JOURNAL OF VIROLOGY, Feb. 1990, p. 757–766
0022-538X/90/020757-10$02.00/0

Vol. 64, No. 2

Sequence Analysis of Amphotropic and 10A1 Murine Leukemia Viruses: Close Relationship to Mink Cell Focus-Inducing Viruses

D. OTT,[1] R. FRIEDRICH,[2] AND A. REIN[1]*

Laboratory of Molecular Virology and Carcinogenesis, BRI-Basic Research Program, NCI-Frederick Cancer Research Facility, P.O. Box B, Frederick, Maryland 21701,[1] and Institute of Medical Virology, University of Giessen, 6300 Giessen, Federal Republic of Germany[2]

Received 13 September 1989/Accepted 1 November 1989

Viral interference studies have demonstrated the existence of four distinct murine leukemia virus (MuLV) receptors on NIH 3T3 mouse cells. The four viral interference groups are ecotropic MuLV; mink cell focus inducing virus (MCF); amphotropic MuLV; and 10A1, a recombinant derivative of amphotropic MuLV that uses a unique receptor but also retains affinity for the amphotropic MuLV receptor. We report here that 10A1 infects rat and hamster cells, unlike its amphotropic parent. We isolated an infectious molecular clone of 10A1 and present here the sequences of the *env* genes and enhancer regions of amphotropic MuLV and 10A1. The deduced amino acid sequences of amphotropic MuLV and 10A1 gp70^{su} are remarkably similar to those of MCF and xenotropic MuLV (for which mouse cells lack receptors), with 64% amino acids identical in the four groups. We generated a consensus from these comparisons. Further, the differences are largely localized to a few discrete regions: (i) amphotropic MuLV has two short insertions relative to MCF, at residues 87 to 92 and 163 to 169, and (ii) amphotropic MuLV and MCF are totally different in a hypervariable region, which is >30% proline, at residues ~253 to 304. 10A1 closely resembles amphotropic MuLV in its N terminus but contains an MCF-type hypervariable region. These results suggest the possibility that receptor specificity is localized in these short variable regions and further that the unique receptor specificity of 10A1 is due to the novel combination of amphotropic MuLV and MCF sequences rather than to the presence of any novel sequences. The Env proteins of ecotropic MuLVs are far more distantly related to those of the other four groups than the latter are to each other. We also found that the enhancer regions of amphotropic MuLV and 10A1 are nearly identical, although 10A1 is far more leukemogenic than amphotropic MuLV.

The *env* gene of murine leukemia viruses (MuLVs) encodes an external glycoprotein, gp70^{su}, which interacts with a receptor molecule on the surface of a host cell in the initial stage of infection. Viral interference studies have shown that MuLVs can be divided into four groups, using four distinct receptors on mouse cells (20, 21) plus a fifth group, xenotropic MuLVs, for which mouse cells lack receptors (3).

One approach to understanding the gp70-receptor interaction in molecular terms is to compare gp70 sequences from different viral groups. Several examples from two of these groups, the ecotropic and mink cell focus-inducing viruses (MCFs), have been sequenced. The gp70s of these groups are quite different in their N-terminal two-thirds; the C-terminal regions, on the other hand, are nearly identical. Thus, receptor specificity is apparently determined by the sequences found in the N-terminal two-thirds of the molecule.

MCFs arise in mice infected with ecotropic MuLVs. The MCF *env* gene is generated by recombination between the ecotropic MuLV and *env* sequences of endogenous MuLVs present in the mouse genome; this recombinational event substitutes the endogenous N-terminal *env* region for the ecotropic *env* sequences. The endogenous MuLV genomes which participate in this recombination have been carefully characterized by Stoye and Coffin (28). One such genome, termed MX27, has been sequenced and appears to be a prototype of the endogenous sequences involved in the creation of MCFs. It seems likely that these endogenous sequences encode gp70s which carry the specificity for interaction with the MCF cell surface receptor, since the gp70s of some MCFs are composed almost entirely of endogenous sequences.

The other two classes of MuLV which can infect mouse cells are amphotropic MuLV and 10A1. The relationship between these two classes appears analogous to that between ecotropic MuLVs and MCFs, since 10A1 was isolated from a mouse infected with amphotropic MuLV (19). In addition, RNase T1 fingerprint analysis of 10A1 has shown that it contains an MCF-like (or xenotropic MuLV-like) substitution near the *pol-env* junction (15). Another parallel with the ecotropic MuLVs and MCFs is that 10A1 is strikingly more leukemogenic than its amphotropic parent, just as many MCFs are more leukemogenic than their ecotropic parents. Nevertheless, 10A1 has unique biological properties; i.e., it interacts with a unique receptor on mouse cells, which is not shared with any other known MuLV, and also appears to interact with the amphotropic receptor. This is the only known case of dual mouse cell receptor specificity among the MuLVs.

We now present the nucleotide sequences of the *env* genes and enhancer regions of amphotropic MuLV and 10A1. The *env* sequence of amphotropic MuLV is remarkably close to that of the MX27 class of endogenous MuLVs, except in three localized regions. In turn, 10A1 is very closely related to both amphotropic and MX27 sequences; inspection of these three divergent regions showed that 10A1 is a recombinant between amphotropic MuLV and MX27-related endogenous sequences. These findings appear to have important implications for the relationship between gp70 structure and receptor specificity.

* Corresponding author.

MATERIALS AND METHODS

Cell culture. NIH 3T3, Rat-2 (a gift of W. Vass, National Cancer Institute, Bethesda, Md.), and CHO cell lines were maintained in Dulbecco Modified Eagle medium supplemented with 10% calf serum, Dulbecco Modified Eagle medium supplemented with 10% fetal bovine serum, and alpha modified minimal essential medium supplemented with 10% fetal bovine serum, respectively. Viral infections were performed with 20 μg of DEAE-dextran per ml. Infectivity was assayed by plating cells in an infectious-center assay on mouse S^+ L^- indicator cells (1).

Hirt DNA isolation and analysis. Unintegrated 10A1 proviral DNA was prepared after cocultivation of a 10A1-producing CHO cell line with CHO cells in 2 μg of Polybrene per ml by using a ratio of 1:4 with a total cell density of $7 \times 10^4/cm^2$. After 2 days of cocultivation, DNA was isolated by the method of Hirt (11). 10A1 Hirt supernatant DNA was then digested with diagnostic restriction enzymes and analyzed by hybridization (25) to retroviral plasmid clone p4070A (5), which had been labeled with ^{32}P by nick translation (22). Hybridization conditions were 10% dextran sulfate–5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–2× Denhardt solution, 0.2% sodium dodecyl sulfate–50 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid; pH 6.8)–40% formamide–10 μg of poly(rA) per ml–50 μg of sheared salmon sperm DNA per ml–1×10^6 to 5×10^6 cpm of ^{32}P-labeled probe per ml. Hybridizations were performed for at least 18 h at 42°C. After hybridization, filters were washed in 0.1× SSC–0.1% sodium dodecyl sulfate at 65°C for 2 to 4 h with changes of wash solution every 30 min.

Cloning and subcloning. Hirt supernatant DNA was digested with *Sal*I to linearize circular 10A1 proviral DNA. The DNA was fractionated on an agarose gel, and 7.5- to 9.5-kilobase (kb) fragments were purified by electroelution with a UEA electroeluter (International Biotechnologies, Inc., New Haven, Conn.). EMBL3 *Sal*I-digested bacteriophage arms were prepared by double digestion of the phage with *Sal*I and *Bam*HI. Following ligation, packaging (Stratagene Cloning Systems, La Jolla, Calif.), and plating of a mixture of insert and vector, screening was accomplished by probing with ^{32}P-labeled p4070A by standard methods (2). DNAs from candidate phage clones were mapped with several restriction enzymes; clones with restriction maps similar to that of amphotropic MuLV (5) were tested for infectivity as follows. The DNAs were digested with *Sal*I to release the circularly permuted viral DNA. These DNAs were then transfected without carrier DNA into CHO cells. The cells were passaged for 2 weeks and then plated as infectious centers in the S^+ L^- focus assay. *env* and long terminal repeat (LTR) regions from both the infectious 10A1 clone and the p4070A clone (5; a gift of Allen Oliff) were subcloned for sequence analysis. 10A1 fragments were subcloned into either a Bluescript KS^+ or KS^- plasmid, while 4070A DNA fragments were subcloned into either M13 phage mp 12 or mp 13.

Sequencing. Single-stranded DNAs from either M13 cultures or helper virus-rescued Bluescript cultures were sequenced by using the dideoxynucleotide chain termination method of Sanger et al. (23) with either DNA *Pol*I Klenow fragment or the Sequenase system (U.S. Biochemicals, Cleveland, Ohio). In all cases, both DNA strands were sequenced. Sequences were assembled and analyzed by using the Stephens software package (27).

Sequence alignment. Nucleotide sequence alignments were made with the NUCALN program (29) by using the following parameters: K-tuple size, 3; window size, 20; gap penalty, 7. Sequence translations and alignments were achieved by using the GAP program (7) with a gap penalty of 5 and a length weight of 0.3 and allowing only identical matches. Other MuLV sequences were obtained from Genbank.

TABLE 1. Host ranges of amphotropic and 10A1 MuLVs[a]

MuLV	No. of infectious centers/2 × 10^4 cells		
	NIH 3T3	Rat-2	CHO
1504A	9,800	1	0
10A1	14,500	900	6,000

[a] Flasks were seeded with 3×10^5 NIH 3T3, Rat-2, or CHO cells. On the following day, they were infected with amphotropic MuLV (isolate 1504A, a kind gift of Janet Hartley, National Institute of Allergy and Infectious Diseases, Bethesda, Md.) or 10A1. Two days later, the cells were trypsinized and plated as infectious centers on S^+ L^- mouse indicator cells (1). Results similar to those shown for 1504A were observed with 4070A MuLV (data not shown).

RESULTS

Host range of 10A1 MuLV. We previously reported (21) that 10A1, a recombinant derived from amphotropic MuLV, uses a receptor different from that of amphotropic MuLV on NIH 3T3 mouse cells. It seemed possible that the 10A1 *env* gene product could also interact with receptors present on cells lacking the amphotropic MuLV receptor. We therefore tested the ability of 10A1 to productively infect CHO cells and Rat-2 cells, which are both almost totally nonpermissive for amphotropic MuLV. 10A1 infected both of these cell lines with nearly the same efficiency as it infected NIH 3T3 cells (Table 1). This expanded host range of 10A1 is fully consistent with its distinct receptor specificity on mouse cells.

Molecular cloning and partial sequence analysis of 10A1 MuLV. Eight molecular clones of 10A1 proviral DNA were isolated from Hirt supernatant DNA and tested as described in Materials and Methods. One of these clones was found to be infectious. The MuLV produced after transfection of this clone appeared to be a faithful replica of 10A1, since it had an expanded host range and the expected interference properties on NIH 3T3 cells (data not shown). It also reproduced the biological properties of 10A1 by inducing lymphomas within 3 months of intraperitoneal inoculation into newborn NIH Swiss mice (data not shown).

Comparison of amphotropic and 10A1 MuLV genomes. As a first step in characterizing the genome of 10A1 MuLV, we generated a restriction map of the 10A1 clone. This map is shown in Fig. 1 along with the map of the 4070A amphotropic MuLV clone (5). It is evident that the two genomes are quite similar, although they show a cluster of restriction site differences at 5 to 6 kb pairs from the 5′ end. The presence of a novel *Bgl*II site at 5.25 kb, together with several restriction site differences immediately downstream, suggested that a recombinational event had occurred 5′ of this *Bgl*II site. This is consistent with the RNase T1 fingerprint analysis of Lai et al. (15). The 3′ limit of this recombination appeared to lie between the 10A1-specific *Bam*HI site at 6.0 kb and the common *Xho*I site at 6.9 kb. The regions which we sequenced are also indicated in Fig. 1.

Sequence analysis of amphotropic and 10A1 *env* genes. To examine the structure of amphotropic *env* and to compare the 10A1 genome with that of amphotropic MuLV, we sequenced the two genomes from just 5′ of the beginning of

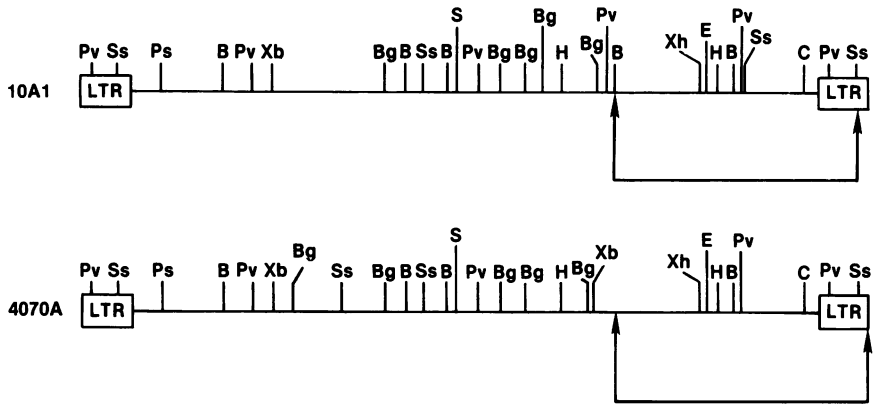

FIG. 1. Restriction enzyme maps of 10A1 and 4070A retroviruses. The two viruses are drawn as complete linear genomes for convenience, although the clones themselves are circularly permuted, single LTR genomes: 10A1 cloned at the *Sal*I site and 4070A cloned at the *Eco*RI site. The regions sequenced are indicated by arrowed brackets. Abbreviations: B, *Bam*HI; Bg, *Bgl*II; C, *Cla*I; E, *Eco*RI; H, *Hin*dIII; Ps, *Pst*I; Pv, *Pvu*II; S, *Sal*I; Ss, *Sst*I; Xb, *Xba*I; Xh, *Xho*I.

the *env* gene to within the LTR (Fig. 1). The nucleotide sequences of the *env* regions of the 4070A and 10A1 clones are shown in Fig. 2.

Comparison of the two sequences showed them to be very similar, except for two stretches of nucleotide sequence differences. Remarkably, the 10A1 sequences that diverged from those of 4070A corresponded almost exactly to the endogenous polytropic sequence MX27 (Fig. 2) (28). Thus, 10A1 apparently contains a stretch of polytropic sequences beginning to the 5′ side of *env*, within the presumed *pol* region, and extending to nucleotides 67 to 87 in the *env* leader region (because of the close homology among the three sequences in this region, the exact crossover point could not be determined). A second pair of crossover events, one at nucleotides 724 to 731 and the other at nucleotide 959, contributed a short polytropic MuLV nucleotide sequence within the gp70-coding region. (It should be noted that there are several distinct isolates of amphotropic MuLV and that the source of 10A1 was a mouse injected with amphotropic MuLV isolate 1504A rather than 4070A [19]. However, the sequence relationship between 4070A and 10A1 is so close, outside of these divergent regions [Fig. 2], that we used 4070A as a prototype amphotropic MuLV for the sequence comparisons described here.)

Deduced amino acid sequences of amphotropic MuLV and 10A1 Env proteins: comparison with other Env proteins. Figure 3A presents the amino acid sequences of the 4070A and 10A1 *env* genes, as deduced from the nucleotide sequences shown in Fig. 2. For comparison, we also show the deduced sequences of MX27 and a xenotropic MuLV derived from an NZB mouse (17). Under the four sequences, we present a polytropic MuLV-related MuLV (PRM) consensus sequence, showing all of the identical amino acids in the four sequences and also indicating positions at which only conservative changes are found. The remaining class of MuLVs is the ecotropic group. We selected two ecotropic MuLVs, one exogenous (Moloney MuLV) and one endogenous (Akv), for comparison with the PRM sequences. Figure 3B presents an alignment of the Moloney and Akv *env* sequences with the PRM consensus. Finally, we compiled an all-MuLV consensus sequence, shown in the last row of Fig. 3B. The implications of these comparisons will be considered in the Discussion.

Partial sequence analysis of amphotropic MuLV and 10A1 LTRs. We also sequenced the 3′-noncoding regions of amphotropic MuLV and 10A1. These sequences are compared in Fig. 4. The 4070A sequence contains two inverted repeats, as found in all MuLV LTRs, and a CCAAT box, a TATAA box, and a poly(A) additional signal. We compared the enhancer regions of 4070A and 10A1 with that of Moloney MuLV, which has been analyzed in detail by Speck and Baltimore (26). Both 4070A and 10A1 contain a sequence, AGAACAGAA, which resembles the canonical glucocorticoid-responsive element sequence AGAACAGATG; a sequence, AAACAG, like the LVa factor-binding sequence GAACAG; sequences like the consensus binding sites for NF-1 (TGA[N_{6-7}]GCCA) and LVb; the core enhancer element TGTGGTAA; and a second glucocorticoid-responsive element sequence (Fig. 4). Both viruses lack the LVc sequence. They differ in that 4070A may have a second NF-1 site (CGGN$_6$GCCA), while the 10A1 sequence diverges slightly from 4070A in this region. There is no duplication of the enhancer region in either 4070A or 10A1. A second difference is that 10A1 contains the negative regulatory consensus sequence, CGCCATTTT, described by Flanagan et al. (9), while in amphotropic MuLV, the sequence is disrupted by a 4-nucleotide duplication, i.e., CGCCATTTA TTTT. Aside from this difference, the two sequences are extremely similar (95% identity).

DISCUSSION

We report here the nucleotide sequences of the 3′ ends of the amphotropic MuLV and 10A1 genomes. Translation of these sequences allowed us, for the first time, to compare the Env sequences of all of the known MuLV receptor specificity groups.

The amino acid sequences from these groups are aligned in Fig. 3. One striking conclusion which emerges from this alignment is the remarkable degree of conservation in these sequences. In particular, four of these groups are extremely similar: amphotropic MuLV, 10A1, endogenous polytropic MuLV, and xenotropic MuLV (Fig. 3A). We will refer to these four groups as the PRM family for simplicity here. We found that the mature gp70s of these four MuLV groups are identical at 296 positions or approximately 64% of the total. Isolates from the fifth group, the ecotropic MuLVs, showed far less homology with PRMs (only 40% identity with the PRM consensus sequence) than the latter do with each other (Fig. 3B); there is also considerably more diversity within

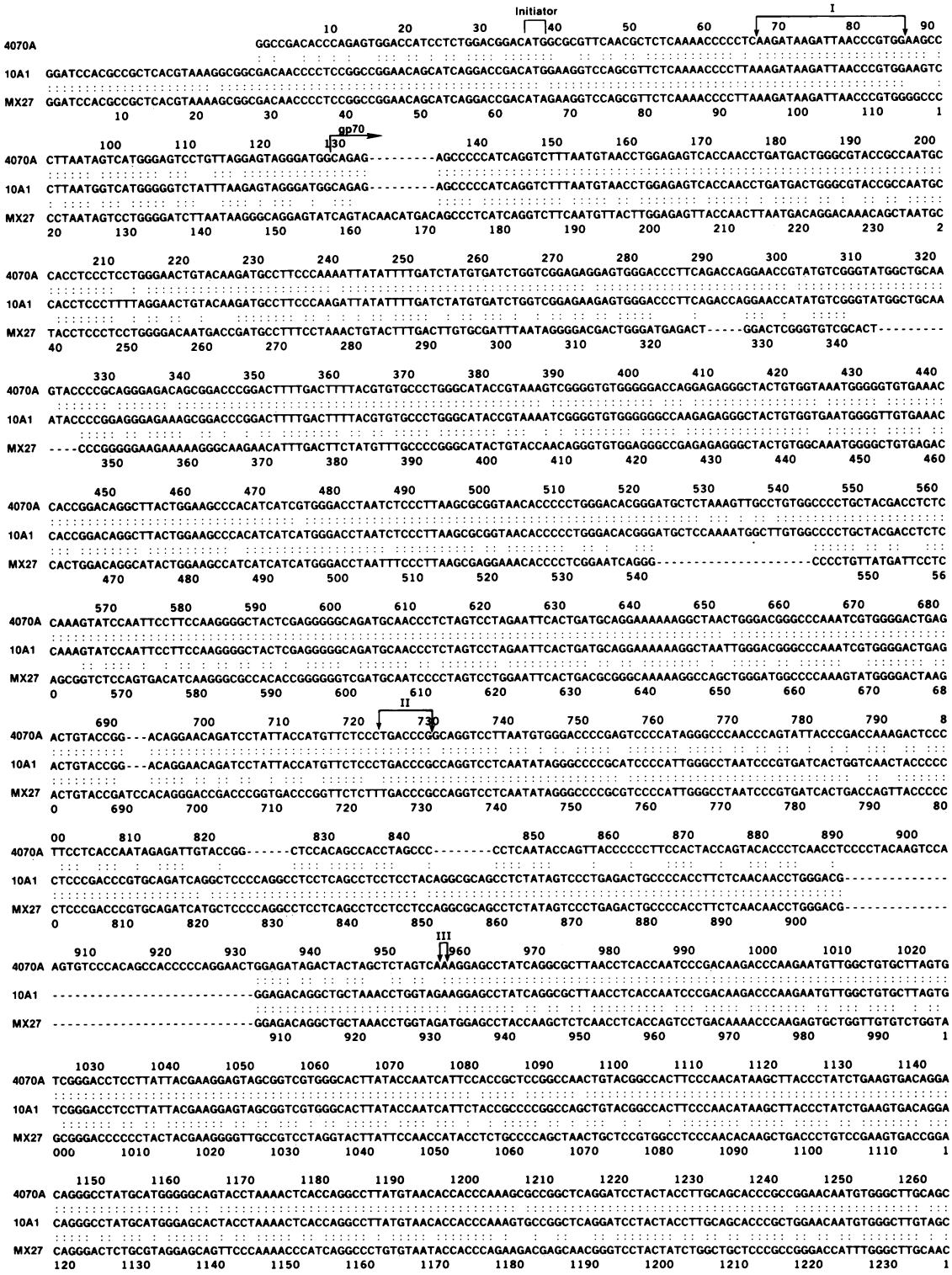

FIG. 2. Nucleotide alignments of 4070A, 10A1, and MX27 *env* genes. The start of the *env* precursor and the inferred boundaries of the mature processed proteins are indicated above the aligned sequences. Additionally, the three areas of probable recombination between amphotropic and polytropic MuLV sequences in 10A1 are denoted by roman numerals. 10A1 is most similar to MX27 to the 5′ side of I, to 4070A between I and II, to MX27 between II and III, and to 4070A on the 3′ side of III. The MX27 sequence is from Stoye and Coffin (28).

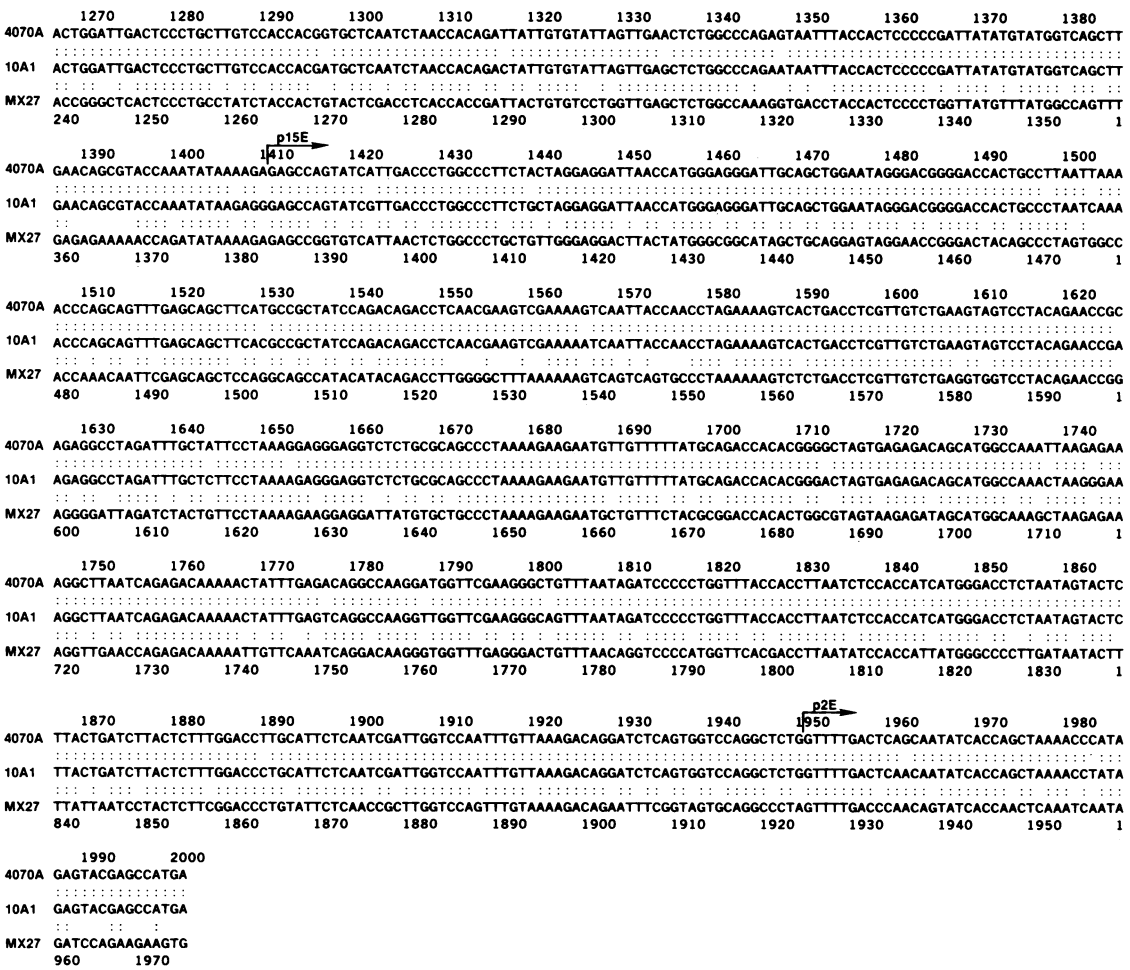

FIG. 2—*Continued*

the ecotropic MuLV group (13) than within any of the PRM groups.

The close alignment of the PRM Env sequences allowed us to derive a consensus PRM Env sequence (Fig. 3A). The degree of identity found among these MuLV Env sequences suggests that all PRM Env proteins share a common overall framework, while the biological differences in receptor specificity and host range among the viruses must result from the minor differences in the Env proteins.

The close resemblance among the different gp70s also raises the possibility that the different receptors are similar to each other, just as a series of similar keys would require a series of similar locks. In this connection, it is interesting to note the findings of Chesebro and Wehrly (6). These investigators made the remarkable observation that some MuLV isolates fall into different interference classes in tests on *Mus dunni* cells than they do on NIH 3T3 *M. musculus* cells. The fact that differences were observed between *M. musculus* and *M. dunni* cells indicates that the gp70-binding sites of some *M. dunni* receptors are somewhat different from those of their *M. musculus* homologs. The fact that some MuLV isolates which use different receptors in one host nevertheless share a common receptor in the other implies that the receptor-binding regions of their gp70s are similar to each other. In turn, since each gp70 must fit its receptor, this conclusion suggests that the gp70-binding regions of some distinct receptors in a given host also resemble each other.

Figure 3A also shows that the differences among the *env* sequences within the PRM group are predominantly localized in a few short regions, as if receptor specificity were carried in these regions of the molecule. In particular, the amphotropic MuLV and 10A1 Env sequences contain two short stretches, at residues 87 to 92 and 163 to 169 (numbering from the initiator methionine of the consensus sequence), which are not present in MX27. Xenotropic MuLV contains a partial copy of the first of these regions but lacks the second. The other region which shows striking diversity among different viruses is the "hypervariable region" (14). This region comprises residues 253 to 304 of the PRM consensus. All MuLV Env proteins have a sequence stretch in this region of the molecule which is >30% proline, and this region has been noted to show significant variation, even within the ecotropic group of MuLVs (13). The sequences immediately flanking this hypervariable region are completely conserved in all MuLVs, and indeed there is very little variation seen to the C-terminal side of this region or in the p15E portion of the Env polyprotein (Fig. 3B). These observations are summarized in Table 2, which indicates the degree of amino acid identity found at different regions of the Env protein. Since the p15Es of all MuLVs are very similar to each other, it seems likely that these transmembrane

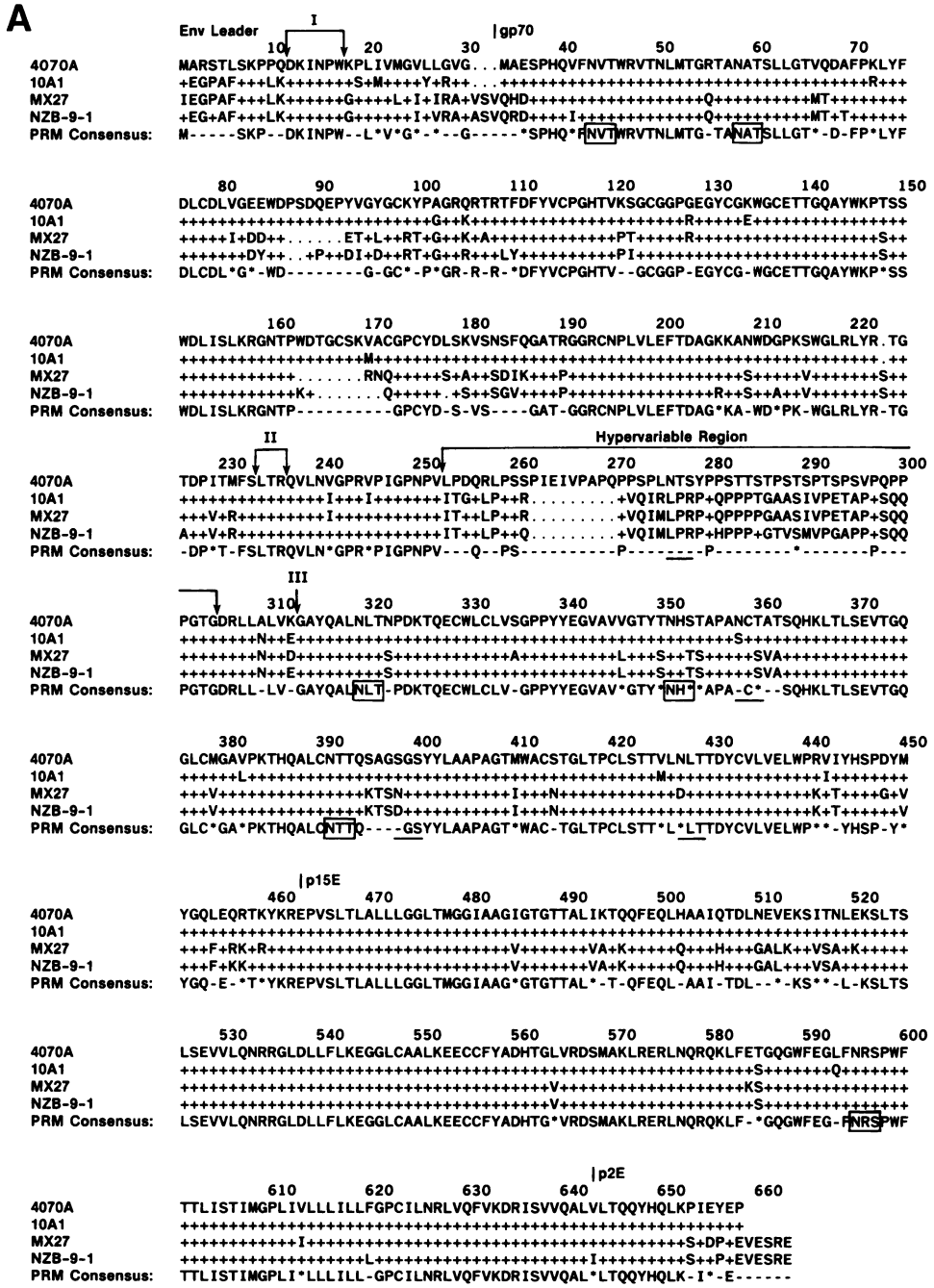

FIG. 3. Amino acid alignment of the PRM groups and all MuLVs. (A) Amino acid sequence of 4070A Env. Periods indicate gaps introduced for maximal alignment. The other sequences, 10A1, MX27, and NZB-9-1 (17), are presented below that of 4070A, and identical amino acids are denoted by +. A PRM consensus is presented below the alignments; identities are shown by letters, and mismatches are indicated by dashes. Conservative amino acid changes (A,G; S,T; F,W,Y; I,L,M,V; K,H,R; and D,E,N,Q) are indicated by asterisks. Recombination points and boundaries of Env proteins are shown above the alignments, as in Fig. 2. Inferred glycosylation sites are boxed on the consensus, and missing or novel sites are underlined. (B) Amino acid sequences of two ecotropic MuLV Env proteins, that of Moloney MuLV (24) and that of Akv (10). Residues in Akv identical to those of Moloney MuLV are indicated by a +. Also shown is the PRM consensus from panel A. Residues which are identical in the PRM consensus, Moloney, and Akv sequences are shown as an MuLV consensus in the bottom line.

proteins interact with the conserved, C-terminal portion of gp70 in the mature Env complex.

Since 10A1 uses a unique receptor on NIH 3T3 cells, we previously suggested (21) that it had acquired unique *env* sequences from an endogenous *env* gene. The distinctive nature of the 10A1 *env* gene is further highlighted by the ability of this virus to infect CHO cells (Table 1). However, our sequence analysis shows that the 10A1 *env* gene contains almost no novel coding sequences. Rather, it appears to be a unique combination of *env* sequences which are also present

B

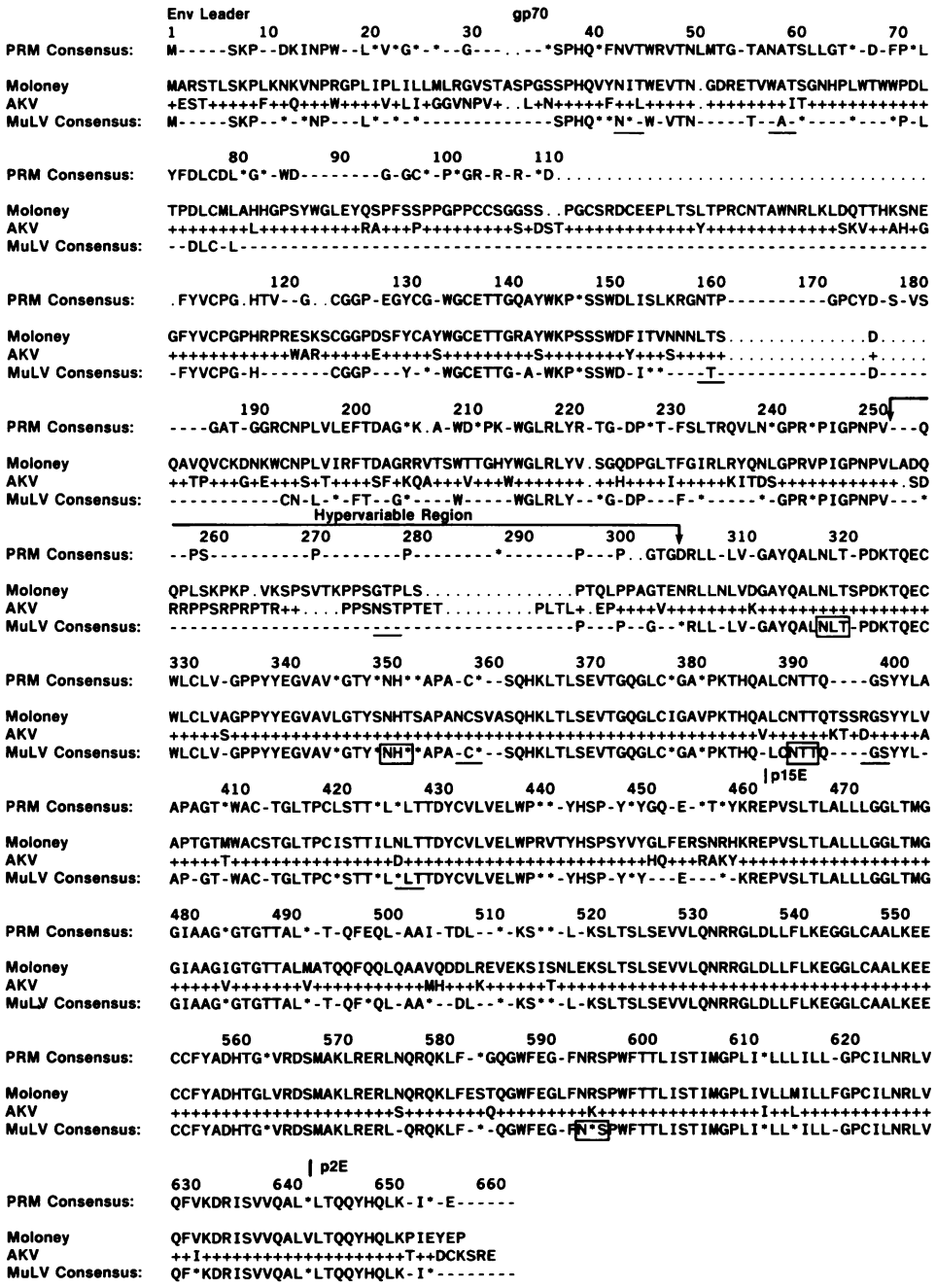

FIG. 3—*Continued*

separately in other MuLVs. Thus, 10A1 contains the two N-terminal insertions noted above as characteristic of amphotropic MuLV, while its hypervariable region is nearly identical to that of MX27 and the polytropic MuLV-derived (28) MCFs. The results suggest that receptor specificity is determined combinatorially, since substitution of an MX27-like hypervariable region in an otherwise amphotropic gp70 has created an Env molecule with a third receptor specificity. The recombinant nature of the 10A1 gp70 is indicated schematically in Fig. 5.

Our previous biological studies indicated that in addition to its affinity for a unique receptor, the 10A1 Env protein also interacts with the amphotropic MuLV receptor on NIH 3T3 cells (21). Our results with 10A1 show that the presence of the amphotropic MuLV hypervariable region is not necessary for amphotropic MuLV receptor specificity. One possibility is that one or both of the N-terminal insertions determine amphotropic MuLV receptor specificity independently of the hypervariable region, while the second, unique receptor specificity of 10A1 reflects the "combinatorial" interaction between these regions of gp70. It seems possible that the dual receptor specificity contributes to the leukemogenicity of 10A1. The ability of an Env protein to interact with more than one receptor may also be important in

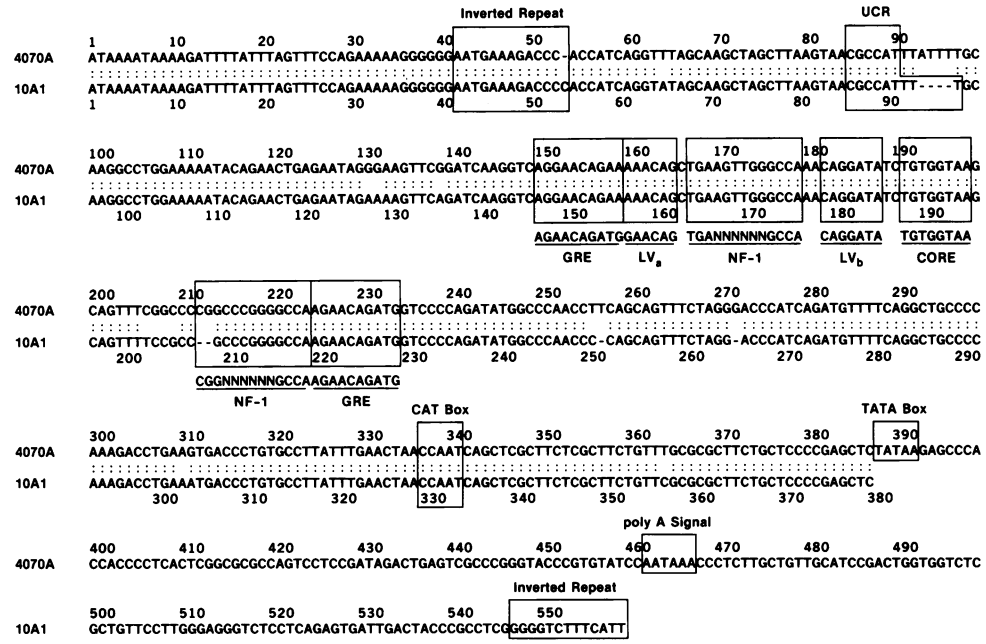

FIG. 4. 4070A LTR sequence and alignment with 10A1 transcriptional signals. The sequences start immediately after the *env* termination codon. Boxes indicate the structural elements of the 4070A LTR and transcriptional regulatory sequences (9, 26).

understanding the pathogenicity of human immunodeficiency virus, since this virus is known to infect CD4$^+$ cells by interaction with CD4 but is also found in a number of CD4$^-$ cell types.

It is also possible that receptor specificity is not entirely determined by the variable-sequence blocks in gp70, since amphotropic MuLV and 10A1 also differ at 15 individual positions within the more highly conserved regions of the molecule. Experiments are under way to establish the role of the variable regions in determining receptor specificity.

Several of our conclusions on receptor specificity in MuLVs are similar to those reached in previous studies on avian retroviruses. Thus, gp85^{su} sequences of avian subgroups A, B, C, and E were aligned by Bova et al. (4), who noted that the differences among these Env molecules with different receptor specificities were confined to discrete regions of the molecule. Dorner and Coffin (8) analyzed a series of recombinants between subgroups B, C, and E viruses and showed that receptor specificity was determined principally by two sequence blocks. The relative positions of these blocks correspond rather closely to those we have described in the murine system; they are found at 36 and 59% of the distance from the N terminus to the C terminus of gp85, while the two short insertions in amphotropic MuLV are at 12 and 30% and the hypervariable region covers 51 to 63%. Avian viruses also show somewhat different interference patterns on chicken and turkey cells. Finally, Dorner and Coffin showed directly that receptor specificity is not always determined in a single region of gp85 but is sometimes a result of the combination of sequences present in a chimeric gp85 molecule. This finding is similar to what we suggest here on the basis of our comparison of the 10A1 sequence with those of 4070A and MX27, although in avian retroviruses, as far as is known, a novel sequence combination does not generate a completely novel receptor specificity or extended host range. The analogies noted here strongly suggest that similar structure-function relationships exist in murine and avian surface glycoproteins and in their interactions with their respective receptors.

Comparison of the 10A1 nucleotide sequence with those of amphotropic MuLV and MX27 also allowed us to determine the sites of recombination involved in the generation of 10A1. As indicated in Results, 10A1 is nearly identical to MX27 over two limited regions: (i) from a site in the *pol* gene to nucleotides 67 to 87 in the *env* leader region and (ii) 724-731 to 959, which encodes the hypervariable region plus 24 to 27 additional amino acids. With the exception of these two regions, the 10A1 *env* gene shows over 96% nucleotide

TABLE 2. Amino acid conservation in Env

Consensus	% Identical amino acids at the indicated amino acid position[a] in:							
	NH_2 one-third of gp70				Middle one-third of gp70		C-terminal one-third of gp70	p15E + p2E (463–663)
	32–86	87–94	95–162	163–172	173–252	253–304	305–462	
PRM	76	0	79	0	78	21	80	85
All MuLV[b]	32	0	26	0	35	6	75	81

[a] Counting from the initiator methionine of the consensus.
[b] PRM plus Moloney MuLV and Akv.

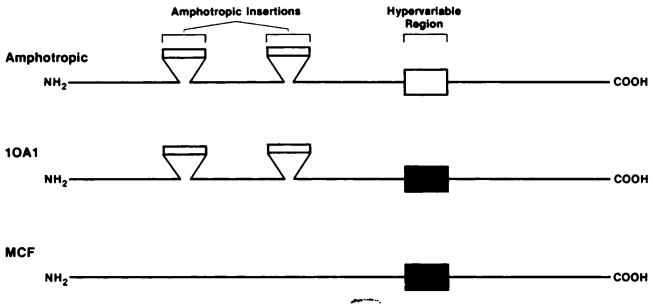

FIG. 5. Schematic of amphotropic MuLV, 10A1, and MCF mature gp70 amino acid sequence comparisons. Black lines indicate regions of all three viruses that are >75% identical. Amphotropic sequence insertions relative to MCF are indicated. MCF hypervariable proline-rich regions are represented by black boxes, and the amphotropic region is shown as a white box.

sequence identity with amphotropic MuLV. Thus, 10A1 appears to have arisen by two substitutions of MX27-related sequences into the parental amphotropic MuLV *env* gene. (Additional recombinations may, of course, have occurred in the *gag* or *pol* region.)

The limited extent of the substitutions that generated 10A1 is in sharp contrast to the recombinations which give rise to MCFs. As far as is known, all viable recombinants between ecotropic MuLVs and endogenous *env* genes use the MCF receptor. In all MCFs which have been analyzed, at least the entire N-terminal approximately two-thirds of the ecotropic gp70-coding region, from the N terminus through the hypervariable region, has been replaced with MX27 (or MX33 [28])-related sequences. The consistency of these findings strongly suggests that if other types of recombinants occur, such as crossovers within the N-terminal two-thirds of gp70, they are not selected for outgrowth in viremic mice. It would be interesting to test the viability and receptor specificity of such chimeras generated in vitro.

O'Neill et al. (18) have subcloned a 133-base-pair fragment of the amphotropic MuLV genome and have demonstrated its utility as an amphotropic MuLV-specific probe. Using this subclone, they showed that this sequence is not present in the DNA of uninfected animals; they concluded that amphotropic MuLV is only transmitted horizontally. This subclone represents nucleotides 191 to 324 of the amphotropic MuLV sequence (Fig. 2), which encode residues 53 to 96 of the Env protein (Fig. 3). It therefore spans the first of the two N-terminal insertions in the amphotropic MuLV protein sequence.

We also present here U3 sequences of amphotropic MuLV and 10A1. The two sequences are extremely similar to each other (94.9% identity) (Fig. 4). Many studies of MuLVs have pointed to the U3 regulatory sequences as major determinants of pathogenicity. On the other hand, the MCF *env* gene has also been shown to contribute to the pathogenicity of Akv- and F-MuLV-derived MCFs (12, 16). It is striking that the two MuLVs analyzed in this study have similar LTRs but differ enormously in leukemogenicity. In light of this difference, it seems possible that the 10A1 *env* gene plays a key role in the pathogenicity of 10A1.

The mechanism by which the MCF *env* gene contributes to the pathogenicity of a recombinant MuLV remains a major unsolved problem in understanding leukemia induction by nonacutely transforming retroviruses. While the sequences presented here suggest the possibility that the 10A1 *env* gene is important in the pathogenicity of this virus, it will be of great importance to test this hypothesis empirically. Experiments testing the leukemogenic potential of chimeric MuLVs are under way.

ACKNOWLEDGMENTS

We thank Janet P. Hanser and Gabriele Maennle for outstanding technical assistance; Suraiya Rasheed for a preliminary restriction map of 10A1; Nancy Hopkins, Nancy Rice, and Jonathan Stoye for helpful discussions and comments on the manuscript; Bob Stephens for indispensable help with computer alignments; and Carol Shawver for preparation of the manuscript.

This research was sponsored in part by the National Cancer Institute under Public Health Service contract NO1-CO-74101 with BRI, Basic Research Program.

LITERATURE CITED

1. **Bassin, R. H., N. Tuttle, and P. J. Fischinger.** 1971. Rapid cell culture assay for murine leukaemia virus. Nature (London) **229**:564–566.
2. **Benton, W. D., and R. W. Davis.** 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. Science **196**:180–182.
3. **Besmer, P., and D. Baltimore.** 1977. Mechanism of restriction of ecotropic and xenotropic murine leukemia viruses and formation of pseudotypes between the two viruses. J. Virol. **21**: 965–973.
4. **Bova, C. A., J. P. Manfredi, and R. Swanstrom.** 1986. *env* genes of avian retroviruses: nucleotide sequence and molecular recombinants define host range determinants. Virology **152**:343–354.
5. **Chattopadhyay, S. K., A. I. Oliff, D. L. Linemeyer, M. R. Lander, and D. R. Lowy.** 1981. Genomes of murine leukemia viruses isolated from wild mice. J. Virol. **39**:777–791.
6. **Chesebro, B., and K. Wehrly.** 1985. Different murine cell lines manifest unique patterns of interference to superinfection by murine leukemia viruses. Virology **141**:119–129.
7. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
8. **Dorner, A. J., and J. M. Coffin.** 1986. Determinants for receptor interaction and cell killing on the avian retrovirus glycoprotein gp85. Cell **45**:365–374.
9. **Flanagan, J. R., A. M. Krieg, E. E. Max, and A. S. Khan.** 1989. Negative control region at the 5′ end of murine leukemia virus long terminal repeats. Mol. Cell. Biol. **9**:739–746.
10. **Herr, W.** 1984. Nucleotide sequence of AKV murine leukemia virus. J. Virol. **49**:471–478.
11. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26**:365–369.
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. **Koch, W., G. Hunsmann, and R. Friedrich.** 1983. Nucleotide sequence of the envelope gene of Friend murine leukemia virus. J. Virol. **45**:1–9.
14. **Koch, W., W. Zimmerman, A. Oliff, and R. Friedrich.** 1984. Molecular analysis of the envelope gene and long terminal repeat of Friend mink cell focus-inducing virus: implications for the functions of these sequences. J. Virol. **49**:828–840.
15. **Lai, M. M. C., S. Rasheed, C. S. Shimizu, and M. B. Gardner.** 1982. Genomic characterization of a highly oncogenic env gene recombinant between amphotropic retrovirus of wild mouse and endogenous xenotropic virus of NIH Swiss mouse. Virology **117**:262–266.
16. **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 erythroleukemia. J. Virol. **46**:718–725.
17. **O'Neill, R. R., C. E. Buckler, T. S. Theodore, M. A. Martin, and R. Repaske.** 1985. Envelope and long terminal repeat sequences

of a cloned infectious NZB xenotropic murine leukemia virus. J. Virol. **53**:100–106.
18. **O'Neill, R. R., J. W. Hartley, R. Repaske, and C. A. Kozak.** 1987. Amphotropic proviral envelope sequences are absent from the Mus germ line. J. Virol. **61**:2225–2231.
19. **Rasheed, S., B. K. Pal, and M. Gardner.** 1982. Characterization of a highly oncogenic murine leukemia virus from wild mice. Int. J. Cancer **29**:345–350.
20. **Rein, A.** 1982. Interference grouping of murine leukemia viruses: a distinct receptor for the MCF-recombinant viruses on mouse cells. Virology **120**:251–257.
21. **Rein, A., and A. M. Schultz.** 1984. Different recombinant murine leukemia viruses use different surface receptors. Virology **136**: 144–152.
22. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling of deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
23. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
24. **Shinnick, T. M., R. A. Lerner, and J. G. Sutcliffe.** 1981. Nucleotide sequence of Moloney murine leukemia virus. Nature (London) **293**:543–548.
25. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
26. **Speck, N. A., and D. Baltimore.** 1987. Six distinct nuclear factors interact with the 75-base-pair repeat of the Moloney murine leukemia virus enhancer. Mol. Cell. Biol. **7**:1101–1110.
27. **Stephens, R. M.** 1985. A sequencers' sequence analysis package for the IBM PC. Gene Anal. Tech. **2**:67–75.
28. **Stoye, J. P., and J. M. Coffin.** 1987. The four classes of endogenous murine leukemia virus: structural relationships and potential for recombination. J. Virol. **61**:2659–2669.
29. **Wilbur, W. J., and D. J. Lipman.** 1983. Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA **80**:726–730.