JOURNAL OF VIROLOGY, Mar. 1984, p. 828–840
0022-538X/84/030828-13$02.00/0

Vol. 49, No. 3

Molecular Analysis of the Envelope Gene and Long Terminal Repeat of Friend Mink Cell Focus-Inducing Virus: Implications for the Functions of These Sequences

WERNER KOCH,[1] WOLFGANG ZIMMERMANN,[1] ALLEN OLIFF,[2] AND ROLAND FRIEDRICH[1]*

Institute of Immunobiology, University of Freiburg, D-7800 Freiburg, Federal Republic of Germany,[1] and DeWitt Wallace Research Laboratory, Memorial Sloan-Kettering Cancer Center, New York, New York 10021[2]

Received 26 October 1983/Accepted 1 December 1983

We sequenced the envelope (*env*) gene and 3′ long terminal repeat of a Friend mink cell focus-inducing virus (F-MCFV). We also sequenced the gp70 coding regions for two cDNA clones of another F-MCFV. The deduced amino acid sequence of the *env* gene products of both F-MCFVs were compared to the corresponding sequences of other MCFVs and of ecotropic viruses. The *env* polypeptides of the different viruses showed long stretches of homology in the carboxy-terminal half of gp70 and in $p15^{env}$ ("constant region"). The amino-terminal half of gp70 was very similar in all MCFVs, but showed extensive variations relative to the ecotropic viruses ("differential region"). This differential region in all MCFVs is of endogeneous origin. We show evidence that this region carries determinants for ecotropic or polytropic host range. No indication could be found that the *env* gene products determine the histological type of disease caused by particular MCFVs. When the long terminal repeats of F-MCFV and Friend murine leukemia virus were compared with those of other viruses causing either lymphatic leukemia or erythroleukemia, several nucleotides were localized which might determine the histological type of disease caused by these viruses.

Mink cell focus-inducing viruses (MCFVs) are polytropic viruses, able to induce cytopathic foci when grown on mink lung fibroblasts (16). The association between a variety of leukemias and MCFVs is becoming increasingly apparent. MCFVs have been implicated in the generation of thymic lymphomas in AKR mice and in the leukemias induced by Moloney murine leukemia virus (M-MuLV), Rauscher MuLV, and Friend MuLV (F-MuLV) (7, 12, 50). Antigenic and tryptic peptide analysis of viral gene products, RNase T_1 oligonucleotide mapping, and heteroduplex, restriction enzyme, and sequence analyses reveal that MCFVs are recombinants between ecotropic MuLVs and endogenous xenotropic virus-like sequences (7, 12, 50).

F-MuLV is an ecotropic, helper-independent, type C retrovirus which induces erythroleukemia at a high rate in newborn mice and lymphatic leukemia with a long latency in adult mice (53). A subgenomic DNA fragment of F-MuLV containing the 3′ portion of the polymerase (*pol*) gene, the *env* gene, a single long terminal repeat (LTR), and the 5′ portion of the *gag* gene carries the pathogenic functions of F-MuLV (33). A 2.4-kilobase-pair DNA fragment ranging from the 3′ portion of the *pol* gene to the 3′ portion of the *env* gene contains sufficient information to convert a nonpathogenic MuLV into a leukemia-inducing virus (34). However, these sequences are not sufficient to impart the high disease incidence and brief latency period of F-MuLV to other MuLVs. These properties must be specified by other sequences in the F-MuLV genome.

The onset of erythroleukemia in mice infected with F-MuLV is invariably associated with the generation of Friend MCFV (F-MCFV) (40, 42, 54). F-MuLV replicates to high titers in both resistant strains of mice and in strains that are susceptible to F-MuLV-induced disease; however, F-MCFV is only generated in susceptible strains of mice (40, 41). F-MCFV itself induces the same type of erythroproliferative disease as F-MuLV. Resistance to this disease correlates with the endogenous expression of a xenotropic-like gp70-related protein that may interfere with the replication or spread of F-MCFV. These results suggest that F-MCFV is a crucial intermediate in the induction of disease by F-MuLV.

Restriction enzyme analysis of a molecularly cloned F-MCFV (clone pFM54B) indicates that a DNA fragment of F-MuLV containing most of the *env* gene has been replaced by endogenous *env*-like sequences (31). To better understand recombined *env* genes, we sequenced this region of the F-MCFV genome. In addition, we sequenced two cDNA clones of the *env* gene of an F-MCFV which we observed in an established Friend virus-infected cell line (45). To identify the sequences which may specify the histological type of leukemia caused by F-MuLV and F-MCFV, the LTR regions of these viruses were also sequenced. These sequences were compared with the LTR sequences of other lymphatic leukemia- and erythroleukemia-inducing murine retroviruses.

MATERIALS AND METHODS

Virus and plasmid. F-MCFV is a helper-independent, polytropic retrovirus which was isolated from a leukemic NIH Swiss mouse after inoculation with a biologically cloned F-MuLV isolate, F-MuLV 201. F-MCFV DNA was molecularly cloned in pBR322 as described (31). One of the resulting clones, pFM54B, which represents a genomic clone of F-MCFV containing a single LTR, was used for sequencing. F-MuLV clone 57 is a helper-independent, ecotropic, highly leukemogenic virus which induces erythroleukemia after inoculation into newborn NIH Swiss mice in the absence of spleen focus-forming virus (SFFV) (32, 33). F-MuLV 57 DNA (pF-MuLV57) was molecularly cloned from viral DNA isolated from fibroblasts infected with F-MuLV clone 201. pF-MuLV57 is a genomic clone which carries a

* Corresponding author.

single LTR (33). pF-MuLV57 was used for all F-MuLV sequence analyses.

Cloning of viral mRNA. Polyadenylate-containing RNA from the Eveline cell line, a Friend virus-infected cell line from an STU mouse (45), was prepared as described by Chirgwin et al. (5). Double stranded cDNA was synthesized and tailed with 3'-dCMP and ligated to vector DNA by the method of Land et al. (21). pBR322 and pBR327 (47) cut with *Pst*I and tailed with 3'-dGMP were used as vectors for constructing recombinant plasmids. Tetracycline-resistant (ampicillin-sensitive) transformants were screened by colony hybridization (14). A nick-translated DNA fragment of pF-MuLV57 (*Hin*dIII-*Kpn*I fragment of 3,300 base pairs [bp]) (20) was used to probe for plasmids containing F-MuLV *env* gene sequences. Two clones (pFM1 and pFM2) were obtained which hybridized to the F-MuLV probe.

Cloning of pFM2 in m13mp8 and dideoxy sequencing. The *Pst*I double-stranded cDNA insert of pFM2 was reinserted into the *Pst*I site of m13mp8 replicative-form DNA (J. Messing, Methods Enzymol., in press), and recombinants were transfected into *Escherichia coli* JM103 (27). Bacteria from white plaques were grown in small cultures and screened for m13mp8 replicative-form DNA containing inserts by using a miniscreen procedure (2). Single-stranded phage DNA was isolated from the cultures and sequenced according to the dideoxy nucleotide chain-terminating procedure developed by Sanger et al. (43), using a primer consisting of 15 nucleotides (Bethesda Research Laboratories, Gaithersburg, Md.).

Restriction enzyme analysis. Preparation of plasmid DNA for restriction enzyme analysis and sequencing was performed as previously described (20). Restriction endonucleases were obtained from Bethesda Research Laboratories, Gaithersburg, Md., Boehringer Mannheim, Mannheim, West Germany, and New England Biolabs, Beverly, Mass., and were used according to the specifications of the manufacturer. Restriction enzyme analysis of DNA was performed as described (20).

DNA sequencing by base-specific chemical cleavages. Sequence analysis of the envelope genes of plasmids pFM54B and pFM1 and of the LTR regions of pFM54B and pF-MuLV57 was performed by base-specific chemical cleavages (26). Restriction enzyme fragments were terminally labeled with [γ-^{32}P]ATP (Amersham Radiochemical Centre, Amersham, England) and T4 polynucleotide kinase (Bethesda Research Laboratories). A+G, G, A>C, C>T, and C+T reactions were used for base-specific modifications.

RESULTS AND DISCUSSION

Sequence determination of the *env* gene and joining regions of F-MCFV. Comparison of the restriction enzyme maps of pFM54B (31) and pF-MuLV57 (20, 31) indicated that the *env* gene and LTR of pFM54B were located within a region ranging from the single *Sph*I site to the *Hin*dIII site in the *gag* gene. A detailed restriction enzyme analysis of this region was performed (Fig. 1C). Nucleotide sequences of the 3' terminus of the *pol* gene, the *env* gene, and the LTR of pFM54B were determined by using the method of Maxam and Gilbert (26). The sequencing strategy is shown in Fig. 1C. The *env*-specific *Pst*I insert of pFM1 (about 1,400 bp) was mapped by digestion with restriction endonucleases (Fig. 1B). Comparison of the restriction enzyme maps of pFM1 (Fig. 1B) and pFM54B (Fig. 1C) suggested that pFM1 contained the region coding for gp70 of F-MCFV. The *Pst*I insert of pFM1 was sequenced according to the method of Maxam and Gilbert (26). The sequencing strategy is shown in Fig. 1B. The *env*-specific *Pst*I insert of pFM2 (about 400 bp) was recloned into the *Pst*I site of m13mp8 replicative-form DNA. Both strands of the insert were sequenced by using the method of Sanger et al. (43). Comparison of the overlapping sequences of pFM2 and pFM1 revealed that both clones were identical in this region (see below). Thus, we assumed that pFM1 and pFM2 are derived from identical F-MCFV genomes. In addition to the sequence coding for the amino terminus of gp70, pFM2 contains the region coding for the peptide leader of the *env* polyprotein. A restriction enzyme map of the *Pst*I insert of pFM2 is shown in Fig. 1A. The nucleotide sequences of the envelope genes and joining regions of pFM54B and pFM1/pFM2 are shown in Fig. 2.

Sequence determination of the LTR of F-MuLV and of a region 5' to the *env* gene. Restriction enzyme maps of the 3'-terminal portion of the *pol* gene, the *env* gene, the LTR of F-MuLV, and the nucleotide sequence of the *env* gene of F-MuLV have been published before (20, 24, 31, 32). We determined the sequence of the 3' terminus of the *pol* gene and the sequence of the LTR of pF-MuLV57. The restriction enzyme maps and the sequencing strategies for these regions are shown in Fig. 1D.

3'-Terminal portion of the *pol* gene of F-MCFV (pFM54B) and F-MuLV. Figure 3 shows the nucleotide sequences of the 3'-terminal regions of the *pol* genes and the 5'-terminal regions of the *env* genes of F-MCFV (pFM54B) and F-MuLV. The *pol* gene and *env* gene overlap by 58 nucleotides (19 1/3 codons) in both viral genomes. Within the overlapping region of *pol* and *env* genes, there are 13 differences in the nucleotide sequences of F-MCFV and F-MuLV. These differences lead to 12 amino acid differences in the frame coding for the *env* polypeptide but to only 3 differences in the frame coding for the DNA polymerase. Thus, the carboxy-terminal region of the polymerase is conserved to a higher degree than the amino terminus of the *env* polypeptide.

LTRs of F-MCFV and F-MuLV. Retroviral LTRs appear to contain sequences that control synthesis of progeny viral RNA and viral mRNA, reverse transcription of viral RNA by DNA polymerase, and integration of viral DNA into cellular DNA and contribute to the pathogenic potential of retroviruses (8, 33, 51, 52). In addition, U_3 sequences can influence the tissue tropism exhibited by some murine leukemia viruses (4, 9). Presumably, these sequences contain tissue-specific transcriptional enhancers which control the level of viral gene expression in different target cells. LTRs of several species of retroviruses differ in size and nucleotide sequence (6, 10, 18, 48, 49, 56). Despite these differences, certain nucleotide sequences which are believed to be important for the functions mentioned above are present in the LTRs of different virus species (8, 51, 52).

We sequenced the LTRs of F-MCFV (pFM54B) and F-MuLV (pF-MuLV57) (Fig. 2 and 4). By comparison with known sequences, we localized the regions U_3-R-U_5, where U_3 and U_5 designate unique sequences originating from the 3' and 5' ends, respectively, of viral RNA and R designates a sequence of viral RNA that is redundant at both ends. The nucleotide sequences of the LTRs of F-MCFV and F-MuLV were very similar to the sequences obtained for other murine (3, 6, 22, 46) and feline (11, 15; M. Wünsch, A. S. Schulz, W. Koch, R. Friedrich, and G. Hunsmann, EMBO J., in press) retroviruses and contained the putative control elements generally present in retroviral LTRs (8, 52). The LTR of F-MCFV consists of 514 bp. The LTR of F-MuLV contains 591 bp. The U_3 regions of both LTRs start with 13 nucleotides

which are found as an inverted repeat at the 3′ termini of the U_5 regions. This inverted repeat is also contained in the LTRs of several other murine (3, 6, 11, 15, 22, 46) and feline (11, 15) retroviruses. The U_3 of F-MuLV contains a stretch of 65 nucleotide pairs which is repeated in tandem. As suggested for similar direct repeat sequences in the U_3 of other retroviruses, this region presumably has an enhancer function for effective expression of the proviral DNA (8, 52). F-MCFV contains no direct repeat. The 3′ portion of the direct repeat of F-MuLV has an insert of nine nucleotide pairs which is part of a small, nearly perfect direct repeat.

A CAT and a TATA box, which presumably are necessary for the initiation of transcription (8, 52), are located in the U_3 of F-MCFV and F-MuLV. A possible Z-DNA segment which is believed to be involved in regulation of transcription (30) is also present in the U_3 region of F-MCFV and F-MuLV. Signals which are hypothesized to be involved in polyadenylation of viral RNA (AATAAA, 22 bp before the start of U_5, and CA, at the very end of R) are found in F-MCFV and F-MuLV.

Sequences in the LTR as possible determinants of the histological type of leukemia. When we aligned the LTRs of F-MCFV and F-MuLV with the LTRs of the lymphatic-inducing viruses Moloney MCFV (M-MCFV) and M-MuLV

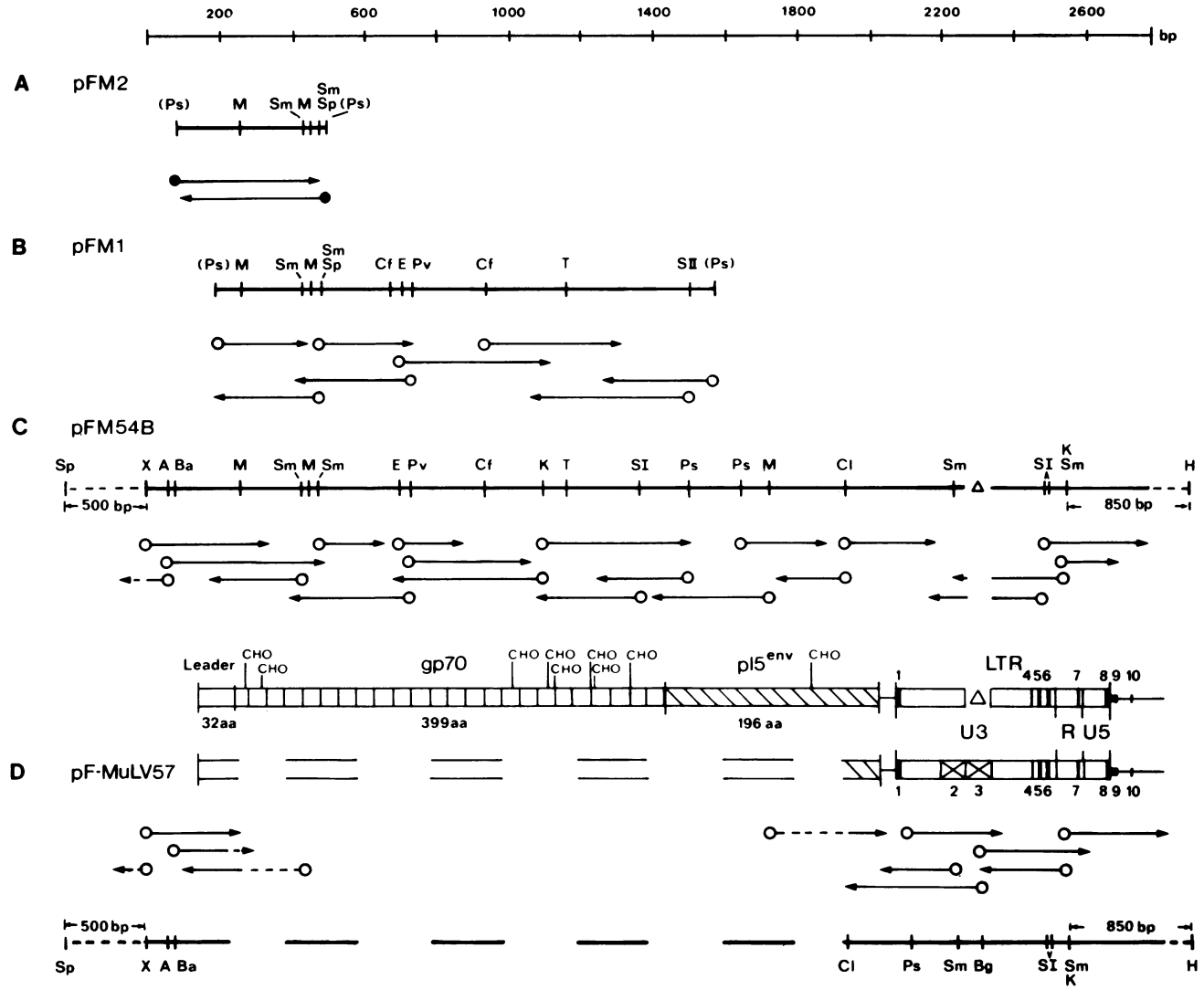

FIG. 1. Physical mapping and use of restriction enzyme fragments for sequence analyses. Closed circles represent oligonucleotide primers used for polymerization reactions (dideoxy nucleotide sequencing). Open circles represent 5′ ends labeled with [γ-^{32}P]ATP (sequencing with base-specific chemical cleavages). Arrows indicate fragments from which unambiguous sequences were obtained. (A) Restriction endonuclease cleavage sites of the *env*-specific *Pst*I insert of clone pFM2. (B) Restriction endonuclease cleavage sites of the *env*-specific *Pst*I insert of clone pFM1. (C) Restriction endonuclease cleavage sites of the *Sph*I-*Hin*dIII fragment of pFM54B containing the *env* gene and LTR of F-MCFV. A triangle indicates the 74-bp deletion in the LTR of F-MCFV with respect to F-MuLV. (D) Restriction endonuclease cleavage sites of the 3′ end of the *pol* gene and the LTR of F-MuLV within the *Sph*I-*Hin*dIII fragment. The numbers in and near the LTRs of F-MCFV (C) and F-MuLV (D) indicate functional regions as follows: 1 and 8, inverted repeats at the ends of the LTRs; 2 and 3, direct repeats in the U_3 of F-MuLV; 4, CAT box; 5, putative Z-DNA segment; 6, TATA box; 7, polyadenylation signal; 9, primer binding site for minus-strand synthesis; 10, splice donor site for generation of subgenomic 21S mRNA. Abbreviations: CHO, potential glycosylation site; A, *Acc*I; Ba, *Bam*HI; Bg, *Bgl*I; Cf, *Cfo*I; Cl, *Cla*I; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; M, *Mbo*II; Ps, *Pst*I; (Ps), *Pst*I site generated by cloning; Pv, *Pvu*II; Sm, *Sma*I; Sp, *Sph*I; SI, *Sst*I; SII, *Sst*II; T, *Tth*111I; X, *Xba*I.

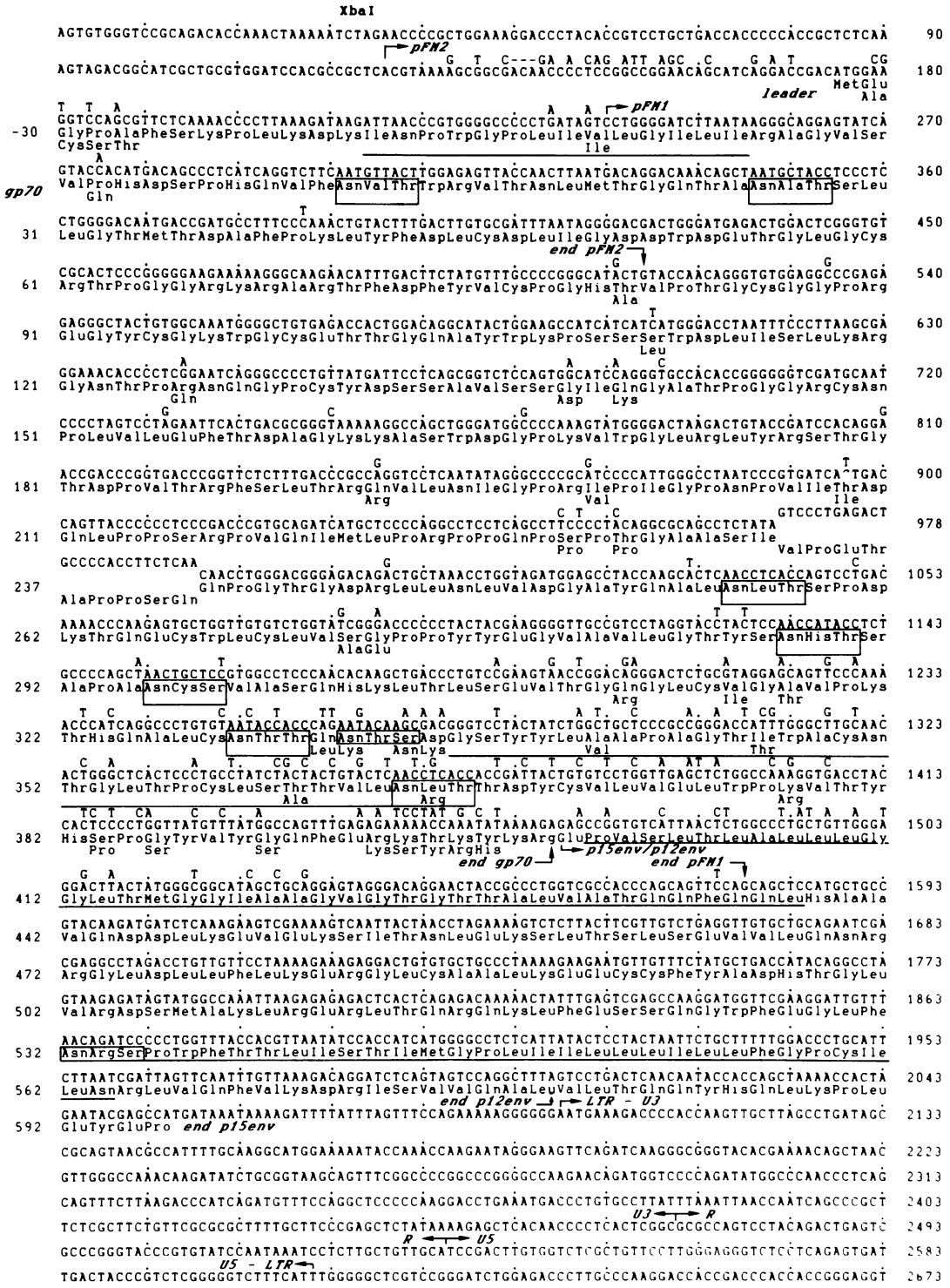

FIG. 2. Nucleotide sequence and deduced amino acid sequence of the *env* gene of F-MCFV (pFM54B). Differing nucleotides and amino acids of pFM1 and pFM2 are also shown. Regions of uncharged amino acids are underlined. Potential glycosylation sites are enclosed in boxes. The sequence presented here differs from the preliminary sequence distributed at the Cold Spring Harbor RNA Tumor Virus Meeting in May 1983 at positions 896, 1539, and 1969. Also given is the nucleotide sequence of the LTR of F-MCFV (pFM54B).

(Fig. 4) we found a high degree of homology between the LTR sequences. However, some nucleotide exchanges as well as insertions and deletions were observed. Most differences were in the U_3 regions. M-MuLV has a perfect major tandem direct repeat flanked by small direct repeats, whereas F-MuLV has a tandem direct repeat which contains a few mismatches and a 9-bp insertion/deletion. Neither F-MCFV nor M-MCFV showed a large direct repeat. Since M-MCFV

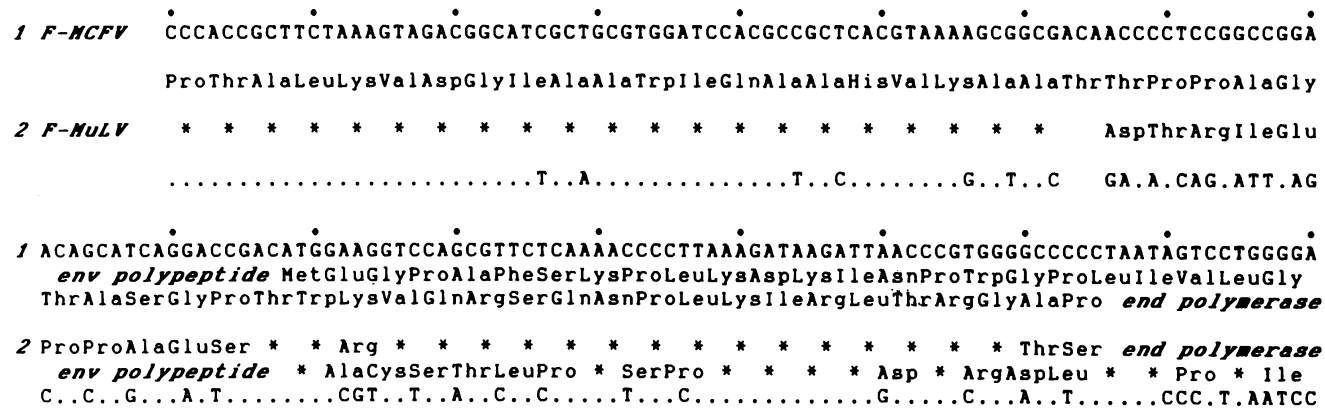

FIG. 3. Nucleotide sequence of the 3′ ends of the *pol* genes and the 5′ ends of the *env* genes of F-MCFV (pFM54B) and F-MuLV. The amino acids of the frames for polymerase and *env* polypeptide are also given. Dots indicate identical nucleotides and asterisks denote identical amino acids in F-MuLV when compared to F-MCFV.

is highly leukemogenic, the presence of a large direct repeat in U_3 apparently is not necessary for the leukemogenic potential of MCFVs. When we compared polytropic viruses F-MCFV and M-MCFV with the ecotropic viruses F-MuLV and M-MuLV, we did not observe a single nucleotide which was specific for either polytropic or ecotropic viruses. Thus, we assume that the sequences responsible for the expanded host tropism of MCFVs are not located in the LTR but in another part of the genome (see below).

We compared the LTRs of the Friend viruses with those of Moloney viruses. We found certain positions (e.g., the region between nucleotides 2093 and 2115 or between nucleotides 2188 and 2192 of F-MCFV) (Fig. 4) which are different between these two groups of viruses. Recent experiments in other laboratories have suggested that the LTR plays a crucial role in determining the type of disease caused by different leukemia viruses (4, 9). The differences in the LTRs described above may be the molecular basis for these properties. Friend SFFV (F-SFFV) is an erythroleukemia-inducing virus. The LTR of F-SFFV (6) is nearly identical to the LTR of F-MCFV (comparison not shown). Therefore, the LTR sequence differences between F-MCFV and the Moloney viruses could be determinants of erythroid versus lymphoid virally induced leukemia. However, when the LTR sequence of AKR MCF247 (a virus which accelerates T-cell leukemia in AKR mice) (19) was also examined (comparison not shown), only some of these nucleotide differences remained. These sequences are indicated by an arrow in Fig. 4.

Comparison of the *env* genes of pFM54B and pFM1/pFM2. Figure 2 shows the composition of the nucleotide sequences obtained for the F-MCFV clones pFM54B, pFM1, and pFM2. The sequence of pFM54B has one large open reading frame consisting of 2,028 nucleotides. This frame starts at nucleotide 28 and terminates at position 2055 with two stop condons. The amino acid sequences for this reading frame starting at the first ATG codon are also shown in Fig. 2.

The sequence coding for the peptide leader of the *env* polyprotein in pFM54B and pFM2 consists of 32 codons. There are seven nucleotide differences in this region of these MCFV clones, five of which give rise to a change in the corresponding amino acid. The open reading frame which encodes gp70 encompasses 1,197 nucleotides (399 codons) in pFM54B and 1,224 nucleotides (408 codons) in pFM1/pFM2. In the sequence coding for the proline-rich region (PR) of gp70 (see below), 27 nucleotides (nine codons) are deleted in pFM54B at position 978. In the gp70's of pFM54B and pFM1, 103 (8.6%) nucleotide differences exist leading to 33 (8.3%) amino acid changes (including the deletion of 27 nucleotides). Nineteen of these nucleotide changes (11 amino acid changes) are situated in the *env* sequence from the initiation codon up to the end of the PR (differential region, see below), whereas 84 nucleotide changes (22 amino acid changes) occur in gp70 downstream of the PR.

The 27 nucleotides in the PR of pFM1 which are deleted from pFM54B are present in other polytropic viruses (3, 25; C. Holland, personal communication) and in F-SFFV (1, 6, 58). Interestingly, this deletion in pFM54B gives a portion of the PRs which are different in F-MCFV and F-MuLV (positions 943 to 963 in pFM54B in Fig. 5A) the exact same length between stretches which are homologous in the two *env* genes. In vitro recombinants of F-MuLV containing the *env* gene of pFM54B yielded infectious but nonpathogenic virus (unpublished data). It is possible that this deletion is responsible for the nonpathogenicity of this clone. Experiments are in progress to test the importance of these deleted sequences in causing disease.

Comparison of the *env* genes and *env* polypeptides of F-MCFV and F-MuLV. Figures 5A and B show the 3′ termini of the *pol* and *env* genes of F-MCFV (pFM54B) and F-MuLV (pF-MuLV57). Also shown are the amino acid sequences of the *env* polypeptides. Alignment was performed to obtain maximal homology of the two nucleotide sequences.

Comparison of the nucleotide sequences of F-MCFV (pFM54B) and F-MuLV show identical sequences in the *pol* gene of both viruses up to nucleotide 36. The two viral sequences then vary up to nucleotide 1509 in F-MCFV. We therefore conclude that the nucleotides between positions 36 and 1650 in F-MuLV have been replaced by endogenous sequences, thereby forming F-MCFV (clone pFM54B). The substitution begins in the 3′ portion of the *pol* gene about 120 nucleotides upstream from the initiator ATG of the *env* gene and ends at or near nucleotide 1510 (in F-MCFV) within the sequence coding for $p15^{env}$. The substituted sequence in F-MCFV (pFM54B) is 141 nucleotides shorter than the corresponding F-MuLV sequence. The two *env* genes share the same initiation and termination codons and are translated in the same reading frame.

Within the region coding for $p15^{env}$, the nucleotide se-

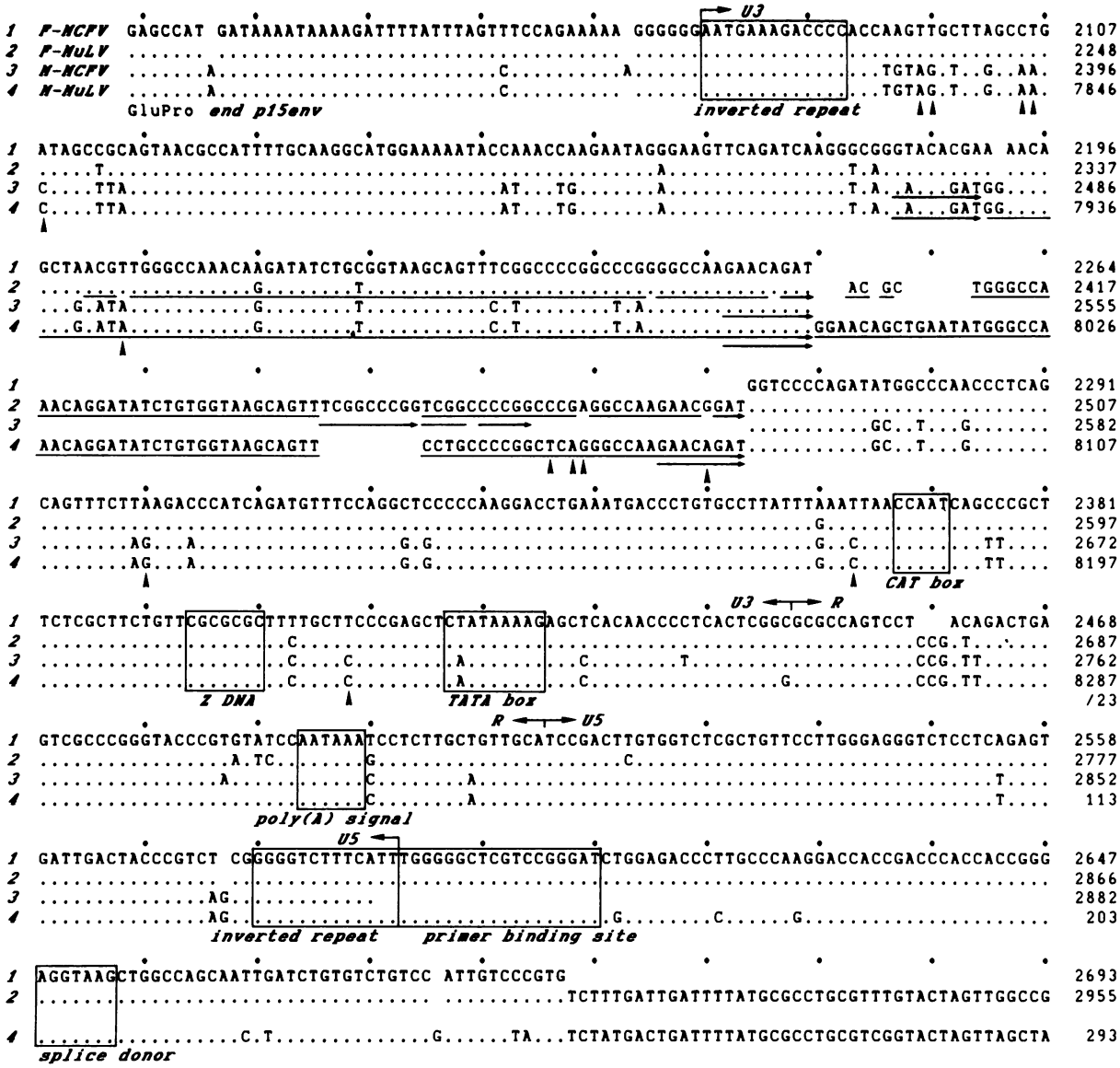

FIG. 4. Nucleotide sequences of the LTRs and adjacent regions of F-MCFV (pFM54B), F-MuLV, M-MCFV (3), and M-MuLV (46). The sequences were aligned to obtain maximal homology. Arrows denote some of the differences between lymphatic leukemia- and erythroleukemia-inducing viruses (see text). Dots indicate identical nucleotides as in F-MCFV. Gaps have been introduced to allow optimal alignment. Regions of functional importance are enclosed in boxes. Direct repeats are underlined. Position numbers for M-MCFV and M-MuLV are taken from references 3 and 46, respectively.

quences of F-MCFV (pFM54B) and F-MuLV are identical from nucleotide 1510 to nucleotide 2112 (5′ terminal portion of the U_3 region in the LTR) with one exception: a change at nucleotide 1693 in the region encoding $p15^{env}$ in F-MCFV leads to an arginine, whereas F-MCFV carries a glycine at the corresponding position. In all MCFVs studied so far, the borders of the substituted regions can be recognized by a change from complete homology (outside the recombination site) to sequence variations (within the recombination sites). Since the change at position 1693 in F-MCFV is located far downstream (180 nucleotides) from the proposed recombination site of F-MCFV and F-MuLV (see below), we suggest that this difference is due to a point mutation which occurred after the generation of F-MCFV.

The substitution in pFM1/pFM2 is considerably shorter than that in pFM54B. It begins within the sequence coding for the peptide leader at or near nucleotide 166 in F-MuLV (Fig. 5A) and ends near position 1211. There are few nucleotide differences between pFM1/pFM2 and F-MuLV clone 57 outside the substituted region (e.g., positions 173, 175, 1152, 1161, 1191, and 1227; Fig. 2). Since pFM1/pFM2 are derived from a Friend virus-infected cell line of an STU mouse (44) and not from F-MuLV clone 201 as are pFM54B and pF-MuLV57, we consider these differences to represent point mutations present in the progenitor MCFV of pFM1/pFM2.

The *env* gene product of F-MCFV (pFM54B), as deduced from the nucleotide sequence, is 636 amino acids long (including the deleted sequence in the PR; Table 1). This is 39 amino acids shorter than the proposed *env* polypeptide

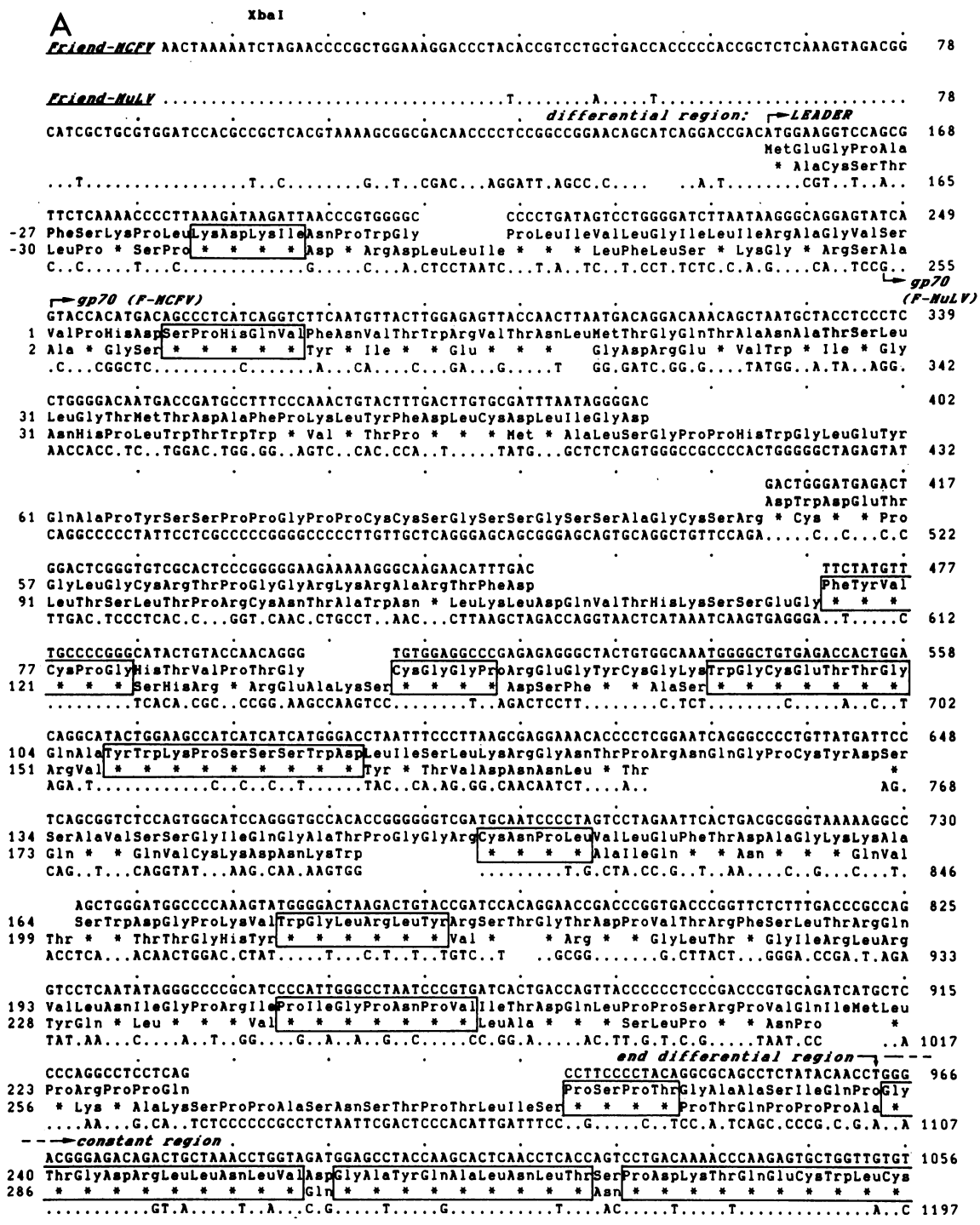

FIG. 5

chain of F-MuLV (675 amino acids). Most of the differences in the *env* polypeptides of F-MCFV and F-MuLV are located in the leader peptide and in gp70, whereas p15^{env} is highly conserved. The proposed signal peptides of the *env* precursor polypeptides of F-MCFV and F-MuLV consist of 32 and 34 amino acids, respectively. The nucleotide and amino acid sequences of the F-MCFV leader peptide are very different from those of the parental F-MuLV. However, in this portion of the *env* polypeptide of both viruses, a stretch of uncharged, mostly apolar amino acids is found. The conservation of this hydrophobic region points to the important function of the signal peptide in the excretion of the *env* polypeptide through cell membranes.

The signal peptide cleavage site of the F-MuLV precursor polypeptide has been determined (23), whereas this site is not known for F-MCFV. In Rauscher MCFV, the mature gp70 polypeptide chain starts with NH_2-Val-Gln-His-Asp (44). We therefore assume that the mature gp70 of F-MCFV

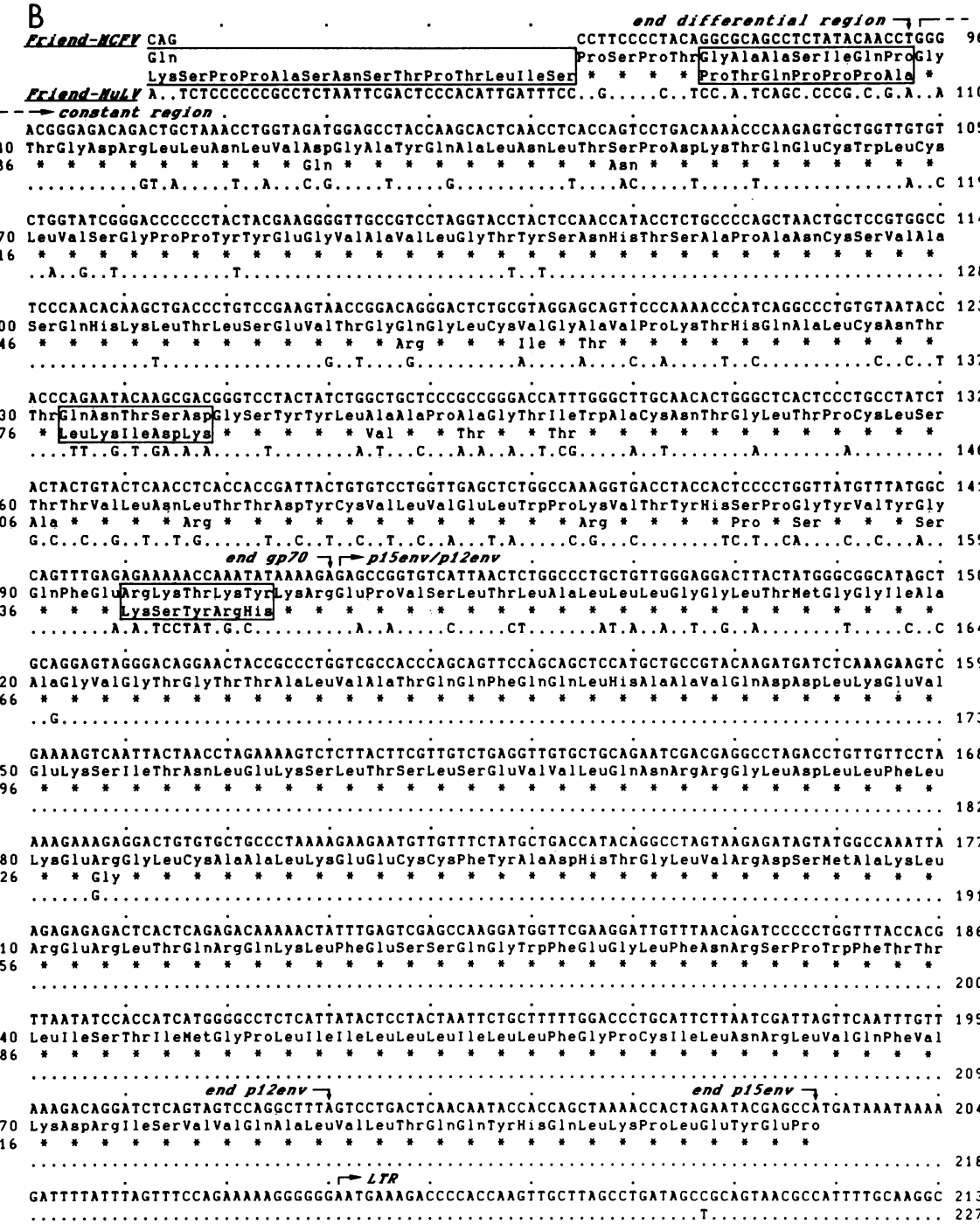

FIG. 5. Nucleotide sequences and deduced amino acid sequences of the *env* genes of F-MCFV (pFM54B) and F-MuLV (pF-MuLV57) (20). The two nucleotide sequences were aligned to obtain maximal homology. Asterisks indicate identical amino acids in F-MuLV with respect to F-MCFV. Gaps have been introduced to allow optimal alignment. (A) 5′ portions of the *env* genes and their products of F-MCFV and F-MuLV representing the differential region (see text). Stretches of identity (at least four identical amino acids) between the two *env* polypeptides are enclosed in boxes. (B) 3′ portions of the *env* genes and carboxy-terminal portion of the *env* polypeptides of F-MCFV and F-MuLV representing the constant region (see text). Stretches of divergence (at least five different amino acids) between the two *env* polypeptides are enclosed in boxes.

begins with the related sequence NH_2-Val-Pro-His-Asp. The carboxy terminus of gp70 of F-MCFV has been located by analogy to F-MuLV (20).

The amino acid sequences of F-MCFV and F-MuLV are very different in the amino-terminal portion of gp70 up to position 238 (in F-MCFV) (Fig. 5A). We call the *env* gene sequence up to position 963 in F-MCFV the differential region, since only a few stretches of identity between F-MCFV and F-MuLV can be detected (boxes in Fig. 5A).

Starting at amino acid 239 in F-MCFV, the carboxy-terminal regions of gp70 in F-MCFV and F-MuLV are very similar (Fig. 5B). In this constant region, only two stretches with significant differences (boxes in Fig. 5B) are present.

The *env* precursor polypeptides of F-MCFV, F-MuLV (20), and other mammalian retroviruses (3, 11, 22, 46; Wünsch et al., in press) contain stretches of uncharged, mostly nonpolar amino acids located in homologous regions. The apolar regions of F-MCFV are underlined in Fig. 2. Some of these regions could allow polypeptides to penetrate through or integrate in cellular and viral membranes according to the model proposed by Lenz et al. (22).

Possible role of the PR of gp70. The gp70's of all MCFVs and MuLVs examined to date and the gp52 of F-SFFV contain a region strongly enriched in proline residues (PR) at the carboxy terminus of the differential region. In Fig. 6, the PRs of two F-MCFVs, M-MCFV (3), F-SFFV (6, 58), F-MuLV (20), M-MuLV (46), and Akv (22) are compared. In these regions the relative amount of proline residues varies between 31 and 36% depending on the virus. MCFVs and F-SFFV have similar amino acid sequences in their PRs, whereas the ecotropic MuLVs have different sequences compared with the MCFVs and F-SFFV. The PRs of all ecotropic MuLVs include a stretch of about 35 amino acids where the gp70's of these viruses are remarkably different with respect to each other (20) (hypervariable region of ecotropic gp70's; see below). As discussed previously (20), the different amino acid sequences of the PRs of the gp70's of ecotropic MuLVs could allow these viruses to interact with specific receptors, thus defining the host range of these viruses.

The PR of clone pMo-MCF-16 (3) of the lymphatic leukemia-inducing M-MCFV and the PR of clone pFM1 of the erythroleukemia-inducing F-MCFV are nearly identical. Therefore, we can exclude the possibility that this region is involved in determining the histological type of disease caused by MCFVs. Rather, the PR might be responsible for the polytropic host range of the MCFVs. It could carry the sequences utilized by xenotropic and polytropic viruses for interaction with cell surface receptors. An examination of the *env* sequences of recombinant MuLVs with ecotropic host ranges (36, 38) would clarify this point.

TABLE 1. Calculated molecular weights of proposed F-MCFV *env* polypeptides[a]

Polypeptide	No. of amino acid residues	Mol wt of the apopolypeptide
Total *env* precursor polypeptide	636	69,155
N-terminal signal peptide	32	3,396
gp70	408[b]	43,944[b]
$p15^{env}$	196	21,834
$p12^{env}$	180	19,867

[a] Clone pFM54B was taken for calculation; however, the sequence of 27 nucleotides which is deleted in the PR of this clone was added before calculation.

[b] Since some amino acids might be cleaved off the carboxy terminus of gp70, the actual number might be slightly smaller.

Comparison of the amino acid sequences of *env* polypeptides of different viruses. Figure 7 shows a comparison of the amino acid sequences of the *env* polypeptides of F-MCFV (pFM54B, pFM1, and pFM2), M-MCFV (3), F-SFFV (6, 58) and F-MuLV (20). The sequences of the *env* polypeptides of the MCFVs and F-SFFV have a high degree of similarity. As described above, the sequence of the F-MuLV *env* polypeptide ranging from the leader peptide to the carboxy terminus of the PR is dissimilar with respect to the MCFV and F-SFFV sequence (differential region; Fig. 5A). F-MuLV and these other viruses have very similar polypeptide sequences from the carboxy terminus of the PR region up to the end of $p15^{env}$ (constant region; Fig. 5B). At the carboxy terminus, the $p15^{env}$-related sequence of F-SFFV gp52 is 33 amino acids shorter than the $p15^{env}$ of F-MCFV, M-MCFV, and F-MuLV. The carboxy-terminal six amino acids of gp52 are different with respect to the $p15^{env}$ of the other viruses, which have a common sequence in this region.

Localization of possible glycosylation sites. Asn-X-Thr/Ser sequences are known to be possible sites for glycosylation of retroviral polypeptides (13, 29). Glycosylation sites within the *env* gene products are highly conserved within different host range classes of mammalian retroviruses (3, 11, 20, 22, 46; Wünsch et al., in press). This argues for a strong selective pressure in favor of their conservation. The gp70's of F-MCFV, M-MCFV, and F-MuLV carry six potential glycosylation sites in homologous positions (boxes in Fig. 7). The three amino-terminal sites are also contained in gp52 of F-SFFV, whereas the three carboxy-terminal sites are deleted in F-SFFV. F-MCFV has one additional possible glycosylation site (at amino acid 26) shared by F-MCFV and F-SFFV and one site (at amino acid 332) not found in M-MCFV or F-SFFV. F-MuLV has two additional sites (at amino acids 168 and 266) not found in these other viruses. Whether all of the potential sites in the gp70's and gp52 are glycosylated is unknown at present. The sequence Asn-Arg-Ser in $p15^{env}$ of the *env* polypeptides of murine retroviruses is not glycosylated (28, 37, 57).

General structure of the *env* gene. Comparison of murine retrovirus *env* genes and their products shows that they are composed of clearly distinguishable regions (Fig. 5A, 5B, and 7). The sequence between the amino terminus of the leader peptide and the end of the PR is highly conserved within the polytropic group of viruses and within the ecotropic viruses but very different between polytropic and ecotropic viruses (differential region). The remainder of the *env* polypeptide is composed of a region which is very similar between polytropic and ecotropic viruses (constant region). A comprehensive view of the elements of the *env* polypeptide is given in Fig. 8.

A few sections of the differential region are homologous between the ecotropic and polytropic viruses (boxes in Fig. 5A, dotted areas in Fig. 8). The constant regions contain two blocks with variable character (boxes in Fig. 5B, white areas in Fig. 8).

As described above, the PR is located at the border of the differential and constant regions. In the *env* polypeptides of all ecotropic mouse viruses studied so far, the PR is the only larger segment with considerable variability. The total *env* polypeptide seems, therefore, to consist of the following elements: differential region (without PR), PR, and constant region. The structure of the *env* polypeptide is reminiscent of the structure of the immunoglobulins which consist of sepa-

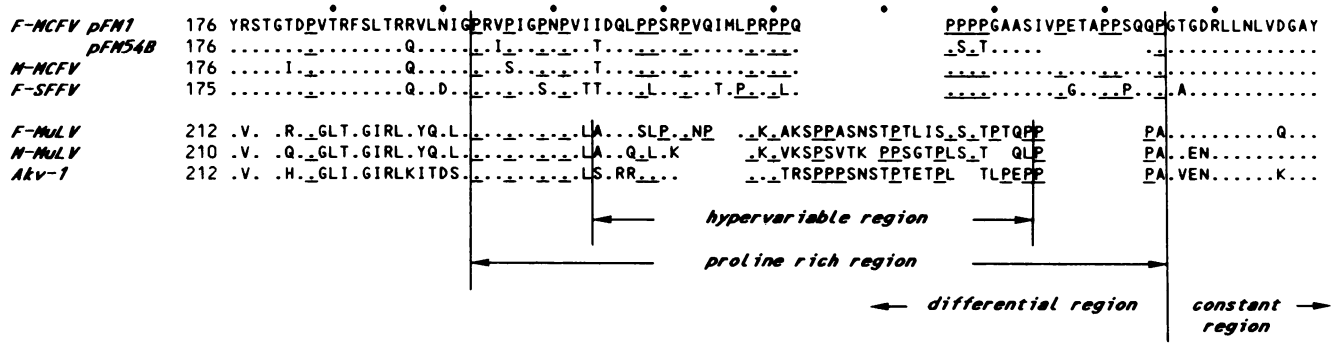

FIG. 6. PRs in the gp70's of F-MCFV (clones pFM1 and pFM54B), M-MCFV (3), F-MuLV (20), M-MuLV (46) and Akv (22) and in the gp52 of F-SFFV (6, 58). The standard one-letter amino acid code is used. Dots indicate identical amino acids with respect to F-MCFV clone pFM1. Proline residues are underlined. Gaps have been introduced to allow optimal alignment. The numbering of the amino acids of the individual sequences refers to their positions in gp70 (or gp52).

rate genetic elements: variable region, PR (hinge), and constant region (17, 55).

It has been suggested recently (Wünsch et al., in press) that the constant portions of gp70 and $p15^{env}$ of mammalian retroviruses are two separate genetic elements which together with the N-terminus of the differential region can be traced to a common short primordial *env*-related gene. We have proposed that the present day *env* gene has been generated by at least two tandem duplications of this primordial *env*-related genetic element and by additional recombinational exchanges and mutations.

Possible functions of the elements of the *env* gene and its products. Comparison of the putative points of recombination which have led to the generation of MCFVs shows that the differential region is the part of the *env* gene in all MCFVs which is replaced by endogenous sequences. In some viruses, e.g., clone pFM1/pFM2 of F-MCFV, little more than the differential region has been substituted. We therefore assume that the differential region is the part of the *env* gene which determines the ecotropic or polytropic host range of the virus. As described above, the PR may contain determinants for this function.

Various parts of the constant region have been replaced by endogenous sequences in the *env* genes of the MCFVs analyzed so far. These endogenous sequences are very similar to their ecotropic partners: pFM1 carries only three and pMo-MCF_1-16 carries only four varying amino acids in this part of the constant region. Thus, we can assume that the constant regions of ecotropic and polytropic viruses carry domains for those functions which are common to all viruses, e.g., determinants for infectivity and structural properties necessary for interaction with cellular membranes.

Do MCFVs contain disease-specific sequences in their *env* polypeptide? Since MCFVs induce the same type of disease as their ecotropic progenitor viruses and since MCFVs are believed to be the ultimate cause of these diseases, we tried to identify sequences which are characteristic for particular histological types of leukemia. It has been shown previously (34) that a 2.4-kilobase-pair fragment of F-MuLV encompassing about 700 bp of the *pol* gene and almost the entire *env* gene contains sequences which contribute to erythroleukemia. Since the *pol* gene is not a likely candidate for determining disease specificity (34), we restricted our analysis to the *env* gene (Fig. 7 and Table 2).

We compared all of the available data on amino acid sequences of the *env* polypeptides of murine retroviruses which cause either lymphatic leukemia or erythroleukemia.

TABLE 2. Comparison of amino acids at selected sites of different viruses

Region	Virus	Reference	Amino acids at the following positions[a]									
			27	139–141	181	200–201	253	282	514	523	550	553
Differential	F-MCFV clone pFM54B	This paper	A	G I Q	T	I P						
	F-MCFV clone pFM1/2	This paper	A	D I K	T	V P						
	F-SFFV	6,58	A	G V L	T	V P						
	F-SFFV	1	A	G V Q	I	I P						
	M-MCFV	3	V	N I K	I	V S						
	MCF247	C. Holland[b]	A	D I K	I	V P						
Constant	F-MCFV clone pFM54B	This paper					Q	V	T	S	I	L
	F-MCFV clone pFM1/2	This paper					Q	V				
	F-SFFV	6,58					Q	V	T	S	I	L
	F-SFFV	1					Q	V	T	S	I	L
	M-MCFV	3					R	I	N	T	V	M
	MCF247	C. Holland[b]					Q	V	S	Q	I	L
	F-MuLV	20					Q	V	T	S	I	L
	M-MuLV	46					Q	V	N	T	V	M
	Akv	22					Q	V	S	Q	I	L

[a] Positions refer to Fig. 7.
[b] Personal communication.

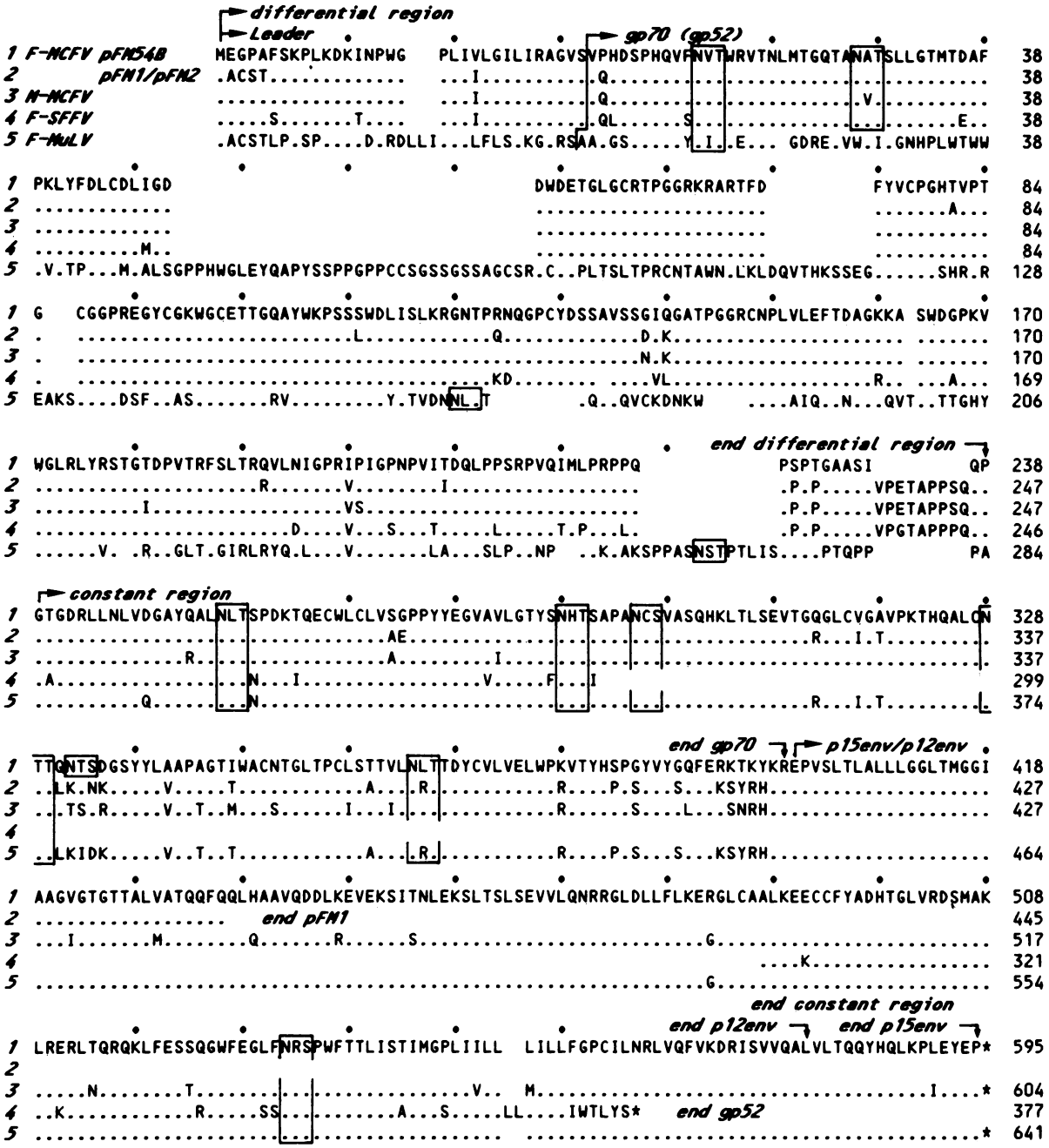

FIG. 7. Comparison of the deduced amino acid sequences of the *env* polypeptides of F-MCFV (clones pFM54B, pFM1, and pFM2), M-MCFV (3), F-MuLV (20), and the *env*-related polypeptide of F-SFFV (6, 58). The standard one-letter amino acid code is used. Sequences are aligned to obtain maximal identity. Comparison of nucleotide sequences helped find proper alignment in some parts. Gaps have been introduced to allow optimal alignment. Dots indicate identical amino acids with respect to F-MCFV (pFM54B). Possible glycosylation sites are enclosed in boxes.

The following differences can be found among the clones listed in Table 2 (differences between the F-MCFV clones pFM1/pFM2 and pFM54B were regarded as not significant and not taken into consideration; positions refer to pFM54B of Fig. 7): positions 27, 139, 141, 181, 201, 253, 282, 292 to 486, 514, 523, 550, 553, and 591. The changes at positions 27, 139, 141, 181, 201, 253, 282, 514, 523, 550, 553, and 591 occur in one or more of the lymphatic leukemia viruses (M-MuLV, M-MCFV, Akv, or AKR MCF247) and in one or more of the erythroleukemia viruses (F-MuLV, F-MCFV, or SFFV). Since these changes occur in both types of viruses, they cannot specify the type of disease produced by these agents. Changes which lie between positions 292 and 486 fall into the region of the F-MCFV genome which is deleted from the SFFV genome.

These comparisons show that there is no site in the amino acid sequence of the *env* polypeptide of different MCFVs which clearly determines their disease specificity. We there-

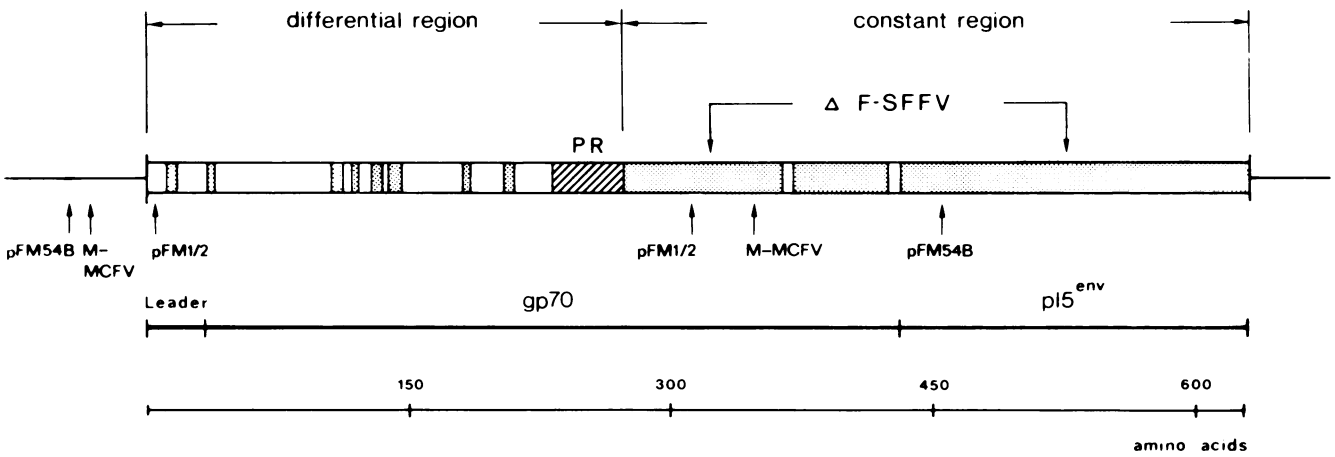

FIG. 8. General structure of *env* polypeptides of MCFVs, MuLVs, and F-SFFV. Differential region, PR, and constant region are explained in the text. Arrows indicate the locations of the putative recombination sites involved in generation of MCFVs. Dotted areas represent stretches of identical amino acids in F-MCFV and F-MuLV. White boxes are stretches of differing amino acids. Also indicated is the major deletion in F-SFFV. The length of the polypeptide is taken from F-MCFV(pFM54B).

fore must assume that the *env* gene of a given MCFV cannot contribute to different histological types of disease and that other sequences in the MCFV genome are responsible for determining what type of disease a particular virus will induce.

We localized several nucleotide positions in the LTRs of different viruses which seem to be characteristic of either lymphatic leukemia- or erythroleukemia-inducing viruses (see above). However, additional biological tests with in vivo- or in vitro-generated recombinant viruses will be necessary to determine the exact role these regions play in determining pathogenicity and disease specificity.

ACKNOWLEDGMENTS

We thank Rita Rossa and Heiko Taube for excellent technical assistance. We thank Stephen Clark, Linda Wolff, Ulf Rapp, and Christie Holland for communication of results before publication.

This research was supported by a grant from the Deutsche Forschungsgemeinschaft and by American Cancer Society grant MV-170.

LITERATURE CITED

1. **Amanuma, H., A. Katori, M. Obata, N. Sagata, and Y. Ikawa.** 1983. Complete nucleotide sequence of the gene for the specific glycoprotein (gp55) of Friend spleen focus-forming virus. Proc. Natl. Acad. Sci. U.S.A. **80**:3913–3917.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
3. **Bosselman, R. A., F. van Straaten, C. van Beveren, I. M. Verma, and M. Vogt.** 1982. Analysis of the env gene of a molecularly cloned and biologically active Moloney mink cell focus-forming proviral DNA. J. Virol. **44**:19–31.
4. **Chatis, P. A., C. A. Holland, J. W. Hartley, W. P. Rowe, and N. Hopkins.** 1983. Role for the 3′ end of the genome in determining disease specificity of Friend and Moloney murine leukemia viruses. Proc. Natl. Acad. Sci. U.S.A. **80**:4408–4411.
5. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
6. **Clark, S. P., and T. W. Mak.** 1983. Complete nucleotide sequences of an infectious clone of Friend spleen focus-forming provirus: gp55 is an envelope fusion glycoprotein. Proc. Natl. Acad. Sci. U.S.A. **80**:5037–5041.
7. **Coffin, J.** 1982. Endogenous viruses, p. 1109–1203. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
8. **Coffin, J.** 1982. Structure of the retroviral genome, p. 261–368. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
9. **DesGroseillers, L., E. Rassart, and P. Jolicoeur.** 1983. Thymotropism of murine leukemia virus is conferred by its long terminal repeat. Proc. Natl. Acad. Sci. U.S.A. **80**:4203–4207.
10. **Dhar, R., W. L. McClements, L. W. Enquist, and G. F. Vande Woude.** 1980. Nucleotide sequences of integrated Moloney sarcoma provirus long terminal repeats and their host and viral junctions. Proc. Natl. Acad. Sci. U.S.A. **77**:3937–3941.
11. **Elder, J. H., and J. I. Mullins.** 1983. Nucleotide sequence of the envelope gene of Gardner-Arnstein feline leukemia virus B reveals unique sequence homologies with a murine mink cell focus-forming virus. J. Virol. **46**:871–880.
12. **Famulari, N. G.** 1983. Murine leukemia viruses with recombinant env genes: a discussion of their role in leukemogenesis. Curr. Top. Microbiol. Immunol. **103**:1–108.
13. **Geyer, R., H. Geyer, G. Hunsmann, J. Schneider, and S. Stirm.** 1982. Oligosaccharides in the surface glycoprotein of Friend murine leukemia virus. Biochim. Biophys. Acta **117**:491–501.
14. **Grunstein, M., and J. Wallis.** 1979. Colony hybridization. Methods Enzymol. **68**:379–389.
15. **Hampe, A., M. Gobbet, J. Even, C. J. Sherr, and F. Galibert.** 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.
16. **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 leukemia. Proc. Natl. Acad. Sci. U.S.A. **74**:789–792.
17. **Honjo, T.** 1983. Immunoglobulin genes. Annu. Rev. Immunol. **1**:499–528.
18. **Ju, G., and A. M. Skalka.** 1980. Nucleotide sequence analysis of the long terminal repeat (LTR) of avian retroviruses: structural similarities with transposable elements. Cell **22**:379–386.
19. **Kelly, M., C. A. Holland, M. L. Lung, S. K. Chattopadhyay, D. R. Lowy, and N. H. Hopkins.** 1983. Nucleotide sequence of the 3′ end of MCF 247 leukemia virus. J. Virol. **45**:291–298.
20. **Koch, W., G. Hunsmann, and R. Friedrich.** 1983. Nucleotide sequence of the envelope gene of Friend murine leukemia virus. J. Virol. **45**:1–9.

21. **Land, H., M. Grez, H. Hauser, W. Lindenmaier, and G. Schutz.** 1981. 5'-Terminal sequences of eucaryotic mRNA can be cloned with high efficiency. Nucleic Acids Res. **9**:2251–2266.
22. **Lenz, J., R. Crowther, A. Stracesci, and W. Haseltine.** 1982. Nucleotide sequence of the Akv *env* gene. J. Virol. **42**:519–529.
23. **Linder, D., S. Stirm, J. Schneider, G. Hunsmann, G. Smythers, and S. Oroszlan.** 1982. Glycoproteins of Friend murine leukemia virus: separation and NH_2-terminal amino acid sequences of gp69 and gp71. J. Virol. **42**:352–355.
24. **Linemeyer, D. L., S. K. Ruscetti, J. G. Menke, and E. M. Scolnick.** 1980. Recovery of biologically active spleen focus-forming virus from molecularly cloned spleen focus-forming virus-pBR322 circular DNA by cotransfection with infectious type C retroviral DNA. J. Virol. **35**:710–721.
25. **Mark, G. E., and U. R. Rapp.** 1984. Envelope gene sequence of two in vitro-generated mink cell focus-forming murine leukemia viruses which contain the entire gp70 sequence of the endogenous nonecotropic parent. J. Virol. **49**:530–539.
26. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
27. **Messing, J., R. Crea, and P. H. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
28. **Moennig, V., H. Frank, G. Hunsmann, J. Schneider, and W. Schafer.** 1974. Properties of mouse leukemia viruses. VII. The major glycoprotein of Friend leukemia virus. Isolation and physicochemical properties. Virology **61**:100–111.
29. **Montreuil, J.** 1980. Primary structure of glycoprotein glycans. Basis for the molecular biology of glycoproteins. Adv. Carbohydr. Chem. Biochem. **37**:157–223.
30. **Nordheim, A., and A. Rich.** 1983. Negatively supercoiled simian virus 40 DNA contains Z-DNA segments within transcriptional enhancer sequences. Nature (London) **303**:674–679.
31. **Oliff, A., L. Collins, and C. Mirenda.** 1983. Molecular cloning of Friend mink cell focus-inducing virus: identification of mink cell focus-inducing virus-like messages in normal and transformed cells. J. Virol. **48**:542–546.
32. **Oliff, A. I., G. L. Hager, E. H. Chang, E. M. Scolnick, H. W. Chan, and D. R. Lowy.** 1980. Transfection of molecularly cloned Friend murine leukemia virus DNA yields a highly leukemogenic helper-independent type C virus. J. Virol. **33**:475–486.
33. **Oliff, A., D. Linemeyer, S. Ruscetti, R. Lowe, D. R. Lowy, and E. Scolnick.** 1980. Subgenomic fragment of molecularly cloned Friend murine leukemia virus DNA contains the gene(s) responsible for Friend murine leukemia virus-induced disease. J. Virol. **35**:924–936.
34. **Oliff, A., and S. Ruscetti.** 1983. A 2.4-kilobase-pair fragment of the Friend murine leukemia virus genome contains the sequences responsible for Friend murine leukemia virus-induced leukemia. J. Virol. **46**:718–725.
35. **Oroszlan, S., L. E. Henderson, T. D. Copeland, A. M. Schultz, and E. M. Rabin.** 1980. Processing and structure of murine leukemia virus *gag* and *env* gene-encoded polyproteins, p. 219–232. *In* G. Koch and D. Richter (ed.), Biosynthesis, modification and processing of cellular and viral polyproteins. Academic Press, Inc., New York.
36. **Pedersen, F. S., R. L. Crother, D. Y. Tenney, A. M. Reimold, and W. A. Haseltine.** 1981. Novel leukaemogenic retroviruses isolated from cell line derived from spontaneous AKR tumour. Nature (London) **292**:167–170.
37. **Pinter, A., and E. Fleissner.** 1977. The presence of disulfide-linked gp70-p15(E) complexes in AKR murine leukemia virus. Virology **83**:417–422.
38. **Rapp, U. R., and G. J. Todaro.** 1978. Generation of oncogenic type C viruses: rapidly leukemogenic viruses derived from C3H mouse cell *in vivo* and *in vitro*. Proc. Natl. Acad. Sci. U.S.A. **75**:2468–2472.
39. **Reddy, E. P., M. J. Smith, E. Canaani, K. C. Robbins, S. R. Tronick, S. Zain, and S. A. Aaronson.** 1980. Nucleotide sequencing analysis of the transforming region and large terminal redundancies of Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. **77**:5234–5238.
40. **Ruscetti, S., L. Davis, J. Feild, and A. Oliff.** 1981. Friend murine leukemia virus-induced leukemia is associated with the formation of mink cell focus-inducing viruses and is blocked in mice expressing endogenous mink cell focus-inducing xenotropic viral envelope genes. J. Exp. Med. **154**:907–920.
41. **Ruscetti, S., J. Feild, L. Davis, and A. Oliff.** 1982. Factors determining the susceptibility of NIH Swiss mice to erythroleukemia induced by Friend leukemia virus. Virology **117**:357–365.
42. **Ruscetti, S., S. D. Linemeyer, J. Feild, D. Troxler, and E. Scolnick.** 1978. Type-specific radioimmunoassays for the gp70s of mink cell focus-inducing murine leukemia viruses: expression of a cross-reacting antigen in cells infected with the Friend strain of the spleen focus-forming virus. J. Exp. Med. **148**:654–663.
43. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.
44. **Schultz, A., A. Rein, L. Henderson, and S. Oroszlan.** 1983. Biological, chemical, and immunological studies of Rauscher ecotropic and mink cell focus-forming viruses from JLS-V9 cells. J. Virol. **45**:995–1003.
45. **Seifert, E., M. Claviez, H. Frank, G. Hunsmann, and W. Schafer.** 1975. Properties of mouse leukemia viruses. XII. Production of substantial amounts of Friend leukemia virus by suspension tissue culture line (Eveline suspension cells). Z. Naturforsch. **30c**:698–700.
46. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukaemia virus. Nature (London) **293**:543–548.
47. **Soberon, X., L. Covarrubias, and F. Bolivar.** 1980. Construction and characterization of new cloning vehicles. IV. Deletion derivatives of pBR322 and pBR325. Gene **9**:287–305.
48. **Sutcliffe, J. G., T. M. Shinnick, I. M. Verma, and R. A. Lerner.** 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. U.S.A. **77**:3302–3306.
49. **Swanstrom, R., W. J. Delorbe, J. M. Bishop, and H. E. Varmus.** 1981. Nucleotide sequence of cloned unintegrated avian sarcoma virus DNA: viral DNA contains direct and inverted repeats similar to those in transposable elements. Proc. Natl. Acad. Sci. U.S.A. **78**:124–128.
50. **Teich, N., J. Wyke, T. Mak, A. Bernstein, and W. Hardy.** 1982. Pathogenesis of retrovirus-induced disease, p. 785–998. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
51. **Temin, H. M.** 1981. Structure, variation and synthesis of retrovirus long terminal repeat. Cell **27**:1–3.
52. **Temin, H. M.** 1982. Function of the retrovirus long terminal repeat. Cell **28**:3–5.
53. **Troxler, D. H., S. K. Ruscetti, and E. M. Scolnick.** 1980. The molecular biology of Friend virus. Biochim. Biophys. Acta. **605**:305–324.
54. **Troxler, D. H., and E. M. Scolnick.** 1978. Rapid leukemia induced by cloned Friend strain of replicating murine type-C virus. Association with induction of xenotropic-related RNA sequences contained in spleen focus-forming virus. Virology **85**:17–27.
55. **Tucker, P. W., J. L. Slightom, and F. R. Blattner.** 1981. Mouse IgA heavy chain gene sequence: implications for evolution of immunoglobulin hinge exons. Proc. Natl. Acad. Sci. U.S.A. **78**:7684–7688.
56. **Van Beveren, C., J. C. Goddard, A. Berns, and I. M. Verma.** 1980. Structure of Moloney murine leukemia viral DNA: nucleotide sequence of the 5' long terminal repeat and adjacent cellular sequences. Proc. Natl. Acad. Sci. U.S.A. **77**:3307–3311.
57. **Witte, O. N., and D. F. Wirth.** 1979. Structure of the murine leukemia virus envelope glycoprotein precursor. J. Virol. **29**:735–743.
58. **Wolff, L., E. Scolnick, and S. Ruscetti.** 1983. Envelope gene of the Friend spleen focus-forming virus: deletion and insertions in 3' gp70/p15E encoding region have resulted in unique features in the primary structure of its protein product. Proc. Natl. Acad. Sci. U.S.A. **80**:4718–4722.