Duplication of U3 Sequences in the Long Terminal Repeat of Mink Cell Focus-Inducing Viruses Generates Redundancies of Transcription Factor Binding Sites Important for the Induction of Thymomas

Nancy L. DiFronzo,¹ Marisa Frieder,^{2†} Scott A. Loiler,^{3‡} Quynh N. Pham,¹
and Christie A. Holland^{1,2*}

Center for Virology and Immunology Research, Children's Research Institute, Children's National Medical Center,¹ and Department of Microbiology and Tropical Medicine, George Washington University School of Medical and Health Sciences,² Washington, D.C., and Program in Immunology and Virology, Graduate School of Biomedical Sciences, University of Massachusetts Medical School, Worcester, Massachusetts³

Received 26 September 2002/Accepted 15 November 2002

The ability of mink cell focus-inducing (MCF) viruses to induce thymomas is determined, in part, by transcriptional enhancers in the U3 region of their long terminal repeats (LTRs). To elucidate sequence motifs important for enhancer function *in vivo*, we injected newborn mice with MCF 1dr (supF), a weakly pathogenic, molecularly tagged (supF) MCF virus containing only one copy of a sequence that is present as two copies (known as the directly repeated [DR] sequence) in the U3 region of MCF 247 and analyzed LTRs from supF-tagged proviruses in two resulting thymomas. Tagged proviruses integrated upstream and in the reverse transcriptional orientation relative to *c-myc* provided the focus of our studies. These proviruses are thought to contribute to thymoma induction by enhancer-mediated deregulation of *c-myc* expression. The U3 region in a tagged LTR in one thymoma was cloned and sequenced. Relative to MCF 1dr (supF), the cloned U3 region contained an insertion of 140 bp derived predominantly from the DR sequence of the injected virus. The inserted sequence contains predicted binding sites for transcription factors known to regulate the U3 regions of various murine leukemia viruses. Similar constellations of binding sites were duplicated in two proviral LTRs integrated upstream from *c-myc* in a second thymoma. We replaced the U3 sequences in an infectious molecular clone of MCF 247 with the cloned proviral U3 sequences from the first thymoma and generated an infectious chimeric virus, MCF ProEn. When injected into neonatal AKR mice, MCF ProEn was more pathogenic than the parental virus, MCF 1dr (supF), as evidenced by the more rapid onset and higher incidence of thymomas. Molecular analyses of the resultant thymomas indicated that the U3 region of MCF ProEn was genetically stable. These data suggest that the arrangement and/or redundancy of transcription factor binding sites generated by specific U3 sequence duplications are important to the biological events mediated by MCF proviruses integrated near *c-myc* that contribute to transformation.

AKR mice succumb to spontaneous thymomas between 6 and 12 months of age. The virological events leading to leukemia include the inheritance and replication of Akv, a non-leukemogenic, ecotropic murine leukemia virus (MuLV), and the sequential recombination between Akv and two other endogenous retrovirus sequences to produce leukemogenic mink cell focus-inducing (MCF) viruses (6, 11, 18, 26). The first recombination event replaces the long terminal repeat (LTR) sequences in the ecotropic virus with those from the inducible xenotropic virus, *Bxv-1*, while the second recombination event replaces a portion of the ecotropic envelope (*env*) gene with polytropic *env* sequences (6, 11, 18, 26). During the *in vivo* spread and replication of the recombinant virus containing both the LTR and *env* substitutions, sequences in the U3 region of the LTR are duplicated, creating the tandemly duplicated sequences commonly referred to as the directly repeated (DR) sequences (18, 26). Both recombination events and the

duplication of U3 sequences are required to generate leukemogenic MCF viruses.

MCF 247, a prototypic leukemogenic MCF virus, induces leukemia following injection into newborn AKR mice (12). The oncogenic potential of MCF 247 is determined, in part, by the two 105-bp DR sequences in the U3 region of the LTR (5, 12, 13). These sequences regulate both the level of transcription from the virus promoter and the level of expression of cellular genes near the integrated provirus. Although the DR sequences of many MuLVs (Moloney, SL3-3, Friend, and MCF 13, among others) have some, if not all, of the characteristics of enhancers (1, 14, 17, 23), the contribution of the DR sequence in the U3 region of MCF 247 to the level of transcription from the viral promoter has not been reported in the literature. Therefore, we compared the transcriptional activity of a reporter containing the majority of the U3 sequences of MCF 247 (MCF 2dr) with the activities of reporters containing the U3 region with one copy of DR sequences (MCF U3 1dr) or only the viral promoter sequences (MCF Promoter) in T cells (Fig. 1). The activity of MCF U3 1dr ($40.2\% \pm 7\%$) was significantly lower ($P < 0.001$) than that of MCF 2dr ($100\% \pm 9\%$) but significantly higher ($P < 0.001$) than that of MCF Promoter ($0.4\% \pm 0.7\%$). These data demonstrate that the

* Corresponding author. Mailing address: Children's National Medical Center, 111 Michigan Ave., N.W., Washington, DC 20010. Phone: (202) 884-3981. Fax: (202) 884-3985. E-mail: chhollan@cnmc.org.

† Present address: Portland VA Medical Center, Portland, Ore.

‡ Present address: Powell Gene Therapy Center and Department of Pediatrics, University of Florida, Gainesville, Fla.

FIG. 1. Reduced transcriptional activity of MCF U3 1dr in Jurkat T cells. MCF 247-derived U3 regions inserted upstream of the *cat* gene in the reporter plasmids are diagrammed on the left. MCF 2dr contains most of the U3 region of MCF 247, including the upstream sequences, two DR sequences, the downstream U3 sequences, and the viral promoter (CCAAT and TATA). Restriction sites are indicated (E, *EcoRI*; BII, *BssHII*; P, *PstI*; D, *DraI*; S, *SmaI*). MCF U3 1dr contains only one DR sequence. MCF Promoter contains only the viral promoter. On the right, the transcriptional activity of each reporter in Jurkat T cells is plotted relative to that of MCF 2dr. Transfections, performed in triplicate in two independent experiments, were performed and analyzed as described previously (4). The error bars indicate the standard deviation about each mean. The asterisks indicate that transcriptional activities of the reporters differ significantly ($P < 0.01$). CAT, chloramphenicol acetyltransferase.

DR sequence in the U3 region of MCF 247 has enhancer activity and that two copies of DR are required for maximal transcriptional activity in T cells.

(Portions of this study were performed in partial satisfaction of the requirements for the degree of Doctor of Philosophy [M.F.] from the Columbian College and Graduate School of Arts and Sciences of George Washington University and the degree of Doctor of Philosophy [S.A.L.] from the University of Massachusetts Medical Center.)

To identify DR sequence motifs important for MCF-induced transformation, newborn AKR mice were inoculated with MCF 1dr (supF), an MCF virus containing a single copy of the two DR sequences in MCF 247 and a molecular tag (supF) in the LTR (5) (Fig. 2). The genome of this virus is similar to those of de novo-generated recombinant viruses containing both the LTR and *env* substitutions but can be differentiated from the genomes of these viruses by the molecular tag (see below). MCF 1dr (supF) induced fewer thymomas (42%) and did so with a delayed onset (mean latency period of 163 ± 3.1 days) compared to the wild-type virus (100% , 97 ± 8.5 days) (Table 1). These data are consistent with earlier studies demonstrating that the full oncogenic activity of MCF 247 requires two copies of the DR sequences (5, 13).

In vitro transcriptional activity (Fig. 1) and in vivo pathogenicity (Table 1) indicate that enhancer function is not optimal in an MCF virus containing only one copy of the DR-homologous sequence present as two copies in the U3 region of MCF 247. The U3 regions of frequently studied MCF isolates MCF 247 and MCF 13 both contain two DR sequences (i.e., 105 and 67 bp, respectively). Each DR sequence contains predicted binding sites for transcriptional regulators important for pathogenesis by other MuLVs, namely, LVb (a *c-ets* family member), core binding factor (CBF; now known as RUNX1),

FIG. 2. MCF 1dr (supF) proviruses integrated upstream of *c-myc*. Diagrammatic representation of MCF 1dr (supF) proviruses integrated in the same (A) and reverse (B) transcriptional orientations relative to that of *c-myc* exon 1. The locations of the molecular tag (supF), the DR sequence, and the *EcoRI* (E) and *BssHII* (BII) restriction sites are indicated. The arrowheads indicate the locations of primers used in PCR and sequencing (supF [5'-GAAGTCGATGAC GGCAGATT-3'], TM [5'-AAGAGGGAGGTTTGTGTGCT-3'], 3 [5'-AGACCACGATTTCGGATGCAAC-3'], s [5'-GCAAGGCATGA AAAAGTACCA-3'], and 614 [5'-CGACTCAGTCTATCGGAGGA C-3']). The diagrams are not drawn to scale. Double slashes and wavy lines indicate mouse genomic DNA.

and nuclear factor 1 (NF-1) (7, 10, 16, 24, 25, 27, 28, 32–34). To determine whether iteration of these specific binding sites is required for full enhancer activity in vivo, we analyzed proviral U3 sequences in MCF 1dr (supF)-induced thymomas. As a first step, we identified thymomas containing supF-tagged proviruses by PCR with primers specific for the 3' portion of supF-tagged proviruses (TM [5'-AAGAGGGAGGTTTGTGTGCT-3'] and supF [5'-GAAGTCGATGACGGCAGATT-3']; Fig. 2). Amplification products were analyzed by Southern blotting with a probe for the supF tag. Products from 12 of the 13 thymoma DNAs tested (92%) hybridized with the probe (Table 1), consistent with the conclusion that MCF 1dr (supF) induced these thymic lymphomas.

MCF viruses replicate in thymocytes and induce thymomas in susceptible mice. MCF 247 integrates near the *c-myc* proto-oncogene in ~20% of MCF 247-induced thymomas (15). In many MCF-induced tumors, proviruses are integrated upstream of and in the reverse transcriptional orientation relative to exon 1 of *c-myc* (2, 15), suggesting that the proviral enhancers contribute to the deregulated expression of this cellular proto-oncogene. To explore the biological basis of the U3 sequence duplication and the interaction between MCF proviral enhancers and *c-myc* in leukemogenesis, we next analyzed thymomas for rearrangements of *c-myc* and characterized the proviral U3 sequences integrated near this cellular proto-oncogene (see below). Three (25%) of the 12 supF-positive thymomas (tumors 92, 95, and 98) contained *c-myc* locus rear-

TABLE 1. Leukemogenicity of MCF 247, MCF 1dr (supF), and MCF ProEn

Virus	No. of diseased mice/no. injected ^a	No. supF positive/total ^b	No. with rearranged <i>c-myc</i> locus/total	Mean latency period (days) ^c ± SEM	Mortality (%) ^d
MCF 247	7/7	NA ^e	6/26 ^f	97 ± 8.5	100
MCF 1dr (supF)	14/33	12/13	3/12 ^g	163 ± 3.1	42
MCF ProEn	15/17	NA	NT ^h	149 ± 4.5	88

^a Number of mice with disease in less than 180 days postinjection.^b Number of supF-positive thymoma DNAs. Of the 14 MCF 1dr (supF)-induced thymomas, only 13 DNAs were available for analysis.^c Number of days from injection to death.^d Number of mice with disease in less than 180 days postinjection as a percentage of the total number of mice injected.^e NA, not applicable.^f Data from reference 15 using an infectious virus derived from MCF 247 and containing a 220-bp fragment from λ bacteriophage inserted into the *Pst*I site in the LTR.^g Number of thymoma DNAs that contain a rearranged *c-myc* locus compared to the number of supF-positive thymoma DNAs.^h NT, not tested.

rangements (Table 1). This is consistent with previous studies evaluating pathogenesis by MCF viruses (2, 3, 15, 21, 22) and other nonacute MuLVs (1, 20, 29).

To confirm that supF-tagged proviruses were integrated in *c-myc*, tumors with rearrangements of this proto-oncogene (thymomas 92, 95, and 98) were analyzed by PCR with one of four *c-myc* primers (primers a, b, c, and d) (20) and primers to detect proviruses inserted in the same (primer TM) or the opposite (primer 3 [5'-AGACCACGATTCGGATGCAAC-3']) transcriptional orientation relative to *c-myc* (Fig. 2). Amplification products were resolved by agarose gel electrophoresis and detected by Southern blotting with a ³²P-labeled, supF-specific probe containing sequences isolated from plasmid PiAN7. As illustrated in Fig. 3A (lanes 1 and 5), the probe hybridized to several amplification products generated from the DNAs extracted from thymomas 92 and 95. Similar results were obtained with amplification products from thymoma 98 (data not shown). On the basis of these results and the location of the primers relative to *c-myc* exon 1, a second primer pair (primers 3 and d) should have produced amplification products that hybridize with the probe; however, these amplification products are predicted to be significantly larger ($\geq 1,000$ bp) than those produced with primers 3 and b and did not amplify under the PCR conditions used. Nonetheless, detection of amplification products produced by primers 3 and b indicated that supF-tagged proviruses were integrated upstream and in the opposite transcriptional orientation relative to *c-myc* exon 1 in three lymphomas. Although we could not explain, on the basis of these data alone, why the supF probe hybridized weakly to the amplification products from these thymomas (Fig. 3A), the intensities of these bands relative to those obtained with control DNA (Fig. 3A, right side) indicated that hybridization was specific.

In order to characterize the U3 regions in the supF-tagged LTRs, we reamplified thymoma DNAs (92 and 95) with primer 3 and a *c-myc*-specific primer (a, b, or d) in the presence of [α -³²P]dCTP, resolved the DNAs in polyacrylamide gels (Fig. 3B), and selectively isolated the radiolabeled amplification products corresponding in size to DNAs that hybridized with supF (Fig. 3A, lanes 1 and 5). One or more amplification products from each thymoma were radiolabeled (Fig. 3B, lanes 2 and 5), suggesting that these tumors contain one or more thymocyte clones with proviruses integrated at different distances from the sequences complementary to primer b in the

c-myc locus. Amplification products larger than the supF-tagged LTR from the injected virus (~620 bp) that hybridized with the supF probe (identified by comparison with those in Fig. 3A) were selected for further analyses. Radiolabeled amplification products (Fig. 3B) that did not correspond to the size of supF-hybridizing amplification products (Fig. 3A, lanes 1 and 5) were not analyzed, since they most likely resulted from hybridization of primer 3 to endogenous MuLV sequences. Likewise, radiolabeled DNAs amplified from normal thymus DNA (Fig. 3B, lane 8) were not analyzed because these probably resulted from mispriming, as evidenced by the absence of hybridization between the supF probe and similarly sized amplification products produced with this primer pair (Fig. 3A, lane 11).

When amplified as described above in the presence of [α -³²P]dCTP, thymoma 92 DNA generated two radiolabeled fragments between 1.1 and 1.3 kbp in size (Fig. 3B, lane 2). These fragments were similar in size to those seen in the Southern blot shown in Fig. 3A (lane 1). To characterize the U3 sequences in these DNAs, the broad band was excised (indicated by the bracket beside Fig. 3B, lane 2) and the DNA was eluted and reamplified with nested primers spanning the U3 sequences (s and 614; Fig. 2). The resulting amplification products were cloned (pCR2.1; Invitrogen, Carlsbad, Calif.), and one of the U3 regions was sequenced (henceforth referred to as 92A [Fig. 4A] and described below). We recognize that either the 92A U3 region was reamplified from one of the products seen in Fig. 3B (lane 2) or that 92A was reamplified from a single amplification product that migrated aberrantly as two fragments because of the effects of secondary structure. To distinguish these possibilities and to confirm that the 92A U3 region was integrated upstream of *c-myc* exon 1, DNA from thymoma 92 was amplified by using primers 3 and b. Amplification products that migrated between 1.1 and 1.3 kbp were separately isolated from agarose gels and cloned independently (pCR4-TOPO; Invitrogen). By using primers that hybridize to the vector and span the amplification products, 15 clones from each ligation were completely sequenced. In each clone, the sequences at the termini of the amplification products were identical to those in primers 3 and b and the cloned DNAs were identical in both length (1,128 bp) and sequence. BlastN analyses revealed that each clone contained 92A sequences joined to 496 bp of murine *c-myc* sequences (nucleotides 131 to 627; GenBank accession no. 199964). These data

FIG. 3. SupF-tagged proviruses are integrated near *c-myc* in thymomas induced with MCF 1dr (supF). (A) To identify supF-tagged LTRs integrated near *c-myc*, DNAs from thymoma 92 (lanes 1 to 3), thymoma 95 (lanes 4 to 6), and the thymus of a naive mouse (lanes 7 to 9) were amplified by using primers specific for provirus sequences (3 and TM) and the *c-myc* locus (a, b, c, and d). Primer pairs are indicated above each lane and correspond to those shown in Fig. 2. Amplification products were resolved by electrophoresis in agarose gels, transferred to membranes, and hybridized with a supF probe. (B) DNAs from thymomas 92 (lanes 1 to 3) and 95 (lanes 4 to 6) and from the thymus of a naive mouse (lanes 7 to 9) were amplified in the presence of [α - 32 P]dCTP by using the primer pairs indicated above each lane and separated by polyacrylamide gel electrophoresis. Amplification products were isolated from the gel (indicated by brackets) and sequenced.

suggested that a single amplification product was generated with primers 3 and b from thymoma 92, but it migrated aberrantly as two fragments. Electrophoresis of the gel-purified amplification products produced with primers 3 and b under denaturing conditions (8 M urea–8% polyacrylamide gel) demonstrated that both DNAs migrated at a length of 1.1 kb. The denaturation and sequencing experiments demonstrate that the two fragments are identical but migrate differently because of secondary structure. These data support the conclusion that a provirus containing the 92A U3 region is integrated upstream of *c-myc* exon 1 in thymoma 92.

In addition, the sequence analyses conclusively demonstrated that 92A proviral LTR sequences (R and U3) were in the inverse transcriptional orientation relative to *c-myc*. Unexpectedly, however, only 19 bp of the supF tag was retained in the U3 region; a computer search with the BlastN algorithm confirmed that these sequences were identical to 19 bp in supF. Thus, only a small portion of the sequence in the supF molec-

ular tag from the U3 region of the injected virus was present in these LTRs, explaining the weak intensity of supF probe hybridization to the amplified DNAs (Fig. 3A).

The sequence of the U3 region from 92A was compared to that of the injected virus. Relative to MCF 1dr (supF), 92A contains a 140-bp insertion between the upstream (thick line) and DR-homologous (open box) sequences (Fig. 4A). The 140-bp insertion contains stretches of nucleotides identical to two portions of the U3 region from the injected virus. The first 90 bp are identical to the 5' portion of the DR-homologous sequences (open box), the next 6 bp (GAGGGG) are unknown in origin, and last 44 bp are identical to U3 sequences immediately upstream (thick line) of the DR-homologous sequences in the U3 region of MCF 1dr (supF). The shaded portions of U3 indicate the 134-bp DR sequences created in the 92A U3 region (Fig. 4A). Predicted binding sites for transcription factors with the potential to regulate MuLV enhancer activity are indicated above the diagram of MCF 1dr (supF) (Fig. 4A). Binding sites include, in 5'-to-3' order, Ikaros, NF-1, Ets (LVb), CBF, NF-1, GRE, nuclear factor of activated T cells, c-Myb, NF- κ B, and Ikaros. The inserted sequences in the U3 region of 92A generate redundancies of many of these binding sites, including Ikaros, NF-1, LVb, CBF, and GRE. Three of these binding sites (LVb, CBF, and NF-1) are part of a constellation of four motifs conserved among nonacute retroviruses that are thought to provide a framework for enhancer function (8). This trio of motifs in the DR sequences of 92A, MCF 247, and MCF 13 (Fig. 4B) raises the possibility that the reiteration of these DNA elements provides a distinct growth advantage to biologically selected MCF viruses and contributes to thymoma induction in vivo.

To test the biological relevance of the 92A U3 sequences to an MCF virus in vivo, we substituted these sequences (*Eco*RI to *Bss*HII) for the corresponding sequences in an infectious molecular clone of MCF 247, generated an infectious virus (MCF ProEn), and tested its pathogenic potential following inoculation into newborn AKR mice (Fig. 4C). MCF ProEn (circles) induced thymomas significantly faster (149 ± 4.5 days versus 163 ± 3.1 days; $P < 0.02$) and in a greater percentage of inoculated mice (88% versus 42%; $P < 0.005$) than did MCF 1dr (supF) (squares). The percent mortality induced by MCF ProEn (88%) did not differ statistically significantly from that induced by MCF 247 (triangles) (100%) at 180 days postinoculation. These results indicate that the 140-bp insertion in the U3 region of MCF ProEn increases the pathogenic potential of a virus otherwise identical to MCF 1dr (supF). Furthermore, molecular analyses of the proviral LTRs in MCF ProEn-induced thymomas indicated that the U3 region of MCF ProEn was genetically maintained in vivo (Table 2). Thymoma DNAs were amplified with primers specific for the U3 region of this exogenous virus. One primer (En5'DR [5'-GGCTGA ACAAACTGGGACA-3']) hybridized to the upstream sequences (thick lines) adjacent to the region corresponding to the DR sequence in MCF 247, and the reverse primer (NovEN [5'-TTCTTTAACTAACTTCCCCCTCA-3']) was complementary to sequences formed by the junction of the unique sequences and sequences in the 140-bp insertion (Fig. 4A). A specific amplification product was detected in 89% of the thymomas (16 of 18) induced within 180 days of MCF ProEn inoculation (Table 2), indicating that this component of the U3

FIG. 4. U3 sequences from tagged provirus integrated upstream from *c-myc* in an MCF 1dr (supF)-induced thymoma increase the pathogenicity of an MCF virus. (A) Diagrams of the U3 regions (*Eco*RI to *Bss*HIII) of MCF 1dr (supF) and 92A, a provirus integrated upstream from *c-myc* in thymoma 92, are shown aligned. The open box represents nucleotides homologous to the DR sequence present as two copies in the U3 region of MCF 247 (see below). Thick and thin lines indicate U3 sequences upstream and downstream of the DR sequences, respectively. The bent arrow indicates the CCAAT box in the promoter region. Identical sequences ($\geq 98\%$) are shown aligned. The dashed line indicates that the two portions of the U3 region are continuous. Shading highlights the DR sequences in 92A. Predicted binding sites for transcriptional regulators are marked above the diagram of the U3 of MCF 1dr (supF); the solid circles indicate the presence of these same motifs in the aligned U3 region. The arrowheads indicate the positions of the primers (En5'DR and NovEn) used to specifically amplify a portion of the 92A U3 region. (B) Diagrams of the MCF 247 (top) and MCF 13 (bottom) U3 regions. Identical sequences ($\geq 98\%$) are shown aligned. The open boxes and dotted lines represent the DR sequences and continuous portions of U3 in MCF 247. Shading highlights the DR sequences in the U3 regions of MCF 247 and MCF 13. Predicted binding sites for transcriptional regulators are indicated in the MCF 247 U3 region and as solid circles in the U3 region of MCF 13. (C) Plot of the incidence of thymomas in mice injected with MCF ProEn, a chimeric virus containing the 92A U3 region (circles), MCF 1dr (supF) (squares), and MCF 247 (triangles) versus the number of days after virus injection.

region was genetically maintained in vivo. As anticipated, these primers did not amplify DNAs from thymomas induced by MCF 247 or thymus DNAs from a control mouse (Table 2). Taken together, these data support the interpretation that the 140-bp stretch of nucleotides contained in the U3 region of MCF ProEn contributes to enhancer function in vivo.

The molecular mechanism by which the MCF ProEn proviral LTR increases the rate and incidence of thymomas relative

to MCF 1dr (supF), the parental MCF virus, is not known. Nonetheless, the difference in the pathogenic potentials of MCF ProEn and MCF 1dr (supF) demonstrates that the 140-bp insertion, or its arrangement within the adjacent U3 sequences, contributes to tumor induction. It is probable that the proteins that bind motifs in the 140-bp insertion cooperate with those that bind in the surrounding U3 sequences and that these protein interactions modulate enhancer topology and

TABLE 2. PCR-detected MCF ProEn U3 sequences in thymoma DNAs

Virus injected	Time from injection to disease ^a	No. of DNAs tested	No. of DNAs with MCF ProEn U3 sequences ^b
MCF ProEn	<180	18	16
MCF ProEn	>180	2	0
MCF 247	<180	2	0
None	NA ^c	1	0

^a Days postinoculation.
^b Primers specific for U3 sequences in the LTR of MCF ProEn were used to amplify thymoma DNAs. The locations of these primers (En5'DR and NovEn) are shown in Fig. 4A.
^c NA, not applicable.

function in vivo. Cooperation between the CBF and Ets proteins (LVb) and between the CBF and *c-myb* proteins contributes significantly to the enhancer function of the Moloney and SL3-3 U3 regions both in vitro and in vivo (9, 24, 25, 27, 30, 34). The contribution of these three motifs to the enhancer activity of the MCF 247 U3 region in vitro and in vivo has not been explored. Nonetheless, the presence of CBF and NF-1 binding sites in the DR sequences of pathogenic MuLVs (8) and the direct demonstration of the oncogenic potential of MCF ProEn suggest that cooperative interactions between

proteins that bind these motifs may be necessary for optimal enhancer function of the U3 sequences in vivo.

We also characterized proviral LTRs integrated upstream of *c-myc* in a second MCF 1dr (supF)-induced tumor (thymoma 95). By using an approach similar to that used to isolate 92A, four amplification products (henceforth called 95A, 95B, 95C, and 95D) were generated and the U3 regions were reamplified and sequenced directly (Fig. 3B, lane 5, and Fig. 5). One U3 region (95D) is identical to the U3 region of the injected virus, and one U3 region (95C) contains a truncation in the DR-homologous sequence. Two of the U3 regions (95A and 95B) contain similar, but not identical, 87-bp insertions at the same location in U3. Each of the 87-bp insertions contains sequences derived from the 5' portion of the DR-homologous sequences (open box) and the upstream U3 sequences (thick line) present in MCF 1dr (supF). The shaded boxes highlight the DR sequences in these two U3 regions. Predicted binding sites for Ikaros, CBF, and NF-1 are redundant in each U3 region and are present in each copy of the DR sequences. The CBF and NF-1 motifs are present in the homologous positions in the DR sequences of 92A (Fig. 4A), as well as in the DR sequences of the pathogenic viruses MCF 247 and MCF 13 (Fig. 4C). Thus, the DR sequences in the LTRs from two lymphomas contain transcription factor binding sites derived from the U3 sequences of the parental virus, MCF 1dr (supF). The presence

FIG. 5. Repeated selection of duplicated U3 sequences in proviruses integrated upstream from *c-myc* in an MCF 1dr (supF)-induced thymoma. At the top and bottom are diagrammatic representations of the U3 regions (*Eco*RI to *Bss*HII) of MCF 1dr (supF) and the wild-type MCF 247 virus. U3 regions from LTRs cloned from sequences upstream of *c-myc* in thymoma 95 are shown aligned relative to MCF 1dr (supF). Open boxes and lines represent portions of U3, as described in the legend to Fig. 4; the zigzag lines represent sequences not homologous to MCF 1dr (supF). Relative to MCF 1dr (supF), 95A and 95B contain insertions of 87 bp. Shading highlights the DR sequences in the U3 regions of 95B (81 bp), 95A (87 bp), and MCF 247 (105 bp).

of four distinct U3 regions (95D, 95C, 95B, and 95A) in the amplification products detected in Fig. 3B suggests that thymoma 95 contained multiple cells with tagged proviruses integrated upstream, but at different distances, from *c-myc* exon 1.

Although the LTRs from proviruses integrated upstream from *c-myc* in two thymomas were similar, the U3 regions did not contain identical insertions in U3 (92A, 95A, and 95B; Fig. 4A and 5). Each insertion (140, 87, and 87 bp, respectively) was derived from the U3 sequences in MCF 1dr (supF) and predominantly from the 5' portion of the sequences homologous to the DR sequence in MCF 247. The inserted sequences created U3 regions with distinct DR sequences (134, 81, and 87 bp, respectively). One U3 region (95B) has DR sequences containing predicted binding sites for CBF and NF-1, while the other two U3 regions (92A and 95A) have DR sequences with predicted binding sites for Ets-1 (LVb), as well as CBF and NF-1. These data raise the possibility that two DR sequences generate redundancies in transcription factor binding sites that modulate U3 enhancer function and the ability of a virus to induce tumors. This idea was tested by using MCF ProEn, a replication-competent MCF virus containing the 92A U3 region. MCF ProEn induced tumors more rapidly and in a greater percentage of animals than the MCF virus with one copy of the DR sequences. The DR sequences in the 92A U3 region have predicted binding sites for Ikaros, LVb, CBF, NF-1, and GRE. Three of these motifs (LVb, CBF, and NF-1) are present in the DR sequences of pathogenic viruses MCF 247 and MCF 13 (8, 26). The maintenance of the LVb, CBF, and NF-1 motifs in the DR sequences of highly pathogenic MCF viruses (MCF ProEn, MCF 247, and MCF 13) supports the idea that the redundancy of this trio of motifs in the U3 region with two DR sequences contributes to tumor induction by MCF viruses.

It is likely that the U3 regions in tagged proviruses cloned from thymomas 92 and 95 were produced by the same molecular mechanism that generates the two DR sequence copies in the LTR of pathogenic MCF viruses, namely, an interstrand template switch occurring as an error during reverse transcription and minus strand DNA synthesis (26). In the present study, we speculated that a similar error of reverse transcription occurred in viruses containing two MCF 1dr (supF) genomic viral RNAs. Duplication of sequences in the U3 regions of pathogenic MCF viruses (MCF 247 and MCF 13) and in the U3 regions of tumor-derived tagged proviruses (92A, 95A, and 95B; Fig. 4A and 5) suggests either that interstrand template switches frequently occur during reverse transcription in vivo or that a rare event has been preferentially selected on the basis of the activity of these resultant U3 sequences in vivo. The presence of multiple copies of DR sequences in other MuLVs and the duplication of sequences in proviral U3 regions generated from viruses that originally contained deletions or mutations in U3 support the former possibility (1, 2, 19, 20, 31). However, the reiteration of an organized constellation of motifs (LVb, CBF, NF-1, and GRE) present in the DR sequences of various MuLVs suggests that the spacing and orientation of specific transcription factor binding sites in U3 are selected biologically for their function in vivo (8). The rate of thymoma induction in mice injected with MCF ProEn, a chimeric virus containing U3 sequences from a supF-tagged provirus in an MCF 1dr (supF)-induced thymoma (Fig. 4C),

and the genetic stability of this U3 region in vivo (Table 2) provide support for the functional significance of the inserted sequences and hence the redundancy of LVb, CBF, and NF-1 transcription factor binding sites in the DR sequences in this U3 region.

We thank Anamaris Colberg-Poley for valuable discussions and critical reading of the manuscript.

This work was supported by Public Health Service grant CA-41510 from the National Institutes of Health to C.A.H. and an Avery Scholar Award from the Children's National Medical Center to N.L.D.

REFERENCES

1. Broussard, D. R., J. A. Mertz, M. Lozano, and J. P. Dudley. 2002. Selection for *c-myc* integration sites in polyclonal T-cell lymphomas. *J. Virol.* **76**:2087–2099.
2. Chen, H., and F. K. Yoshimura. 1994. Identification of a region of a murine leukemia virus long terminal repeat with novel transcriptional regulatory activities. *J. Virol.* **68**:3308–3316.
3. Corcoran, L. M., J. M. Adams, A. R. Dunn, and S. Cory. 1984. Murine T lymphomas in which the cellular *myc* oncogene has been activated by retroviral insertion. *Cell* **37**:113–122.
4. DiFronzo, N., C. T. Leung, M. K. Mammel, K. G. Georgopoulos, B. J. Taylor, and Q. N. Pham. 2002. Ikaros, a lymphoid-specific transcription factor, contributes to the leukemogenic phenotype of a mink cell focus-inducing murine leukemia virus. *J. Virol.* **76**:78–87.
5. DiFronzo, N. L., and C. A. Holland. 1999. Sequence-specific and/or stereospecific constraints of the U3 enhancer elements of MCF 247-W are important for pathogenicity. *J. Virol.* **73**:234–241.
6. Elder, J. H., J. W. Gautsch, F. C. Jensen, R. A. Lerner, J. W. Hartley, and W. P. Rowe. 1977. Biochemical evidence that MCF murine leukemia viruses are envelope (env) gene recombinants. *Proc. Natl. Acad. Sci. USA* **74**:4676–4680.
7. Ethelberg, S., J. Lovmand, J. Schmidt, A. Luz, and F. S. Pedersen. 1997. Increased lymphomagenicity and restored disease specificity of AML1 site (core) mutant SL3-3 murine leukemia virus by a second-site enhancer variant evolved in vivo. *J. Virol.* **71**:7273–7280.
8. Golemis, E. A., N. A. Speck, and N. Hopkins. 1990. Alignment of U3 region sequences of mammalian type C viruses: identification of highly conserved motifs and implications for enhancer design. *J. Virol.* **64**:534–542.
9. Gunther, C. V., and B. J. Graves. 1994. Identification of ETS domain proteins in murine T lymphocytes that interact with the Moloney murine leukemia virus enhancer. *Mol. Cell. Biol.* **14**:7569–7580.
10. Hallberg, B., J. Schmidt, A. Luz, F. S. Pedersen, and T. Grundström. 1991. SL3-3 enhancer factor 1 transcriptional activators are required for tumor formation by SL3-3 murine leukemia virus. *J. Virol.* **65**:4177–4181.
11. Hartley, J. W., N. K. Wolford, L. J. Old, and W. P. Rowe. 1977. A new class of murine leukemia virus associated with development of spontaneous lymphomas. *Proc. Natl. Acad. Sci. USA* **74**:789–792.
12. Holland, C. A., J. W. Hartley, W. P. Rowe, and N. Hopkins. 1985. At least four viral genes contribute to the leukemogenicity of murine retrovirus MCF 247 in AKR mice. *J. Virol.* **53**:158–165.
13. Holland, C. A., C. Y. Thomas, S. K. Chattopadhyay, C. Koehne, and P. V. O'Donnell. 1989. Influence of enhancer sequences on thymotropism and leukemogenicity of mink cell focus-forming viruses. *J. Virol.* **63**:1284–1292.
14. Hollon, T., and F. K. Yoshimura. 1989. Mapping of functional regions of murine retrovirus long terminal repeat enhancers: enhancer domains interact and are not independent in their contributions to enhancer activity. *J. Virol.* **63**:3353–3361.
15. Li, Y., C. A. Holland, J. W. Hartley, and N. Hopkins. 1984. Viral integration near *c-myc* in 10–20% of MCF 247-induced AKR lymphomas. *Proc. Natl. Acad. Sci. USA* **81**:6808–6811.
16. LoSardo, J. E., A. L. Boral, and J. Lenz. 1990. Relative importance of elements within the SL3-3 virus enhancer for T-cell specificity. *J. Virol.* **64**:1756–1763.
17. Lovmand, S., N. O. Kjeldgaard, P. Jorgensen, and F. S. Pedersen. 1990. Enhancer functions in U3 of Akv virus: a role for cooperativity of a tandem repeat unit and its flanking DNA sequences. *J. Virol.* **64**:3185–3191.
18. Lung, M. L., C. Hering, J. W. Hartley, W. P. Rowe, and N. Hopkins. 1980. Analysis of the genomes of mink cell focus-inducing murine type-C viruses: a progress report. *Cold Spring Harbor Symp. Quant. Biol.* **44**:1269–1274.
19. Martiney, M. J., K. Rulli, R. Beaty, L. S. Levy, and J. Lenz. 1999. Selection of reversions and suppressors of a mutation in the CBF binding site of a lymphomagenic retrovirus. *J. Virol.* **73**:7599–7606.
20. Morrison, H. L., B. Soni, and J. Lenz. 1995. Long terminal repeat enhancer core sequences in proviruses adjacent to *c-myc* in T-cell lymphomas induced by a murine retrovirus. *J. Virol.* **69**:446–455.
21. O'Donnell, P. V., E. Fleissner, H. Lonial, C. F. Koehne, and A. Reicin. 1985.

- Early clonality and high-frequency proviral integration into the *c-myc* locus in AKR leukemias. *J. Virol.* **55**:500–503.
22. **Selten, G., H. T. Cuypers, M. Zijlstra, C. Melief, and A. Berns.** 1984. Involvement of *c-myc* in MuLV-induced T cell lymphomas in mice: frequency and mechanisms of activation. *EMBO J.* **3**:3215–3222.
 23. **Short, M. K., S. A. Okenquist, and J. Lenz.** 1987. Correlation of leukemogenic potential of murine retroviruses with transcriptional tissue preference of the viral long terminal repeats. *J. Virol.* **61**:1067–1072.
 24. **Speck, N. A., B. Renjifo, E. Golemis, T. N. Fredrickson, J. W. Hartley, and N. Hopkins.** 1990. Mutation of the core or adjacent LVb elements of the Moloney murine leukemia virus enhancer alters disease specificity. *Gene Dev.* **4**:233–242.
 25. **Speck, N. A., B. Renjifo, and N. Hopkins.** 1990. Point mutations in the Moloney murine leukemia virus enhancer identify a lymphoid-specific viral core motif and 1,3-phorbol myristate acetate-inducible element. *J. Virol.* **64**:543–550.
 26. **Stoye, J. P., C. Moroni, and J. M. Coffin.** 1991. Virological events leading to spontaneous AKR thymomas. *J. Virol.* **65**:1273–1285.
 27. **Sun, W., B. J. Graves, and N. A. Speck.** 1995. Transactivation of the Moloney murine leukemia virus and T-cell receptor β -chain enhancers by *cbf* and *ets* requires intact binding sites for both proteins. *J. Virol.* **69**:4941–4949.
 28. **Thornell, A., B. Hallberg, and T. Grundström.** 1991. Binding of SL3–3 enhancer factor 1 transcriptional activators to viral and chromosomal enhancer sequences. *J. Virol.* **65**:42–50.
 29. **Tsichlis, P. N., P. G. Strauss, and M. A. Lohse.** 1985. Concerted DNA rearrangements in Moloney murine leukemia virus-induced thymomas: a potential synergistic relationship in oncogenesis. *J. Virol.* **56**:258–267.
 30. **Wotton, D., J. Ghysdael, S. Wang, N. A. Speck, and M. J. Owen.** 1994. Cooperative binding of Ets-1 and core binding factor to DNA. *Mol. Cell. Biol.* **14**:840–850.
 31. **Yoshimura, F. K., and T. Wang.** 2001. Role of the LTR region between the enhancer and promoter in mink cell focus-forming murine leukemia virus pathogenesis. *Virology* **283**:121–131.
 32. **Zaiman, A. L., and J. Lenz.** 1996. Transcriptional activation of a retrovirus enhancer by CBF (AML1) requires a second factor: evidence for cooperativity with c-Myb. *J. Virol.* **70**:5618–5629.
 33. **Zaiman, A. L., A. F. Lewis, B. E. Crute, N. A. Speck, and J. Lenz.** 1995. Transcriptional activity of core binding factor α (AML1) and β subunits on murine leukemia virus enhancer cores. *J. Virol.* **69**:2898–2906.
 34. **Zaiman, A. L., A. Nieves, and J. Lenz.** 1998. CBF, Myb, and Ets binding sites are important for activity of the core I element of the murine retrovirus SL3–3 in T lymphocytes. *J. Virol.* **72**:3129–3137.