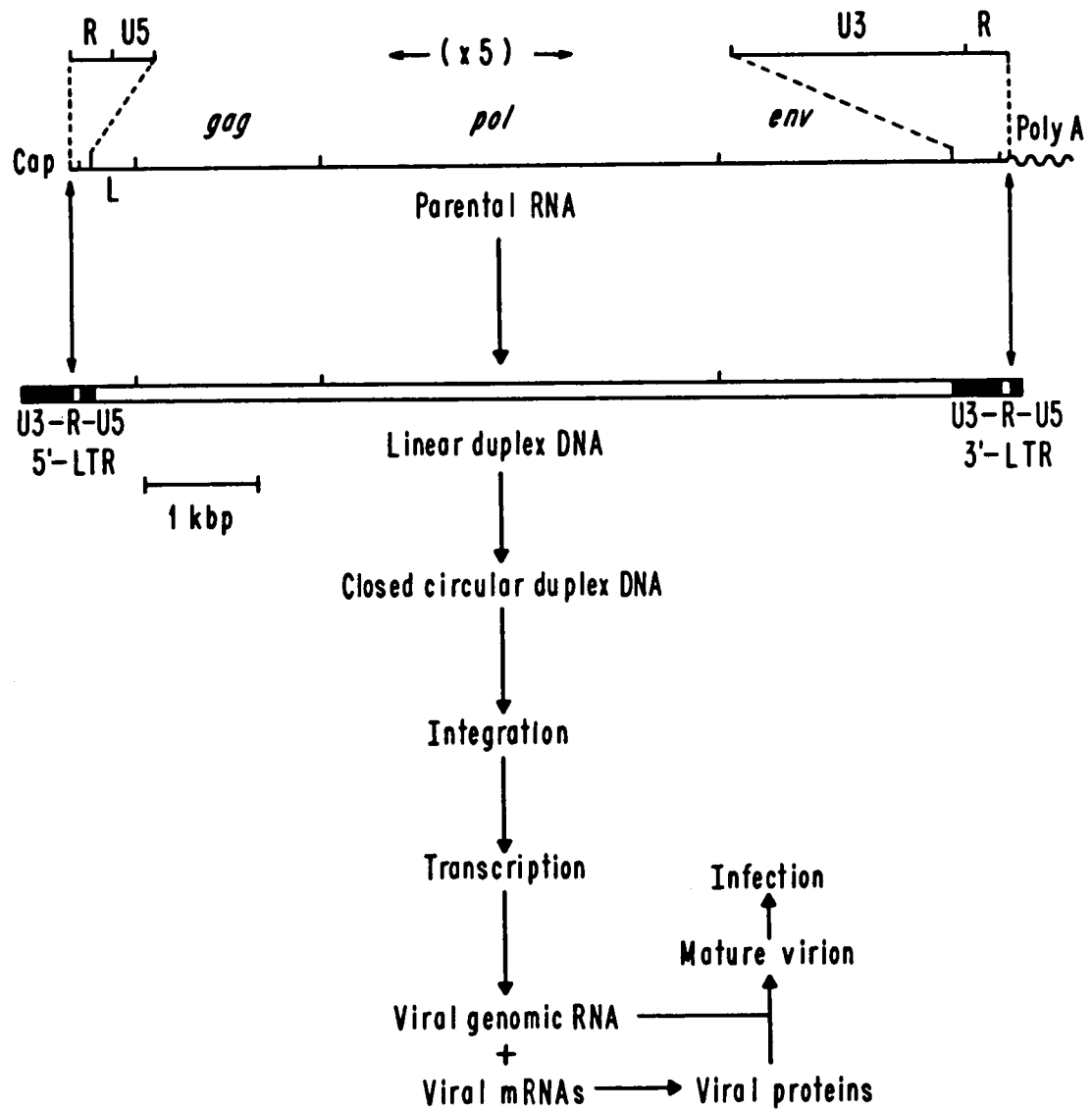


Figure 1

(13). The U5 region is a unique sequence of 77 nucleotides present at the 5' terminus of the genome. The unique sequence present at the 3' terminus (U3) is 449 nucleotides long and contains most of the transcriptional regulatory signals. In addition to these discrete terminal units of noncoding sequences, an untranslated region (L) of ≈ 0.48 Kb is also present between the U5 and initiation codon for the gag gene. Two functions have been attributed to this region; it contains the putative 5' splice donor site for env gene mRNA (12,14,15); and it provides a signal for packaging of the viral RNA during replication (16).

The genomic RNA of murine leukemia viruses shows post-transcriptional modifications similar to those of eukaryotic mRNA molecules. These modifications include a poly(A) chain of approximately 200 residues at the 3' terminus (17,18) and a cap with the structure $m^7G^{5'}ppp^{5'}Gm$ at the 5' end (19,20). The genomic RNA also contains internal methyl groups, a feature shared by eukaryotic RNAs (21). The physiological importance of these chemical modifications in MuLVs and other retroviruses is not well understood. Studies in an avian retrovirus (RSV) indicate that neither the intact cap structure nor the internal methylations are absolutely necessary for viral replication (22). The capping group however, might be important for correct binding of viral mRNA to the ribosomes for initiation of translation.

In addition to the 70S diploid genomic RNA complex, virions of murine leukemia and other retroviruses contain RNA species of host cell origin which may comprise as much as 50% by weight of the total viral RNA (23,24,25,26). These include the 18S and 28S ribosomal RNAs as well

as smaller species of 4S, 5S and 7S RNAs. The 4S RNA is present in a large copy number (>100 copies/virion) and is composed mainly of host tRNA.

During the initial step of replication, which takes place in the cytoplasm of permissive cells, a linear duplex DNA molecule is synthesized from the 35S genomic RNA template by the viral encoded reverse transcriptase (27). In most MuLVs this DNA synthesis is primed by a proline tRNA molecule (28,29) which is of cellular origin and is tightly bound to the genomic RNA near its 5' end, as in the case of tryptophan tRNA in avian sarcoma virus genome (30). Except for the tRNA primer, it is not known that any of the other host-specific RNAs or the small amount of host DNA present in the virion play any significant role in the virus replication cycle. After synthesis of the first DNA strand, reverse transcriptase, which also possesses RNase H activity, processively degrades the genomic RNA of the RNA-DNA hybrid (31,32,33). After migration to the nucleus the linear DNA duplex is converted to a covalently closed circular DNA and subsequently integrates into the host genome. The integration process probably takes place randomly, but the number of available sites might be limited.

The integrated viral DNA known as a "provirus" (27) is flanked by short direct repeats of cellular origin which are generated by some unknown mechanism during integration (34,35). The provirus has a structure almost identical to that of the linear duplex DNA (36). It is, however, longer than the 35S genomic RNA due to the addition of a U3 at the 5' and a U5 at the 3' terminus (37). These additional sequences

are generated during reverse transcription and are part of large direct terminal repeats (LTRs) present at both ends of the provirus. Each LTR therefore, consists of two unique (U) and one repeat (R) discrete regions in the order of 5'-U3-R-U5-3'. LTRs contain all the eukaryotic transcriptional regulatory elements including the Goldberg-Hogness "TATA" box and the "CCAAT" sequence which regulate the initiation of transcription, an mRNA polyadenylation signal and sequences important in the termination of transcription (38,39,40,41,42). Additionally, LTRs contain inverted repeat sequences of 11 nucleotides at their edges which are essential for the integration of the circularized proviral DNA (43,44,45).

Not only can LTRs regulate the expression of viral genes during the replication cycle but under certain circumstances they can also activate the expression of a cellular gene leading to tumorigenesis. In some tumors induced by avian leukosis virus (ALV), integration of provirus immediately upstream of the c-myc gene has been shown (46,47). Studies on tumor induction by ALV carried out by other investigators have shown that the provirus was present upstream, downstream, or in opposite orientation in relation to the c-myc gene (48,49).

The provirus serves as a template for the synthesis of viral mRNA as well as progeny virion RNA. Biosynthesis of viral RNA transcripts is catalyzed by a normal cellular enzyme, most likely RNA polymerase II (50,51). After 2 identical 35S mRNA molecules containing caps and poly(A) tails are packaged by viral encoded proteins, the mature virion will be released by budding into the extracellular space, completing the

life cycle. Excreted virions can be transmitted horizontally to other permissive cells and thus starting a new round of replication.

Proteins of MuLVs

Proteins comprise about 60% by weight of the overall chemical composition of MuLVs. The internal structural proteins are mainly the products of the gag gene. This gene encodes a precursor polyprotein, Pr65^{gag} which upon proteolytic cleavage forms individual gag proteins (52,53). Pr65^{gag} is the translational product of the genomic size 35S mRNA (54). The gag polypeptides include p30 ($\approx 30,000$ daltons) which is the major antigenic determinant for group-specific antigenicity (55,56) and is the main constituent of the core shell; p15 ($\approx 15,000$ daltons) which is a hydrophobic protein (57); ppl2 ($\approx 12,000$ daltons), a phosphoprotein (58) and p10 ($\approx 10,000$ daltons) which is a basic protein (59). p15 and most of ppl2 are constituents of the viral inner membrane (between the core shell and envelope) whereas p10 and some ppl2 are located within the viral core and are bound to RNA (60). The physical order of gag gene products is NH₂-p15-ppl2-p30-p10-COOH.

Reverse transcriptase (80,000 daltons) a protein found within the viral core structure (≈ 40 copies/virion) in a ribonucleoprotein complex, is the cleavage product of a precursor protein, Pr180^{gag-pol} (61,62). Like Pr65^{gag}, this precursor molecule is also synthesized from a virus specific genomic size mRNA (63) and contains most if not all of the Pr65^{gag} peptide sequences. Whether genomic size mRNAs coding for Pr180^{gag-pol} and Pr65^{gag} are identical or differ by a small splice(s) is not clear. Genetic analyses using mutant MuLVs have shown that the

3' portion of the pol gene encodes a polypeptide with an endonuclease activity which is presumably required for integration of the proviral DNA (64,65). Western blot analysis using an endonuclease-specific antibody indicates a size of ≈ 46 Kd for this polypeptide (Rebecca Furkas, personal communication).

The primary encoded product of the env gene is a polyprotein known as Pr90^{env} or Pr80^{env} (depending on MuLV strain) (66). This polyprotein is translated from a subgenomic 24S mRNA (54,67) which contains nucleotides from the 5' end of the genomic 35S RNA spliced to the 5' end of the env gene coding region. Processing of the polypeptide precursor molecule involves both carbohydrate addition and proteolytic cleavage. As a result, two major products are generated (68,69), a glycoprotein (70,000 daltons) known as gp70 and a nonglycosylated species (15,000 daltons) referred to as p15[E]. In addition to these two major species, a small quantity of another protein of 12,000 daltons (p12[E]) is also present in the virion envelope. This protein is presumably a cleavage product of p15[E] (70). The map order of envelope proteins within the env gene is NH₂-gp70-p15[E]-COOH. Since gp70 is the major glycoprotein component on the surface of virion particles, it is the main determinant against which antibodies are elicited after viral infection of the host animal. The antigenicity of this glycoprotein is determined primarily by the polypeptide core structure with the carbohydrate moiety being of less significance (71).

Host Range of MuLVs

One of the most important biological functions of gp70 is its involvement in the adsorption of virion particles to the cell surface. This adsorption and subsequent penetration of the particles into the host cell involves recognition of the cell surface receptors by the virion gp70 molecules (72). Based on this receptor recognition, several host-range groups can be defined: a) ecotropic MuLVs, which can infect cells only from murine origin; b) xenotropic MuLVs, which infect cells from nonmurine origin, c) amphotropic MuLVs which have the host-range of both eco- and xenotropic viruses. A fourth group of MuLVs known as dualtropic or MCF (mink cell focus-forming) viruses have a host-range more or less similar to that of amphotropic viruses but can be distinguished from the latter based on their neutralization and interference patterns (73). MCF viruses are recombinants between ecotropic MuLVs and some unidentified endogenous nonecotropic env sequences of the host cell (74,75,76,77). MCF viruses are generally more pathogenic than their parental ecotropic MuLVs and some can accelerate the development of naturally occurring leukemia in high-incidence strains of mice such as AKR (73,78).

Based on their ability to propagate in different inbred mice, ecotropic MuLVs can be divided further into three classes: N-tropic, B-tropic and NB-tropic (79). The host gene which governs the susceptibility of various mouse cells to these three classes is located at the Fv-1 locus on chromosome 4 of the mouse, adjacent to the Gpd gene (80). Fv-1^{b/b} cells are mostly susceptible to infection by B-tropic whereas

Fv-1^{n/n} cells are primarily infected by N-tropic viruses. NB-tropic viruses on the other hand grow efficiently in cells of both genotypes. The Fv-1 restriction affects a postpenetration stage, most likely at the level of proviral integration (81,82,83). It is believed that the target for the host Fv-1 restriction is located within the p30 region of the gag gene (84,85,86). Nucleotide sequence analysis of this region (87) revealed two amino acid changes in residues 109 and 110 of p30 leading to significant conformational alteration of this protein in BALB/c N-tropic and B-tropic murine leukemia viruses.

Endogenous MuLV and MuLV-Related Proviral Sequences

The genomes of all inbred strains of Mus musculus contain cellular DNA sequences homologous to murine leukemia viruses. There are 20-50 copies of these endogenous sequences per haploid genome of the laboratory mouse (88). These sequences apparently replicate in concert with the cellular genome and are vertically transmitted from generation to generation by classical Mendelian inheritance. Accumulating evidence suggests that these sequences were acquired by germ line infections of the mouse (89,90). This acquisition has probably taken place relatively recently during subspeciation and inbreeding as various lines of inbred mice contain these sequences at different loci (91,92). Despite the complex distribution pattern of the endogenous MuLV-related sequences, independent studies by Steffen et al. in NIH/Swiss mice (91) and Dolberg et al., in several other inbred strains (e.g. BALB/c, NZB, DBA) (92) have shown that, based on their restriction endonuclease maps, the majority of these sequences fall into two organizational classes, each

repeated several times. They also showed that both classes are similar to standard MuLV proviral DNA in size as well as sequence organization.

Some of the endogenous MuLV proviruses are spontaneously expressed giving rise to infectious ecotropic viruses. The level of spontaneous expression of these viruses is directly correlated with the incidence of leukemia in various inbred strains. For instance, high-virus strains such as C58, AKR and C3H/Fg produce the ecotropic virus early in life and have a high incidence of leukemia, whereas low virus strains such as BALB/c, C3H/He and C57BL/6 produce virus only late in life and exhibit a low incidence of leukemia (93). Some strains (e.g. C57L) are virus negative and do not produce any infectious virus spontaneously throughout their life. Spontaneous expression of xenotropic MuLVs has also been shown in some inbred strains. The NZB mouse is the prototype for high-virus strains whereas the BALB/c mouse exhibits only a low level of spontaneous expression in vivo (93).

In addition to spontaneous expression, both endogenous ecotropic and xenotropic proviruses can be induced from cultured cells of the virus positive strains. A wide variety of agents such as halogenated pyrimidines (IUdR or BUdR), inhibitors of protein synthesis (e.g. cycloheximide, puromycin), 5-azacytidine and irradiation can be used for induction (94,95,96,97). Studies by Aaronson and Dunn (98) on BALB/c mouse cells in culture have shown that protein synthesis inhibitors induce only the xenotropic virus whereas halogenated pyrimidines induce both eco- and xenotropic MuLVs. High-virus strains contain a larger number of inducible loci than low-virus or virus-negative strains. The

AKR mouse carries at least two unlinked inducible loci for ecotropic virus production, Akv-1 and Akv-2 mapped to chromosomes 7 and 16, respectively (99,100,101). Two other high-virus strains, C3H/Fg and C58 carry 3 to 4 inducible loci for ecotropic virus. Like Akv-1, the Fgv-1 locus of C3H/Fg mice has been mapped to chromosome 7 but occurs at a different chromosomal location (101). Several low virus strains derived from the Bagg albino, including BALB/c, C3H/He and A, carry a single inducible locus Cv mapped to chromosome 5 (101).

In addition to spontaneous or induced expression of complete infectious viruses, various inbred mice carry a number of genetic loci which spontaneously express only some viral gene products but not others. The best example of this partial virus expression is the gp70-associated G_{IX} antigen, which was first described on thymocytes (102). Further immunofluorescent studies have also demonstrated a high degree of localization of this antigen to epithelial cells of mouse epididymis (103). Another antigen, XenCSA, which is displayed on the surface of spleen cells and thymocytes of the mouse (104) is related to the envelope glycoproteins of xenotropic viruses. Although at the moment it is difficult to infer a major role for these virus-related proteins, their striking tissue specific expression might suggest a direct role for them in normal developmental processes.

The majority of DNA sequences found in the mouse genome homologous to exogenous MuLVs are related to MCF-type viruses (105). This class of endogenous sequences might be present in more than 30 copies per haploid genome in some strains. MCF-related sequences are found in all inbred

strains of mice which are either virus inducible or noninducible. Despite the large copy number, there is no evidence that these sequences can produce viral particles even after induction of the cell cultures. This may indicate that MCF-related sequences are either defective or are under some kind of negative control which renders them inactive.

Until recently, structural characterization of mouse DNA sequences related to MuLVs, involved only those loci with the potential of producing infectious viral particles. Using molecular cloning technique, Ou et al., have isolated and characterized an endogenous MuLV-related DNA sequence (designated AL10) from the BALB/c mouse (106). This clone exhibited several distinct structural features when compared with infectious MuLVs. For instance, the U3 region of the LTR contained a novel ≈ 170 bp sequence segment which made this LTR correspondingly larger than that of infectious MuLVs. Furthermore, nucleotide sequence of the region homologous to the site for tRNA primer binding showed 17/18 complementary match with that of a major rat glutamine tRNA isoacceptor. The same region, showed dispersed 6 base differences when matched with proline tRNA, a primer found in all infectious MuLVs. This isolate as well as several others studied subsequently by Khan et al. (from AKR mice) (107,108) are representatives of a large number of endogenous MuLV-related sequences of the mouse genome which are presumably distinct from infectious ecotropic and xenotropic murine leukemia viruses.

Objectives

The long term goal of this study is to understand the mechanism of interaction of the endogenous murine leukemia virus related sequences with the host animal. The large copy number of these sequences present naturally in all inbred strains of mice might point out the relative capability of these elements to establish themselves in the mouse germ-line. In this research project, studies have been confined to one mouse strain, that is, RFM/Un. Interest in this strain stems from the high incidence of thymic lymphoma and myeloid leukemia developed after whole-body irradiation of the mouse. Chemical treatment of the mouse cell cultures with 5-iododeoxyuridine induces a biologically homogenous N-tropic ecotropic virus population, most likely from a single locus. Furthermore, this strain of mouse spontaneously expresses a low level of an endogenous type C retrovirus in many tissues during its life span. Throughout this project, recombinant DNA clones of endogenous murine leukemia proviruses have been used to define the structure and biological properties of individual proviral genomes. The use of recombinant clones will eliminate many of the complications which could arise in trying to study a single copy amidst a large family of related sequences. In the following chapters, the results of studies on the structure of specific regions of the proviral genomes as well as the biological properties of some selected clones will be presented.

CHAPTER II

CHARACTERIZATION OF THE PRIMER BINDING SITES FOR MINUS STRAND DNA SYNTHESIS IN ENDOGENOUS MuLV-RELATED PROVIRAL DNA SEQUENCES

INTRODUCTION

An early step in the replication cycle of retroviruses involves conversion of the plus sense viral genomic RNA into DNA by RNA-dependent DNA polymerase (reverse transcriptase). This DNA synthesis starts at the 3'-OH end of a tRNA primer molecule which is bound near the 5' end of the genomic RNA (30). The primer tRNA molecule is hydrogen-bonded to the viral genomic RNA through its 3' terminal 18 nucleotides (109). With the 3'-A-OH of the primer molecule forming a phosphodiester bond with the first deoxynucleotide, the synthesis of the first DNA strand (minus sense DNA) proceeds until it reaches the 5' terminal nucleotide of the genomic RNA (13) (Figure 2). The newly synthesized DNA strand referred to as "minus strand strong stop" is a short sequence spanning only the U5 and R regions. This DNA with its covalently linked tRNA primer molecule will then dissociate from the genomic RNA and hybridize to terminally repeated sequences at the 3' end of the viral genomic RNA. Further elongation of minus strand strong stop will generate a full-length (genomic size) DNA strand which is hydrogen-bonded to the genomic RNA throughout most of its length. Following ribonuclease H action on the genomic RNA, the synthesis of plus strand DNA is initiated from RNA primers along the genome length in a discontinuous manner and eventually linked by the ligase action. The DNA synthesis primed by a short

Figure 2: A proposed model for reverse transcription (13). The sequences in the viral RNA molecule (plus-strand sequences) are shown with uppercase letters whereas minus-strand sequences are denoted by lowercase letters. Reverse transcription is primed by a tRNA molecule which is H-bonded to the viral genomic RNA at the primer binding site (P) through its 3'-terminal 18-20 nucleotides. The minus strand strong stop is a short sequence spanning only U5 and R regions.

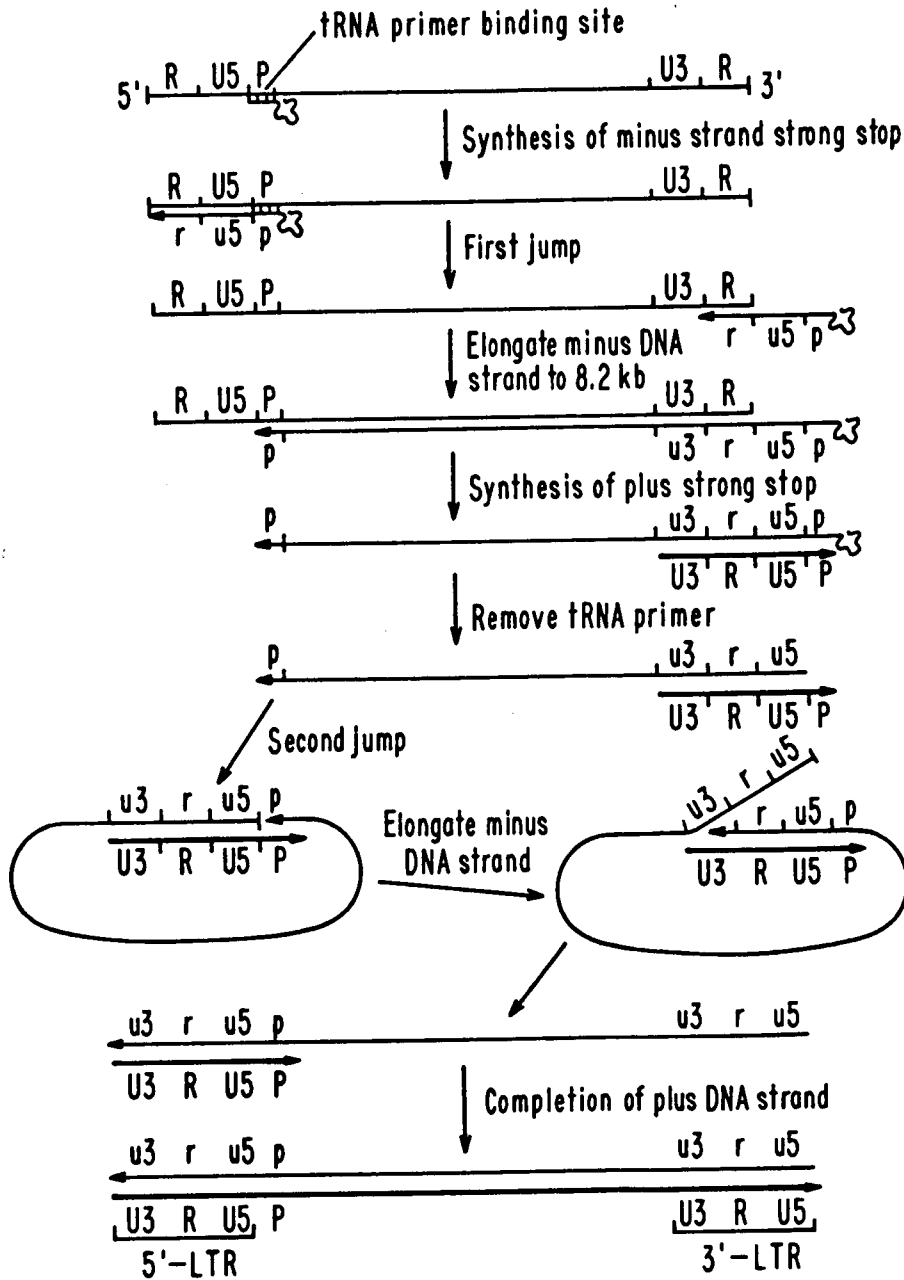


Figure 2

(12-14 nucleotides) purine-rich RNA fragment located immediately 5' to the U3 region of viral genomic RNA (110,111), leads to the formation of "plus strand strong stop" DNA that contains a complementary copy (P) of the 18 nucleotides at the 3' end of the tRNA primer. After removal of tRNA primer by the RNase H action, "plus strand strong stop" DNA forms partially duplex DNA structure with the minus strand DNA as a result of base pairing between the regions designated as p and P in Figure 2. The synthesis of minus strand DNA is extended by copying the "plus strand strong stop" DNA, thus eventually generating a completely double-stranded DNA with two long terminal direct repeats (LTRs) at both ends. In a proviral DNA, the 18-nucleotide DNA segment homologous to the primer binding site of viral genomic RNA is located on the plus strand on the immediate 3' side of the 5' LTR (P in Figure 2) (34,112). Several different tRNA primer molecules have been identified among retroviruses. These include proline tRNA in ecotropic murine leukemia, feline leukemia, and avian reticuloendotheliosis viruses (28,29,113, 114), lysine tRNA in mouse mammary tumor viruses (115,116), and tryptophan tRNA in avian sarcoma and avian myeloblastosis viruses (117,118). In a previous study by Ou et al. (106) on a cloned endogenous MuLV-related proviral DNA from the BALB/c mouse, it was found that the tRNA primer binding site was complementary (17/18 match) to that of a major glutamine tRNA molecule (119). Since this structure was clearly different from the previously reported primer site for proline tRNA, I have extended these studies to several other cloned endogenous MuLV-related proviral DNAs from the RFM/Un mouse in an effort to elucidate the structure of this region in endogenous sequences.

MATERIALS AND METHODS

Molecular cloning of DNA fragments containing the endogenous MuLV-related sequences

Initial molecular cloning experiments were carried out by P. L. Glover and Dr. L. R. Boone (then of this laboratory) as briefly described in the following. HindIII restricted DNA fragments from spleen of a RFM/Un mouse suffering from radiation-induced myeloid leukemia were used to prepare a recombinant genomic library. The phage plaque hybridization procedure of Benton and Davis (120) was used to screen the library with molecular probe of LTR sequences of an ecotropic MuLV (121). Phage plaques showing positive hybridization with the LTR sequence probe were isolated by two more rounds of plaque purification. The inserted RFM mouse DNA sequences in individual phage clones, after further examination with other structural gene sequence probes of MuLV, were subsequently transferred into the HindIII site of the plasmid pBR322. The pBR322 recombinant plasmids harboring individual LTR-containing RFM mouse DNA sequences were designated as pRFM#1, pRFM#2 ... etc.

Preparation of phage DNA

Phage amplification and DNA extraction was carried out according to the following protocol:

1. 100 λ of a phage suspension containing 10^5 - 10^7 pfu were preadsorbed to 150 λ of overnight-grown DP50supF E. coli cells at 37°C for 15 min.
2. After addition of 80 ml NZC broth [supplemented with thymidine (.004%), diaminopimelic acid (.01%) and MgSO₄ (10 mM)], the

suspension was incubated overnight at 35°C with continuous shaking (250 rpm).

3. Lysis of E. coli cells became apparent as a clear lysate with suspended cell debris was obtained after overnight incubation. At this point, 4 drops of chloroform were added to the lysate and shaken for an additional 15' (at 35°C) to ensure that all cells were lysed.
4. Cell debris was removed by centrifugation at 12,000 rpm (5 min) using a Sorvall rotor. The supernatant was transferred into a 125 ml flask.
5. To the supernatant, 10 λ of DNase I (10 mg/ml stock) and 10 λ of RNase A (10 mg/ml stock) were added and incubated at 37°C for 20 min with shaking.
6. The suspension was brought to 6% NaCl, held at 4°C for 1-2 hrs and then centrifuged at 8,000 rpm (30 min) using a Sorvall rotor.
7. The supernatant was adjusted to 7% PEG 6000, held at 4°C for 2-4 hrs and then spun at 8,000 rpm (30 min) as before.
8. The pellet was suspended in 200-600 λ TE (10 mM Tris-HCl pH 7.6, 1 mM EDTA) and extracted once with chloroform, twice with phenol and once with chloroform/isoamylalcohol (50/1).
9. To the aqueous phase, 1/10 volume of 2 M potassium acetate and 2 volumes of cold 95% ethanol were added, held at -70°C for a few minutes and then spun in a microfuge for 10 min. The pellet was rinsed 2 or 3 times with 80% ethanol, lyophilized briefly and finally suspended in 100-200 λ TNE.

10. Construction of pRFM#16 from recombinant phage DNA was carried out as follows. ≈ 10 μ g of the phage DNA were digested to completion with HindIII restriction enzyme and subsequently electrophoresed in 0.7% agarose gel. After staining with EtBr, the DNA insert band was purified and ligated to the HindIII site of the pBR322 vector. This recombinant plasmid (pRFM#16) was used for transfection of competent E. coli (C 600 strain) cells, followed by large scale preparation of the plasmid DNA. Preparation of competent E. coli cells and details of DNA transfection experiments will be described in the following chapter.

Large scale preparation of plasmid DNA by hydroxylapatite chromatography

Large quantities of recombinant plasmid DNAs were prepared according to the following protocol:

1. 40 ml of L-broth containing 0.1 ml of ampicillin (20 mg/ml stock) were inoculated with a single colony of positive transformant and incubated overnight in a 125-ml Erlenmeyer flask at 37°C.
2. 50 ml of L-broth containing ampicillin were inoculated with 200 λ of fresh overnight 40 ml culture and incubated at 37°C until OD₅₆₀ ≈ 0.4 (usually took 5-6 hrs).
3. 200 ml of L-broth containing ampicillin were inoculated with 10 ml of 0.4 OD₅₆₀ culture and incubated in a 2 L Erlenmeyer flask for 2.6 hr (or until OD₅₆₀ ≈ 0.4) at 37°C. The culture was then amplified overnight with 1.2 ml chloramphenicol (34 mg/ml stock).
4. The cells were sedimented at 6000 rpm for 5 minutes in a 250 ml centrifuge bottle and the pellet was washed once in 50 ml of cold

PBS (136 mM NaCl, 2 mM KCl, 0.9 mM CaCl₂, 0.5 mM MgCl₂, 10 mM Na₂HPO₄, 1.5 mM KH₂PO₄).

5. After final sedimentation (6000 rpm, 5 minutes), the pellet was resuspended in 4 ml of cold 15% sucrose, 50 mM EDTA, 50 mM Tris (pH 7.8) solution and transferred to a 30 ml Oak Ridge tube.
6. 2-4 mg of lysozyme crystals (Sigma) were added to the suspension and after mixing was allowed to stand on ice for 5-10 minutes. The cells were then frozen at -80°C and quickly thawed at 37°C water bath.
7. An equal volume of cold 0.5% Triton X 100, 50 mM EDTA, 50 mM Tris (pH 7.8) was added and mixed gently on ice for 10 minutes.
8. The cell lysate was centrifuged at 24,000 rpm (4°C) for 90 minutes using Beckman type 30 rotor. The supernatant was carefully drawn off and transferred into another 30 ml Oak Ridge tube.
9. 1/10 volume of 10% SDS was added and the mixture was heated at 60°C water bath for 5 minutes. After incubation, 1/3 volume of 5 M NaCl was added and mixed gently on ice by inversion and then allowed to stand overnight on ice.
10. The suspension was centrifuged at 25,000 rpm for 30 minutes and the supernatant was transferred to a 30 ml Oak Ridge tube.
11. To the supernatant, solid PEG 6000 was added to 10% (w/v) concentration and mixed by vortexing until all PEG was in solution. The mixture was allowed to stand on ice overnight or greater than 4 hrs and then centrifuged at 20,000 rpm for 20 minutes. The supernatant was carefully decanted and the pellet was dissolved in 2-3 ml of

7 M urea, 0.2 M Na-PO₄ (pH 6.8), 0.7% SDS solution. Heating at 45°C greatly facilitated solubilization of the pellet.

12. For hydroxylapatite chromatography, 0.5 g of hydroxylapatite [DNA grade Bio-Gel HTP (Bio-Rad Laboratories)] were mixed with 50 ml H₂O in a graduated cylinder and allowed to stand for 30 minutes. The excess water was poured off and hydroxylapatite (settled to the bottom of cylinder) was transferred to a 10 ml syringe containing 2 layers of filter cloth (all purpose miracloth) fitted tightly above the needle hub. The suspension was allowed to settle down for 30 minutes after which the excess water was drained off by removing the rubber stopper seal from the needle.
13. The column was equilibrated with 30 ml of 7 M urea, 0.2 M NaPO₄ (pH 6.8), 0.7% SDS solution and sample was then loaded into the column allowing to bind for 30 minutes before the stopper seal was removed.
14. After draining off thoroughly, the column was washed with 50 ml or more of 7 M urea, 0.2 M Na-PO₄ (pH 6.8) solution until A₂₆₀ = 0. The column was then further washed with 20 ml of 0.01 M Na-PO₄ (pH 6.8).
15. Plasmid DNA was eluted from the column with 10 ml of 0.3 M Na-PO₄ (pH 6.8) solution and 1 ml fractions were collected and A₂₆₀ determined. Peak fractions were pooled together and dialyzed against 10 mM Tris (pH 7.6), 10 mM NaCl, 5 mM EDTA solution. Usually 50-150 µg of plasmid DNA were recovered from a 200 ml culture. The plasmid DNA preparations were essentially free of RNA when checked on agarose gel.

Digestion of DNA with restriction enzymes, electrophoresis and gel blotting

All restriction enzymes were purchased from New England Biolabs (Beverly, MA) and Bethesda Research Laboratories (Rockville, MD) and digestions were carried out according to the specifications recommended by the supplier. Reactions were terminated by addition of 0.1% SDS and 10 mM EDTA and heating at 60°C for 2 minutes. Before loading on gels, samples were brought to final concentrations of 5% (v/v) glycerol and 0.05% (w/v) bromophenol blue. Restriction fragments were separated by horizontal submerged agarose gel electrophoresis using 0.7-1% gels. Electrophoresis was carried out at room temperature using Tris-acetate running buffer (0.04 M Tris-HCl pH 7.8, 0.001 M EDTA, 0.005 M sodium acetate). To minimize any pH change, the running buffer was continuously circulated between two buffer compartments. After visualization by ultraviolet illumination of ethidium bromide stained gels, the DNA fragments were transferred to 0.45 μ m nitrocellulose membranes (Schleicher and Schuell) according to the method described by Southern (122).

Hybridization of DNA fragments to specific probes

The nitrocellulose filter membranes were prehybridized for several hrs at 42°C in a solution containing 50% formamide, 0.075 M Na Citrate, 0.75 M NaCl, 0.02% polyvinylpyrrolidone, 0.02% Ficoll, 200 μ g/ml bovine serum albumin, 0.05 M Na phosphate pH 7.0, 0.025% calf thymus DNA and 1% glycine. After completion of prehybridization, the solution was

squeezed from the plastic bag and replaced by probe-containing hybridization solution, which contained all the components of the prehybridization solution except that 1% glycine was omitted and dextran sulfate (final conc. 10%) was added. Hybridization was carried out at 42°C for a minimum of 24 hrs (maximum 50 hrs) after which the membrane was washed free of unbound probe. The following washing procedure was used:

1. 3 times with 2X SSC, 0.1% SDS (1X SSC = 0.15 M NaCl, 0.015 M NaCitrate) at room temperature.
2. 3 times with 2X SSC, 0.1% SDS at 50°C (20-30 minutes each).
3. once with 0.2X SSC, 0.1% SDS at 50°C (20 minutes).
4. once with 2X SSC at room temperature.

After the final wash, the filter membranes were air-dried and used for autoradiography. The autoradiograms were usually exposed at -80°C with an intensifying screen for several hrs to a few days.

Preparation of molecular probes

The ecotropic MuLV LTR probe used in screening of recombinant phage plaques was prepared from pRFV105, the molecular clone of RFV, the endogenous retrovirus chemically induced from RFM/Un mouse cell cultures (123). The LTR probe used in the analysis of selected pRFM clones was prepared from a 0.6 kbp PstI/PvuI fragment of pWNB5 (121), the molecular clone of an endogenous B-tropic virus (WN1802B) derived from BALB/c mice. This virus was originally obtained from Dr. W. P. Rowe and Dr. J. Hartley of the National Institute of Allergy and Infectious diseases, NIH. The [IS]-specific probe was prepared from a Sau3A/Sau3A

internal LTR fragment isolated from the AL10 clone (106). All DNA fragments used in the preparation of molecular probes were labeled with ^{32}P -dCTP (800-1000 Ci/mmol) (obtained from Amersham or ICN radiochemicals) by nick-translation (124). All other reagents for nick-translation were purchased in standard kits either from Amersham or New England Nuclear. The specific activities obtained ranged from 5×10^7 to 3×10^8 cpm/ μg DNA.

5'-End labeling of cloned DNA fragments

In order to remove 5'-phosphate groups, DNA fragments (20-40 μg) were treated with calf intestine alkaline phosphatase (CIAP) (Boehringer-Mannheim) in a reaction mixture containing 50 mM Tris pH 9.0, 1 mM MgCl_2 , 0.01 mM ZnSO_4 and 2-10 units of the enzyme. Incubation was carried out at 37°C for 1-2 hrs after which the enzyme was inactivated by phenol (twice) and chloroform-isoamylalcohol (25/1) (once) extraction. The kinase reaction mixture contained 50 mM Tris-HCl pH 7.5, 10 mM MgCl_2 , 5 mM dithiothreitol, 0.01 mM spermidine, $[\gamma\text{-}^{32}\text{P}]\text{ATP}$ (2000-5000 Ci/mmol) (ICN radiochemicals) and 1-3 units of T4 polynucleotide kinase (BRL). Incubation was at 37°C for 1-2 hrs.

Restriction endonuclease mapping

For restriction endonuclease mapping of cloned DNA fragments, the partial enzyme digestion technique of Smith and Birnstiel was used (125). DNA fragments labeled at both termini were digested by appropriate restriction enzymes to generate two single end labeled fragments. These fragments were separated on 0.7-1% agarose gel and then purified by electroelution. The partial enzyme digestion mixture

usually contained 0.1 unit of the restriction enzyme and 2000-5000 cpm of single end labeled DNA fragment in a total volume of 30 λ digestion buffer. Different size restriction fragments were separated either on 0.7% agarose or on 5% acrylamide gel.

Nucleotide sequence determination

Restriction fragments to be sequenced were subcloned into the pUC9 cloning vector (126). DNA sequencing was performed according to the method described by Maxam and Gilbert (127), usually for both DNA strands.

RESULTS

Cloning of endogenous MuLV-related proviral DNA sequences

According to initial analysis by P. L. Glover and Dr. L. R. Boone on their 20 isolates of LTR-containing RFM/Un mouse DNA recombinant clones, 6 showed also positive hybridization to a gag-pol specific probe indicating the presence of at least the 5' half of a proviral structure in these clones. Inserts of five of these isolates were transferred into the HindIII site of pBR322 and designated as pRFM#1, pRFM#6, pRFM#9, pRFM#16 and pRFM#17. The LTRs of these clones hybridized strongly to the [IS] (insertion-sequence)-specific probe indicating that they were structurally related to those clones previously described by Ou et al. (106). Figure 3 shows hybridization of the [IS]-specific probe in the plaque purification procedure, applied to a recombinant phage isolate from which pRFM#16 was constructed. Plaque hybridization of the AL10 clone (106) to the same probe is also shown.

Figure 3: Hybridization of recombinant phage plaques to [IS]-specific probe (106). Transfer of phage DNA to nitrocellulose filters was carried out according to the method described by Benton and Davis (120). The [IS]-specific probe had a specific activity of $\approx 5 \times 10^7$ cpm/ μ g DNA and the autoradiogram was exposed overnight at -80°C with an intensifying screen. a): plaque hybridization of a purified recombinant phage from which pRFM#16 was constructed and b): plaque hybridization of AL10 clone (106) used as a positive control.

(a)

(b)



Figure 3

Restriction endonuclease maps

To determine restriction endonuclease maps of selected recombinant clones, DNA inserts of pRFM#1, pRFM#9 and pRFM#17 were labeled at both ends and then digested with the single-site restriction enzyme, Sall, to generate two single end labeled fragments. Each fragment was then mapped separately. In the case of pRFM#16 which lacked the Sall restriction site, the entire recombinant plasmid DNA was digested with BamHI and all three restriction fragments were labeled simultaneously at both termini as previously described. The labeled fragments were purified on agarose gels and subsequently digested with XhoI restriction enzyme to generate single end labeled fragments. These fragments were purified and mapped separately.

Restriction endonuclease maps of four endogenous MuLV-related proviral DNAs are shown in Figure 4 (restriction map of pRFM#6 will be shown later in Chapter III). For comparison purposes, the restriction map of AL10, the endogenous MuLV-related proviral DNA cloned from the BALB/c mouse (106) is also presented. The approximate sizes of the inserts were 10.9, 6.1, 13.6, 5.2 and 10.5 kbp for pRFM#1, pRFM#9, pRFM#16, pRFM#17 and AL10 clones, respectively. The sizes of the MuLV-related regions were 5.7 kbp (pRFM#1 and pRFM#9), 7.6 kbp (pRFM#16), 3.8 kbp (pRFM#17) and 6.9 kbp (AL10). Several restriction sites frequently found among the endogenous ecotropic and xenotropic murine leukemia proviruses (107,112,128,129) were also present in most of these recombinant clones. These sites (shown in approximate kilobases from the 5'-terminus of the 5' LTR) were 0.04 - PstI, 0.5 - KpnI, 0.5 - SmaI sites in the LTR, 0.9-BglII site in the gag gene, 3.3 - KpnI,

Figure 4: Restriction endonuclease maps of four endogenous MuLV-related proviral DNAs cloned from RFM/Un mouse DNA. For comparison purposes, the restriction map of AL10, the endogenous MuLV-related proviral DNA isolated from the BALB/c mouse, is also shown. Restriction sites were determined by the partial digestion technique of Smith and Birnstiel (125). The 5' and 3' cellular flanking sequences are shown as dotted lines. The blackened area in each LTR represents the unique 170-bp [IS] sequence segment present in the U3 region. A 0.6 kbp PstI-PvuI fragment (shown at the bottom) was from pWNB5, the molecular clone of an endogenous B-tropic ecotropic virus WN1802B derived from the BALB/c mouse; the LTR probe prepared from this fragment was used in the gel blots of Figures 6 and 7. Restriction endonuclease sites are SmaI (□), KpnI (■), BamHI (◆), PstI (○), EcoRI (△), XhoI (▽), HindIII (●), SalI (◇), and BglII (▲).

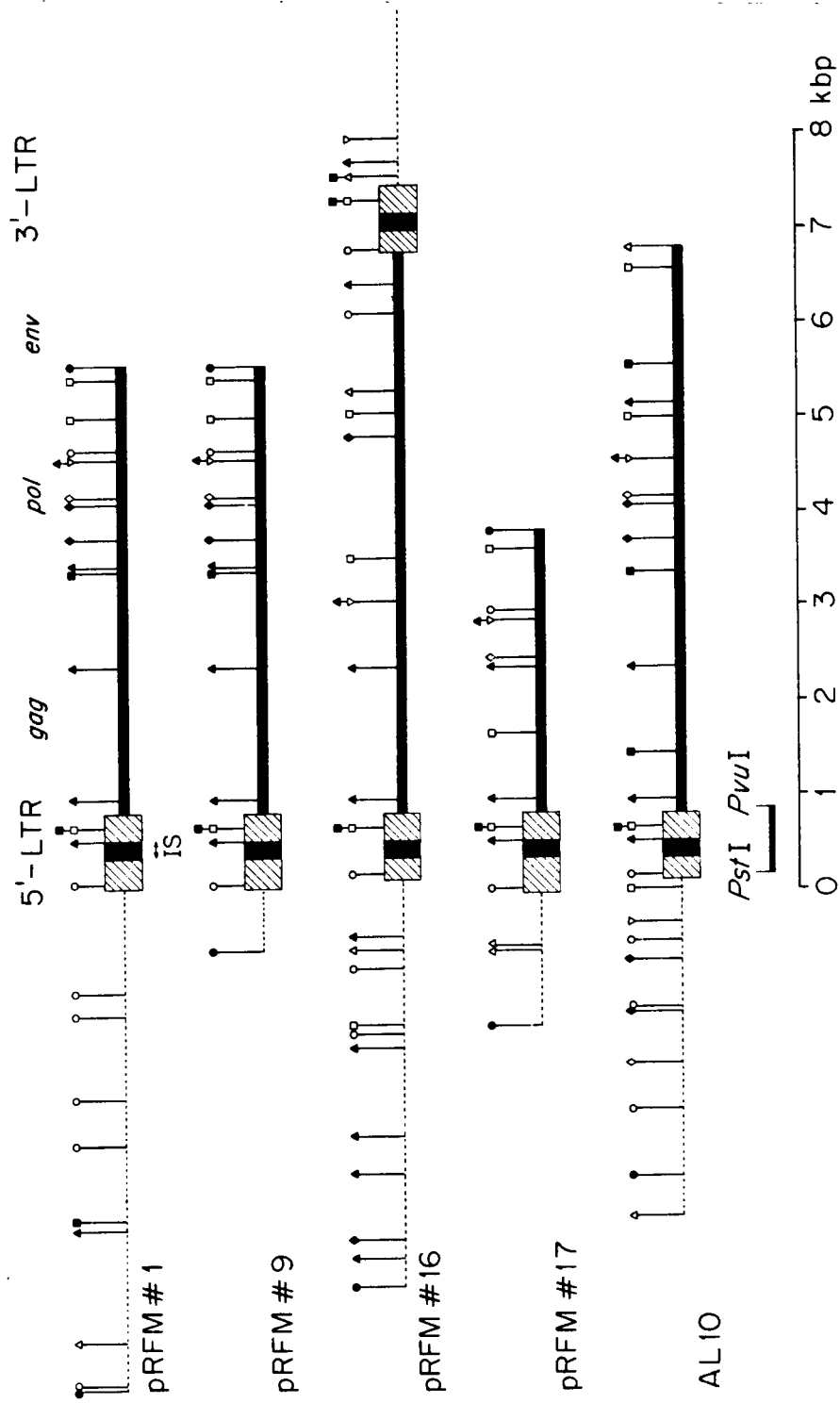


Figure 4

3.7 - BamHI, 4.1 - BamHI, 4.2 - SalI, 4.5 - BglII, 4.5 - XhoI and 5.1 SmaI sites in the pol gene and 6.2 - BamHI, 6.5 - SmaI, 6.8 - EcoRI and 7.9 - BglII sites in the env gene. Although two recombinant clones, pRFM#1 and pRFM#9, contained identical restriction sites in their MuLV-related regions, they represented two different and independent isolates since cellular sequences flanking these proviruses were distinctly different. The presence of a single internal HindIII site in the proviral DNA sequences of pRFM#1, pRFM#9, pRFM#17 and a single internal EcoRI site in AL10 prevented the isolation of the 3' portions of the proviral genomes. All these recombinant clones therefore, contained only one LTR (i.e. the 5' LTR). The absence of the internal HindIII restriction site in pRFM#16, on the other hand, allowed the isolation of a proviral genome which contained both 5' and 3' LTRs. The presence of these two LTRs was also demonstrated by gel blot analysis of digested pRFM#16 clone DNA using the [IS]-specific probe. As shown in Figure 5, digestions with three restriction enzymes, EcoRI, PstI and BamHI generated two fragments both of which carried an [IS]-containing sequence segment. Despite the presence of two LTRs, the 7.6 kbp MuLV-related DNA sequence of pRFM#16 was shorter than the reported 9.0 kbp proviral DNA sequences of the infectious murine leukemia viruses (128). From this restriction mapping analysis, it was concluded that this discrepancy was due to the deletion of ≈ 1.5 kbp in the pol gene between the BglII site at 2.3 kbp and XhoI/BglII sites at 4.5 kbp map positions. As a result of this deletion, the XhoI/BglII restriction sites, which are commonly present at 4.5 kbp map position in other MuLV-related proviral genomes, had been shifted to the 3 kbp map position in pRFM#16.

Figure 5: Demonstration of the presence of two LTRs in pRFM#16 by gel blot analysis (1% agarose). $\approx 0.5 \mu\text{g}$ of the pRFM#16 plasmid DNA were double-digested with HindIII to release the insert from the vector, and another restriction enzyme shown on the top of each lane. The [IS]-specific probe (106) used for hybridization had a specific activity of $\approx 5 \times 10^7$ cpm/ μg DNA and the autoradiogram was exposed for 10 hrs at -80°C with an intensifying screen. The numbers along the left side of the figure show the positions of HindIII-digest fragments of λ DNA used as MW markers. H = HindIII, E = EcoRI, P = PstI, A = AvaI, S = SalI, B = BamHI.

H E P A S B

Figure 5

The deleted fragment contained five restriction sites in the order of 5'-KpnI-BglII-BamHI-BamHI-SalI-3'. A similar, if not identical, deletion was also detected in pRFM#17. This type of deletion (roughly at the same map position) has also been reported in several other MuLV-related proviral DNAs isolated from the AKR/J mouse (107).

Size heterogeneity of the LTRs in pRFM recombinant clones

It has been reported previously (106,107) that when mouse chromosomal DNA is double-digested with PstI/KpnI, two enzymes with internal restriction sites in the LTR, three major fragments with sizes of ≈ 600 , ≈ 550 and ≈ 370 bp are generated which hybridize to a MuLV-specific LTR probe. This is indicative of a size heterogeneity among the endogenous murine leukemia virus related LTR sequences. LTRs of the known endogenous infectious ecotropic or xenotropic murine leukemia viruses such as AKV (112) and RFV (130) are generally shorter than those of endogenous MuLV-related sequences of the mouse genome. These LTRs are represented by the small ≈ 370 bp fragment (c band, Figure 6) upon combined PstI/KpnI digestion. The majority of endogenous MuLV-related sequences contain distinct LTRs with a novel 170-200 bp sequence segment [IS] inserted within the U3 region (106,108). These LTRs therefore, are represented by the larger ≈ 600 and ≈ 550 bp fragments (a and b bands, Figure 6) upon PstI/KpnI digestion. Digestion of pRFM recombinant clones with these enzymes demonstrated a size heterogeneity among the LTRs. pRFM#1, pRFM#9 and pRFM#17 belonged to the 600 bp class whereas pRFM#16 belonged to the 550 bp class (Figure 6). This also showed that the LTRs of the selected pRFM recombinant clones were generally larger

Figure 6: Gel blot analysis of the internal LTR PstI-KpnI fragments in endogenous MuLV-related proviral DNA clones. Cloned DNA (0.5 μ g) and chromosomal DNA (10 μ g) were double-digested with PstI and KpnI and electrophoresed at 60 V for 16 hrs in 1.5% agarose gel. The LTR-specific probe (described in the legend of Figure 4) used for hybridization had a specific activity of $\approx 3 \times 10^8$ cpm/ μ g DNA. The autoradiogram of the cloned DNA was exposed for 1 hr; chromosomal DNAs were exposed for 24 hrs at -80°C with an intensifying screen. The weakly hybridizing bands seen in some lanes and a large number of high MW DNA fragments in chromosomal DNAs represented those fragments containing the 3' ≈ 0.1 kbp portion of the LTR. The clones examined were pRFM#1 (lane 1), pRFM#9 (lane 2), pRFM#16 (LTR), a subclone of pRFM#16 containing the BglII fragment with 5'-LTR (lane 3), pRFM#17 (lane 4), pAL10B3, a subclone of AL10 (106) containing the 5.1 kbp BamHI fragment with 5'-LTR (lane 5), pWN41 clone of the BALB/c mouse derived N-tropic ecotropic MuLV strain WN1802N (lane 6). The chromosomal DNAs were those of RFM/Un (lane 7) and BALB/c (lane 8) mice.

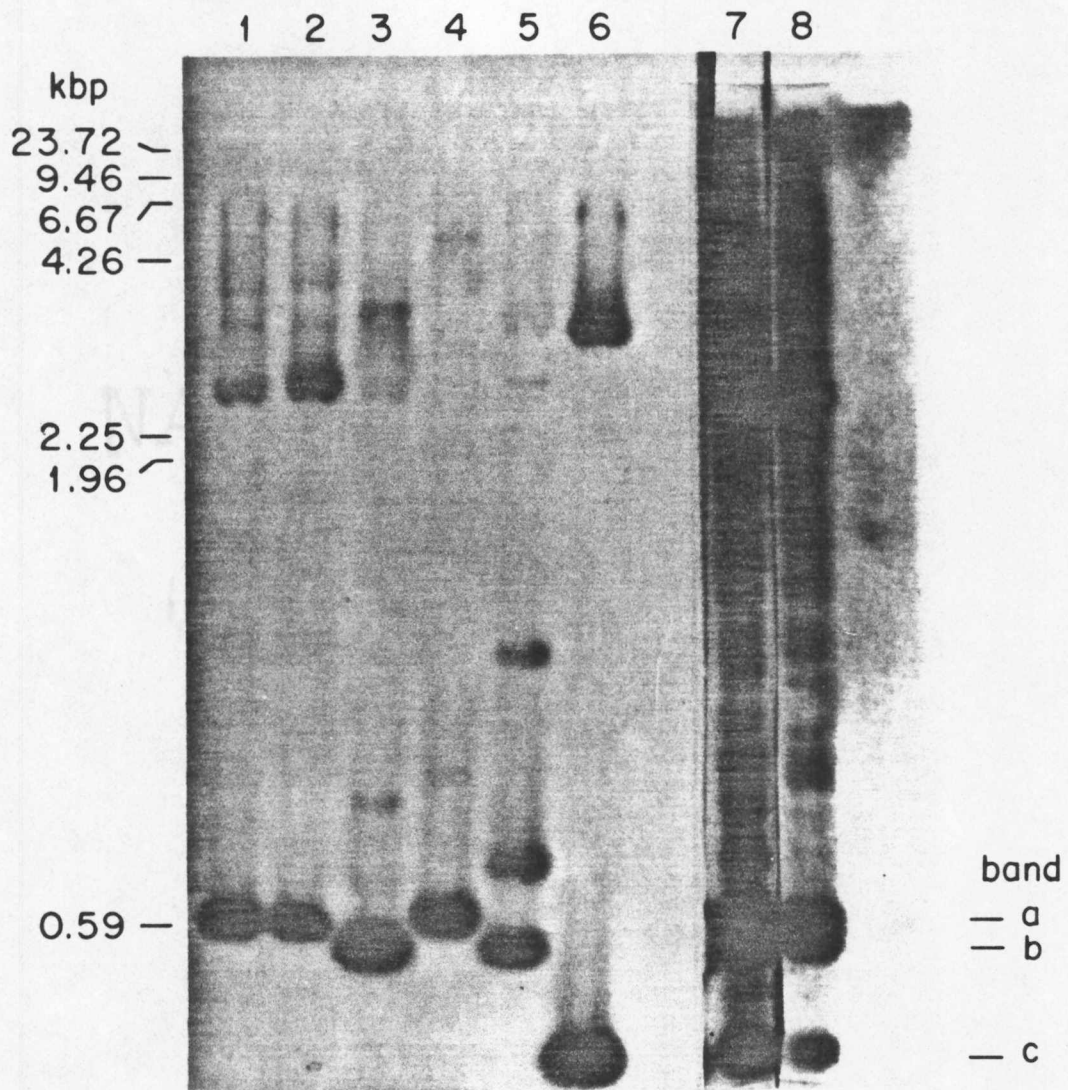


Figure 6

than those of the endogenous infectious MuLV sequences. Hybridization studies using the [IS]-specific probe (Figure 3) as well as nucleotide sequence analysis (next chapter) provided further evidence that the larger LTRs were present in the isolated pRFM recombinant clones.

Identification of DNA fragments carrying the tRNA primer binding site

As was described above, in a provirus, the 18-nucleotide DNA region homologous to the primer binding site is located immediately on the 3' side of the 5' LTR. Examination of the restriction maps shown in Figure 4 indicated that the tRNA primer binding site was included in a 0.46 kbp BglII fragment from pRFM#1, pRFM#9 and pRFM#17 and in a 1.4 kbp BglII fragment from pRFM#16. These recombinant clones were therefore digested with BglII and primer binding site-containing fragments were identified after separation on agarose gels using a LTR-specific probe (Figure 7). After purification, these fragments were subcloned into the pUC9 vector and used for nucleotide sequence analysis in the Maxam and Gilbert procedure.

Nucleotide sequence analyses of tRNA primer binding sites and adjacent regions in pRFM recombinant clones

Nucleotide sequences of tRNA primer binding sites and LTR U5 regions in pRFM recombinant clones are shown in Figure 8. For comparison, homologous sequences of the AL10 clone (106) are also shown. The 3'-terminal 18-nucleotide sequence of a rat liver major glutamine tRNA (119) (shown at the bottom) shows an 18/18 complementary match to primer binding sites of pRFM#1, pRFM#9 and pRFM#16 recombinant clones. The

Figure 7: Gel blot analysis of the BglII restriction fragments from cloned endogenous MuLV-related proviral DNAs of the RFM/Un mouse. A 0.5 µg sample of cloned DNAs was digested with BglII restriction enzyme and electrophoresed at 75 V for 12 hrs in 0.7% agarose gel. The LTR probe used for hybridization was the same as that described previously in the legends of Figures 4 and 6. The BglII fragments containing the tRNA primer binding site appeared as 0.46 kbp strongly hybridizing bands in pRFM#1 (lane 1), pRFM#9 (lane 2) and pRFM#17 (lane 4) and as a 1.4 kbp band (the upper band of the two strongly hybridizing bands) in pRFM#16 (lane 3).

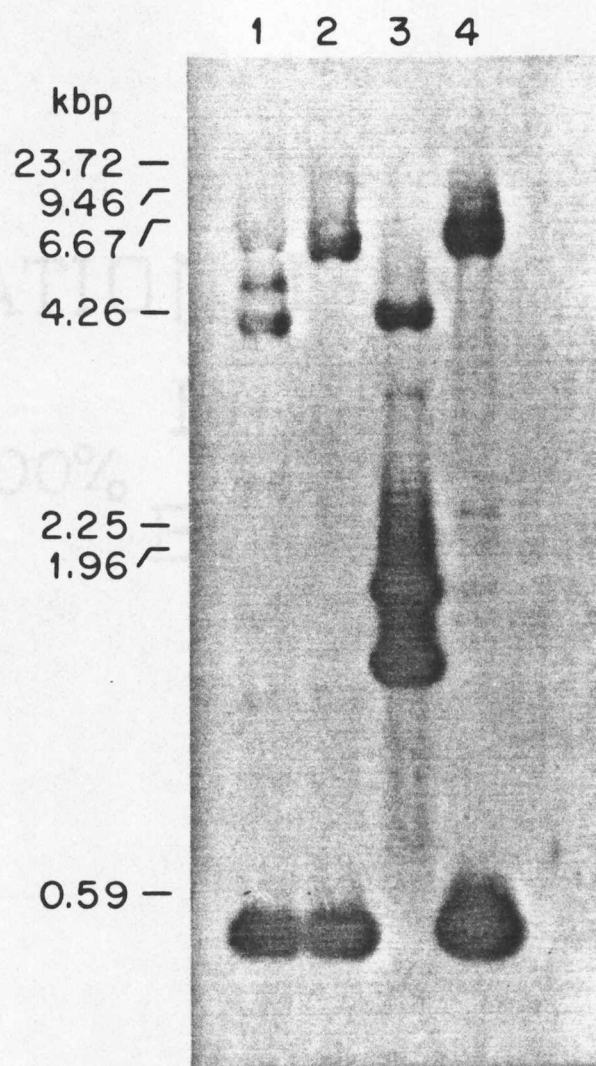


Figure 7

Figure 8: Plus-strand nucleotide sequences of the LTR U5, inverted repeat (IR), tRNA primer binding site and the immediate downstream variable (V) and conserved regions in four endogenous MuLV-related proviral DNAs from the RFM/Un mouse. The complete LTR structure and primer binding site of pRFM#6 will be shown in the following chapter. The nucleotide sequences of the corresponding regions of the AL10 clone (106) are also presented for comparative purposes. The 3'-terminal 18-nucleotides of a major glutamine tRNA isoacceptor from rat liver (119) and proline tRNA appear at the bottom of the figure. With the pRFM#1 sequence as reference, identical bases of other sequences are shown as dots. Base deletions are indicated by minus signs.

	U5	IR	tRNA Primer Site	V	3'
RFM#1 p	ATCCGAATCGTGGCTGGCTGATCCTTGGGAGGGTCTCCTCAGATTGATTGACTGCCACCTCG	GGGGTCTTTTCATT	TGGAGGTTCCACCCAGAT	-TGGAGACCCC	
RFM#9 p					
RFM#16 p	G.....A.....A.....G.....			T.....	
RFM#17 p	A.....ACTCG.T.....A.....CA.....	A.....C.....	C.....G.....	T.....	
AL10	G.....A.....A.....G.....		C.....	-C.....	

tRNA ^{Gln}	3'	5'
HO	ACCUCCAAGGUGGUCUA	→
tRNA ^{Pro}	HO	ACCCCCGAGCAGGCCCUA

Figure 8

primer binding sites of pRFM#17 and AL10 were respectively 16/18 and 17/18 complementary to the 3'-terminal sequence of glutamine tRNA. Matching with other published tRNA sequences (131) showed yet smaller degrees of complementation in all cases for these two clones. When primer binding sites of pRFM clones were matched with the 3'-terminal 18-nucleotide sequence of proline tRNA (also shown at the bottom of Figure 8), a primer identified to date in all ecotropic MuLVs, dispersed mismatches of 5 to 7 bases were observed. This indicates that the primer binding sites of these DNAs could not bind tightly to proline tRNA by base pairing.

Among the isolated pRFM clones, pRFM#17 showed the most nucleotide sequence variations in the U5, inverted repeat (IR) and the primer binding site. This clone contained not only two base changes in the tRNA primer binding site, but also showed two base variations in the highly conserved IR region. The T to C transition in the primer binding site and G to A transition in the IR region had generated restriction endonuclease sites for Sau96I (GGNCC) and HinfI (GANTC) enzymes, respectively. Both restriction sites were absent in the homologous regions of other MuLV-related recombinant clones, with the exception of AL10 which contained the Sau96I restriction site. Digestion of pRFM#17 clone in the IR and primer binding site regions by these two enzymes further substantiated the accuracy of sequence data (not shown).

Although the primer binding sites of recombinant pRFM DNAs were distinctly different from that of the RFM/Un mouse endogenous ecotropic MuLV (RFV) (130), comparison of the LTR U5 regions showed strong homologies between them. The 76-nucleotide U5 regions of pRFM#1,

pRFM#9, pRFM#16 and pRFM#17 differed from RFV provirus only by 3 (pRFM#1 and pRFM#9), 7 and 8 bases, respectively. On the other hand, the endogenous pRFM clones contained a short region of 1 or 2 nucleotides (designated V in Figure 8) located immediately on the 3' side of the primer binding site which was highly variable among these clones. The sequences following this variable region appeared to be conserved again. This pattern; 5'-primer binding site-V-conserved region-3'; has been also reported for other MuLV proviruses (40).

DISCUSSION

These results demonstrate that four out of five recombinant MuLV-related proviral DNAs isolated from RFM/Un and BALB/c mice possess a primer binding site for a major glutamine tRNA isoacceptor from rat liver. The tRNA primer binding site of pRFM#17, which is only 16/18 complementary match to glutamine tRNA, may show better complementation to the 3'-terminal sequences of three or four other glutamine tRNA isoacceptors of mouse cells (132) whose structures have not been elucidated. This clone may also represent a provirus with a defective tRNA primer binding site. From the data presented here, we may conclude that the endogenous MuLV-related sequences of RFM/Un, BALB/c and most likely several other inbred mouse strains, if replicate, will be distinct from the ecotropic murine leukemia viruses in using glutamine tRNA rather than proline tRNA to initiate reverse transcription. The presence of this specific tRNA primer binding site in endogenous MuLV-related sequences also provides supportive evidence for the hypothesis that these sequences were originally introduced into the mouse germ line through retroviral infection.

How was a primer binding site specific for glutamine tRNA generated in these endogenous sequences? The possibility of a relatively large number of point mutations in an original proline tRNA binding site leading to the generation of a new primer binding sequence is remote for two reasons. First, nucleotide sequences of the regions adjacent to the primer binding site (that is, U5 sequences shown in Figure 8 and the "gag" gene sequence in AL10; unpublished data) show much stronger homologies to corresponding regions in ecotropic MuLVs than the primer binding site itself. Second, the chances of random point mutations leading to the generation of an 18-nucleotide region with perfect complementary sequence match to a "specific" tRNA is extremely low. The existence of these novel sites may be explained by at least two more plausible mechanisms. First, according to a proposed model for reverse transcription (13), the mispaired binding of a new tRNA primer to the viral genomic RNA in place of the original tRNA or replacement of the original tRNA already bound to the "minus strand strong stop" by a new tRNA species can eventually lead to the creation of a new primer binding site in the integrated proviral genome. The second alternative is that the ecotropic MuLVs, possessing a binding site for proline tRNA, and the endogenous MuLV-related sequences with a putative binding site for glutamine tRNA may represent two distinct and independent retroviral lineages which have acquired similar structural genes through recombination. Although the exact genetic origin of glutamine tRNA binding site in the endogenous MuLV-related sequences is not clear at the moment, a recent report by Colicelli and Goff (133) has shown that this site can be utilized for primer binding during the initiation of reverse

transcription. This observation indicates that the switch to a new primer binding sequence in the endogenous MuLV-related genomes should not be considered as a "shut-off" mechanism to render these sequences replication incompetent, despite the fact that there is no evidence that indicates the infectivity of these proviral genomes.

In order to understand the nature of deletions detected in the pol gene of pRFM#16 and pRFM#17, nucleotide sequence analysis was performed at the site of deletion in pRFM#16. Through comparison with the published sequences of Moloney MuLV (12) and AKV (112), it was found that direct repeats of ACTGCCC sequence were present at both sides of the deleted region. Furthermore, unpublished data from our laboratory have indicated that proviral sequences carrying deletions, are preferentially selected during cloning of the endogenous DNA sequences. This conclusion was based on analysis of the genomic DNA which showed that by far the majority of the endogenous proviral genes exist in apparently intact non-deleted forms in the host animal (L. Y. Chang, personal communication).

CHAPTER III

CHARACTERIZATION OF A MOLECULAR CLONE OF A FULL-LENGTH ENDOGENOUS

MuLV-RELATED PROVIRAL DNA

INTRODUCTION

The large number of mouse endogenous DNA sequences homologous to murine leukemia viruses (MuLVs) fall into two main categories based on their structure and infectious properties. One category involves the endogenous DNA sequences which represent the few loci (1-3 copies/cell) of the infectious xeno- or ecotropic murine leukemia viruses, which have the potential of becoming induced to produce infectious virus upon treatment of cell cultures with chemical or physical agents or by cocultivation with appropriate cell lines. The second category of endogenous MuLV-homologous sequences are usually referred to as "MuLV-related" sequences, since at the moment it is not clear whether they can produce infectious retrovirus. The members of this group constitute the majority of the endogenous proviral genomes and are represented by the AL10 (106) or pRFM recombinant clones described in the previous Chapter. At least two distinct structural features may distinguish this group of endogenous sequences from the infectious viruses: 1) the presence of a primer binding site for glutamine tRNA instead of proline tRNA; 2) the presence of a novel sequence segment [IS] in the LTR U3 region. The [IS] sequence is also present in a distinct family of retroviral long terminal repeats in the mouse genome. These LTRs, called LTR-IS because of their presence mainly in a solitary form in mouse DNA (134), resemble

the LTRs of the endogenous MuLV-related sequences in that they contain a similar [IS] element, constituting the apparent U3 region of the LTR structure. The remainder of the LTR structure, however, shows little sequence homology to the LTRs of endogenous MuLV-related or infectious proviruses. Approximately 5% of these LTRs are associated with putative retroviral structural genes in a proviral form; a prototype of this group of endogenous proviruses is the IS23 clone which was isolated in our laboratory and is currently under investigation (Ou et al., in preparation).

Of the endogenous sequences described above, only members of the first group, i.e. infectious viruses, have been well characterized with respect to their structure and infectious properties. Members of other groups have not been well characterized both in terms of structure and biology. The present work was undertaken in an effort to elucidate the structure of one of the endogenous MuLV-related proviral DNAs, designated pRFM#6. The major goal was to determine the structure of those regions of the proviral genome involved in the regulation of retroviral transcription and the putative pathogenesis function. This recombinant clone has been of particular interest to us since it harbors an apparently complete proviral genome as determined by restriction enzyme map analysis. Other pRFM recombinant clones described earlier contain either partial proviral structure (pRFM#1, pRFM#9), or defective genomes (pRFM#16, pRFM#17). In this Chapter, results of the nucleotide sequence analyses of both the 5' and 3' LTRs as well as cellular and viral sequences flanking each LTR in the pRFM#6 provirus are presented. The structures of several other regions of the proviral genome and

distribution of this proviral insertion in genomes of some mouse strains will also be described.

MATERIALS AND METHODS

Molecular cloning and restriction enzyme mapping of pRFM#6

This recombinant clone was isolated together with other pRFM recombinants from cloning of spleen chromosomal DNA of a RFM/Un mouse according to the procedure described in Chapter II. The proviral genome was originally cloned into Charon 9 lambda phage and subsequently transferred to a pBR322 vector for further characterization. For restriction mapping, the partial enzyme digestion technique of Smith and Birnstiel (125) was used.

Preparation of DNA fragments containing the LTRs or specific structural genes from pRFM#6

Preparation of specific DNA fragments for nucleotide sequence analysis usually involved digestion of pRFM#6 DNA with appropriate restriction enzymes, separation of restricted fragments on 0.7 or 1% horizontal agarose gels, further identification of the fragments using specific probes and final purification and subcloning into pUC9 cloning-sequencing vector. The recombinant plasmids containing these fragments were subsequently used for transformation of E. coli competent cells, followed by large scale preparation of plasmid DNAs.

Preparation of transformation-competent E. coli cells by CaCl_2 treatment

E. coli cells (strains RR1, HB101 or C600) were made transformation-competent using the procedure of Morrison (135): a 500 ml portion of

L-broth (10 g Bacto-tryptone, 5 g Bacto-yeast extract (both from Difco), 5 g NaCl, 1 g glucose, 3 ml 1 M NaOH per liter) was inoculated with E. coli cells and incubated in a 2 L Erlenmeyer flask with aeration at 37°C. When the culture reached OD₅₆₀ ≈ 0.5, it was chilled in an ice-water bath by swirling until the temperature reached 0-5°C. The culture was then divided among three 250 ml centrifuge bottles and sedimented in a Sorvall rotor in the cold at 6000 rpm for 10 minutes. The pellets were gently suspended by repeated flushing (using plastic transfer pipettes) in a total volume of 125 ml ice-cold 0.1 M MgCl₂ and combined in one bottle. The cells were sedimented as before, and the pellet was resuspended in 125 ml cold 0.1 M CaCl₂ and kept at 0°C for 20 minutes. After final sedimentation, the pellet of CaCl₂-treated cells was resuspended in 22 ml cold 0.1 M CaCl₂ mixed with 3.5 ml of glycerol. The suspension of competent cells was stored at -80°C in 0.5 ml aliquots until use.

Transformation of competent E. coli cells

A tube of frozen competent E. coli cells was thawed in an ice-water bath for 10-15 minutes. Approximately 0.1 ml of the cells was transferred into two chilled clear 5-ml Falcon tubes (with snap caps) one of which contained 5 λ of the DNA sample (0.1-0.5 μ g DNA) in TNE buffer. After gentle mixing, both control and sample tubes were kept on ice for 30 minutes and then heat-shocked at 42°C for 2 minutes. To each tube 4 ml of warm (37°C) L-broth were added and the culture was incubated at 37°C for 1-2 hrs. After incubation, 2 ml of YT top agar were added to 0.3, 0.7 and 1.5 ml portions of the sample and plated on YT bottom agar

(6 g Bactoagar, 3.2 g Bacto-tryptone, 2 g yeast extract, 2 g NaCl per 400 ml). YT top agar was the same as bottom agar except for lower agar concentration. Since pUC9 (126) was used as the cloning vector, both top and bottom agars were supplemented with the following: 0.05 mg/ml selective drug ampicillin, 0.02 mg/ml color indicator X gal (BRL) and 0.04 mM IPTG (BRL). After overnight incubation at 37°C, transformants appeared as blue or white colonies on the agar plates. White colonies indicating positive transformants (inactivation of the lacZ gene by DNA insertion into the vector) were selected for large scale preparation of recombinant plasmids.

DNA gel blot analysis

Methods used in enzymatic digestion of cloned or genomic DNA, agarose gel electrophoresis, transfer to filter membranes, hybridization to specific probes, membrane washing and autoradiography were essentially the same as those described in Chapter II.

Extraction of 5'-end labeled DNA fragments from acrylamide gels

For sequencing purposes, DNA fragments labeled at one end, ranging in size from several hundred nucleotides to 2-4 kbp were usually separated on 5% polyacrylamide gels. To extract the EtBr stained DNA bands from the acrylamide gel, the bands were cut out with a razor blade and crushed in a 1.5 ml Eppendorf tube with the aid of a small teflon pestle. 400-500 λ of DNA extraction buffer (0.5 M NH_4Ac , 5 mM MgAc_2 , 1 mM EDTA, 0.1% SDS) were added to the tube and vortexed for a few seconds. The gel suspension was incubated for 30 minutes at 37°C and then spun for 5 minutes in a microfuge. The supernatant was drawn off

and transferred into another 1.5 ml Eppendorf tube. To the supernatant, 1 ml of cold 95% EtOH was added and the mixture was kept at -80°C for 5 minutes, then spun for 10 minutes as before. After spinning, the supernatant was carefully drawn off with a Pasteur pipett and discarded; the DNA pellet was kept on ice. To extract the residual DNA still in the gel, another portion of extraction buffer was added and extraction was carried out as before. This DNA was sedimented on the top of the first pellet and the combined DNA pellets were washed 3 $\times$ with 80% cold EtOH, then lyophilized. Dried DNA (salt free) was suspended in H_2O and used for sequencing by the Maxam and Gilbert procedure.

Preparation of molecular probes

The following probes were used in different stages of pRFM#6 DNA analyses: a LTR specific probe was made by nick-translation of a 372 bp PstI/KpnI internal LTR fragment isolated from pWN41; the molecular clone of WN1802N; the N-tropic ecotropic virus of the BALB/c mouse (121). The gag gene specific probe was prepared by nick-translation of a 1.5 kbp BglII/BglII restriction fragment isolated from the AL10 clone (106). The ecotropic-specific probe was prepared in Dr. L. Boone's laboratory (Research Triangle Park, NC) from a 310 bp AvaI (SmaI) env fragment located 1.8-2.2 kbp from the 3'-end of the ecotropic RFV proviral DNA (130,136). The MCF-specific probe was prepared by Dr. R. O'Neill (National Institute of Allergy and Infectious Diseases, Bethesda, Maryland) from a 100 bp BstNI/XmaI env fragment located 275 bp upstream from the single EcoRI site in the AKR MCF247 provirus (137); a homologous 112 bp BstNI/XmaI env fragment from NZB9-1 provirus (138) was

used to prepare the xenotropic-specific probe. The cellular sequence probe was prepared by nick-translation of a 1.2 kbp EcoRI/HindIII restriction fragment present at the 3'-cellular flanking sequence of pRFM#6 DNA.

RESULTS

Restriction endonuclease map of pRFM#6

The restriction endonuclease map of the mouse DNA cloned in pRFM#6 for 9 different restriction enzymes is shown in Figure 9. The total length of the cloned DNA fragment was estimated to be 12.4 kbp. The region of the DNA related to murine leukemia viruses contained all diagnostic restriction sites of the LTR (i.e. PstI, KpnI and SmaI) present at both ends of the proviral genome. The presence of these restriction sites as well as hybridization studies using a LTR-specific probe (shown in the following figures) demonstrated the occurrence of two LTRs in this cloned DNA fragment. Comparison of pRFM#6 and pRFM#16 (Figure 4, Chapter II) restriction maps showed that all the frequently observed restriction sites of endogenous infectious eco- and xenotropic MuLVs present in pRFM#16, were also present in pRFM#6. These included (in approximate kilobases from the 5'-terminus of the 5'-LTR) 0.9 BglII in the gag gene; 5.1 SmaI, 4.6 BglII, 4.6 XhoI in the pol gene and 6.7 SmaI, 8.0 BglII in the env gene. Additionally, pRFM#6 contained several restriction sites which were absent in pRFM#16 as a result of $\approx$ 1.5 kbp deletion in the 5' portion of the pol gene in the latter. These included 3.3 KpnI, 3.7 BamHI, 4.1 BamHI and 4.2 SalI restriction sites. With respect to the pol gene structure, therefore, pRFM#6 was similar

Figure 9: Restriction endonuclease map of the mouse DNA cloned in pRFM#6. The 12.4 kbp DNA insert of this clone apparently contained a full-length endogenous MuLV-related proviral genome of ≈ 9.3 kbp. DNA sequences of cellular origin are shown as thin lines flanking each LTR. The blackened area in each LTR represents the novel ≈ 170 bp sequence segment present in the U3 region. Also shown in the figure is the 1.5 kbp BglIII/BglIII restriction fragment isolated from the AL10 clone (106) and used as a gag-specific probe to identify the fragments containing the 5'-LTR (Figure 11). Map distances in kbp are shown at the bottom. Restriction sites were determined by the partial digestion technique of Smith and Birnstiel (125).

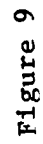


Figure 9

(but not identical) to pRFM#1, pRFM#9 and AL10 and different from pRFM#16 and pRFM#17 recombinant clones. The 6.25 BamHI restriction site in the env gene is not present in any known infectious ecotropic or xenotropic proviruses and is characteristically observed only in endogenous MuLV-related and MCF proviral structures (107,139). The single 6.9 EcoRI site also in the env gene is present in many non-ecotropic proviral genomes studied thus far (107). Based on restriction enzyme map analysis, the length of the 5' and 3' cellular flanking sequences were estimated to be 0.02 and 3.1 kbp, respectively. Size comparison of the estimated 9.3 kbp proviral-related genome of pRFM#6 with the genomic length of other endogenous eco- or xenotropic proviruses (128,130,140) suggested that the pRFM#6 clone harbored a complete full-length endogenous MuLV-related provirus. In Table 1, a list of the restriction fragments expected upon digestion of pRFM#6 clone DNA with various restriction enzymes is presented. Digestion of pRFM#6 clone DNA with the restriction enzymes described above is shown in Figure 10.

Identification of DNA fragments carrying intact 5' and 3' LTRs

For sequencing analysis, several restriction fragments carrying the intact 5' and 3' LTRs were identified on agarose gel after hybridization to specific probes. From the pRFM#6 restriction map (Figure 9), it was apparent that several restriction enzymes such as BamHI, Sall, XhoI, EcoRI and BglII should generate DNA fragments of various sizes carrying the intact 5' or 3' LTR. Restriction with several other enzymes, on the other hand, such as SmaI, KpnI and PstI, with internal LTR restriction

Table 1: Approximate sizes of restriction fragments generated upon digestion of pRFM#6 recombinant plasmid DNA with various enzymes^a:

Enzyme	Fragment size ^b							
BamHI	0.4	2.2	(3.7) ^c	(6.1)				
PstI	0.7	1.0	2.8	7.8				
KpnI	0.6	2.7	3.4	5.7				
SmaI	0.6	1.5	1.7	1.8	2.3	4.5		
XhoI	(4.6)	(7.7)						
SalI	(4.3)	(8.1)						
EcoRI	1.2	(4.3)	(6.9)					
BglII	0.4	0.8	(0.9)	1.2	1.5	(1.9)	2.2	3.4

- a) These fragments were generated from 12.4 kbp DNA insert of pRFM#6 after the plasmid DNA was double-digested with HindIII (to release the insert) and another restriction enzyme (see Figure 10).
- b) Fragment sizes were estimated by comparison with the electrophoretic migration pattern of HindIII-digested λ DNA fragments in 1% agarose gel with the aid of a computer.
- c) The figures in parenthesis represent those fragments containing a complete 5' or 3' LTR sequence.

Figure 10: Digestion of pRFM#6 recombinant clone DNA with various restriction enzymes. Except for PvuI, restriction sites are shown in the previous figure. Each digestion involved a single enzyme (shown on the top of each lane) as well as combined digestion by that enzyme plus HindIII (to release the pBR322 vector). DNA fragments were electrophoresed for 16 hrs at 50 V in 1% agarose gel and then visualized by staining with EtBr. 2 μ g of cloned DNA were used in each digestion. MW markers are shown on the left. A list of DNA fragments generated upon each digestion appears in Table 1. Bm = BamHI, Sm = SmaI, H = HindIII, E = EcoRI, K = KpnI, X = XhoI, Bg = BglII, Pv = PvuI, Sl = Sall, Ps = PstI.

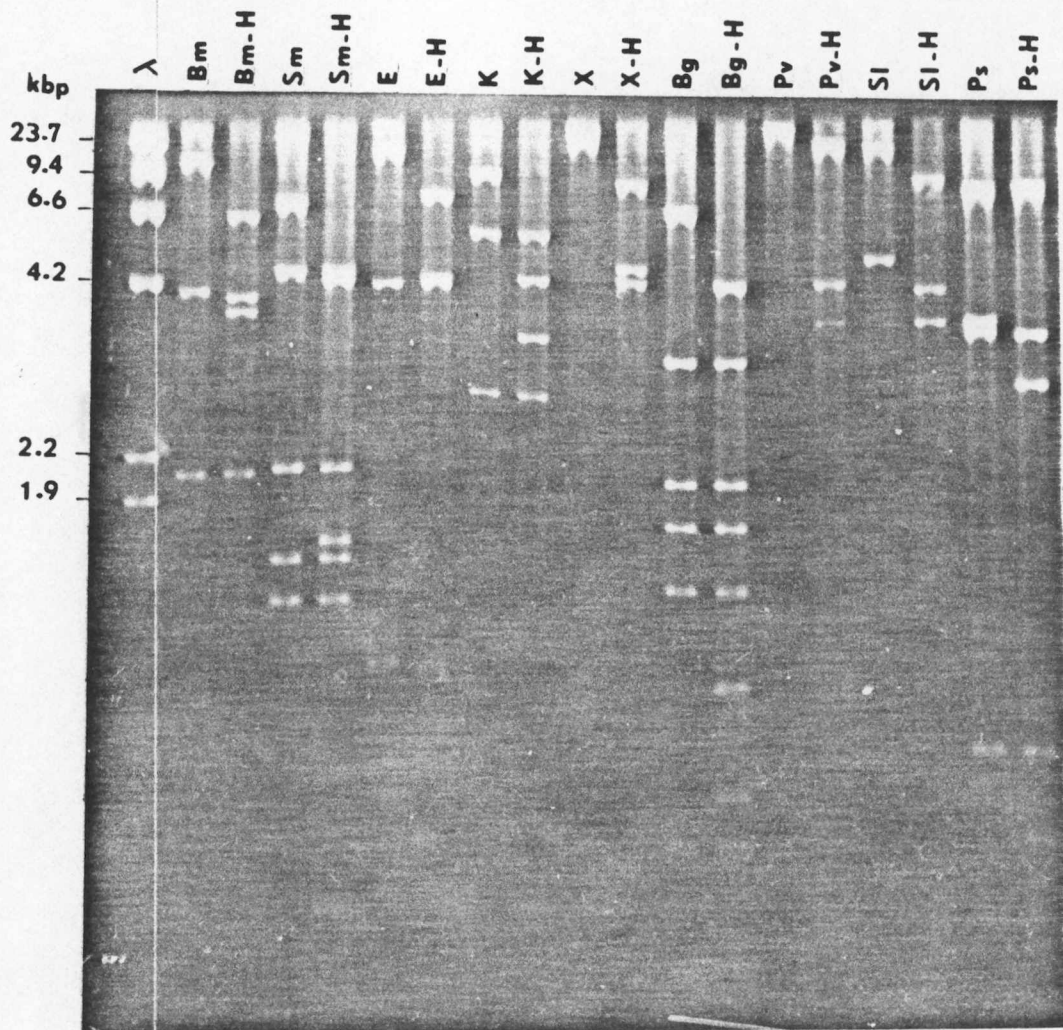


Figure 10

sites, could not generate intact LTR-containing fragments. With this knowledge in mind, pRFM#6 clone DNA was digested with the appropriate restriction enzymes and LTR-containing fragments were detected using specific probes. The nitrocellulose filter membrane containing bound restriction fragments was initially probed with a gag gene specific probe, since the 5' LTR is located adjacent to the gag gene, and with this approach it was possible to distinguish it from the DNA fragments carrying the 3' LTR. DNA fragments containing the intact 5' LTR were identified as 3.7, 4.3, 4.6 and 6.9 kbp bands after digestion with BamHI, SalI, XhoI and EcoRI restriction enzymes, respectively (Figure 11). The strongly hybridizing 7.5 kbp band seen with PstI digestion contained most of the 5' LTR whereas bands of 4.6 kbp in SmaI and 2.7 kbp in KpnI digestion carried only a small portion (≈ 100 bp) of the 5' LTR. The 1.5 kbp band observed in BglII digestion did not contain any 5' LTR sequence and represented a segment of pRFM#6 clone DNA homologous to the 1.5 kbp gag fragment isolated from the AL10 clone (106). In order to identify DNA fragments carrying the 3' LTR, the same filter membrane of Figure 11 was hybridized (without removal of the bound gag probe) to a LTR specific probe (Figure 12). DNA fragments containing the 3' LTR in its entirety were identified as 6.1, 4.3, 7.7, 1.9 and 8.1 kbp hybridizing bands upon digestion with BamHI, EcoRI, XhoI, BglII and SalI enzymes, respectively. The 0.9 kbp band seen with BglII digestion, contained an intact 5' LTR which could not be detected previously by the gag gene specific probe. To facilitate nucleotide sequencing, intact LTR-containing restriction fragments of the smallest possible sizes were selected. These included a BglII fragment of

Figure 11: Identification of DNA fragments containing the 5'-LTR of the proviral genome in pRFM#6. A 0.5 μg sample of cloned DNA was double digested with HindIII to release the insert from the vector, and another restriction enzyme shown on the top of each lane. Restriction fragments were analyzed by horizontal agarose gel electrophoresis using 1% gels. The gag-specific probe (described in the legend of Figure 9) used for hybridization had a specific activity of $\approx 3 \times 10^8$ cpm/ μg DNA. The autoradiogram was exposed for 1 hr at -80°C with an intensifying screen. MW markers are shown on the left side of the figure. The two additional fainter bands of ≈ 3.3 and 10 kbp detected in KpnI digestion (also seen in Figure 12) probably resulted from incomplete digestion by this enzyme.

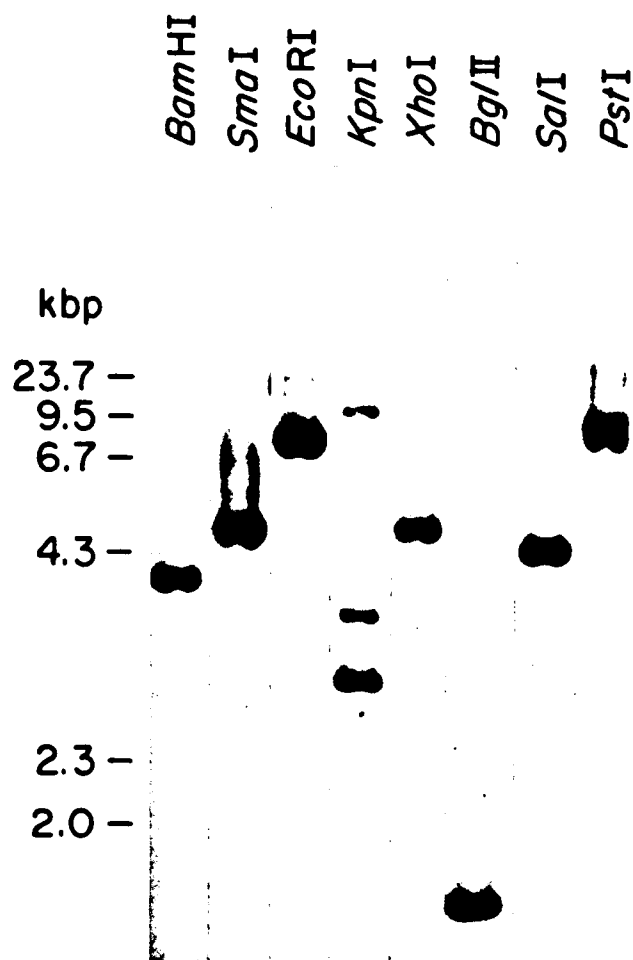


Figure 11

Figure 12: Identification of DNA fragments containing the 3'-LTR of the proviral genome in pRFM#6. The same filter membrane used for identification of the 5'-LTR-containing fragments (Figure 11) was used, without deprobing, for hybridization to a LTR-specific probe. The LTR-specific probe had a specific activity of $\approx 3 \times 10^8$ cpm/ μ g DNA and was prepared from nick-translation of a 372 bp PstI/KpnI internal LTR fragment from pWN41, the molecular clone of WN1802N, the N-tropic ecotropic virus of the BALB/c mouse (121). Two additional bands of ≈ 3.3 and 10 kbp seen in KpnI digestion probably resulted from the incomplete digestion by this enzyme.

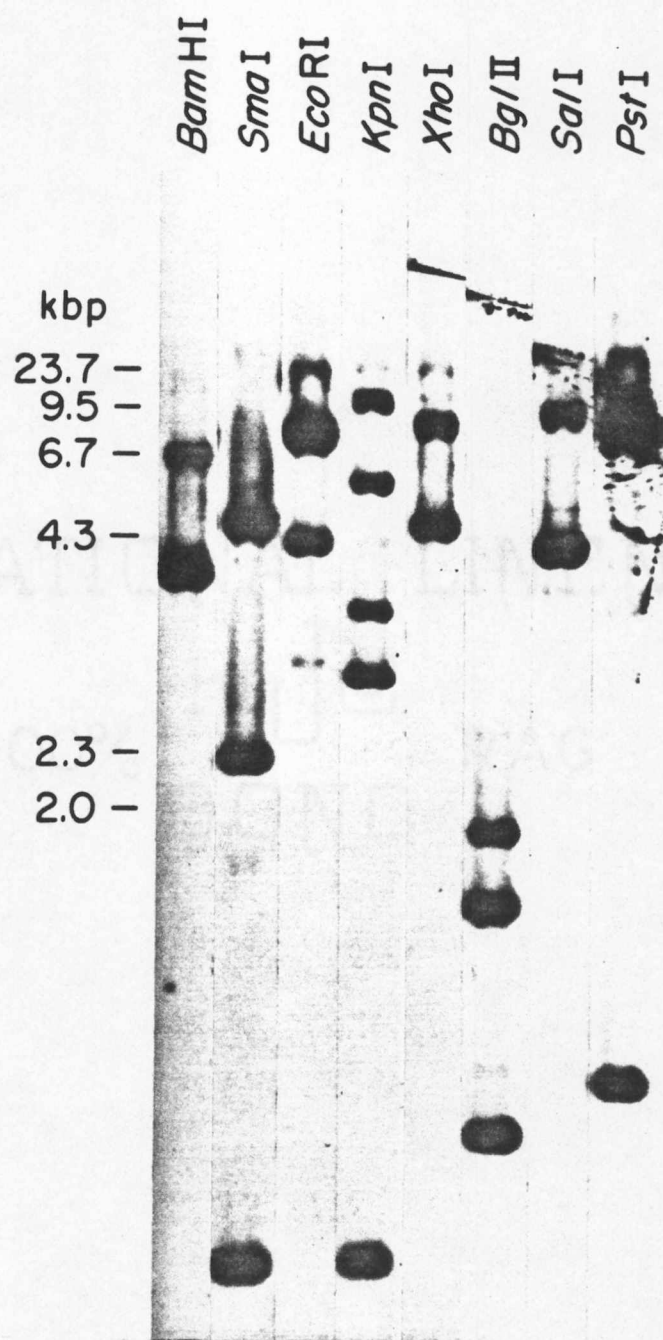


Figure 12

0.9 kbp (5'-LTR) and a BglII fragment of 1.9 kbp (3' LTR) both of which were subcloned into the pUC9 sequencing vector.

Nucleotide sequences of the 5' and 3' LTRs

Sequencing strategy and nucleotide sequences of the 5' and 3' LTRs of pRFM#6 DNA are shown in Figures 13 and 14, respectively. As described previously, each LTR which is formed by the process of reverse transcription contains three distinct regions in the order of U3-R-U5, where U3 and U5 represent unique sequences derived from the 3' and 5' ends of viral RNA, and R designates a relatively short, terminally redundant sequence present at both ends of viral genomic RNA. Sequence determination of the LTRs showed that 5' and 3' LTRs were 699 and 700 bp long, respectively. Furthermore, the two LTRs had identical structures, except for one base deletion and two base transitions occurring within the U3 region. As defined by comparison with Moloney murine leukemia virus LTR (10,12,34) the sizes of the U3, R and U5 regions were determined to be 554 (555), 69 and 76 bp, respectively. Two transcriptional regulatory signals common to eukaryotic regulatory elements were identified in the U3 region; the presumed initial recognition site for RNA polymerase II, CCAAC or the "CAT" box, and the presumed binding site for RNA polymerase II, TATAAAAA or the "TATA" box. The short nucleotide sequence commonly believed to be the polyadenylation signal, the AATAAA sequence, was present 47-52 bp downstream from the mRNA cap site.

Sequences of viral origin adjacent to 5' and 3' LTRs

As was shown in Chapter II, the proviral DNA sequence of 18-20 nucleotides immediately 3' to the U5 of the 5' LTR is homologous to the

Figure 13: Strategy used in the sequencing of the LTRs in pRFM#6. The LTR was 0.7 kbp long and is represented by the hatched and blackened [IS] areas in this figure. Two internal LTR restriction sites used for 5'-end labeling and sequencing are indicated. Thin and bold arrows show the direction of sequencing in 5' and 3' LTRs respectively. Broken lines indicate that sequencing was started or terminated outside the LTR.

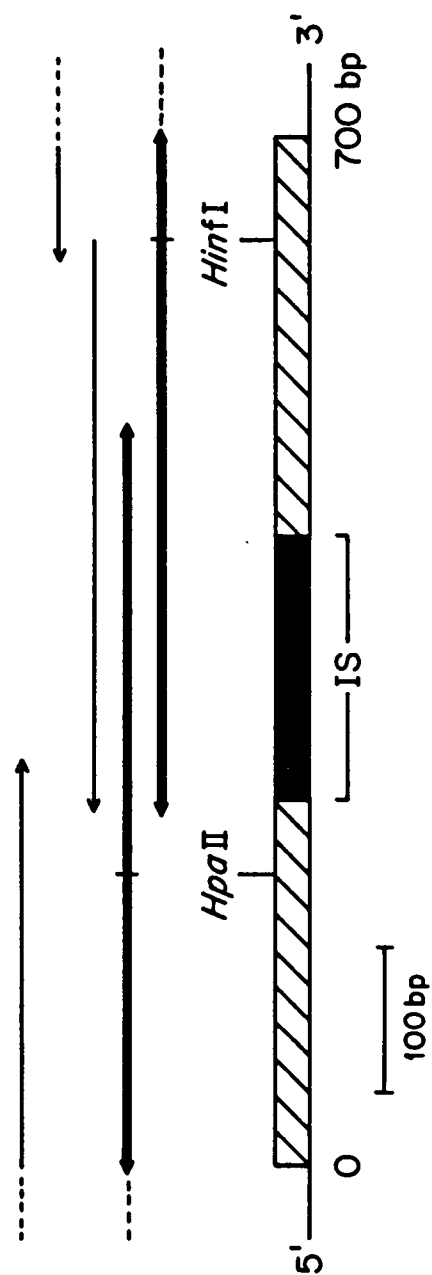


Figure 13

Figure 14: Nucleotide sequences of the 5' and 3' LTRs and adjacent cellular and viral junctions in pRFM#6. For sequencing purposes, a 0.9 kbp BglII fragment, containing the intact 5'-LTR (Figure 11) and a 1.9 kbp BglII fragment, containing the intact 3'-LTR (Figure 12) were subcloned into the sequencing vector pUC9 (126). Sequencing was performed by the chemical cleavage method of Maxam and Gilbert (127) (for sequencing strategy see Figure 13). The proviral sequences shown are those of the positive strand, i.e. corresponding to retroviral genomic RNA. Sequences of structural or regulatory significance are boxed. Abbreviations: "TATA" and "CAT", RNA polymerase II consensus sequences; pA, polyadenylation signal; PBS, primer binding site for (-) strand synthesis; PU, primer site for (+) strand synthesis; IR, inverted repeat; DR, direct repeat. The U3/R junction corresponds to the cap site for the 5' end and R/U5 junction corresponds to the polyadenylation site or the 3' end of the viral genomic RNA. -, absence of a nucleotide.

Figure 14

binding site of a tRNA primer molecule. Nucleotide sequence analysis of pRFM#6 at this region showed the primary structure TGGAGGTTCCACCGAGAT (Figure 14) which is a perfect complementary sequence match to the 3' terminal 18 nucleotides of a glutamine tRNA isoacceptor. This provirus, therefore, is similar to other MuLV-related proviral DNAs in possessing a glutamine tRNA primer binding site for initiation of reverse transcription.

The purine-rich sequence (PU), 12-14 nucleotides long and located immediately 5' to the U3 region of the 3' LTR, is the putative initiation site for plus (+) strand DNA synthesis (13,110,111). A stretch of 13 nucleotides; AGAAAGAGGGGGG was present at this region in pRFM#6 (Figure 14) which showed considerable homology to corresponding regions of AKV and RFV, the endogenous ecotropic retroviruses of AKR and RFM/Un mice, respectively (112,130). From this observation it was concluded that in contrast to a distinctly different primer binding site for minus strand DNA synthesis, the primer site for plus strand DNA synthesis in MuLV-related proviral genomes is not unique but rather is closely related to that of other type C mammalian retroviruses (40).

Sequences of cellular origin adjacent to 5' and 3' LTRs (cell-provirus junctions)

From nucleotide sequence analyses of proviral integration sites, it is now well established that integration of exogenous retroviral DNA into the host genome results in the same DNA structural features as those associated with prokaryotic and eukaryotic transposable gene elements (10,34,35). In both, the nucleotide sequence at each host DNA-insert

junction consists of an inverted repeat at the terminus of the inserted DNA and a short direct repeat in the host DNA flanking the insert. As a result of integration, two dinucleotides, usually AA/TT present at the termini prior to integration are lost by some unknown mechanism. When pRFM#6 DNA was examined at the cell-provirus junctions, similar structural features were revealed (Figure 14). The proviral genome was bounded by inverted repeats (IR) of 11 nucleotides; TGAAAGACCCC at the 5' end and GGGGTCTTTCA at the 3' end. When compared with the canonical AATGAAAGACCCC/GGGGTCTTTCATT inverted repeats present in all characterized unintegrated linear DNA intermediates of murine leukemia viruses (40), loss of the dinucleotides AA/TT at both termini could be detected. Furthermore, each inverted repeat at the terminus of pRFM#6 provirus was linked to a short direct repeat of the four nucleotides, CCAA, of cellular origin. It was, therefore, concluded that the proviral genome of pRFM#6 was most likely generated by an integration mechanism similar to that of exogenous murine leukemia proviruses. Nucleotide sequence analyses of cellular flanking sequences adjacent to the 5' LTR in AL10 (106) and pRFM#16 (not shown) revealed 4 bp sequences of CCAC and AGGT, respectively. Comparison of these sequences with that of pRFM#6 (i.e. CCAA) clearly demonstrated that integration of these proviral genomes into the host cell DNA was not site-specific. The autoradiograms of sequencing gels showing cell-provirus junctions of pRFM#6 are presented in Figure 15.

Figure 15: The autoradiograms of sequencing gels showing cell-provirus junctions of pRFM#6. In both autoradiograms, the nucleotide sequence of the (+) sense strand is shown. The 5' to 3' direction of each sequence is indicated. DR = direct repeat, IR = inverted repeat.



Figure 15

Comparison of LTR structures in pRFM#6 provirus and RFV, the endogenous ecotropic MuLV of the RFM/Un mouse

It was of interest to determine the structural relationship between the LTRs of pRFM#6 provirus and RFV, the infectious ecotropic murine leukemia virus induced from the RFM/Un mouse cell cultures by halogenated pyrimidines (130). As is shown in Figure 16, both LTRs are bounded by identical inverted repeats, i.e. TGAAAGACCCC/GGGGTCTTTCA. Furthermore, in both, the same "TATA" box sequence, TATAAAAA, was present 23-30 nucleotides upstream from the mRNA cap site. The polyadenylation signals were also identical and were found 18-23 nucleotides upstream from the R/U5 junction. The "CAT" box of pRFM#6 LTR, while showing one base difference (i.e. CCAAC vs. CCAAT), was 6 nucleotides further upstream from the cap site than that of RFV because of a CAGCCC duplication on the 5' side of the TATA box -- a feature also found in LTRs of some other MuLV-related proviral clones (106,108). In general, greater than 90% sequence homology was observed in the U5 and R regions, whereas the U3 region was distinctly different in these two proviral genomes. The presence of a unique 170 to 200 bp internal sequence segment [IS] in the mid-U3 region of pRFM#6 LTR made this LTR correspondingly larger than that of the RFV provirus. As a result, the total length of the pRFM#6 proviral genome (≈ 9.3 kbp) is somewhat larger than the proviruses of known infectious retroviruses (123). In addition to the presence of this novel sequence segment, the remainder of the U3 region in pRFM#6 also showed considerably less homology to the corresponding sequences of RFV provirus LTR than its U5 and R regions.

Figure 16: Comparison of the LTR structures and primer binding sites for (-) strand DNA synthesis in pRFM#6 and RFV proviruses. The nucleotide sequence of the RFV provirus is matched to pRFM#6 at homologous regions, therefore showing insertions or deletions. The novel ≈ 170 bp sequence segment in the U3 region is underlined. Abbreviations are as described in the legend of Figure 14.

	U3		
#6	IR TGAAAGACCCOACCATCAGGCTTAGCAAGCTAGCTGCAGTAACGCCATTTGCAAGGCATGAAAAAGTACCAGAGCTGA G	80	
RFV	TGAAAGACCCOCTCATAAGGCTTAGCCAGCTAACTGCAGTAACGCCATCTTGAAGGCATGGGAAAAATACCAGAGCTGATG	81	
#6	TTCTCAAAAGTTACAAGAAAGTTCAAGTTAAAGATTAACAGTTAAAGATTAAGGCTGAATAATACTGGGACAGGGGCCAAA	160	
RFV	TTCTCAGAA AAACAAGAACAGGAAGTACAGAGAGGCTGG AAAG TACCGGACTAGGGCCAAA	144	
#6	TATCGTGGTCAAGCACTTGGGCCCCGGCTCAGGGCCAAGAACAGATGGCTCTC	214	
RFV	CAGGATATCTGTGGTCAAGCACTAGGGCCCCGGCGCAGGGACCAGAACAGATGGTCCCAGAAATAGCTAAAAACAAAAAC	224	
#6	AGACGTCAGTGTAGCAGAACTAGCTTCACTGATTTAAAAAATAGAGGTGCACAGTCTCTGGCCACTCCTTGAACCTG	294	
#6	TGTGTCTGCCAATGTTCTGACCAAGTGTGTGCCATTGTTGAACCTTCATTAGACCCTTCTCTCGTACCCCTCCCATACC	374	
#6	CATTCTTAAAAATAGACATTGTTAGAACTAAAAAGTCCC	431	ACCTC AGTTCCCAA
RFV	AGTTTCAAGAGACCCAGAAACTGTCTCAAGGTCCCAG	263	
#6	ATGACCGAGAAA TACCCCAAGCCTTATTGAACTAA"CAT"CAAGCAGCTCGCTTCTCGCTTCTGTAACGGCGCTTTTGCTC	510	
RFV	ATGGCCGGGGATCAACCCCAAGCCTCATTAACTAA"CAAT"CAAGTCTCGCTTCTCGCTTCTGTACCGCGCTTATTGCTG	343	
#6	CCCAGCCCCAGCCCTATAAAAAAGGTAAAACTCCACACTCGGCGCGCCAGTCTCCGATAGACTGCGTCGCCCCGGGTAC	590	
RFV	CC CAGCTCTATAAAAAAGGTAAAAACCCACACTCGGCGCGCCAGTCTCCGATAGACTGAGTCGCCCCGGGTAC	417	
#6	CCGTGTTCCCAATAAAACCTCTTGCTGTTTGCA"U3"R"TCGGAATCGTGGTCTCGCTGGTCTCTGAGAGGGTCTCCTCAGATTGA	670	
RFV	CCGTGTATCCATAAAACCTTTTGCTG TTGCATCCGAATCGTGGTCTCGCTGATCCTTGGAGGGTCTCCTCAGAGTGA	496	
#6	TTGACTACCCACGTCGGGGCTTTTCATT"IR"U5"TGAGGTTCCACCGAGATTG	720	
RFV	TTGACTGCCACGCTTCGGGGCTTTTCATT"IR"U5"TGAGGTTCCCGGATTG	546	

Figure 16

Endonuclease gene structure of pRFM#6

During the replication cycle of murine leukemia viruses, the unintegrated linear duplex DNA is converted into a covalently closed circular DNA duplex before it integrates into the host genomic DNA. Recent studies with mutant MuLVs have shown that the 3' portion of the pol gene encodes a polypeptide of ≈ 46 KD with an endonuclease activity which is presumably required for integration of the circular proviral DNA (64,65,141). This activity has been purified from Moloney MuLV (Y. Yoshinaka et al., unpublished data). Since infectivity of the endogenous murine leukemia proviruses apparently depends on a functional endonuclease, the 3' region of the pol gene in pRFM#6 was sequenced and the predicted amino acid sequence compared with the corresponding region of AKV (112). For sequencing purposes, a 1.5 kbp SmaI/SmaI restriction fragment of pRFM#6 (Figure 9, map positions 5.1 to 6.6 kbp) that includes the 3' portion of the pol gene and the 5' portion of the env gene was subcloned into the pUC9 vector. The entire fragment was then sequenced by the method described by Maxam and Gilbert (127) after cleavage into smaller overlapping pieces and labeling. The sequencing strategy is shown in Figure 17. Nucleotide sequences of the endonuclease genes in AKV and pRFM#6 with the predicted polypeptide structures are shown in Figure 18. For sequencing analysis, the numbering system of Herr (112) was adopted. The 5' SmaI site of the cloned 1.5 kbp fragment corresponded to nucleotides 4579-4584 in AKV. The position of the NH₂-terminal amino acid (i.e. Ile) in both polypeptides was deduced from the unpublished sequence data of Y. Yoshinaka et al. on Moloney MuLV endonuclease. This amino acid is encoded by an

Figure 17: Strategy used in sequencing of 1.5 kbp SmaI/SmaI restriction fragment of pRFM#6 (Figure 9, map positions 5.1 to 6.6 kbp) that includes the 3' portion of the pol gene (the endonuclease) and the 5' portion of the env gene. The internal restriction sites used for 5'-end labeling and sequencing are shown. Broken lines indicate that sequencing was started from outside this fragment, that is, in the cloning vector, pUC9.

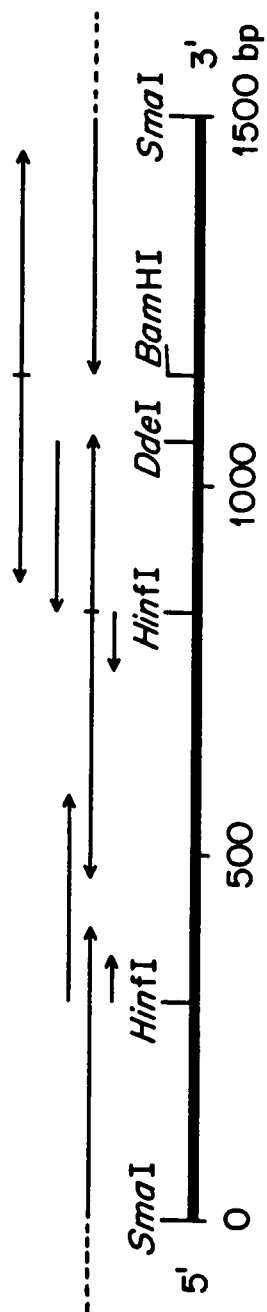


Figure 17

Figure 18: Comparison of nucleotide and predicted amino acid sequences of the corresponding 3' portion of pol genes in pRFM#6 and AKV. For sequencing purposes, a 1.5 kbp SmaI/SmaI restriction fragment of pRFM#6 (Figure 9, map positions 5.1-6.6 kbp) that includes the 3' portion of pol gene (endonuclease) and the 5' portion of env gene was subcloned into sequencing vector pUC9. Using the chemical cleavage method of Maxam and Gilbert (127), the entire piece was sequenced after cleavage into smaller overlapping fragments and 5'-end labeling (for sequencing strategy see Figure 17). Only nucleotide sequences of the endonuclease genes are shown here. The numbering system of Herr (112) for the AKV sequence was adopted. With the elucidated pRFM#6 sequences as the reference, dots (.) and asterisks (*), respectively, represent identical nucleotide bases and identical predicted amino acid residues found in AKV sequences. (ooo) = termination codon.

5' Endonuclease

```

#6 nuc : ATAGAGACCTCAACCCCGTACACTCCAGACTTCTTCCATTATACTGAGACAGATATAAAGAACCTACAAGAGTTGGGAGCC 4706
a.a.: IleGluThrSerThrProTyrThrProAspPhePheHisTyrThrGluThrAspIleLysAsnLeuGlnGluLeuGlyAla
AKV nuc : .....GAT.....A..T..G..T..C..AT.....A.....C.....A..CAG....C.T..G... 4706
a.a.: * * Asp * * * * * AlaTyr * * * * * Leu * Lys * Arg * * * *

#6 nuc : ACATATGATAGAGAGAAAAAATATTGGGTCTGCAAGGTAACCTGTGATGCCTGACCAGTTACACCTTTGAATTATTAGAC 4787
a.a.: ThrTyrAspArgGluLysLysTyrTrpValLeuGlnGlyLysProValMetProAspGlnPheThrPheGluLeuLeuAsp
AKV nuc : ..C...A.CCAGAGC...GG...C.....T.C.....C..G..G.....C.....A..TGTA.....C.G..... 4787
a.a.: * * * GlnSer * Gly * * * Phe * * * * * Val * * * *

#6 nuc : TTCCTTCACAGCTCACCCACCTTAGCTATCAGAAGATGAGGGCACTTCTAGACAGGAAAGAAAGCCCTATTACATGCTA 4868
a.a.: PheLeuHisGlnLeuThrHisLeuSerTyrGlnLysMetArgAlaLeuLeuAspArgLysGluSerProTyrTyrMetLeu
AKV nuc : .CA..C....G.....CG....C.....A....AA.....C..T....AGG.....C..... 4868
a.a.: Ser * * Arg * * * * Gly * * * * Lys * * * * Gly * * * *

#6 nuc : AATAAGATAAGATCTCTCCACGAGGTGGCGGAATCATGCCAAGCCTGTGTCCAAGTAAATGCCAGTAAAGACTAAGATCAGG 4949
a.a.: AsnLysAspLysIleLeuHisGluValAlaGluSerCysGlnAlaCysValGlnValAsnAlaSerLysThrLysIleArg
AKV nuc : ..CCCC..C..A.C.....AT.T....A..T..C....ACG.T.....C.....C..AG....A....G... 4949
a.a.: * Arg * * Thr * GlnTyr * * Asp * * ThrVal * Ala * * * * Ala * * Gly

#6 nuc : GCCGGAACACGAGTAAGAGGACATCGACTAGGCACCCATTGGGAAATTGACTTTACTGAAGTGAAGCCCGGACTGTATGGG 5030
a.a.: AlaGlyThrArgValArgGlyHisArgLeuGlyThrHisTrpGluIleAspPheThrGluValLysProGlyLeuTyrGly
AKV nuc : ..A...GTGA.....C.....G.C....T.....G..C..T....A....C.....A..G..... 5030
a.a.: * * Val * * * * * Pro * Ser * * * * *

#6 nuc : TACAAGTATCTCTTGGTATTCGTGGACACGTTCTCTGGCTGGGTTGAAGCCTTCCCAACCAAAACATGAGACTGCCAAAATA 5111
a.a.: TyrLysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaPheProThrLysHisGluThrAlaLysIle
AKV nuc : .....C....C.....C.....T.....G.....T.....GAGA..A..A..A.G.G.C 5111
a.a.: * * * * * Arg * * * ArgVal

#6 nuc : GTGACCAAGAAACTTCTGGAAGAAATATTCCAAAGTTTGAATGCCCAAGTGTGGGGACTGATAATGGGCCTGCCTTC 5192
a.a.: ValThrLysLysLeuLeuGluGluIlePheProArgPheGlyMetProGlnValLeuGlyThrAspAsnGlyProAlaPhe
AKV nuc : ...T.....T.G.....C..G..A..C.....A..G..A.....AT.....C..... 5192
a.a.: * Ser * * * * * Ser * * * *

#6 nuc : GTCTCCCAAGTAAGTCAGTCGGTGGCCAAGCTACTGGGATTGATTGGAACCTACATTGTGCTTACAGACCCACAGTTCA 5273
a.a.: ValSerGlnValSerGlnSerValAlaLysLeuLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerSer
AKV nuc : AC.....G.....G.TT.....C.....T.....T..... 5273
a.a.: Thr * * * * * Asp * * * *

```

Figure 18

```

#6  nuc : GGTCAAGTAAAGAATGAATAGGACAATCAAGGAGACTTTGACCAAATTAACGCTTGCAACTGGCACTAGAGACTGGGTA 5354
    a.a.: GlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrLeuAlaThrGlyThrArgAspTrpVal
AKV  nuc : .....A..C.....C.A..T.....G..... 5354
    a.a.: * * * * * Ala * * * * *

#6  nuc : CTCCTACTTCCCTTAGCCCTCTACCGAGCCCGCAACACTCCGGGCCCCCATGGACTCACTCCGTATGAAATCCTGTATGGG 5435
    a.a.: LeuLeuLeuProLeuAlaLeuTyrArgAlaArgAsnThrProGlyProHisGlyLeuThrProTyrGluIleLeuTyrGly
AKV  nuc : .....C.....T..T.....G.....G.....T....C... 5435
    a.a.: * * * * *

#6  nuc : GGGCCCCCGCCCTTGTAAATTTCCATGATCCTGAAATGTCAAAGTTTACTAATAGCCCTCTCTCCAAGCTCACTTACAG 5516
    a.a.: AlaProProProLeuValAsnPheHisAspProGluMetSerLysPheThrAsnSerProSerLeuGlnAlaHisLeuGln
AKV  nuc : .....C.....C..C..C.....G..A..A.....A..... 5516
    a.a.: * * * * * Asp * * * * * GluLeu * * * * *

#6  nuc : GCCCTCCAAGCAGTACAACGAGAGGTCTGGAAGCCACTGGCCGCTGCCTATCAGGACCAGCTGGACCAGCCAGTGATACCA 5597
    a.a.: AlaLeuGlnAlaValGlnArgGluValTrpLysProLeuAlaAlaAlaTyrGlnAspGlnLeuAspGlnProValIlePro
AKV  nuc : .....A.G..G..G.....AA..T....A.....AG.....C.G.....A..A.....A..... 5597
    a.a.: * * * Thr * * * * * Ile * * * * * Glu * * * * * Arg * * * * *

#6  nuc : CACCCCTTCCGTGTCGGCGACACCGTGTGGGTACGCCGACCAGACTAAGAAGTGGAACTCGCTGGAAAGACCCCTAC 5678
    a.a.: HisProPheArgValGlyAspThrValTrpValArgArgHisGlnThrLysAsnLeuGluProArgTrpLysGlyProTyr
AKV  nuc : .....GA..T..A...T.....G.....C..A.....A.....G..... 5678
    a.a.: * * * * * Ile * * * Ser * * * * *

#6  nuc : ACCGTCCTGCTGACCACCCACCCTCTCAAAGTAGACGGCATCGCTGCGTGGATCCAGCCGCTCACGTAAAGCGGGCG 5759
    a.a.: ThrValLeuLeuThrThrProThrAlaLeuLysValAspGlyIleAlaAlaTrpIleHisAlaAlaHisValLysAlaAla
AKV  nuc : .....A.....G.....T....A.....A.....C.....C..G..A... 5759
    a.a.: * * * * * Ser * * * * *

#6  nuc : ACAACCCCTCCGCGCGGAACAGCATCAGGACCGACATAG 5798
    a.a.: ThrThrProProAlaGlyThrAlaSerGlyProThr...
AKV  nuc : ..C..G..C...ATAAA..C..T...GGA..GTACA.CGCTCTCAAAACCCITTAATAATCAGGTTAACCCGTGGGGCCCC 5840
    a.a.: * * * * * IleLysProSerTrpArgValGlnArgSerGlnAspProLeuLysIleArgLeuThrArgGlyAlaPro

AKV  nuc : TAA 5843
    a.a.: ...

```

Figure 18 (Continued)

ATA codon which starts at nucleotide 4626 in AKV. As shown in Figure 18 the open reading frame of the endonuclease gene in pRFM#6 is 1170 nucleotides long (390 amino acids) and terminated by an amber (TAG) codon whereas that of AKV is 1215 nucleotides long (405 amino acids) and terminated by an ochre (TAA) termination codon. As a result, the deduced polypeptide sequence for pRFM#6 endonuclease is 15 amino acids shorter than that of AKV. Except for these additional amino acids, which were all present at the carboxy-terminal of the AKV endonuclease, the remainder of the two polypeptides were 86% homologous. The two endonuclease genes were 80% homologous with respect to their nucleotide sequences.

Characterization of the env gene - predicted host-range for pRFM#6 provirus

It was described in Chapter I that various host-range type specificities of murine leukemia viruses (i.e. ecotropic, xenotropic, amphotropic and MCF types) are primarily determined by a 70,000 Dalton envelope glycoprotein referred to as gp70. In particular, amino acids present at the N-terminal region of gp70 seem to differ most significantly in different host-range types of MuLVs (137,138). Characterization of the env gene of pRFM#6 provirus was carried out by DNA gel blot analysis as well as by nucleotide sequence analysis of the 5'-env gene. The non-ecotropic type nature of pRFM#6 provirus was evident by its lack of hybridization to an ecotropic-specific env probe both in the initial phage plaque screening and in the DNA gel blot analysis of this recombinant clone (data not shown). When two other molecular probes,

one specific for MCF and the other specific for xenotropic MuLVs were used, significant hybridization was observed with the MCF (Figure 19) but not with the xenotropic-specific probe (data not shown). This indicated that the pRFM#6 provirus is more related to MCF type viruses than to eco- or xenotropic MuLVs.

Nucleotide sequence analysis of the 5' env gene further substantiated the data obtained by gel blot analyses. The 5' env gene of pRFM#6 provirus is located between the BamHI site at 6.25 and SmaI site at 6.7 kbp map positions (Figure 9). This short region of ≈ 0.45 kbp was contained at the 3' end of the same 1.5 kbp SmaI/SmaI fragment used previously for endonuclease gene sequencing (see Figure 17 for sequencing strategy). The nucleotide sequence data presented in Figure 20 show that the 5' env gene of the pRFM#6 provirus was more related to the corresponding region of AKR MCF247 (137) than to NFS/N xenotropic (137) or NZB₉₋₁ xenotropic MuLVs (138). In the particular 100-nucleotide sequence segment bounded by BstNI and XmaI sites, where MCF and xenotropic viruses apparently differed the most (in fact, xeno- and MCF-specific probes were synthesized from this segment), the pRFM#6 provirus was identical to the AKR MCF247 virus with the exception of a G to A base transition and hence the absence of BstNI site. Within the same segment, the pRFM#6 provirus differed from the xenotropic NFS/N endogenous virus by as many as 21 nucleotides, in addition to a conspicuous 12-nucleotide sequence deletion/addition. With respect to polypeptide sequences at the N-terminal region of gp70, the predicted sequence of the pRFM#6 provirus differed from the AKR MCF247 virus by only one amino acid (Gln in place of Arg) whereas it differed from the

Figure 19: a): Hybridization of a MCF-specific probe, to the env gene of pRFM#6. A 0.5 μ g sample of the cloned DNA was double-digested with HindIII and another restriction enzyme shown on the top of each lane. DNA fragments were separated in 1% agarose gel and transferred to nitrocellulose filter membrane. The MCF-specific probe was prepared from a 100 bp BstNI/XmaI env fragment isolated from the AKR MCF247 provirus (137). MW markers are shown on the left side of the figure. b): Control experiment showing tropism-specificity of the MCF probe used in (a). Molecular clones tested were those of the ecotropic AKV 623, amphotropic 4070, AKR MCF247 and NZB₉₋₁ xenotropic proviruses with bands at 12.0, 8.2, 6.7 and 8.2 kbp respectively. Strong hybridization was observed only with the AKR MCF247 env gene. Experimental conditions were as in (a).

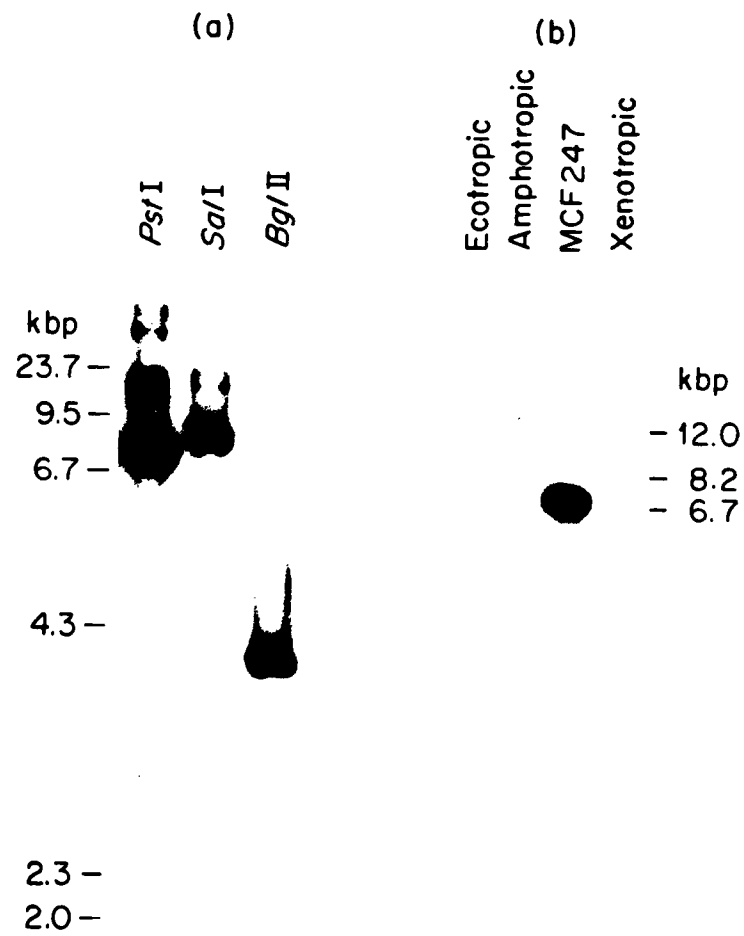


Figure 19

Figure 20: Comparison of nucleotide and predicted amino acid sequences of the corresponding 5' portion of env genes in pRFM#6, AKR MCF247 MuLV and NFS xenotropic MuLV. This portion of the env gene is located at 3' end between BamHI site at 6.25 and SmaI site at 6.7 kbp (Figure 9) of the 1.5 kbp SmaI/SmaI fragment described in the legend of Figure 18 (for sequencing strategy see Figure 17). With the elucidated pRFM#6 sequences as the reference, dots (.) and asterisks (*) represent identical nucleotide bases and identical predicted amino acid residues found in AKR MCF247 and NFS xenotropic MuLV sequences. The underlined sequence in pRFM#6 indicates the region where most marked similarity to, or differences from, the two MuLV types was noted.

```

      5'
      |env
#6       nuc : ATAGAAGGTCCAGCGTTCTCAAAACCCCTTAAAGATAAGATTAAACCCGTCGGGGCCCCCTAATA
      a.a. : IleGluGlyProAlaPheSerLysProLeuLysAspLysIleAsnProTrpGlyProLeuIle

MCF247   nuc : ..G.....
      a.a. : Met * * * * *

XENO     nuc : ..G.....T.....
      a.a. : Met * * Ser * * * * *

      90          |gp70
#6       nuc : GTCCTAGGGATCTTAATAAGGGCAGGAGTATCAGTACAACATGACAGCCCTCATCAGGTCTTC
      a.a. : ValLeuGlyIleLeuIleArgAlaGlyValSerValGlnHisAspSerProHisGlnValPhe

MCF247   nuc : .....G..A.....G.....
      a.a. : * * * * * Arg * * * * *

XENO     nuc : ..TA.G.....GG.G.....CC..G.....G.....C...A.....
      a.a. : * Met * * * Val * * * Ala * * * Arg * * * * * Ile *

      150         |
#6       nuc : AATGTTACTTGAGAGTTACCAACTTAATGACAGGACAAAAGCTAATGCTACCTCCCTCCTA
      a.a. : AsnValThrTrpArgValThrAsnLeuMetThrGlyGlnThrAlaAsnAlaThrSerLeuLeu

MCF247   nuc : .....BstNl
      a.a. : * * * * *

XENO     nuc : .....C.....C..C.....G
      a.a. : * * * * * C * * * * *

      210         |           240
#6       nuc : GGGACAATGACCGATGCCCTTTCCTAAACTGTACTTTGACTTGTGCGATTTAATAGGGGACCAC
      a.a. : GlyThrMetThrAspAlaPheProLysLeuTyrPheAspLeuCysAspLeuIleGlyAspAsp

MCF247   nuc : .....
      a.a. : * * * * *

XENO     nuc : .....G.....A..CA....C.....A..T.....C....T.....G....A....T..
      a.a. : * * * * * Thr * * * * * Val * * Tyr

      270         |           300    3'
#6       nuc : TGGGATGAG             ACTGGACTCGGGTGTGCGACTCCCGGG
      a.a. : TrpAspGlu             ThrGlyLeuGlyCysArgThrProGly

MCF247   nuc : .....XmaI
      a.a. : * * * * *

XENO     nuc : .....CCCAGAACCCGAT.T...GGAT..T..C.....
      a.a. : * * AspProGluProAspile * Asp * * * * *
```

Figure 20

NFS/N xenotropic MuLV by 12 residues. All these data clearly demonstrated the MCF-type nature of the env gene in the pRFM#6 provirus.

Strain distribution of the pRFM#6 provirus

The cellular DNA site harboring the pRFM#6 provirus in RFM/Un mice was examined in several other inbred strains of mouse for the presence or absence of the provirus. Genomic DNAs from these strains were digested to completion with HindIII restriction enzyme and DNA fragments were separated on agarose gel and transferred to filter membranes as described previously. Using a 3'-cellular flanking sequence from pRFM#6 cloned DNA as the molecular probe, the following strains were shown to harbor the provirus at the same genomic site; BALB/c, AKR, C3H and CBA (Figure 21). The proviral DNA genomes in these strains apparently had not incurred any detectable alterations in their structures as all hybridizing bands comigrated with the 12.4 kbp band detected in the RFM/Un mouse genomic DNA. At least three strains, BALB/c, C3H and CBA, are derived from the same origin, i.e. Bagg albino (142,143). The RFM mouse may also be related to this lineage since the RFM/Un strain apparently shares a common ecotropic virus inducible locus on chromosome 5 with the strains described above. The hybridizing bands detected in the genomic DNAs of C57BL/6 and NFS/N strains did not comigrate with the 12.4 kbp band of the RFM/Un mouse genomic DNA but rather appeared in lower MW regions. This indicates the absence of the proviral genome at this site in these two inbred strains. The lack of the proviral genome could be either due to the fact that this provirus never appeared in the genomes of these strains, or a previously integrated provirus came out

Figure 21: Detection of the endogenous MuLV-related proviral genome of pRFM#6 in various inbred mice. A 5-10 μg sample of genomic DNA was digested to completion with HindIII and analyzed by horizontal agarose gel electrophoresis using 0.7% gels. The molecular probe used had a specific activity of $\approx 3 \times 10^8$ cpm/ μg DNA and was prepared by nick-translation of a unique 1.2 kbp EcoRI/HindIII restriction fragment (Figure 9, map positions 11.2-12.4 kbp) present in the 3'-cellular flanking sequence of pRFM#6 provirus. Different inbred strains examined are shown on the top of each lane. A HindIII digested pRFM#6 clone DNA sample was also run as a control. The autoradiogram was exposed for 48 hrs at -80°C with an intensifying screen.

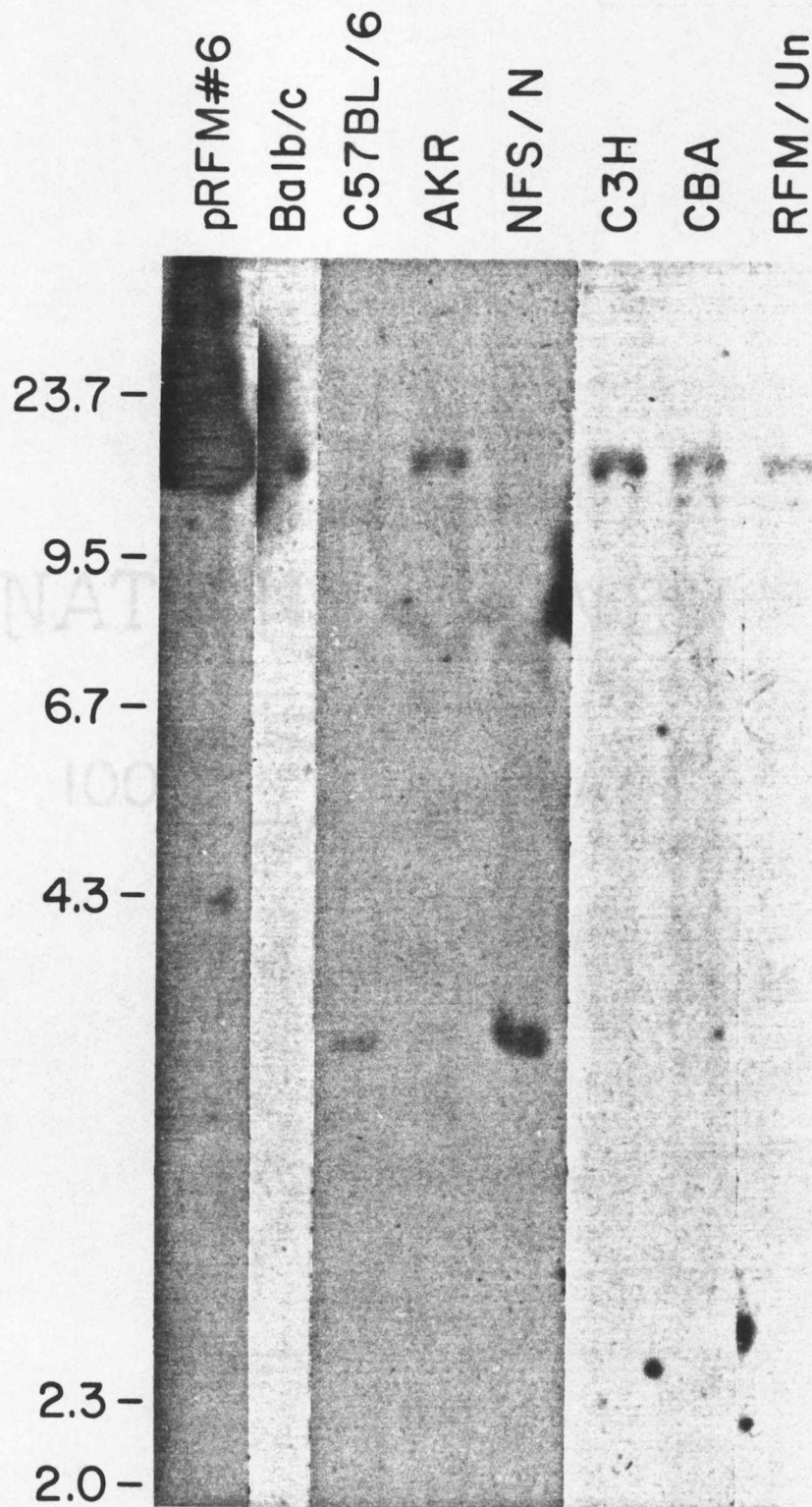


Figure 21

of the genome as a result of a homologous recombination between the two LTRs. In the former case, where no trace of the provirus should be present in the mouse genome, the cellular flanking sequence probe (described above) should be able to detect a band of ≈ 3.2 kbp (i.e. the size of the 5'+3' cellular sequences in pRFM#6) whereas in the latter case, where a single LTR always remains in the genome after homologous recombination, a band of ≈ 3.9 kbp [i.e. 3.2 kbp (cellular) + 0.7 kbp (LTR)] is expected. The gel blot of Figure 22 clearly shows that the cellular DNA probe detected a band of ≈ 3.2 kbp rather than 3.9 kbp. This proved that the proviral genome never appeared at this particular site in the genomes of NFS/N and C57BL/6 mice. This conclusion is also consistent with the fact that neither strain originated from Bagg albino, the common progenitor of most of the strains harboring the provirus (142,143). The strongly hybridizing band of 3.9 kbp shown in the Figure (lane 6), was isolated from the rec^+ RRI strain of E. coli cells in which the pRFM#6 plasmid was highly unstable and underwent a homologous recombination between the two LTRs giving rise to the above fragment. This fragment was therefore used as a control 3.9 kbp size marker.

DISCUSSION

Based on the restriction enzyme map analysis, it was concluded that the MuLV-related proviral DNA sequence of pRFM#6 is a complete genome and apparently without any detectable structural deletion(s). Despite the ubiquity of endogenous MuLV-related proviral sequences in the chromosomal DNA of laboratory mice, isolation of these sequences in intact

Figure 22: Gel blot demonstrating the absence of pRFM#6 LTR in the genomic DNAs of C57BL/6 and NFS/N mice. The following samples were used, pRFM#6 clone DNA (lane 1), C57BL/6 genomic DNA (lane 2), AKR genomic DNA (lane 3), NFS/N genomic DNA (lanes 4 and 5). All DNA samples were digested to completion with HindIII enzyme (the NFS/N genomic DNA in lane 5 was double-digested with HindIII/BamHI). The strong band seen in lane 6 represents a 3.9 kbp fragment isolated from rec⁺ RRI *E. coli* cells in which pRFM#6 plasmid underwent a homologous recombination between the two LTRs giving rise to above fragment. A dark band seen at ≈ 8 kbp in NFS/N digests occurred in the space between lanes 4 and 5 and therefore may be considered as an artifact. The molecular probe used and other experimental conditions were the same as those described in the legend of Figure 21.

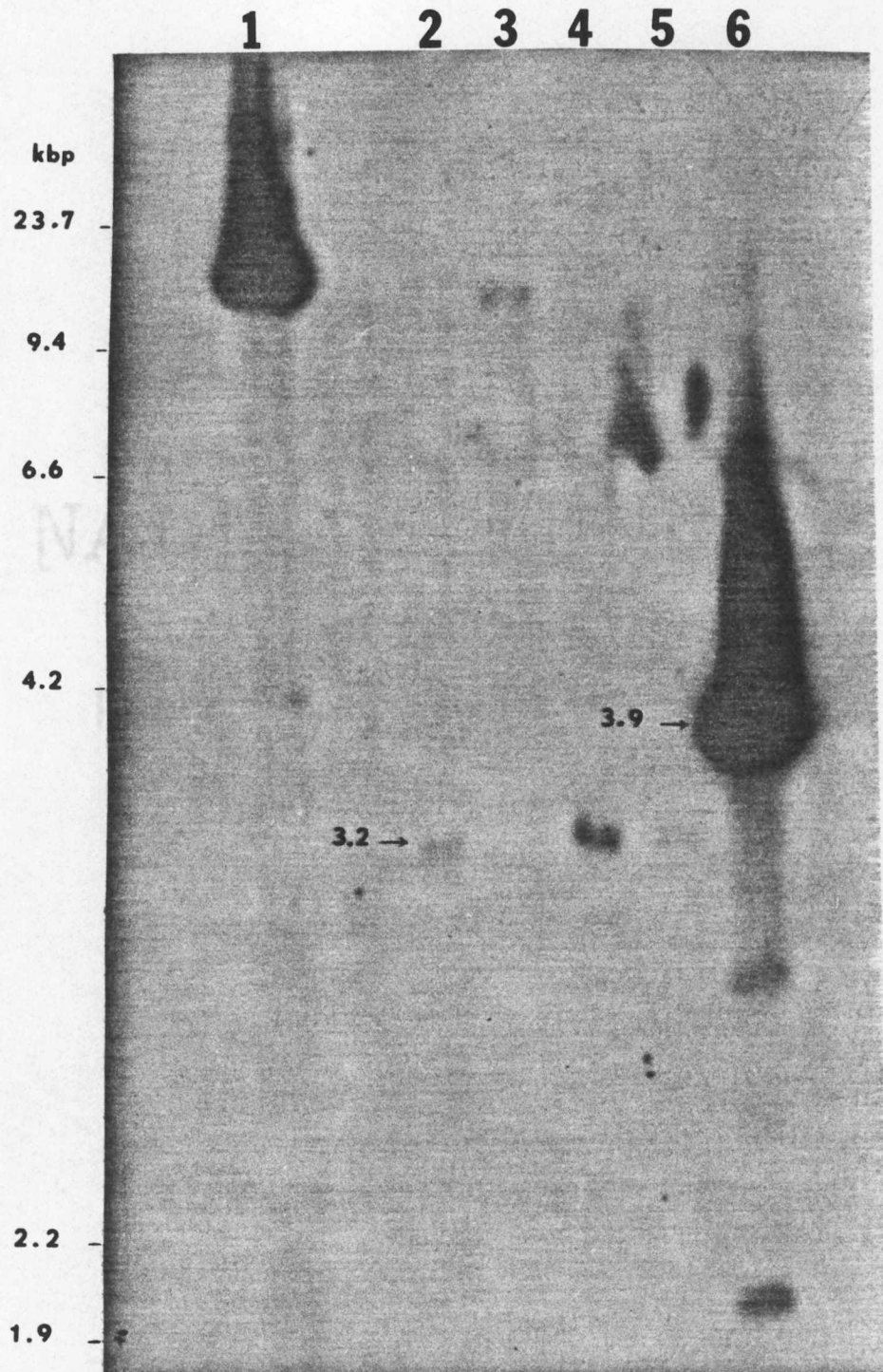


Figure 22

form has been hindered in several cases, either by the choice of the restriction enzymes for molecular cloning or by deletions of various sizes in the structural genes (91,106,107,144). At least some of these deletions seem to take place during cloning of the endogenous sequences. The choice of HindIII restriction enzyme in this work was based on a previous observation that some MuLV-related sequences in the chromosomal DNA do not possess an internal site for this enzyme (106). The laboratory of L. Bachelier (145) has also reported the isolation of a complete endogenous MuLV-related proviral genome from the BALB/c mouse. Although no sequence data are at present available on this provirus for comparison, the pRFM#6 provirus is apparently different from it as evidenced by different restriction enzyme sites present in the cellular sequences flanking each proviral genome.

Based on nucleotide sequence analyses of various regions of the proviral genome, it was concluded that with respect to replicative and integration mechanisms, the pRFM#6 provirus is similar to other exogenously acquired proviral genomes (111). For example, the presence of a primer binding site for glutamine tRNA and a poly-purine primer site, respectively, on the immediate 3' side of the 5' LTR and the immediate 5' side of the 3' LTR suggests that this proviral genome was originally synthesized by the reverse transcription mechanism. Furthermore, the presence of 4 bp direct repeats flanking the LTR inverted repeats at the termini of this provirus suggests that this proviral genome was integrated into the germ line of the mouse by an insertional mechanism known to be characteristic of prokaryotic and eukaryotic transposable gene elements. These intact structural

features, as well as the virtually identical 5' and 3' LTR sequences, suggest that this MuLV-related proviral genome either was acquired rather recently by the mouse germ line or has been subjected to minimal mutational changes in the host genome.

The limited structural analysis at the 3' region of the pol gene (endonuclease) showed 86% homology between the deduced polypeptide sequences of pRFM#6 and AKV. Unless some mutation(s) somewhere upstream (in the pol or gag gene) from the endonuclease gene of pRFM#6 cause a shift in the reading frame, the endonuclease gene of this provirus, if expressed, seems to utilize the same reading frame as that of AKV. Whether the observed significant degree of homology between the two polypeptides predicts a functional endonuclease for pRFM#6 requires further investigation.

The proviral genome of pRFM#6 contains all the structural features characteristic of MuLV-related sequences which distinguish them from the proviruses of ecotropic and xenotropic MuLVs. These included a distinct primer binding site for glutamine tRNA, presence of a novel insertion sequence segment [IS] in the LTR, and distinct restriction sites in the structural genes. One of the most important structural features of pRFM#6 provirus is its MCF-related env gene (gp70) which is characteristically different from the env genes of infectious eco-, xeno-, or amphotropic murine leukemia viruses. If the parental endogenous env gene sequences which participate in the recombinational events leading to the generation of MCF viruses originate from pRFM#6 or similar MuLV-related genomes, this study should provide strong evidence that such sequences themselves may be present in a complete proviral

structure in the mouse genome. It is therefore, a matter of curiosity that these endogenous sequences may have the potential of producing certain type of leukemogenic viruses through recombination with other MuLVs but by themselves cannot "directly" express the infectious viral particles despite their apparently complete genomic structure. Further discussion on the role of cellular and viral factors which may influence the expression from these sequences will be presented in the next Chapter.

CHAPTER IV

EXPRESSION AND INFECTIVITY OF THE ENDOGENOUS MuLV-RELATED PROVIRAL DNA SEQUENCES

INTRODUCTION

Although the number of murine leukemia virus related sequences in the mouse genome is relatively large, the number of genetic loci bearing the potential of "spontaneous" infectious virus production is small and seen only in certain mouse strains (e.g. AKR). This observation may indicate that most of the inherited proviruses are under some kind of cellular regulation which renders them transcriptionally silent. If this negative control is exerted by some cellular sequences located in close proximity to integration sites of proviruses, then one might expect that the appearance of a provirus at a new integration site away from these inhibitory sequences (146) should be free from such control. When susceptible mouse cells are exogenously infected by an ecotropic MuLV, usually new copies of the provirus appear at various sites in the cellular genome which leads to the production of infectious virus particles. Another possible indication of the presence of genetic control is the induction of retrovirus production following exposure of the cells to physical (e.g. irradiation) or chemical (e.g. halogenated pyrimidines) agents which cause damages to cellular DNA (94,95,96,97). Despite the fact that physical or chemical induction can cause the production of eco- or xenotropic viruses from otherwise nonproducing cells, the number of inducible loci still remains very small (1-3 copies/cell)

as compared to the total number of endogenous MuLV-related sequences. This may indicate that most of the endogenous MuLV-related sequences contain irreversible structural alterations such as point-mutations or large deletions in the structural genes which permanently prevent them from producing viral particles even after cellular induction. Another important factor which may contribute to the low level expression of endogenous MuLVs is the methylation state of these sequences. Several studies have implicated direct involvement of methylation in the expression of endogenous retroviral sequences (147,148,149). For instance, Harbers et al. (147), have demonstrated that a provirus of Moloney MuLV introduced into the mouse germ-line and apparently became silent as a result of methylation, could become infectious following molecular cloning which removed the methyl groups.

Although the endogenous MuLV-related DNA sequences of the mouse apparently cannot give rise to infectious virus particles, several lines of evidence suggest that these sequences are indeed transcriptionally active and generate RNA species which are distinctly different from those transcribed from the infectious inducible loci. Unpublished results from our laboratory have shown that these transcripts characteristically contain the [IS] element (106) and are expressed in various mouse organs. The overall level of expression was not the same in different organs of the NFS/N and the RFM/Un strain mouse. Liver and kidney of these mice showed the highest level of expression whereas the spleen contained lower activity. Only trace amounts of these transcripts could be detected in the brain, testes and thymus. Recent studies by Levy et al. (150) have demonstrated normal expression of

polymorphic endogenous retroviral RNA in strain 129G_{IX}⁺ mice containing segments identical to mink cell focus-forming (MCF) virus. It was described in Chapters I and III that the majority of the endogenous MuLV-related sequences including our pRFM#6 are of the MCF-type. This Chapter describes the expression and infectious properties of selected clones of MuLV-related sequences after transfection into a heterologous cell line lacking any endogenous MuLV genomes. With this approach, not only are these studies done in the absence of presumed cellular negative control but also they bypass the problems of attempting to examine a specific copy amidst a large family of related sequences.

MATERIALS AND METHODS

Cell lines and recombinant clones

Mouse NIH3T3 and human fibrosarcoma HT 1080 cell lines were originally obtained from the laboratories of Dr. G. Todaro and Dr. R. M. McAlister, respectively. Isolation of pRFM#16 and pRFM#6 recombinant clones (both of which carry two LTRs in their genomes) was described previously. The recombinant plasmid, p0.28-9.5 obtained from Dr. L. R. Boone (NIEHS), contains the newly acquired provirus of an ecotropic MuLV isolated from thymic lymphoma induced in a RFM/Un mouse by 5-azacytidine treatment (pBR322 was the cloning vector). pSV2-NEO, a recombinant plasmid carrying the neomycin-resistant gene under control of the SV40 early promoter (151) was obtained from BRL.

DNA transfection and clone selection

DNA transfection of the cells with p0.28-9.5 and selected pRFM plasmids was carried out by D. M. Yang (Oak Ridge National Laboratory) using the CaCl_2 precipitation procedure (152,153). Briefly, these cells were co-transfected with pSV2-NEO and one of the recombinant plasmids described above at a ratio of 1 to 20 and in a final DNA concentration of 10 $\mu\text{g}/\text{ml}$. 0.3 ml of DNA solution, following calcium chloride precipitation (final conc. of $\text{CaCl}_2 = 0.125 \text{ M}$) was applied to overnight cultures of 5×10^5 cells in 35 mm wells. DNA-containing medium was removed 15 hrs later and the cells were trypsinized at 48 hrs into 3 100 mm dishes for selection of transfected cell clones. Cellular clone selection was carried out by addition of neomycin [G418 (Gibco), final conc. 375 $\mu\text{g}/\text{ml}$] to the culture media of growing cells. Selected neomycin-resistant cell colonies were subsequently isolated by ring-cloning, expanded and at third passages, stored in liquid N₂ freezer.

Reverse transcriptase assay

Sixteen hour culture medium of near-confluent 100 mm dishes was used for assay of particle-associated reverse transcriptase activity. After removing cell debris by centrifugation at 6000 rpm (15 min) in a Sorvall type SS-34 rotor, the supernatant was spun at 45,000 rpm (60 min) using a Ti50 rotor in a Beckman Spinco ultracentrifuge. After removing the supernatant, the pellet (often not visible) was suspended in 100 λ of lysis buffer (0.05 M Tris-HCl pH 8.0, 0.5 M KCl, 10 mM MgCl_2 , 0.2% nonidet P-40, 12% glycerol and 5 mM dithiothreitol). The suspension was kept on ice for at least 30 minutes with occasional vortexing before use

in the reverse transcriptase assay. The enzyme reaction mixture contained the following:

BSA (1.7 mg/ml stock)	5	μl
[(rA)n-(dT) ₉][100 μg/ml poly(A)/10 μg/ml(dT) ₉]	5	μl
dithiothreitol (0.1 M stock)	2.5	μl
(10 mM MgCl ₂ + 5 mM MnCl ₂ stock)	5	μl
³ H-dTTP diluted with unlabeled dTTP to 4 Ci/mmmole	10	μl
H ₂ O	12.5	μl
enzyme (virus lysate)	10	μl
Total =		50 μl

The reaction mixture was incubated at 37°C in a water bath for 60 minutes after which 30 λ of the mixture were applied uniformly onto a Whatman 3 MM filter paper disk. After air-drying, disks were washed successively 4 times in cold acid solution (5% glacial acetic acid, 0.7% HCl, 1% sodium pyrophosphate), once in cold 95% ethanol, once in ethanol/ether (1/1) and finally once in ether (each wash lasted 10-15 minutes). After the final wash, disks were air-dried and placed in scintillation vials containing 5 ml of 0.4% BBOT and counted in a LKB liquid scintillation counter.

DNA extraction from monolayer culture cells

The following protocol was used for DNA extraction from monolayer culture cells of the selected neomycin-resistant cell clones:

1. Monolayer cells on 100 mm dishes were rinsed once with PBS (without calcium and magnesium) and once with TE (10 mM Tris-HCl pH 7.6, 1 mM EDTA) buffer.
2. After draining the residual TE buffer, 1.5 ml of Hirt's solution (10 mM Tris-HCl pH 7.6, 10 mM EDTA, 1% SDS) (kept at 65°C) were added and the dish was swirled gently. The cells lysed quickly and became gelatinous.
3. Cell lysates from four dishes were combined and residual extracts were rinsed with 4 ml of $T_{10}N_{100}E_{10}$ solution (10 mM Tris-HCl pH 7.6, 100 mM NaCl, 10 mM EDTA) and added to combined extracts. The total extract (≈ 10 ml) was transferred into a 50 ml capped, sterile polypropylene tube (Falcon).
4. 0.5 ml of self-digested pronase (20 mg/ml stock) (Boehringer-Mannheim) were added to the cell extract, mixed and incubated at 37°C overnight.
5. 1 ml of 5 M NaCl was added to the pronase-digested cell extract and mixed briefly at 37°C.
6. After incubation, 12 ml of chloroform/isoamyl alcohol (25/1) were added and mixed by gentle inversion of the tube (at least 100 $\times$).
7. The mixture was centrifuged at 2,000 rpm (IEC centrifuge, model PR-6) for 10 minutes. The top aqueous layer was removed and transferred into another 50 ml sterile Falcon tube.

8. 25 ml of cold 95% ethanol were layered gently over the aqueous phase (≈ 10 ml) and DNA fibers formed at the interface were spooled out using a 1 ml polystyrene pipet.
9. The DNA (wound around the pipet tip) was soaked first in 75% and then in 95% ethanol for at least 30 minutes and subsequently transferred into a polypropylene tube (Falcon, 17×100 mm).
10. After a brief lyophilization, 2 ml of $T_{10}N_{100}E_{10}$ buffer were added to the tube and allowed to stand until the DNA fibers loosened from the pipet tip.
11. 30 λ of ribonuclease (5 mg/ml stock) (Boehringer-Mannheim) were added to the DNA mixture and incubated at 37°C for 1 hr with occasional shaking.
12. At this point, the pipet tip (free from DNA) was removed from the tube and 20 λ of 10% SDS solution and 20 λ of proteinase K (10 mg/ml stock) (Boehringer-Mannheim) were added and incubated at 37°C for 1-2 hrs.
13. After incubation, 200 λ of 5 M NaCl were added and mixed at 65°C for a few minutes to clarify the solution.
14. DNA solution was extracted once with 4 ml of phenol/chloroform/isoamylalcohol (25/25/1) by gentle inversion of the tube and then centrifuged at 2,000 rpm for 10 minutes as before.

15. The top aqueous phase (≈ 2 ml) was removed and aliquoted into four 1.5 ml Eppendorf tubes. The DNA was precipitated by adding 1 ml of 95% cold ethanol and the pellet was washed 2 $\times$ with 75% ethanol. After lyophilization, the pellet was suspended in H₂O and used for A₂₆₀ determination and restriction enzyme digestion as previously described.

DNA gel blot analyses

Except for the hybridization buffer (50% formamide, 6 $\times$ SSC, 0.25% nonfat dry milk; the same buffer was used for prehybridization) other methods used in restriction enzyme digestion, electrophoresis, transfer, washing and autoradiography were essentially the same as described in Chapters II and III.

Molecular probes

Preparation of the LTR and gag gene specific probes were described in Chapters II and III, respectively. An env gene specific probe was prepared by nicktranslation of a 1.2 kbp SmaI/PstI env fragment (Figure 4, map positions 5-6.2 kbp) from the pRFM#16 clone.

Total RNA extraction from neomycin-resistant cell cultures

The modified procedure of Chirgwin et al. (154) was used for total RNA extraction:

1. Approximately 0.5 g of cells were homogenized in 9 ml of guanidine thiocyanate stock solution [5 M guanidine thiocyanate (Kodak), 1% sodium-N-lauroyl sarcosine (Sigma), 25 mM sodium citrate pH 7.0,

- 0.1 M 2-mercaptoethanol (Sigma)] using a motor-driven homogenizer with a teflon pestle (5-10 strokes).
2. The homogenate was layered over a 3 ml cushion of 5.7 M CsCl, 0.1 M EDTA pH 7.4 solution in a cellulose nitrate tube (size: 9/16" dia. $\times$ 3 1/2") and centrifuged at 35,000 rpm for 20 hrs using a Beckman SW41 rotor (21°C).
 3. The supernatant was removed carefully and after complete draining of the tube, the pellet was suspended in 1 ml of 1 mM EDTA pH 7.4 solution.
 4. The suspension was transferred into a sterile polystyrene (17 $\times$ 100 mm) (Costar) tube and 71 μ l of 3 M sodium acetate pH 5.2 solution and 2.5 ml of cold absolute ethanol were added and the mixture was kept at -20°C overnight or greater than 4 hrs.
 5. After spinning at 8,000 rpm (Sorvall type SS-34 rotor) for 30 minutes (4°C), the ethanol was drained off and the pellet was suspended in 0.5 ml of cold 3 M sodium acetate pH 5.2 solution. The mixture was allowed to stand on ice for 10 minutes and then centrifuged as before.
 6. After spinning, the supernatant was drawn off and the pellet was resuspended in 1 ml sterile H₂O. This suspension was sedimented again by adding sodium acetate and ethanol, followed by centrifugation as before. The RNA pellet from the last centrifugation step was lyophilized and resuspended in 1 ml of H₂O.

7. The RNA solution was aliquoted in 1.5 ml Eppendorf tubes (0.5 ml per tube), ethanol precipitated, and then stored at -20°C until use. Each time before use the pellet was rinsed $2\times$ with 75% cold ethanol, lyophilized and suspended in a small volume of H_2O to yield concentrations of 20-40 μg RNA/10 λ . The total RNA yield from a 0.5 g cell culture varied from several hundred micrograms to a few milligrams.

RNA gel blot analysis

The RNA samples used for gel blot analysis were usually prepared by mixing 10 λ of the total RNA ($\approx 20-40$ μg RNA in H_2O) with 30 λ of a sample buffer prepared by mixing 150 λ deionized formamide, 54 λ formaldehyde (37%) and 3 λ 1 M sodium phosphate pH 6.5. This buffer was prepared fresh each time. After mixing, the samples were heated at 50°C for 15 minutes and then 4 λ of $10\times$ sample loading buffer (0.25% xylene cyanol, 0.25% bromophenol blue, 25% ficoll 400) were added. RNA samples were run in 300 ml formaldehyde-agarose submerged horizontal gels (2.2 M formaldehyde, 10 mM sodium phosphate pH 6.5, 1.2% agarose) at 100 volts for 8-10 hrs. The electrophoresis running buffer was 10 mM sodium phosphate at pH 6.5 and was circulated during the run between the two buffer compartments. After completion of the electrophoresis, the gels were soaked in $15\times$ SSC for 15-20 minutes after which RNAs were transferred to nitrocellulose filters using the same procedure described earlier for DNA transfer. Methods used in hybridization and washing of the filter membranes were also the same as described before (Chapters II and III). The hybridization (and prehybridization) buffer was 50%

formamide, 6× SSC and 0.25% nonfat dry milk. The autoradiograms were usually exposed for 7 to 10 days with an intensifying screen at -80°C.

RESULTS

Transient expression of p0.28-9.5 and pRFM#6 plasmids in cell lines from heterologous species

To examine transient expression of the recombinant plasmids described above, the plasmid DNAs were transfected into the following heterologous cell lines; mink CCL64, normal rat kidney (NRK), Chinese hamster V79, rat XC and human F160-5 (amniotic cells). 48-72 hrs following the transfection, the culture media from transfectants were collected and tested for particle-associated reverse transcriptase activity as described above. The culture medium of the mouse NIH3T3 cells transfected with p0.28-9.5 was also tested as a positive control. The results of reverse transcriptase assays are shown in Table 2. These data indicate that none of the transfected heterologous cell lines were capable of producing viral particles. This experiment, therefore, showed that neither the endogenous MuLV (p0.28-9.5), nor the endogenous MuLV-related (pRFM#6) provirus was expressed as viral particles in the heterologous cells examined. As expected, the ecotropic provirus of p0.28-9.5 plasmid transfected into the NIH3T3 cells produced viral particles which were released into the culture medium.

Table 2: Reverse transcriptase assay of the culture media from various cell lines transfected with p0.28-9.5 and pRFM#6 plasmids

Cell line		cpm ^c
CCL64 (mink)	p0.28-9.5	167
	pRFM#6	150
	(-) control ^a	164
F160-5 (human)	p0.28-9.5	155
	pRFM#6	148
	(-) control	161
XC (rat)	p0.28-9.5	247
	pRFM#6	237
	(-) control	222
V79 (Chinese hamster)	p0.28-9.5	113
	pRFM#6	149
	(-) control	144
NRK (rat)	p0.28-9.5	151
	pRFM#6	161
	(-) control	155
NIH3T3 (mouse)	p0.28-9.5 ^b	3451
	pRFM#6	158
	(-) control	159

a) Calf thymus DNA was substituted for the plasmid in transfection experiments.

b) The culture medium from NIH3T3 cells transfected with p0.28-9.5 ecotropic provirus served as a positive control. These cells continuously release the infectious viral particles into the medium.

c) Each vial containing 5 ml of 0.4% BBOT was counted for 10 minutes in a liquid scintillation counter.

Stable transfection of human fibrosarcoma HT 1080 cell line with
p0.28-9.5, pRFM#6 and pRFM#16 recombinant plasmids

After transfection of the HT 1080 cells with the recombinant plasmids described above, a total of 20 neomycin-resistant cell colonies were selected from each transfection assay (total = 60 colonies). Genomic DNAs from all these transfectants were extracted and subsequently digested to completion with HindIII restriction enzyme. After electrophoresis of the DNA fragments on agarose gel and transfer to nitrocellulose filters, integration of the supercoiled recombinant plasmids into the recipient cells could be demonstrated after hybridization to a LTR specific probe. Since human cells did not contain any background MuLV specific sequences, all detected bands were those of integrated plasmids (Figure 23). Approximately 30% of all transfectants contained one or more copies of the integrated plasmids. The remaining 70%, represented those cell clones which had picked up the neomycin-resistant gene but not any of the other recombinant plasmid. Figure 23 shows the analysis of the genomic DNAs from 20 (out of 60) neomycin-resistant cell colonies; 5 colonies transfected with p0.28-9.5, 5 transfected with pRFM#6 and 10 colonies transfected with pRFM#16 plasmid. Of those clones transfected with p0.28-9.5, only 3 cell clones p[2-5], p[3-1] and p[2-3] contained a relatively large number of integrated copies of the plasmid; the other two clones, p[1-2] and p[2-1] seemed not to harbor any integrated copies of the recombinant plasmid. From 5 pRFM#6 transfectants only two, 6[4-3] and 6[2-2] contained multiple copies of the integrated plasmid; colonies 6[3-2] and 6[2-1] apparently did not contain any integrated copy whereas 6[3-3] showed a faint band at the

Figure 23: Gel blot analysis of genomic DNAs isolated from the human HT 1080 transfectants. 10 μ g of genomic DNA extracted from each neomycin-resistant colony were digested to completion with HindIII restriction enzyme and subsequently electrophoresed in 0.7% agarose gel. The LTR-specific probe used for hybridization had a specific activity of $\approx 3 \times 10^8$ cpm/ μ g DNA and the autoradiogram was exposed overnight at -80°C with intensifying screen. Analysis of the genomic DNAs from 5 colonies transfected with p0.28-9.5; p (.....), 5 transfected with pRFM#6; #6 (.....) and 10 colonies transfected with pRFM#16; #16 (.....) are shown here.

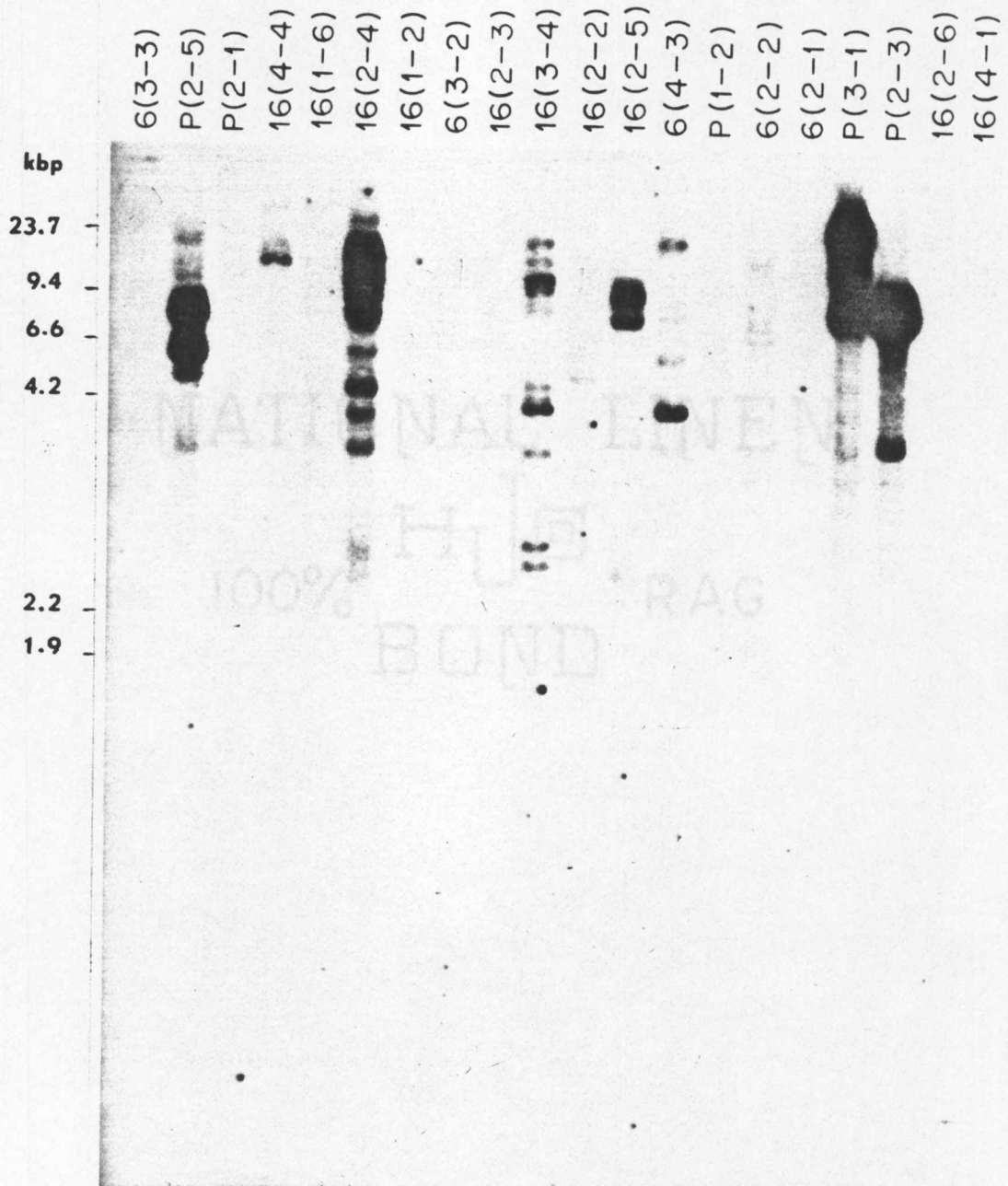


Figure 23

high MW region. Out of 10 pRFM#16 transfectants, three, 16[2-4], 16[3-4] and 16[2-5] contained multiple bands whereas 16[4-4] had a single strong band at the high MW region. Three transfectants, 16[4-1], 16[2-6] and 16[1-2] lacked any detectable bands whereas 16[1-6] and 16[2-2] contained only a single faint band. The fact that a relatively large number of the hybridizing bands seen in pRFM#6 and pRFM#16 transfectants appeared in the regions below 12.4 and 13.6 kbp (the DNA insert sizes of pRFM#6 and pRFM#16, respectively) indicated that in these cases, the supercoiled plasmid DNA opened up within the DNA insert rather than the vector (pBR322) before integration. In the latter case, HindIII digestion of the genomic DNAs should generate the expected 12.4 and 13.6 kbp bands if no other deletions or rearrangements of the insert DNA took place. This apparently led to the generation of some defective copies of the integrated MuLV-related genomes. The overall results from the transfection assays indicated random integration of the recombinant plasmid DNAs into the host genome as was predicted.

Reverse transcriptase assay of the culture media from HT 1080 transfectants

To examine whether infectious virus particles were produced from the integrated recombinant plasmids, the culture media from all 60 HT 1080 transfectants described above were tested for the presence of reverse transcriptase, an enzymatic activity present in all retroviruses. Table 3 shows the results of these assays for the 20 transfectants described in Figure 23 (the remaining transfectants were not noticeably different in their enzymatic activities from the ones listed here). It

Table 3: Reverse transcriptase assay of the culture media from human
HT 1080 stable transfectants

HT 1080 clone ^a	cpm ^d
16[1-2]	148
16[1-6]	148
16[2-2]	139
16[2-3]	145
16[2-4]	145
16[2-5]	149
16[2-6]	156
16[3-4]	139
16[4-1]	160
16[4-4]	158
6[2-1]	150
6[2-2]	152
6[3-2]	141
6[3-3]	147
6[4-3]	144
p[1-2]	138
p[2-1]	147
p[2-5]	151
p[3-1]	145
(-) Control ^b	142
(+) Control ^c	4846

- a) Only those transfectants described in Figure 23 are shown here. Enzymatic activities of other clones were not significantly different from the ones listed here.
- b) H₂O was substituted for the enzyme in mock assays.
- c) The culture medium from NIH3T3 cells transfected with p0.28-9.5 ecotropic provirus served as (+) control. These cells continuously release the infectious viral particles into the medium.
- d) Each vial containing 5 ml of 0.4% BBOT was counted for 10 minutes in a liquid scintillation counter.

can be observed that the enzymatic activities in the culture media of these transfectants were not any higher than the background level (H_2O was substituted for the enzyme in mock assays). On the other hand the medium collected from the cultures of NIH3T3 cells transfected with p0.28-9.5 plasmid [this transfectant is known to produce infectious N-tropic ecotropic MuLV and thus served as a positive control] contained significant amounts of reverse transcriptase activity. These results clearly showed that neither p0.28-9.5 nor pRFM recombinant plasmids (all of murine origin) were capable of giving rise to infectious MuLV in the heterologous human cell line.

Transcriptional activities of integrated pRFM plasmids in HT 1080 transfectants

The absence of infectious particles in the culture media of HT 1080 transfectants could be the result of several factors such as lack of retroviral transcription, translation or packaging of various retroviral components into a mature virion. Since in infectious viruses, transcriptional activity of the provirus is under the control of the LTR promoter, it was of interest to examine the potential of the LTRs in pRFM genomes with their distinct structure (i.e. presence of the [IS] element) to direct transcription of the viral RNA in a heterologous host. The Northern gel blot of the total RNA transcripts isolated from three HT 1080 transfectants is shown in Figure 24. Clone 16 [2-4] which exhibited the highest amount of MuLV-specific RNA transcripts (as detected by a gag gene specific probe), also contained the largest number of integrated copies of the recombinant plasmid (see Figure 23).

Figure 24: RNA blot analysis of three human HT 1080 transfectants and producer NIH3T3 cells (transfected with p0.28-9.5 ecotropic provirus). Clones 16[2-4] and 16[4-4] were transfected with pRFM#16 and 6[4-3] was transfected with pRFM#6 supercoiled plasmid DNAs. All clones contained one or more copies of the transfected plasmid integrated into their genomes (as shown in previous figure). $\approx 35 \mu\text{g}$ of the total RNA from human or mouse cell transfectants were electrophoresed in 1.2% agarose gel under denaturing conditions. The gag-specific probe used for hybridization had a specific activity of $\approx 5 \times 10^7$ cpm/ μg DNA and the autoradiogram was exposed for one week at -80°C with intensifying screen. At shorter exposure (≈ 16 hrs), three distinct bands could be detected in the total RNA from producer NIH3T3 cells (marked with arrows); the two lower bands comigrated with the bands seen most visibly in 16[2-4] clone whereas the top band migrated more slowly than that of 16[2-4]. Size markers are given in the left margin.

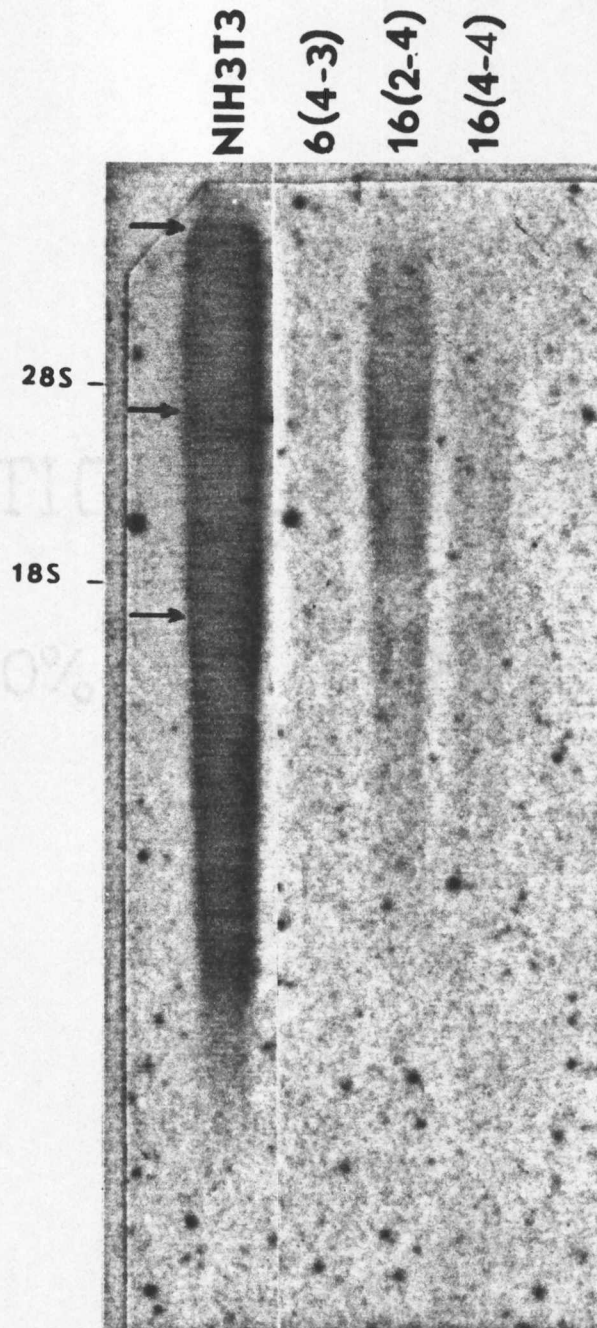


Figure 24

Little amounts of RNA transcripts were detected in two other transfectants, 16[4-4] and 6[4-3]. For comparative purposes, an equal amount of the total RNA isolated from the producer p0.28-9.5-transfected NIH3T3 cells was also run along with the RNA from HT 1080 transfectants. With a shorter exposure of the autoradiogram (≈ 16 hrs), at least 3 distinct bands (their relative positions marked with arrows) could be detected in RNA from the producer NIH3T3 cells. Two of these, $\approx 27S$ and $\approx 16S$ transcripts, were of subgenomic sizes and comigrated with the bands seen most clearly in RNA from the 16[2-4] transfectant. The largest size transcript of the 16[2-4] cell clone (a diffuse band at $\approx 32S$), however, was somewhat smaller than the genomic size ($\approx 35S$) ecotropic virus transcript seen in RNA of the producer NIH3T3 cells. This was expected from the integrated pRFM#16 plasmid whose provirus-related genome carried a deletion of ≈ 1.5 kbp in the 5' pol gene (Chapter II, Figure 4). If this band truly represented a full-size genomic transcript of pRFM#16, it would indicate that at least some of the integrated copies of the pRFM#16 genome in the 16[2-4] transfectant contained unaltered provirus (except for the 1.5 kbp deletion). This was demonstrated after digestion of the genomic DNA from the 16[2-4] transfectant with PstI restriction enzyme which generated an internal ≈ 6.3 kbp fragment detectable by gag, pol and env gene specific probes, as might have been expected from an intact pRFM#16 provirus (data not shown). From these studies, it was concluded that the LTR of these endogenous MuLV-related sequences can function in a heterologous host, but with very low efficiency.

Effect of 5-azacytidine on transcriptional activities of pRFM plasmids:
methylation state of pRFM LTRs in heterologous cells

To find out whether the low level expression of MuLV-related transcripts detected in HT 1080 transfectants was caused, at least partially, by the methylation state of integrated pRFM recombinant plasmids, the cells were treated with an inhibitor of DNA methylation, 5-azacytidine (3.5 μ M for 24 hrs) (149). When the total RNA isolated from 5-azacytidine treated cells was compared with that of untreated cells, no significant difference could be detected in the amounts of MuLV-related transcripts in the two HT 1080 transfectants examined (Figure 25). This suggested that the integrated MuLV-related proviral genomes in HT 1080 cells were probably not hypermethylated. This conclusion was further substantiated by direct examination of a specific site in the LTR prone to methylation. In eukaryotic cells, DNA is almost exclusively methylated at m^5C in the dinucleotide CpG (155), and a large portion of this dinucleotide is methylated in all cells examined so far. The presence of a SmaI site (CCCGGG) in the LTR of endogenous MuLV-related sequences (Figures 4 and 9), made it possible to examine the methylation state of CG dinucleotide at this site. When this site is not methylated, combined PstI/SmaI digestion would generate an internal LTR fragment of ≈ 600 bp. On the other hand, since SmaI is sensitive to methylation at the CG dinucleotide, methylation of this site would prevent enzyme digestion. A second restriction enzyme, KpnI which is located immediately next to the SmaI site in the LTR is not sensitive to DNA methylation and combined PstI/KpnI digestion would generate the same ≈ 600 bp fragment whether the KpnI site is methylated or not. As shown in Figure 26, combined

Figure 25: Effect of 5-azacytidine on transcriptional activities of the integrated pRFM plasmids in human HT 1080 transfectants. $\approx 35 \mu\text{g}$ of total RNA from control and 5-azacytidine-treated cells were electrophoresed in 1.2% agarose gel under denaturing conditions. The gag-specific probe used for hybridization was described in the legend of Figure 24. The autoradiogram was exposed for one week at -80°C with intensifying screen. The HT 1080 transfectants examined are shown on the top of the lanes. Size markers are given in the margin.

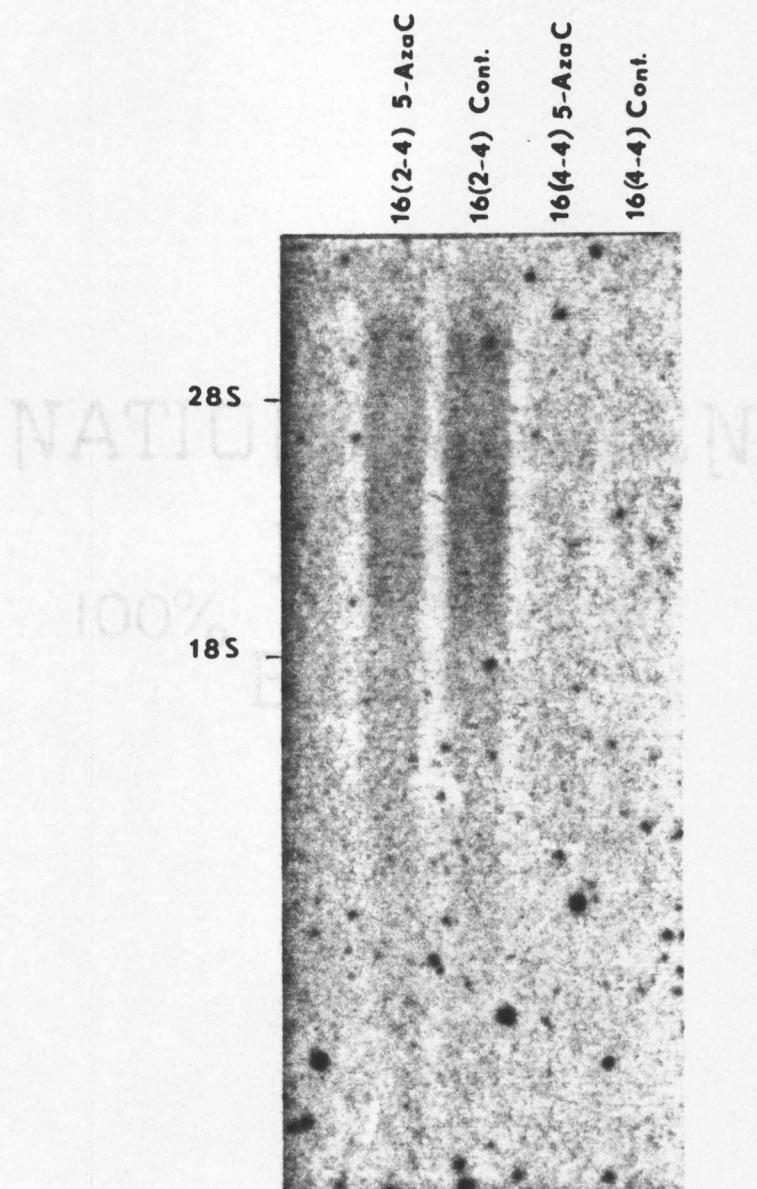


Figure 25

Figure 26: The methylation state at SmaI site in the LTRs of integrated pRFM#16 provirus in HT 1080 transfectants. 10 µg of genomic DNA from three HT 1080 transfectants (shown on the top) were digested to completion with PstI, PstI/SmaI and PstI/KpnI and analyzed in 0.7% horizontal agarose gel. The LTR probe had a specific activity of $\approx 5 \times 10^7$ cpm/µg DNA and the autoradiogram was exposed for 3 days at -80°C with intensifying screen. The comigrating bands of ≈ 600 bp generated upon PstI/SmaI and PstI/KpnI digestions are marked with arrows. The MW markers are shown on the left.

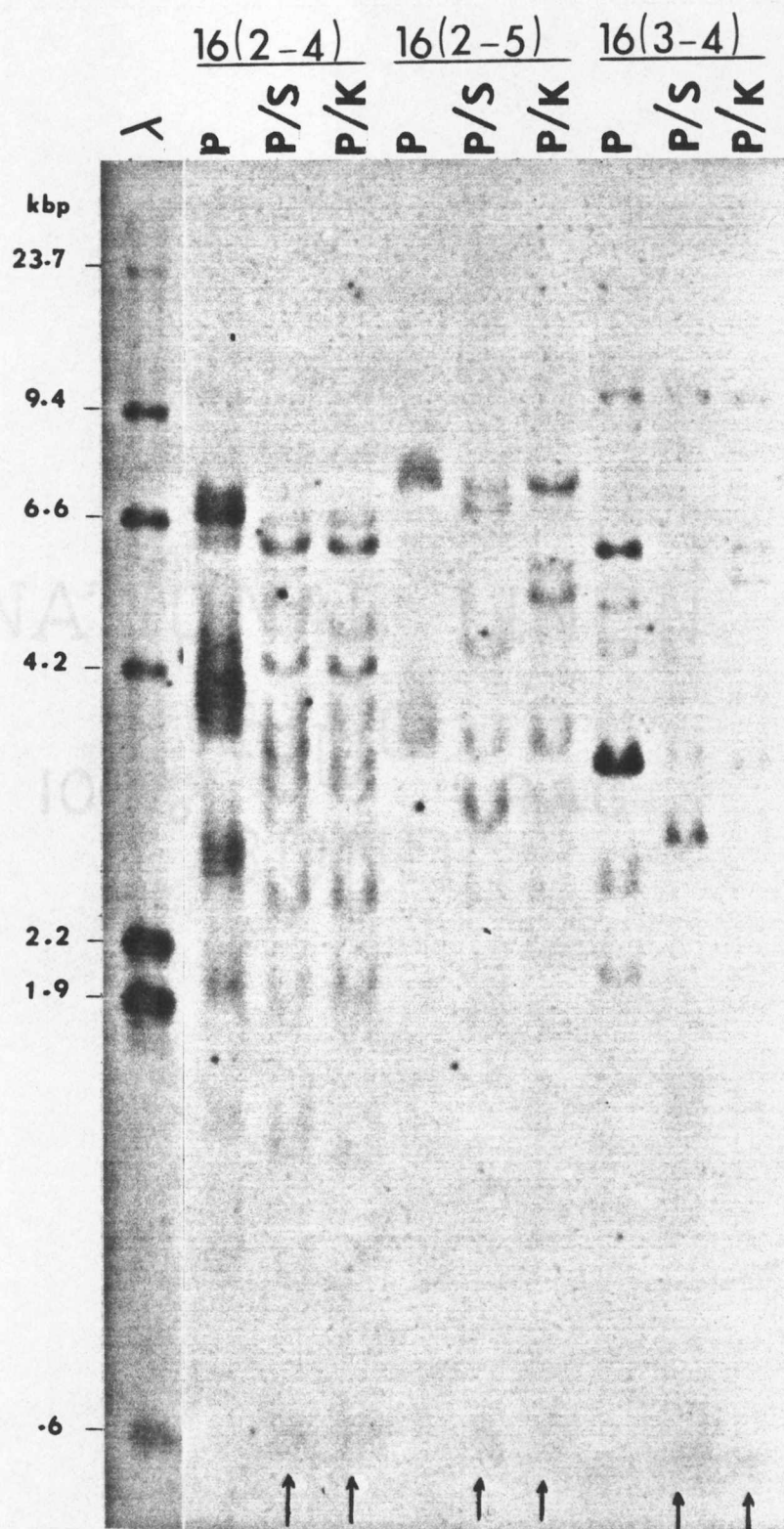


Figure 26

enzymatic digestion of the genomic DNAs from three HT 1080 transfectants, 16[2-4], 16[2-5] and 16[3-4] with PstI/SmaI and PstI/KpnI generated comigrating ≈ 600 bp bands almost with the same intensities in both digestions. This showed that the SmaI site in the LTR was not methylated, a conclusion consistent with the fact that 5-azacytidine treatment did not induce the LTR activity in HT 1080 transfectants. These results, therefore, suggest that the low level expression of pRFM plasmids in HT 1080 cells is not caused by the hypermethylation of integrated genomes but rather other factors are probably involved.

Splicing of MuLV-related transcripts in HT 1080 transfectants

As described earlier (Chapter I), proteins encoded by the gag and pol genes are translational products of genomic size 35S mRNA transcripts, whereas env proteins are synthesized from a subgenomic 24S RNA. This latter transcript (size ≈ 3.2 kb) is a spliced molecule containing a small segment from the 5'-terminus of 35S RNA transposed to the 3' half of the MuLV genome carrying the env gene. As a result of splicing, the entire gag gene and most of the pol gene are absent in the 24S transcript. In order to determine whether the 24S transcript was produced in HT 1080 transfectants, total RNAs isolated from several transfectants were hybridized to an env gene specific probe. As shown in Figure 27, the discrete 24S band of ecotropic MuLV (from producer NIH3T3 cells) comigrated with fainter bands detected in the RNA from HT 1080 transfectants. This transcript was not detectable by the gag gene specific probe (Figure 24). Additionally, a smaller band of ≈ 16 S was detected in both HT 1080 and NIH3T3 transfectants. The intensity of

Figure 27: Detection of a MuLV-specific 24S transcript in human HT 1080 transfectants. The same nitrocellulose filter of Figure 24 was deprobed from the bound gag-specific probe and reprobbed with an env-specific probe. The env-specific probe had a specific activity of $\approx 1 \times 10^7$ cpm/ μ g DNA and the autoradiogram was exposed at -80°C for one week with intensifying screen. The 24S RNA transcript is marked.

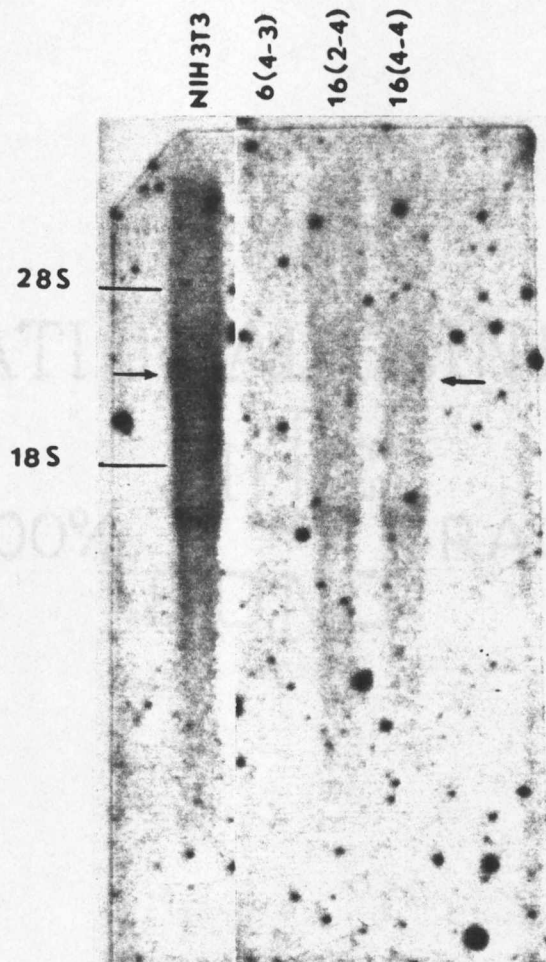


Figure 27

this band, relative to that of 24S transcript, was higher in HT 1080 than NIH3T3 transfectants. This small transcript, apparently carrying the env sequences, has not been reported previously and therefore, requires further characterization. From these data, it was concluded that MuLV-related transcripts of pRFM genomes were spliced in the HT 1080 cells apparently in the same way as the ecotropic virus transcripts in murine cells.

DISCUSSION

The experimental evidence presented here, indicate that selected pRFM clones are capable of expressing (although with low efficiency) MuLV-related transcripts in heterologous host cells. The choice of human cells for these studies was for several reasons. First, these cells do not contain any background MuLV-specific sequences, making the detection of viral transcripts rather easy. Second, it was of interest to find out how these MuLV-related sequences would function once integrated into the human genome. Third, given the fact that human genome contains both type B and type C endogenous retroviral sequences (156,157,158), these studies could possibly provide some clues as to the role and functionality of the human endogenous sequences.

Although analysis of the LTR structure in two previous studies (106,108) and in the present work on pRFM#6 (Chapter III) and pRFM#16 (data not shown) have demonstrated that endogenous MuLV-related proviral genomes do contain presumptive transcriptional signals in their intact forms (as compared with infectious virus LTRs), there is no evidence from in vivo studies to indicate that infectious retroviruses are

produced directly from these sequences. The data presented here suggest that the apparent incapability of these sequences to produce infectious retrovirus cannot be solely attributed to inherent structural alterations of the LTR, since the LTR of pRFM clones can be functional in human cells as well as in mouse cells. Furthermore, the observation that none of the HT 1080 transfectants carrying multiple copies of p0.28-9.5 or pRFM plasmids produced infectious particles would suggest that, in this case, host factors rather than viral factors may have contributed to noninfectivity of these genomes. This is indicated by the fact that p0.28-9.5 expressed significant amounts of viral-specific transcripts and infectious virus particles following transfection into NIH3T3 cells. As shown in Figures 23 and 24, among the different HT 1080 transfectants, the 16[2-4] cell clone, with the largest number of integrated copies, also showed the highest content of MuLV-specific RNA transcripts. Another cloned cell line, 16[3-4], also with relatively large number of integrated MuLV-related copies, showed a higher content of RNA transcripts than others (except 16[2-4]) (data not shown). It seemed therefore, that the level of transcriptional activities in these transfectants primarily reflected the number of integrated copies of MuLV-related sequences. Clone 16[4-4], however, with a single band in its genomic blot, contained more MuLV-specific transcripts than 6[4-3] with multiple copies of integrated genomes. This difference could not be attributed to structural differences in the LTRs of pRFM#6 and pRFM#16 clones, since except for a few bases (but not in the regulatory sequence boxes) the two LTRs had identical nucleotide sequence structures. This may therefore indicate that in addition to

the copy number, other factors such as the site of integration, unaltered structure of the integrated provirus, and the turn-over rate of transcripts could also affect the MuLV-specific RNA contents of various HT 1080 transfectants.

Although restriction map analyses of the endogenous MuLV-related sequences (Figures 4 and 9) showed that at least some of these genomes, for example pRFM#6 and possibly AL10, pRFM#1 and pRFM#9, occur as full-length proviruses in the host genome, it is curious that under normal (non-induced) conditions, no full-length transcripts are expressed. This observation may implicate a possible differential regulation of these sequences by the host cell, namely, that only those genomes which do not pose any immediate threat to the cell are allowed to be expressed. Transcriptional control of endogenous retroviral genomes, in particular full-length proviruses, with the potential of infecting the host cell may be achieved not only by nearby cellular inhibitory sequences but also by differential methylation of these genomes. In fact, recent studies in our laboratory have shown that some of these endogenous MuLV-related sequences are methylated at the SmaI site in the LTR. Whether this methylation involved only proviruses with full-length genomes has not yet been determined. If methylation is actually one of the factors rendering these sequences transcriptionally inactive, then cloned MuLV-related genomes with minimal or no methylation may exhibit more activity once introduced into a new host, provided they remain undermethylated in the new environment. In Figure 26, it is shown that in all three HT 1080 transfectants examined; 16[2-4], 16[2-5] and 16[3-4], the LTRs of the integrated MuLV-related genomes are

unmethylated at the specific SmaI site. Furthermore, when two of these clones, 16[2-4] and 16[3-4], were examined for the expression of transfected proviral genomes, not only low but significant levels of RNA transcripts was observed in both (Figure 24; for 16[2-4]), but also similar type of transcripts as those expressed by an infectious ecotropic murine leukemia virus were detected (Figures 24 and 27). One of these transcripts was a spliced 24S RNA molecule for the synthesis of envelope proteins. The methylation state of the LTR in one transfectant, 6[4-3], with low content of RNA transcripts (Figures 24 and 27) has not been determined.

In addition to possible involvements of inhibitory cellular sequences and the methylation state in the regulation of the expression of endogenous MuLV-related sequences, point mutations can also contribute to the noninfectivity of these genomes by altering the reading frame of viral-specific transcripts. This possibility can be tested in several ways; for example, by examining the translational products of the MuLV-related transcripts such as those described above, or by determining the complete nucleotide sequence of a full-length MuLV-related genome such as that of pRFM#6 provirus.

CHAPTER V

SUMMARY

This dissertation research involved the collection of experimental data on the structure and biological properties of the selected endogenous MuLV-related DNA sequences from the RFM/Un mouse. Structural analyses of these DNAs involved specific regions of the provirus important in the regulation of transcription (i.e. the LTRs), retroviral replication (the tRNA primer binding site and the endonuclease gene) and host-range type specificity (the 5'-env gene). Studies on the structure of tRNA primer binding site for minus strand DNA synthesis showed that four out of five pRFM proviruses contained sequences with perfect complementary matches (18/18) to the 3'-terminal 18-nucleotide sequence of a major glutamine tRNA isoacceptor from the rat liver. The primer binding site of one cloned DNA, i.e. pRFM#17 was only 16/18 complementary to the same tRNA. From these results, it was concluded that the tRNA primer binding sites in these genomes are distinct from those of infectious murine leukemia viruses which utilize a proline tRNA molecule to initiate (-) strand DNA synthesis. These structural differences apparently had not been caused by random point mutations, since strong nucleotide sequence homologies with infectious proviruses were observed in the regions adjacent to the primer binding site in these genomes. Two possible mechanisms explaining the switch from proline tRNA to glutamine tRNA binding sites were presented.

Restriction endonuclease analyses of the pRFM recombinant clones showed that three clones contained partial proviral sequences due to the

presence of an internal HindIII restriction site in their structural genes. Two of these, pRFM#1 and pRFM#9, were probably present originally as full-length proviruses in their host genome whereas the third, pRFM#17 contained a deletion of ≈ 1.5 kbp in the 5'-pol gene. Two other clones, pRFM#16 and pRFM#6 harbored MuLV-related genomes with structural arrangements similar to those of the integrated infectious proviruses, that is, 5'-LTR-gag-pol-env-LTR-3'. Of these two, pRFM#6 seemed to carry a full-size MuLV-related genome with no detectable structural deletions; pRFM#16 contained a deletion of ≈ 1.5 kbp roughly at the same map position as that of pRFM#17. The MuLV-related genomes of both pRFM#16 and pRFM#17 recombinant clones, therefore, represented defective proviruses.

Since the provirus-related DNA sequence of pRFM#6 contained a full-length genome, further structural characterization of this genome was carried out. Nucleotide sequence analysis of the 0.7 kbp long terminal repeats revealed identical structures for 5' and 3' LTRs except for two base transitions and one nucleotide deletion/addition all in the U3 region. When compared with the LTR of RFV, the endogenous infectious ecotropic virus induced from the RFM/Un mouse, strong homologies were observed only in the U5 and R regions; the U3 region of pRFM#6 LTR was distinctly different from that of the RFV provirus, mainly as a result of the insertion of a novel 170-200 bp sequence segment in this region. The transcriptional regulatory elements on the other hand, were identical (except for one base change in the "CAT" box) in RFV and pRFM#6 LTRs and were present at nearly the same distances from the mRNA cap site. Structural analysis of the proviral integration site

indicated that the provirus of pRFM#6 had integrated into the host genome through the same insertional mechanism as that of exogenous murine leukemia proviruses or the retrotransposable gene elements.

The host-range type of the pRFM#6 provirus was defined through the analysis of nucleotide and predicted amino acid sequences of the 5'-env gene (gp70). The results obtained from hybridization studies using host-range type specific probes as well as those from nucleotide sequence analysis, predicted a host-range of the MCF-type for the pRFM#6 provirus. Similar results were obtained from studies on the pRFM#16 provirus. This indicated that the env genes of pRFM or similar MuLV-related genomes (of which 20-50 copies per haploid genome are present in mice) may represent endogenous mouse sequences which participate in the recombinational events with exogenous or endogenous ecotropic MuLVs, leading to the generation of the highly leukemogenic MCF viruses.

The provirus-related genome of pRFM#6 is also present in the same cellular DNA site in four other inbred strains (i.e. BALB/c, AKR, C3H and CBA) but is absent in C57BL/6 and NFS/N mice. Analysis of the genomic DNAs from the latter two strains indicated that this provirus had never integrated at this particular DNA site in the germ line of these mouse strains.

Preliminary studies on transcriptional activities and infectious properties of two pRFM recombinant clones; pRFM#6 and pRFM#16, demonstrated that the LTRs of both proviruses were functional in a heterologous human cell line, but presumably not as efficient as the LTR of an ecotropic provirus present in a murine system. Furthermore, integrated pRFM plasmids expressed subgenomic transcripts comparable in

size to those of the ecotropic MuLV provirus, indicating a similar splicing mechanism in the murine and human cells. The levels of transcriptional activities seemed to be determined primarily by the number of integrated copies of the MuLV-related provirus. Although the selected pRFM clones were transcriptionally active in human cells, no viral particles were detected in the culture media of transfected cells. The lack of virus production could be in part due to host factor(s), as judged by the observation that the same ecotropic provirus, highly infectious in murine cells, did not give rise to viral particles in human cells.

Based on the experimental data presented here as well as those by others on the structure and biological properties of endogenous murine leukemia viruses, the following distinctions may be made between "endogenous infectious" and MuLV-related proviral DNA sequences.

1. The LTR and structural genes of endogenous MuLV-related genomes contain a number of common restriction sites which are absent in infectious proviruses.
2. MuLV-related genomes, if they replicate, will utilize glutamine tRNA rather than proline tRNA to prime minus strand DNA synthesis.
3. The LTRs of endogenous MuLV-related genomes contain a novel 170-200 bp sequence segment in the mid-U3 region which is absent in infectious proviruses.
4. In contrast to infectious ecotropic and xenotropic proviruses, the env gene of endogenous MuLV-related genomes is of the MCF-type.

5. The LTR of endogenous MuLV-related genomes is transcriptionally active in both murine and nonmurine cells. The lack of virus production under normal or induced conditions might be a result of the differential control of full-size genomes by the host cell. Whether viral factors are also involved in such regulation remains to be determined.

The data presented here should aid us to understand the physiological role of endogenous MuLV-related sequences. Future studies should involve more detailed structural and functional characterization of these genomes. The association of endogenous MuLVs with the spontaneous development of leukemia in mice is now well documented. Whether endogenous MuLV-related and LTR-IS families are associated with tumors involving other tissues such as the liver, kidney and ovary remains to be explored.

LIST OF REFERENCES

LIST OF REFERENCES

1. Fenner, F. (1976). The classification and nomenclature of viruses. Summary of results of meetings of the international committee on taxonomy of viruses in Madrid, September 1975. *Virology* 71, 371-378.
2. Fenner, F. (1975). The classification and nomenclature of viruses. *Intervirology* 6, 1-12.
3. Bernhard, W. (1958). Electron microscopy of tumor cells and tumor viruses. A review. *Cancer Res.* 18, 491-509.
4. Bernhard, W. (1960). The detection and study of tumor viruses with the electron microscope. *Cancer Res.* 20, 712-727.
5. Dalton, A. J. (1972). RNA tumor viruses - Terminology and ultra-structural aspects of virion morphology and replication. *J. Natl. Cancer Inst.* 49, 323-328.
6. Bender, W. and Davidson, N. (1976). Mapping of polyA sequences in the electron microscope reveals unusual structure of type C oncornavirus RNA molecules. *Cell* 7, 595-607.
7. Beemon, K. L., Faras, A. J., Haase, A. T., Duesberg, P. H., and Maisel, J. E. (1976). Genomic complexities of murine leukemia and sarcoma, reticuloendotheliosis and visna viruses. *J. Virol.* 17, 525-537.
8. Duesberg, P. H. (1968). Physical properties of Rous sarcoma virus RNA. *Proc. Natl. Acad. Sci. USA* 60, 1511-1518.
9. Duesberg, P. H., Vogt, P. K., Beemon, K., and Lai, M. (1974). Avian RNA tumor viruses: mechanism of recombination and complexity of the genome. *Cold Spring Harbor Symp. Quant. Biol.* 39, 847-857.
10. Van Beveren, C., Goddard, J. G., Berns, A., and Verma, I. (1980). Structure of Moloney murine leukemia viral DNA: Nucleotide sequence of the 5' long terminal repeat and adjacent cellular sequences. *Proc. Natl. Acad. Sci. USA* 77, 3307-3311.
11. Sutcliffe, J. G., Shinnick, T. M., Verma, I., and Lerner, R. A. (1980). Nucleotide sequence of Moloney leukemia virus: 3' end reveals details of replication, analogy to bacterial transposons, and an unexpected gene. *Proc. Natl. Acad. Sci. USA* 77, 3302-3306.
12. Shinnick, T. M., Lerner, R. A., and Sutcliffe, J. G. (1981). Nucleotide sequence of Moloney murine leukemia virus. *Nature* 293, 543-548.

13. Gilboa, E., Mitra, S. W., Goff, S., and Baltimore, D. (1979). A detailed model of reverse transcription and tests of crucial aspects. *Cell* 18, 93-100.
14. Van Zaane, D., Gielkens, A. L. J., Hesselink, W. G., and Bloemers, H. P. J. (1977). Identification of Rauscher murine leukemia virus-specific mRNAs for the synthesis of gag- and env-gene products. *Proc. Natl. Acad. Sci. USA* 74, 1855-1859.
15. Rothenberg, E., Donoghue, D. J., and Baltimore, D. (1978). Analysis of a 5' leader sequence on murine leukemia virus 21S RNA: Heteroduplex mapping with long reverse transcriptase products. *Cell* 13, 435-451.
16. Mann, R., Mulligan, R. C., and Baltimore, D. (1983). Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell* 33, 153-159.
17. Gillespie, D., Marshall, S., and Gallo, R. C. (1972). RNA of RNA tumor viruses contains poly A. *Nat. New Biol.* 236, 227-231.
18. Lai, M. M. C. and Duesberg, P. H. (1972). Adenylic acid-rich sequence in RNAs of Rous sarcoma virus and Rauscher mouse leukemia virus. *Nature* 235, 383-386.
19. Furuichi, Y., Shatkin, A. J., Stavnezer, E., and Bishop, J. M. (1975). Blocked, methylated 5'-terminal sequence in avian sarcoma virus RNA. *Nature* 257, 618-620.
20. Rose, J. K., Haseltine, W. A., and Baltimore, D. (1976). 5'-terminus of Moloney murine leukemia virus 35S RNA is m⁷G^{5'}GmpCp. *J. Virol.* 20, 324-329.
21. Stoltzfus, C. M. and Dimock, K. (1976). Evidence for methylation of B77 avian sarcoma virus genome RNA subunits. *J. Virol.* 18, 586-595.
22. Dimock, K. and Stoltzfus, C. M. (1978). Cycloleucine blocks 5'-terminal and internal methylations of avian sarcoma virus genome RNA. *Biochemistry* 17, 3627-3632.
23. Bishop, J. M. (1978). Retroviruses. *Annu. Rev. Biochem.* 47, 35-88.
24. Sawyer, R. C. and Dahlberg, J. E. (1973). Small RNAs of Rous sarcoma virus: characterization by two-dimensional polyacrylamide gel electrophoresis and fingerprint analysis. *J. Virol.* 12, 1226-1237.

25. Bishop, J. M., Levinson, W. E., Quintrell, N., Sullivan, D., Fanshier, L., and Jackson, J. (1970). The low molecular weight RNAs of Rous sarcoma virus. 1. The 4S RNA. *Virology* 42, 182-195.
26. Walker, T. A., Pace, N. R., Erikson, R. L., Erikson, E. and Behr, F. (1974). The 7S RNA common to oncornaviruses and normal cells is associated with polyribosomes. *Proc. Natl. Acad. Sci. USA* 71, 3390-3394.
27. Temin, H. M. (1976). The DNA provirus hypothesis. *Science* 192, 1075-1080.
28. Harada, F., Peters, G. G. and Dahlberg, J. E. (1979). The primer tRNA for Moloney murine leukemia virus DNA synthesis. Nucleotide sequence and aminoacylation of tRNA^{Pro}. *J. Biol. Chem.* 254, 10979-10985.
29. Waters, L. C. (1975). Transfer RNAs associated with 70S RNA of AKR murine leukemia virus. *Biochem. Biophys. Res. Commun.* 65, 1130-1136.
30. Taylor, J. M. and Illmensee, R. (1975). Site on the RNA of an avian sarcoma virus at which primer is bound. *J. Virol.* 16, 553-558.
31. Grandgenett, D. P., Gerard, G. F., and Green, M. (1972). Ribonuclease H: A ubiquitous activity in virions of ribonucleic acid tumor viruses. *J. Virol.* 10, 1136-1142.
32. Verma, J. M. (1975). Studies on reverse transcriptase of RNA tumor viruses. III. Properties of purified Moloney murine leukemia virus DNA polymerase and associated RNase H. *J. Virol.* 15, 843-854.
33. Verma, I. M. (1977). The reverse transcriptase. *Biochim. Biophys. Acta* 473, 1-38.
34. Dhar, R., McClements, W. L., Enquist, L. W., and Vande Woude, G. F. (1980). Nucleotide sequences of integrated Moloney sarcoma provirus long terminal repeats and their host and viral junctions. *Proc. Natl. Acad. Sci. USA* 77, 3937-3941.
35. Ju, G. and Skalka, A. M. (1980). Nucleotide sequence analysis of the long terminal repeat (LTR) of avian retroviruses: structural similarities with transposable elements. *Cell* 22, 379-386.
36. Hughes, S. E., Shank, P. R., Spector, D. H., Kung, H. J., Bishop, J. M., Varmus, H. E., Vogt, P. K., and Breitman, M. L. (1978). Proviruses of avian sarcoma virus are terminally redundant, co-extensive with unintegrated linear DNA and integrated at many sites. *Cell* 15, 1397-1410.

37. Taylor, J. M., Hsu, T. W., Yeater, C., and Mason, W. S. (1979). Synthesis and integration of avian sarcoma virus DNA. Cold Spring Harbor Symp. Quant. Biol. 44, 1091-1098.
38. Temin, H. M. (1982). Function of the retrovirus long terminal repeat. Cell 28, 3-5.
39. Temin, H. M. (1981). Structure, variation and synthesis of retrovirus long terminal repeat. Cell 27, 1-3.
40. Chen, H. R. and Barker, W. C. (1984). Nucleotide sequences of the retroviral long terminal repeats and their adjacent regions. Nucl. Acids Res. 12, 1767-1778.
41. Devare, S. G., Reddy, E. P., Law, J. D., and Aaronson, S. A. (1982). Nucleotide sequence analysis of the long terminal repeat of integrated simian sarcoma virus: Evolutionary relationship with other mammalian retroviral long terminal repeats. J. Virol. 42, 1108-1113.
42. Reddy, E. P., Smith, M. J., Canaani, E., Robbins, K. C., Tronick, S. R., Fain, S., and Aaronson, S. A. (1980). Nucleotide sequence analysis of the transforming region and large terminal redundancies of Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. USA 77, 5234-5238.
43. Colicilli, J. and Goff, S. P. (1985). Mutants and pseudorevertants of Moloney murine leukemia virus with alterations at the integration site. Cell 42, 573-580.
44. Panganiban, A. T. and Temin, H. M. (1983). The terminal nucleotides of retrovirus DNA are required for integration but not virus production. Nature 306, 155-160.
45. Duyk, G., Longiaru, M., Cobrinik, D., Kowal, R., DeHaseth, P., Skalka, A. M., and Leis, J. (1985). Circles with two tandem terminal repeats are specifically cleaved by pol gene-associated endonuclease from avian sarcoma and leukosis viruses: Nucleotide sequences required for site-specific cleavage. J. Virol. 56, 589-599.
46. Hayward, W. S., Neel, B. G., and Astrin, S. M. (1981). Activation of cellular onc gene by promoter insertion in ALV-induced lymphoid leukosis. Nature 290, 475-480.
47. Neel, B. G., Hayward, W. S., Robinson, H. L., Fang, J., and Astrin, S. M. (1981). Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. Cell 23, 323-334.

48. Nusse, R. and Varmus, H. E. (1982). Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* 31, 99-109.
49. Payne, G. S., Bishop, J. M., and Varmus, H. E. (1982). Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas. *Nature* 295, 209-214.
50. Dinowitz, M. (1975). Inhibition of Rous sarcoma virus by -amanitin: Possible role of cell DNA-dependent RNA polymerase form II. *Virology* 86, 1-9.
51. Rymo, L., Parsons, J. T., Coffin, J. M., and Weissmann, C. (1974). In vitro synthesis of Rous sarcoma virus-specific RNA is catalyzed by a DNA-dependent RNA polymerase. *Proc. Natl. Acad. Sci.* 71, 2782-2786.
52. Arcement, L. J., Karshin, W. L., Naso, R. B., and Arlinghaus, R. B. (1977). "gag" polyprotein precursors of Rauscher murine leukemia virus. *Virology* 81, 284-297.
53. Evans, L. H., Dresler, S., and Kabat, D. (1977). Synthesis and glycosylation of polyprotein precursors to the internal core proteins of Friend murine leukemia virus. *J. Virol.* 24, 865-874.
54. Van Zaane, D., Gielkens, A. L. J., Hesselink, W. G., and Bloemers, H. P. J. (1977). Identification of Rauscher murine leukemia virus-specific mRNAs for the synthesis of gag- and env-gene products. *Proc. Natl. Acad. Sci. USA* 74, 1855-1859.
55. Gregoriades, A. and Old, L. J. (1969). Isolation and some characteristics of a group-specific antigen of the murine leukemia viruses. *Virology* 37, 189-202.
56. Schafer, W., Anderer, F. A., Bauer, H., and Pister, L. (1969). Studies on mouse leukemia viruses. I. Isolation and characterization of a group-specific antigen. *Virology* 38, 387-394.
57. Barbacid, M. and Aaronson, S. A. (1978). Membrane properties of the gag gene-coded p15 protein of mouse type-C RNA tumor viruses. *J. Biol. Chem.* 253, 1408-1414.
58. Pal, B. K., McAllister, R. M., Gardner, M. B., and Roy-Burman, P. (1975). Comparative studies on the structural phosphoproteins of mammalian type C viruses. *J. Virol.* 16, 123-131.
59. Schulein, M., Burnette, W. N., and August, J. T. (1978). Stoichiometry and specificity of binding of Rauscher oncovirus 10,000-dalton (p10) structural protein to nucleic acids. *J. Virol.* 26, 54-60.

60. Bolognesi, D. P., Luftig, R., and Shaper, J. H. (1973). Localization of RNA tumor virus polypeptides. I. Isolation of further virus substructures. *Virology* 56, 549-564.
61. Jamjoom, G. A., Naso, R. B., and Arlinghaus, R. B. (1977). Further characterization of intracellular precursor polyproteins of Rauscher leukemia virus. *Virology* 78, 11-34.
62. Witte, O. N. and Baltimore, D. (1978). Relationship of retrovirus polyprotein cleavages to virion maturation studied with temperature-sensitive murine leukemia virus mutant. *J. Virol.* 26, 750-761.
63. Murphy, E. C., Jr., Kopchick, J. J., Watson, K. F., and Arlinghaus, R. B. (1978). Cell-free synthesis of a precursor polyprotein containing both gag and pol gene products by Rauscher murine leukemia virus 35S RNA. *Cell* 13, 359-369.
64. Schwartzberg, P., Colicelli, J., and Goff, S. P. (1984). Construction and analysis of deletion mutations in the pol gene of Moloney murine leukemia virus: A new viral function required for productive infection. *Cell* 37, 1043-1052.
65. Levin, J. G., Hu, S. C., Rein, A., Messer, L. I., and Gerwin, B. I. (1984). Murine leukemia virus mutant with a frame-shift in the reverse transcriptase coding region: Implications for pol gene structure. *J. Virol.* 51, 470-478.
66. Witte, O. N. and Wirth, D. F. (1979). Structure of the murine leukemia virus envelope glycoprotein precursor. *J. Virol.* 29, 735-743.
67. Fan, H. and Verma, I. M. (1978). Size analysis and relationship of murine leukemia virus specific mRNA's: Evidence for transposition of sequences during synthesis and processing of subgenomic mRNA. *J. Virol.* 26, 468-478.
68. Famulari, N. G., Buchhagen, D. L., Klenk, H. D., and Fleissner, E. (1976). Presence of murine leukemia virus envelope proteins gp70 and p15(E) in a common polyprotein of infected cells. *J. Virol.* 20, 501-508.
69. van Zaane, D., Dekker, Michielsen, M. J. A., and Bloemers, H. P. J. (1976). Virus-specific precursor polypeptides in cells infected with Rauscher leukemia virus: Synthesis, identification and processing. *Virology* 75, 113-129.
70. Van de Ven, W. J. M., Vermorken, A. J. M., Onnekink, C., Bloemers, H. P. J., and Bloemendal, H. (1978). Structural studies on Rauscher murine leukemia virus: Isolation and characterization of viral envelopes. *J. Virol.* 27, 595-603.

71. Bolognesi, D. P., Collins, J. J., Leis, J. P., Moenig, V., Schafer, W., and Atkinson, P. H. (1975). Role of carbohydrate in determining the immunochemical properties of the major glycoprotein (gp71) of Friend murine leukemia virus. *J. Virol.* 16, 1453-1463.
72. DeLarco, J. and Todaro, G. J. (1976). Membrane receptors for murine leukemia viruses: Characterization using the purified viral envelope glycoprotein, gp71. *Cell* 8, 365-371.
73. Hartley, J. W., Wolford, N. K., Old, L. J., and Rowe, W. P. (1977). A new class of murine leukemia virus associated with development of spontaneous lymphomas. *Proc. Natl. Acad. Sci. USA* 74, 789-792.
74. Elder, J. H., Gautsch, J. W., Jensen, F. C., Lerner, R. A., Hartley, J. W., and Rowe, W. P. (1977). Biochemical evidence that MCF murine leukemia viruses are envelope (env) gene recombinants. *Proc. Natl. Acad. Sci. USA* 74, 4676-4680.
75. Shih, T. Y., Weeks, M. O., Troxler, D. H., Coffin, J. M., and Scolnick, E. M. (1978). Mapping host range-specific oligonucleotides within genomes of the ecotropic and mink cell focus-inducing strains of Moloney murine leukemia virus. *J. Virol.* 26, 71-83.
76. O'Donnell, P. V. and Nowinski, R. C. (1980). Serological analysis of antigenic determinants on the env gene products of AKR dualtropic (MCF) murine leukemia viruses. *Virology* 107, 81-88.
77. Chien, Y. H., Verma, I. M., Shih, T. Y., Scolnick, E. M., and Davidson, N. (1978). Heteroduplex analysis of the sequence relations between the RNAs of mink cell focus-inducing and murine leukemia viruses. *J. Virol.* 28, 352-360.
78. Cloyd, M. W., Hartley, J. W., and Rowe, W. P. (1980). Lymphomagenicity of recombinant mink cell focus-inducing murine leukemia viruses. *J. Exp. Med.* 151, 542-551.
79. Hartley, J. W., Rowe, W. P., and Huebner, R. J. (1970). Host-range restrictions of murine leukemia viruses in mouse embryo cell cultures. *J. Virol.* 5, 221-225.
80. Rowe, W. P. and Sato, H. (1973). Genetic mapping of the Fv-1 locus of the mouse. *Science* 180, 640-641.
81. Yang, W. K., Kiggans, J. O., Yang, D. M., Ou, C. Y., Tennant, R. W., Brown, A., and Bassin, R. H. (1980). Synthesis and circularization of N- and B-tropic retroviral DNA in Fv-1 permissive and restrictive mouse cells. *Proc. Natl. Acad. Sci. USA* 77, 2994-2998.

82. Jolicoeur, P. and Baltimore, D. (1976). Effect of Fv-1 gene product: on proviral DNA formation and integration in cells infected with murine leukemia viruses. *Proc. Natl. Acad. Sci. USA* 73, 2236-2240.
83. Jolicoeur, P. and Rassart, E. (1980). Effect of Fv-1 gene product on synthesis of linear and supercoiled viral DNA in cells infected with murine leukemia virus. *J. Virol.* 33, 183-195.
84. Hopkins, N., Schindler, J., and Hynes, R. (1977). Six NB-tropic murine leukemia viruses derived from a B-tropic virus of BALB/c have altered p30. *J. Virol.* 21, 309-318.
85. Tennant, R. W., Otten, J. A., Brown, A., Yang, W. K., and Kennel, S. J. (1979). Characterization of Fv-1 host range strains of murine retroviruses of titration and p30 protein characteristics. *Virology* 99, 349-357.
86. Boone, L. R., Myer, F. E., Yang, D. M., Ou, C. Y., Koh, C. K., Roberson, L. E., Tennant, R. W., and Yang, W. K. (1973). Reversal of Fv-1 host range by in vitro restriction endonuclease fragment exchange between molecular clones of N-tropic and B-tropic murine leukemia virus genomes. *J. Virol.* 48, 110-119.
87. Ou, C. Y., Boone, L. R., Koh, C. K., Tennant, R. W., and Yang, W. K. (1983). Nucleotide sequences of gag-pol regions that determine the Fv-1 host range property of BALB/c N-tropic and B-tropic murine leukemia viruses. *J. Virol.* 48, 779-784.
88. Chattopadhyay, S. K., Lowy, D. R., Teich, N. M., Levine, A. S., and Rowe, W. P. (1974). Qualitative and quantitative studies of AKR-type murine leukemia virus sequences in mouse DNA. *Cold Spring Harbor Symp. Quant. Biol.* 39, 1085-1101.
89. Jaenisch, R. (1976). Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. *Proc. Natl. Acad. Sci. USA* 73, 1260-1264.
90. Weinberg, R. A. and Steffen, D. L. (1981). Regulation of expression of the integrated retrovirus genome. *J. gen. Virol.* 54, 1-8.
91. Steffen, D. L., Mural, R., Cowing, D., Mielcarz, J., Young, J., and Roblin, R. (1982). Most of the murine leukemia virus sequences in the DNA of NIH/Swiss mice consist of two closely related proviruses, each repeated several times. *J. Virol.* 43, 127-135.
92. Dolberg, D. S., Bacheler, L., and Fan, H. (1981). Endogenous type C retroviral sequences of mice are organized in a small number of virus-like classes and have been acquired recently. *J. Virol.* 40, 96-106.

93. Kozak, C. A. (1985). Retroviruses as chromosomal genes in the mouse. *Adv. Cancer Res.* 44, 295-336.
94. Aaronson, S. A., Anderson, G. R., Dunn, C. Y., and Robbins, K. C. (1974). Induction of type-C RNA virus by cycloheximide: increased expression of virus-specific RNA. *Proc. Natl. Acad. Sci. USA* 71, 3941-3945.
95. Lowy, D. R., Rowe, W. P., Teich, N., and Hartley, J. W. (1971). Murine leukemia virus: High-frequency activation in vitro by 5-iododeoxyuridine and 5-bromodeoxyuridine. *Science* 174, 155-156.
96. Rowe, W. P., Hartley, J. W., Lander, M. R., Pugh, W. E., and Teich, N. (1971). Noninfectious AKR mouse embryo cell lines in which each cell has the capacity to be activated to produce infectious murine leukemia virus. *Virology* 46, 866-876.
97. Otten, J. A., Quarles, J. M., and Tennant, R. W. (1976). Cell division requirement for activation of murine leukemia virus in cell culture by irradiation. *Virology* 70, 80-87.
98. Aaronson, S. A. and Dunn, C. Y. (1974). High-frequency C-type virus induction by inhibitors of protein synthesis. *Science* 183, 422-424.
99. Kozak, C. A. and Rowe, W. P. (1980). Genetic mapping of the ecotropic virus-inducing locus Akv-2 of the AKR mouse. *J. Exp. Med.* 152, 1419-1423.
100. Rowe, W. P., Hartley, J. W., and Bremner, T. (1972). Genetic mapping of a murine leukemia virus-inducing locus of AKR mice. *Science* 178, 860-862.
101. Kozak, C. A. and Rowe, W. P. (1982). Genetic mapping of ecotropic murine leukemia virus-inducing loci in six inbred strains. *J. Exp. Med.* 155, 524-534.
102. Stockert, E., Old, L. J., and Boyse, E. A. (1971). The G_{IX} system. A cell surface allo-antigen associated with murine leukemia viruses; implications regarding chromosomal integration of the viral genome. *J. Exp. Med.* 133, 1334-1355.
103. Lerner, R. A., Wilson, C. B., Del Villano, B. C., McConahey, P. J., and Dixon, F. J. (1976). Endogenous oncornaviral gene expression in adult and fetal mice: Quantitative, histologic, and physiologic studies of the major viral glycoprotein, gp70. *J. Exp. Med.* 143, 151-166.

104. Morse, H. C., III, Chused, T. M., Boehm-Truitt, M., Mathieson, B. J., Sharrow, S. O., and Hartely, J. W. (1979). XenCSA: cell surface antigens related to the major glycoproteins (gp70) of xenotropic murine leukemia viruses. *J. Immunol.* 122, 443-454.
105. O'Neill, R. R., Khan, A. S., Hoggan, M. D., Hartley, J. W., Martin, M. A., and Repaske, R. (1986). Specific hybridization probes demonstrate fewer xenotropic than mink cell focus-forming murine leukemia virus env-related sequences in DNAs from inbred laboratory mice. *J. Virol.* 58, 359-366.
106. Ou, C. Y., Boone, L. R., and Yang, W. K. (1983). A novel sequence segment and other nucleotide structural features in the long terminal repeat of a BALB/c mouse genomic leukemia virus-related DNA clone. *Nucl. Acids Res.* 11, 5603-5620.
107. Kahn, A. S., Rowe, W. P., and Martin, M. A. (1982). Cloning of endogenous murine leukemia virus-related sequences from chromosomal DNA of BALB/c and AKR/J mice: identification of an env progenitor of AKR-247 mink cell focus-forming proviral DNA. *J. Virol.* 44, 625-636.
108. Kahn, A. S. and Martin, M. A., (1983). Endogenous murine leukemia proviral long terminal repeats contain a unique 190 base-pair insert. *Proc. Natl. Acad. Sci. USA* 80, 2699-2703.
109. Eiden, J. J., Quade, K., and Nichols, J. L. (1976). Interaction of tryptophan tRNA with Rous sarcoma virus 35S RNA. *Nature* 259, 245-247.
110. Finston, W. I. and Champoux, J. J. (1984). RNA-primed initiation of Moloney murine leukemia virus plus strand by reverse transcriptase in vitro. *J. Virol.* 51, 26-33.
111. Varmus, H. and Swanstrom, R. (1985). Replication of retrovirus. In *RNA Tumor Virus*, edited by R. Weiss, N. Teich, H. Varmus, and J. Coffin. pp. 75-134, Cold Spring Harbor Laboratory, New York.
112. Herr, W. (1984). Nucleotide sequence of AKV murine leukemia virus. *J. Virol.* 49, 471-478.
113. Hampe, A., Gobet, M., Even, J., Sherr, C. J., and Galibert, F. (1983). Nucleotide sequences of feline sarcoma virus long terminal repeats and 5' leaders show extensive homology to those of other mammalian retroviruses. *J. Virol.* 45, 466-472.
114. Peters, G. G. and Glover, C. (1980). Low-molecular-weight RNAs and initiation of RNA-directed DNA synthesis in avian reticulo-endotheliosis virus. *J. Virol.* 33, 708-716.

115. Peters, G. G. and Glover, C. (1980). tRNA's and priming of RNA-directed DNA synthesis in mouse mammary tumor virus. *J. Virol.* 35, 31-40.
116. Waters, L. C. (1978). Lysine tRNA is the predominant tRNA in murine mammary tumor virus. *Biochem. Biophys. Res. Commun.* 81, 822-827.
117. Dahlberg, J. E., Sawyer, R. C., Taylor, J. M., Faras, A. J., Levinson, W. E., Goodman, H. M., and Bishop, J. M. (1974). Transcription of DNA from the 70S RNA of Rous sarcoma virus. I. Identification of a specific 4S RNA which serves as primer. *J. Virol.* 13, 1126-1133.
118. Waters, L. C., Mullin, B. C., Ho, T., and Yang, W. K. (1975). Ability of tryptophan tRNA to hybridize with 35S RNA of avian myeloblastosis virus and to prime reverse transcription in vitro. *Proc. Natl. Acad. Sci. USA* 72, 2155-2159.
119. Yang, J. A., Tai, L. W., Agris, P. F., Gehrke, C. W., and Wong, T. W. (1983). The nucleotide sequence of a major glutamine tRNA from rat liver. *Nucl. Acids Res.* 11, 1991-1996.
120. Benton, W. D. and Davis, R. W. (1977). Screening λ gt recombinant clones by hybridization to single plaques in situ. *Science* 196, 180-182.
121. Boone, L. R., Myer, F. E., Yang, D. M., Kiggans, J. O., Koh, C., Tennant, R. W., and Yang, W. K. (1983). Analysis of recombinant DNA clones of the endogenous BALB/c murine leukemia virus WN1802N: Variation in long terminal repeat length. *J. Virol.* 45, 484-488.
122. Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 38, 503-517.
123. Tennant, R. W., Otten, J. A., Wang, T. W., Liou, R. S., Brown, A., and Yang, W. K. (1983). Control of RFM strain endogenous retrovirus in RFM mouse cells. *J. Virol.* 45, 47-54.
124. Rigby, P., Dieckman, M., Rhodes, C., and Berg, P. (1977). Labeling deoxyribonucleic acid to high specific activity in vitro, by nick-translation with DNA polymerase I. *J. Mol. Biol.* 113, 237-251.
125. Smith, H. O. and Birnstiel, M. L. (1976). A simple method for DNA restriction site mapping. *Nucl. Acids Res.* 3, 2387-2398.
126. Ruther, U. (1982). pUR250 allows rapid chemical sequencing of both DNA strands of its inserts. *Nucl. Acids Res.* 10, 5765-5772.

127. Maxam, A. M. and Gilbert, W. (1977). A new method for sequencing DNA. *Proc. Natl. Acad. Sci. USA* 74, 560-564.
128. Joseph, D. R. (1981). Comparison of the restriction endonuclease maps of unintegrated proviral DNAs from four xenotropic murine leukemia viruses. *J. Virol.* 40, 963-970.
129. Rands, E., Lowy, D. R., Lander, M. R., and Chattopadhyay, S. K. (1981). Restriction endonuclease mapping of ecotropic murine leukemia viral DNAs: size and sequence heterogeneity of the terminal repeat sequence. *Virology* 108, 445-452.
130. Liou, R. S., Boone, L. R., Kigans, J. O., Yang, D. M., Wang, T. W., Tennant, R. W., and Yang, W. K. (1983). Molecular cloning and analysis of the endogenous retrovirus chemically induced from RFM/Un mouse cell cultures. *J. Virol.* 46, 288-292.
131. Sprinzl, M. and Gauss, D. H. (1984). Compilation of tRNA sequences. *Nucl. Acids Res.* 12 (Suppl.), r1-r57.
132. Yang, W. K. and Novelli, G. D. (1969). Multiple isoaccepting transfer RNA's in a mouse plasma cell tumor. *Proc. Natl. Acad. Sci. USA* 59, 208-215.
133. Colicelli, J. and Goff, S. P. (1986). Isolation of a recombinant murine leukemia virus utilizing a new primer tRNA. *J. Virol.* 57, 37-45.
134. Wirth, T., Glogler, K., Baumruker, T., Schmidt, M., and Horak, I. (1983). Family of middle repetitive DNA sequences in the mouse genome with structural features of solitary retroviral long terminal repeats. *Proc. Natl. Acad. Sci. USA* 80, 3327-3330.
135. Morrison, D. A. Transformation and preservation of competent bacterial cells by freezing. *Methods in Enzymology* 68, 326-331.
136. Chattopadhyay, S. K., Lander, M. R., Rands, E., and Lowy, D. R. (1980). Structure of endogenous murine leukemia virus DNA in mouse genome. *Proc. Natl. Acad. Sci. USA* 77, 5774-5778.
137. Khan, A. S. (1984). Nucleotide sequence analysis establishes the role of endogenous murine leukemia virus DNA segments in formation of recombinant mink cell focus-forming murine leukemia viruses. *J. Virol.* 50, 864-871.
138. O'Neill, R., Buckler, C. E., Theodore, T. S., Martin, M. A., and Repaske, R. (1985). Envelope and long terminal repeat sequences of a cloned infectious NZB xenotropic murine leukemia virus. *J. Virol.* 53, 100-106.

139. Chattopadhyay, S. K., Cloyd, M. W., Linemeyer, D. L., Lander, M. R., Rands, E., and Lowy, D. R. (1982). Cellular origin and role of mink cell focus-forming viruses in murine thymic lymphomas. *Nature* 295, 25-31.
140. Lowy, D. R., Rands, E., Chattopadhyay, S. K., Garon, C. F., and Hager, G. L. (1980). Molecular cloning of infectious integrated murine leukemia virus DNA from infected mouse cells. *Proc. Natl. Acad. Sci. USA* 77, 614-618.
141. Donehower, L. A. and Varmus, H. E. (1984). A mutant murine leukemia virus with a single missense codon in pol is defective in a function affecting integration. *Proc. Natl. Acad. Sci. USA* 81, 6461-6465.
142. Morse, H. C., III. (1978). The genealogy of inbred mice. *In* *Origins of Inbred Mice*, edited by H. C. Morse III, pp. 19-20. Academic Press.
143. Staats, J. (1980). Origins of inbred strains. *In* *Handbook on Genetically Standardized Jax Mice*, edited by Hans-Jorg Heiniger and J. T. Dorey, pp. 2.3-2.9. The Jackson Laboratory, Bar Harbor, Maine.
144. Roblin, R., Yang, J. M., Mural, R. J., Bell, T. E., and Ihle, J. N. (1982). Molecular cloning and characterization of murine leukemia virus-related DNA sequences from C3H/HeN mouse DNA. *J. Virol.* 43, 113-126.
145. Bacheler, L. T. (1984). Molecular clones of endogenous murine leukemia virus-related DNA sequences from BALB/c mice: Characterization of integration sites. *Virology* 138, 129-142.
146. Cooper, G. M. and Silverman, L. (1978). Linkage of the endogenous avian leukosis virus genome of virus producing chicken cells to inhibitory cellular DNA sequences. *Cell* 15, 573-577.
147. Harbers, K., Schneike, A., Stiehlmann, H., Jahner, D., and Jaenisch, R. (1981). DNA methylation and gene expression: Endogenous retroviral genomes become infectious after molecular cloning. *Proc. Natl. Acad. Sci. USA* 78, 7609-7613.
148. Groudine, M., Eisenman, R., and Weintraub, H. (1981). Chromatin structure and endogenous retroviral genes and activation by an inhibitor of DNA methylation. *Nature* 292, 311-317.
149. Hsiao, W. L. W., Gattoni-Celli, S., and Weinstein, I. B. (1986). Effects of 5-azacytidine on expression of endogenous retrovirus-related sequences in C3H 10T1/2 cells. *J. Virol.* 57, 1119-1126.

150. Levy, D. E., Lerner, R. A., and Wilson, M. C. (1985). Normal expression of polymorphic endogenous retroviral RNA containing segments identical to mink cell focus-forming virus. *J. Virol.* 56, 691-700.
151. Southern, P. J. and Berg, P. (1982). Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. *J. Mol. Appl. Genet.* 1, 327-341.
152. Hsu, I. C. and Yang, W. K. (1977). DNA transfection of ecotropic murine leukemia viruses in mouse cell cultures. *Cancer Res.* 37, 1709-1714.
153. Graham, F. L. and van der Eb, A. J. (1973). A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 52, 456-467.
154. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* 18, 5294-5299.
155. Razin, A. and Riggs, A. D. (1980). DNA methylation and gene function. *Science* 210, 604-610.
156. Noda, M., Kurihara, M., and Takano, T. (1982). Retrovirus-related sequences in human DNA: detection and cloning of sequences which hybridize with the long terminal repeat of baboon endogenous virus. *Nucl. Acids Res.* 10, 2865-2878.
157. Bonner, T. I., O'Connell, C., and Cohen, M. (1982). Cloned endogenous retroviral sequences from human DNA. *Proc. Natl. Acad. Sci. USA* 79, 4709-4713.
158. Callahan, R., Drohan, W., Tronick, S., and Schlom, J. (1982). Detection and cloning of human DNA sequences related to the mouse mammary tumor virus genome. *Proc. Natl. Acad. Sci. USA* 79, 5503-5507.

VITA

Kave Noghrei Nikbakht was born in Abadan, a port city in Southern IRAN, on October 10, 1951. He received his high school diploma in 1970 from Kharazmi High School and his Bachelor of Science degree with a major in Biology from Tehran University in 1974. From 1974 until 1976, he worked toward a Master of Science degree in Biology at Texas Southern University in Houston, Texas, on a scholarship fund awarded by the National Iranian Oil Company. In the fall of 1980, he received a Master of Science degree with a major in Biochemistry at The University of Tennessee in Knoxville. In the summer of 1981, he entered The University of Tennessee-Oak Ridge Graduate School of Biomedical Sciences at the Oak Ridge National Laboratory, Oak Ridge, Tennessee, where he received a Ph.D. degree with a major in Biomedical Sciences (Biochemistry) in the summer of 1986.