JOURNAL OF VIROLOGY, May 1994, p. 3163–3174
0022-538X/94/$04.00+0

Vol. 68, No. 5

Naturally Occurring Genotypes of the Human Immunodeficiency Virus Type 1 Long Terminal Repeat Display a Wide Range of Basal and Tat-Induced Transcriptional Activities

NELSON L. MICHAEL,[1]* LISA D'ARCY,[2] PHILIP K. EHRENBERG,[2] AND ROBERT R. REDFIELD[1]

Division of Retrovirology, Walter Reed Army Institute of Research,[1] and Henry M. Jackson Foundation,[2] Rockville, Maryland

Received 29 November 1993/Accepted 16 February 1994

The primary body of information on the structure of human immunodeficiency virus type 1 (HIV-1) long terminal repeat (LTR)/*gag* leader genotypes has been determined from the analysis of cocultivated isolates. Functional studies of this regulatory portion of the provirus have been derived from the study of in vitro-generated mutations of laboratory-adapted molecular clones of HIV-1. We have performed a longitudinal analysis of molecular clones from the LTR/*gag* leader region amplified directly from the peripheral blood of four patients over three years. We have found a remarkable number of point mutations and length polymorphisms in *cis*- and *trans*-acting regulatory elements within this cohort. Most of the length polymorphisms were associated with duplications of Sp1 and TCF-1α sequences. These mutations were associated with a wide range of transcriptional activities for these genotypes in a reporter gene assay. Mutations in conserved Sp1 sequences correlated with a diminished capacity of such genotypes to bind purified Sp1 protein. Although no generalized trend in transcriptional activity was seen, a single patient accumulated mutations in NF-κB, Sp1, and TAR elements over this period. The analysis of naturally occurring mutations of LTR genotypes provides a means to study the molecular genetic consequences of virus-host interactions and to assess the functional impact of HIV therapeutics.

Transcripts with positive-strand polarity are initiated from the 5′ long terminal repeat (LTR) of human immunodeficiency virus type 1 (HIV-1) under the complex control of viral regulatory sequences (2, 7, 11, 20–22, 37, 49). The LTR can be divided into modulatory, core promoter, and transactivating regions (Fig. 1). The modulatory regions include elements with limited homology to AP-1 enhancer sequences at positions −347 to −329 relative to the cap site (14), two NF-AT sites at positions −292 to −255 (51), a USF site at positions −159 to −173 (15, 17, 30), a TCF-1α site at positions −139 to −124 (53), and two NF-κB enhancer elements at positions −104 to −81 (40). The core positive-strand promoter is composed of three Sp1 binding sites located at positions −78 through −47 and a TATA box at positions −28 to −24 (38, 46). A 59-bp region which confers transcriptional transactivating potential to the core promoter elements by the viral Tat protein is located within the R region of the LTR at positions +1 to +59 (3, 52). This transactivating region (TAR) folds into alternate RNA stem-loop structures that are found in the 5′ termini of all viral transcripts. The TAR element mediates a substantial increase in transcriptional initiation and elongation through interactions with the viral Tat protein and other cellular factors (11, 25, 27, 29, 31, 50). Other regulatory sequences that govern HIV-1 replication lie immediately downstream of the LTR and include the primer binding site, major 5′ splice donor, and genomic RNA packaging signal (ψ) (Fig. 1).

Most work on the HIV-1 LTR has come from the dissection of laboratory-adapted molecular clones (16). In contrast, studies on natural LTR genotypes have received relatively little attention in the literature. Previous genetic analyses of naturally occurring LTRs have included an intensive study of a single patient by Delassus and coworkers (8, 9). They reported numerous point mutations in the base of the TAR stem-loop and additional mutations 5′ to NF-κB site II. These mutations had no effect on basal-level transcription but a wide range of transactivation potentials when tested by transient transfection of reporter constructions into SW480 cells (8). Longitudinal study of the TAR/*tat* genotypes from this patient over 4 years (inclusive of progression from asymptomatic to symptomatic disease) showed no temporal enrichment for transcriptionally more active genotypes (9). Differential growth characteristics from multiple genotypes derived from single parental isolates by other workers were found to be associated with differential basal transcriptional activities of their cognate LTRs (10, 18). Studies by Pomerantz and coworkers using LTRs amplified by PCR directly from brain samples from four patients showed no mutations in NF-κB, Sp1, and TATA elements, but the study was limited to only two clones per patient (42). Length polymorphisms were found in the LTR region encompassed by the TCF-1α and Sp1 sites in 5 of 17 noncultivated patient samples analyzed by Koken and colleagues (26). One of these variants, containing a fourth Sp1 site, outgrew wild-type viruses containing three Sp1 sites when the two viruses were cocultivated.

Previous studies have shown an increasing viral nucleic acid burden associated with progression of HIV disease in the peripheral blood of infected individuals (4, 5, 35, 41, 47, 48). Such quantitative nucleic acid-based measurements of viral burden are now being applied to the assessment of HIV therapeutics in the attempt to define useful surrogate markers of therapeutic efficacy. We have undertaken a longitudinal study of LTR/*gag* leader region genotypes recovered directly

* Corresponding author. Mailing address: 13 Taft Ct., Suite 200, Rockville, MD 20850. Phone: (301) 762-0089. Fax: (301) 762-4177. Electronic mail address: mich9009@hjfhiv.hjf.org.

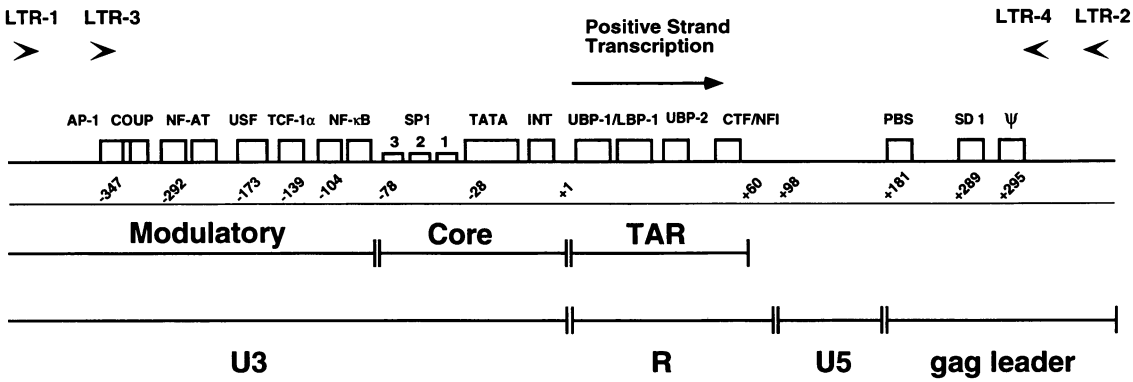


FIG. 1. Schematic drawing of the LTR/*gag* leader region of HIV-1. The locations of the amplifying primers used in the nested PCR strategy to obtain this amplicon are shown at the top of the figure, with arrowheads indicating polarity. The positions of *cis*-acting regulatory sequences are given on the heavy line, and genomic positions relative to the initiation of positive-strand transcription (arrow) are given on the thinner line below. The functional and structural regions of the HIV-1 LTR are given at the bottom of the figure. The schematic is not drawn to scale. PBS, primer binding site; SD 1, splice donor 1 (major 5′ splice donor); ψ, packaging signal.

from the peripheral blood mononuclear cells (PBMC) from four early-stage asymptomatic patients over 3 years to better define the role of viral transcriptional control regions in the pathogenesis of HIV disease. We have discovered a considerable number of length polymorphisms and point mutations in control sequences, resulting in a wide range of transcriptional activities for these genotypes.

MATERIALS AND METHODS

Description of patients. All samples were obtained from the first four early-stage asymptomatic patients enrolled in a phase I recombinant gp160 (VaxSyn; MicroGeneSys, Inc.) vaccine therapy trial (45). None of these patients were undergoing concurrent therapy with antiretroviral drugs.

Amplification, molecular cloning, and nucleotide sequencing of LTR DNA from patient PBMC. DNA was prepared from cryopreserved PBMC essentially as described previously (34). Cell equivalents (5×10^5) of the lysate in 25-μl aliquots were then assembled in a 50-μl reaction mixture containing 60 pmol of each of the outer primers LTR-1 and LTR-2 and 0.2 mM of each deoxynucleoside triphosphate (dNTP) in the presence of 55 mM Tris-HCl (pH 8.3), 50 mM KCl, and 2.75 mM $MgCl_2$. The reaction mixture was held for 10 min at 94°C prior to the addition of 5 U of *Taq* DNA polymerase (Perkin-Elmer) and subsequent thermal cycling (DNA thermal cycler; Perkin-Elmer). The first-round conditions were 94°C (1 min), 55°C (1 min), and 72°C (2.5 min) for 2 cycles followed by 94°C (45 s), 55°C (30 s), and 72°C (2.5 min) for 25 cycles followed by a 5-min incubation at 72°C. Ten microliters of the first-round PCR mixture was used as a template for the second round in a 100-μl reaction volume containing 60 pmol each of the inner primers LTR-3 and LTR-4 and 0.2 mM of each dNTP in the presence of 50 mM Tris-HCl (pH 8.3), 50 mM KCl, and 1.6 mM $MgCl_2$. The assembled second reaction mixture was held at 94°C for 10 min prior to the addition of 10 U of *Taq* DNA polymerase and amplification using the cycling conditions for the first round. The primers used were as follows: LTR-1, 5′ CACACAAGGCTAYTTCCCTGA (positions 59 to 69); LTR-2, 5′ TCCYCYTGGCCTTAACCGAAT (positions 863 to 843); LTR-3, 5′ actgtctagaTGGATGGTGCTWCAAGY TAGT (positions 128 to 148); LTR-4, 5′ actgctcgagTCCT TCTAGCCTCCGCTAGTC (positions 783 to 763). Sequences in lowercase letters represent clamp restriction sites, with the restriction sites underlined, Y representing pyrimidine, and W indicating A+T. The numbers in parentheses refer to HXB2 sequence positions (44).

Amplified fragments were extracted once each with phenol-chloroform-isoamylalcohol (25:24:1) and chloroform prior to purification by spin dialysis with Chromaspin C-100 columns (Clontech, Inc.). The fragments were then subjected to digestion with *Xho*I and *Xba*I prior to a second round of spin dialysis and directional cloning into pBluescript KS II− (Stratagene Cloning Systems, Inc.). Plasmid DNA from multiple clones was prepared by rapid column purification (Qiagen, Inc.) and then subjected to cycle sequencing with fluorescent-labeled oligonucleotide primer and dideoxynucleotide terminator approaches (Applied Biosystems, Inc., Foster City, Calif.). Nucleotide sequences were aligned by using MegAlign version 1.01 software (DNAStar, Madison, Wis.).

Preparation and analysis of reporter gene constructions. Cloned HIV-1 LTR fragments were subcloned upstream of the bacterial chloramphenicol acetyltransferase (CAT) gene in order to assess their transcriptional activities. The plasmid pSV_2CAT was double digested with *Hin*dIII and *Bam*HI, and the resulting 1.6-kb fragment containing the CAT-poly(A) region was gel purified. Plasmids containing cloned LTR fragments were double digested with *Xba*I and *Hin*dIII to release a 409-bp LTR fragment corresponding to HXB2 sequence positions 128 to 536. Gel-purified LTR fragments were then ligated in a tripartite reaction with the CAT-poly(A) fragment and an *Xba*I-*Bam*HI double-digested pBluescript KS II− plasmid. pSKCAT, a negative control construction, was constructed by cloning the CAT-poly(A) fragment from pSV_2CAT into pBluescript II SK− at the *Hin*dIII and *Bam*HI sites. pU3R-III, a clone containing the U3/R sequences from HXB2 cloned upstream of the CAT-poly(A) sequence (52), was obtained from the AIDS Reagent Repository. A previously described Tat expression plasmid, pgTat (32), was used in cotransfection experiments. Constructions were recovered by transformation of *Escherichia coli* DH5α cells (Gibco/BRL) and subsequent restriction enzyme screening. All constructions were purified by double banding in ethidium bromide-cesium chloride gradients and confirmed by nucleotide sequencing.

Reporter gene assays were performed essentially as described previously (36). For each transfection, 1.5×10^7 SupT1 cells, maintained under standard culture conditions

(13), were mixed with plasmid DNAs and subjected to 960 μF/0.25 kV of electricity from a Gene-Pulser (Bio-Rad). SupT1 cells are $CD4^+$-TdT^+-$CALLA^-$-DR^--transformed T lymphocytes derived from the pleural effusion of a child with non-Hodgkin's lymphoma. The DNA concentration in all transfections was kept constant with pBluescript KS II−. Cells were harvested 48 h posttransfection, washed once with phosphate-buffered saline, resuspended in 85 μl of 0.25 M Tris-HCl (pH 7.6), and subjected to three successive freeze-thaw cycles in baths held at −80 and 37°C. Total protein concentrations of the lysates were determined by a commercial dye-binding method (Pierce), and equal amounts of protein from each lysate were used in the standard CAT assay (19). Chromatograms were quantitated by a storage phosphor system (Molecular Dynamics, Inc.). Percent conversion was determined by dividing the amount of radioactivity in acetylated forms of [^{14}C]chloramphenicol by the amount of radioactivity in acetylated and nonacetylated forms of [^{14}C]chloramphenicol multiplied by 100.

Electrophoretic mobility shift assays. Oligodeoxyribonucleotides (42 bp) corresponding to both strands of the region of U3 containing the three Sp1 sites (HXB2 sequence positions 370 to 411) from wild-type and mutant clones were synthesized and gel purified with denaturing acrylamide gels. One member of each set was 5′ end labeled with ^{32}P prior to annealing in stoichiometric amounts with its cognate complementary strand at 70°C for 30 min. For noncompetitive reactions, sufficient labeled double-stranded probe to give a final concentration of 1.57×10^{-11} M was mixed with 4.0 μl of 5× binding buffer [1.0 mg of poly(dI)-poly(dC) per ml, 100 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 7.8), 3 M KCl, 50% glycerol, 10% Nonidet P-40] and 30 ng of purified Sp1 protein (Promega) for 25 min at 10°C in a final volume of 20 μl. For competitive reactions, unlabeled double-stranded DNA was preincubated with the Sp1 protein and binding buffer at concentrations of 1.4×10^{-8}, 2.9×10^{-8}, and 5.7×10^{-8} M for 5 min at 10°C prior to the addition of labeled probes. Two microliters of loading buffer (48% glycerol, 0.05% bromophenol blue) was then added prior to electrophoresis through a 5% acrylamide gel (acrylamide-to-bisacrylamide ratio, 50:1) at 14.4 V/cm for 1 h. The gel was cast and run with a buffer containing 40 mM Tris base, 307 mM glycine, and 0.16% Nonidet P-40. The gel was dried down prior to storage phosphor imaging (Molecular Dynamics, Inc.) and standard autoradiography.

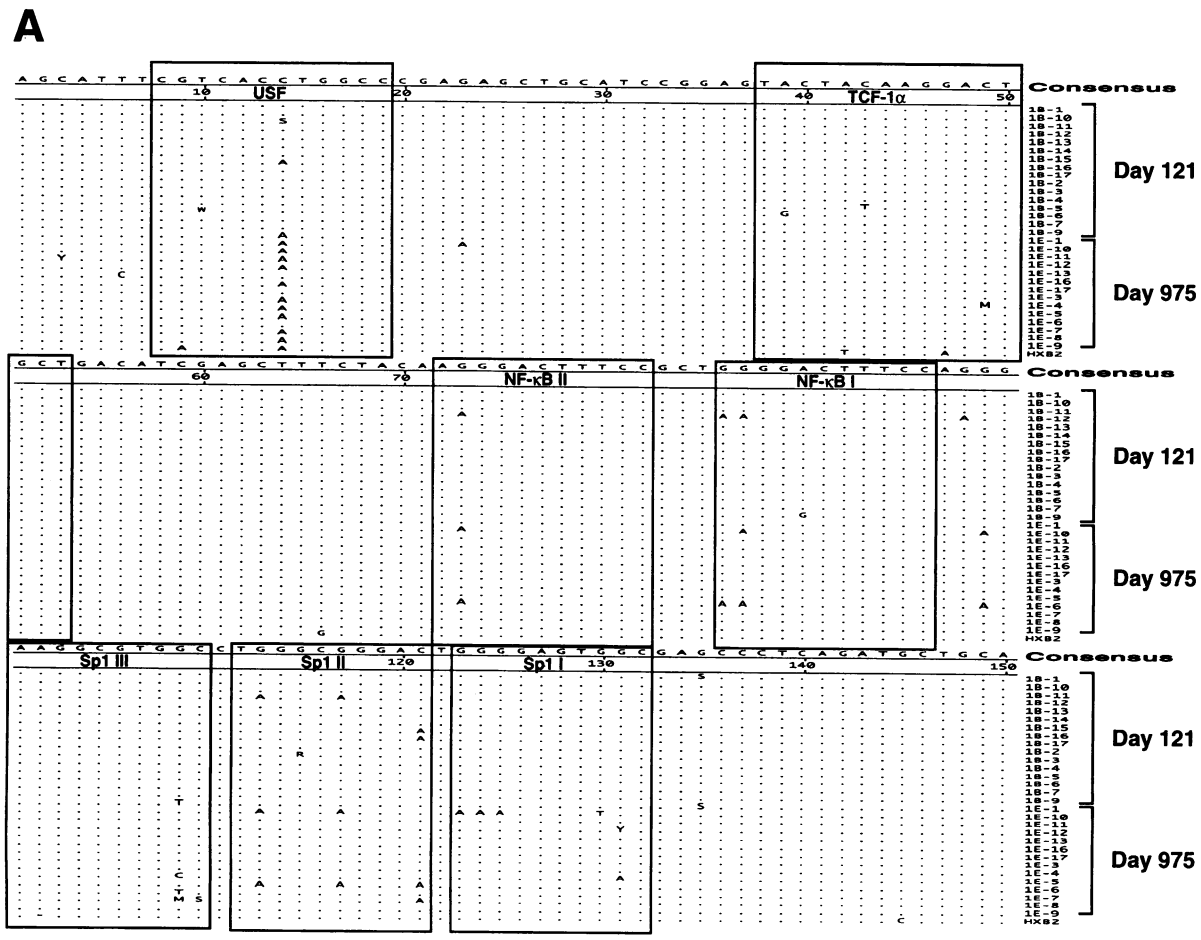
The sequences for the oligodeoxynucleotides used are given here for the sense strand in each pair: 21.1.B-7, 5′ TTCCAGG GAAGGCGTGGGCTGGGCGGGACTGGGGAGTGGCG AG; 21.1.B-6, 5′ TTCCAGGGAAGGCGTGGGCTaGGCa GGACTGGGGAGTGGCGAG; 21.1.B-10, 5′ TTCCAGG GAAGGCGTGGGCTaGGCaGGACTaaaGAGTGGCGAG. Mutations are indicated by lowercase designations. Sp1 binding sites are underlined.

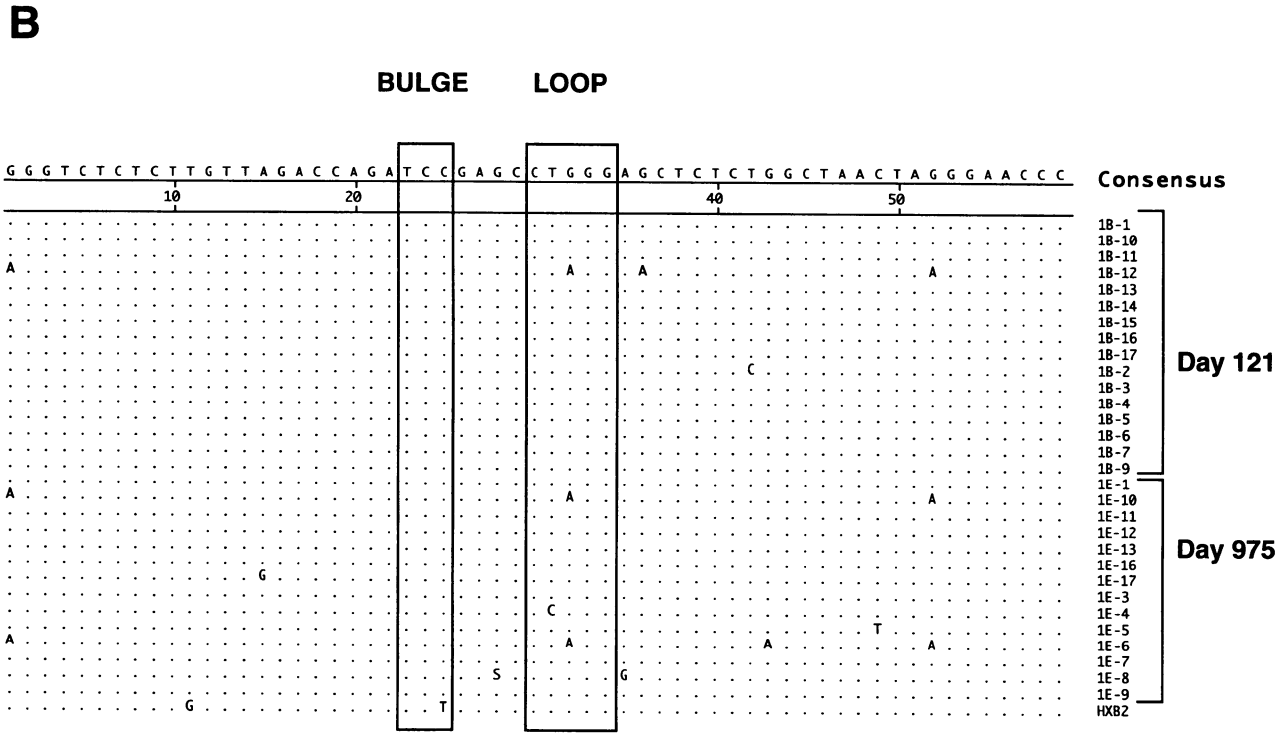
Nucleotide sequence accession numbers. The sequences described here have been deposited at GenBank under the accession numbers L28841 through L28917.

RESULTS

LTR sequences display considerable point mutations and length polymorphisms. LTR and *gag* leader sequences were recovered from sequential samples of noncultivated PBMC DNA from four patients over approximately 3 years by a nested PCR technique. The day designation refers to the time of entry into the clinical trial. We have chosen to examine proviral genotypes in this work, as opposed to cell or plasma RNA genotypes, in order to assess the full range of transcriptional activities that would be obscured by sequencing from RNA templates. The amplicon, representing HXB2 sequence positions 128 to 783, extends from the second AP-1 binding site through the ψ site (Fig. 1). Multiple molecular clones from each time point were completely sequenced and aligned. Since the majority of mutations were within either the core enhancer/promoter or TAR sequences, these subsets of sequences alone are shown in Fig. 2. Sequences from days 121 (clones 1B) and 975 (clones 1E) from patient 21.1.1 are shown for the enhancer/promoter region in Fig. 2A and for the TAR element in Fig. 2B. The USF and TCF-1α sites are shown upstream of the two NF-κB and three Sp1 sites. Sp1 site III had an insertion in all clones of an A nucleotide at position 2 not present in most published sequences (39). Clone 1B-12 contains G-to-A mutations in the second position of NF-κB site II, in the first and second positions of NF-κB site I, and in the second and sixth positions of Sp1 site II (Fig. 2A). This clone has two additional G-to-A mutations at positions +32 and +36 in the conserved heptanucleotide loop and stem, respectively, of the TAR element (Fig. 2B). Clone 1B-9 has an A-to-G transition in the 5th position of NF-κB site I and clones 1B-16 and 1B-17 have C-to-A transversions in the last position of Sp1 site II. These last two mutations would be expected to have minimal functional significance, as they are not within the central core of the Sp1 site (24). Thus, 2 of 16 of the day 121 clones and 4 of 14 of the day 975 clones from patient 21.1.1 bear mutations in conserved portions of the NF-κB promoter and/or TAR regions. Eleven of the day 975 clones had a C-to-A transversion in the conserved hexanucleotide core of the USF site not found in the day 121 clones. Three clones from this patient at the later time point also displayed mutations in the primer binding site (HXB2 positions 636 to 653). Clones 1E-6, 1E-18, and 1E-12 had 5, 2, and 1 mutations, respectively, in this highly conserved 18-bp element (data not shown). No mutations were found in the TATA box, major 5′ splice donor sequence, or ψ site sequences from this patient. However, all of the clones from this patient contained the TAR bulge sequence of TCC instead of the more typical TCT sequence reported in the data base (Fig. 2B) (24).

Figures 2C and D contain the enhancer/promoter and TAR sequences, respectively, from days 75 (clones 2A), 431 (clones 2B), and 919 (clones 2C) from patient 21.1.2. All clones from this patient, with the exception of 2C-6, display a 29-bp insertion that essentially duplicates the TCF-1α site. This insertion can be dissected into 15- and 10-bp direct repeats separated by a 4-bp spacer sequence. The 15-bp direct repeat is made imperfect by virtue of three point mutations and a single-base pair insertion. This group of clones is also notable for the presence of an additional G residue between NF-κB site I and Sp1 site III. Once again, Sp1 site III also has a single-base pair insertion not reported in most data base sequences, although this involves a G nucleotide. Sp1 site I contains a G-to-A mutation in the first position in all but a single clone. There are polymorphisms in the penultimate position of Sp1 site III as well as G-to-A mutations in the fourth position of Sp1 site III in clone 2B-5 and in the second position of the Sp1 site I of clone 2C-4. Unlike the patient 21.1.1 group, these mutations do not show a temporal trend. The USF, TATA box, primer binding site, and ψ site sequences from this patient do not diverge significantly from those of HXB2. Similarly, the TAR regions from this patient have conserved loop and TCT bulge sequences (Fig. 2D). Clones 2C-1 and 2C-5 displayed deletions of the sequence TG | GTGA (HXB2 positions 741 to 746) representing the major 5′ splice


A
USF
TCF-1α
NF-κB II
NF-κB I
Sp1 III
Sp1 II
Sp1 I
Consensus
Day 121
Day 975


B
BULGE
LOOP
Consensus
Day 121
Day 975
HXB2

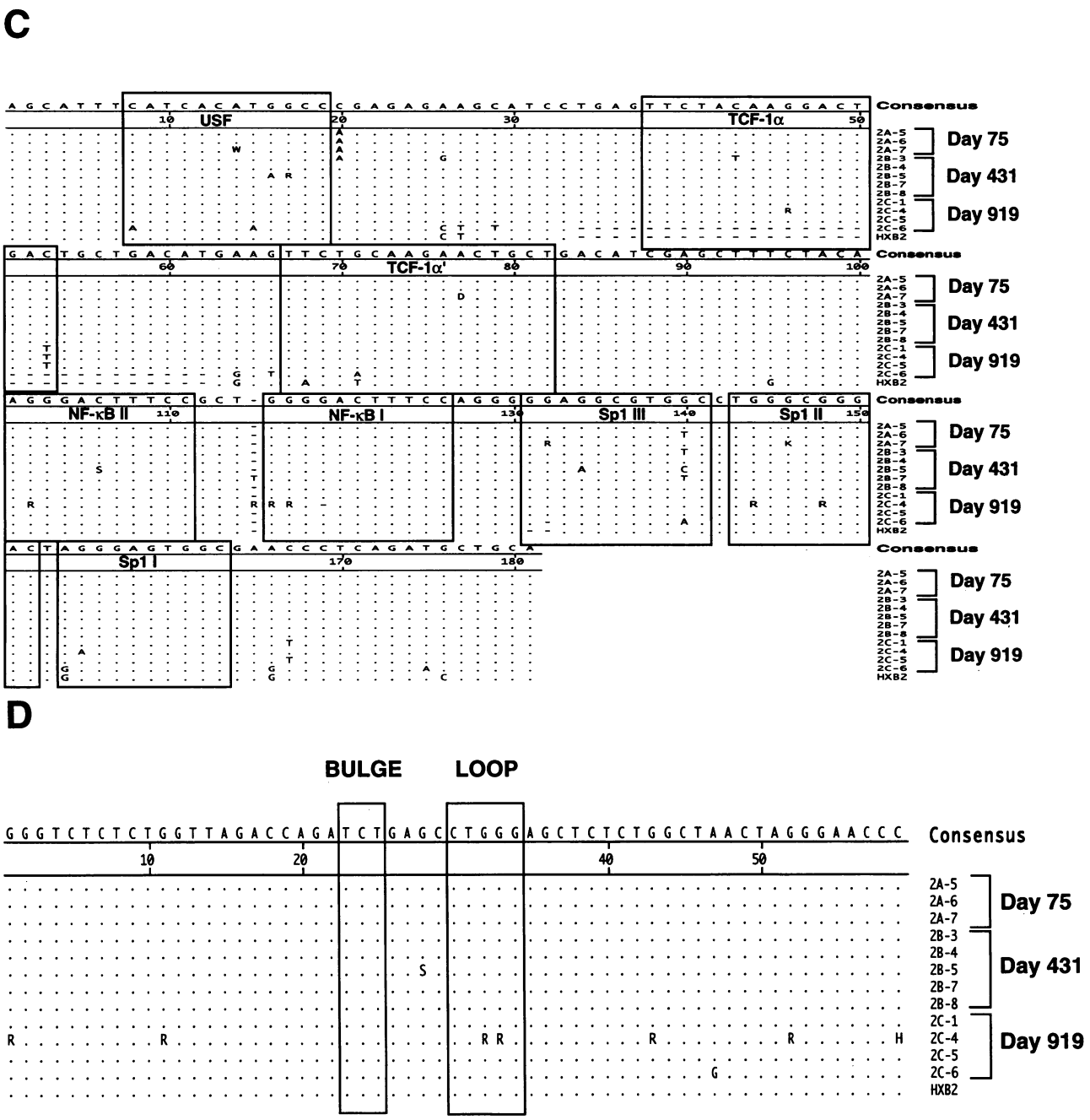


FIG. 2. Alignment of U3 and TAR sequences. The locations of control region sequences are indicated by boxes. The compiled consensus nucleotide sequence of each data set is given on the top line, and the cognate region from HXB2 is given on the bottom line. Dots connote identity with the consensus sequence, deviations from the consensus are indicated by the cognate nucleotide assignment, and dashes indicate gaps introduced to effect the alignment. Clones are identified at the right of each sequence and are grouped by days. The numbering scheme for the U3 sequences, which begins at the 5′ end of the sequence shown, bears no relationship to either the HXB2 numbering scheme (44) or to the position relative to the cap site. The TAR sequences are numbered from the cap site (+1). (A) Patient 21.1.1 (U3); (B) patient 21.1.1 (TAR). (C) Patient 21.1.2 (U3); (D) patient 21.1.2 (TAR). (E) Patient 21.1.3 (U3); (F) patient 21.1.3 (TAR). (G) Patient 21.1.4 (U3); (H) patient 21.1.4 (TAR). R, purine; W, A or T; S, C or G; M, A or C; K, G or T; D, not C.

donor sequence. However, these two clones possessed the sequence AGC | GT in its place (data not shown) that has recently been described as a cryptic 5′ splice donor with markedly delayed effects on viral protein production (43).

Figures 2E and F contain the enhancer/promoter and TAR sequences, respectively, from days −46 (clones 3A) and 1,079 (clones 3B) from patient 21.1.3. All clones except 3A-4 have a 32-bp insertion in the TCF-1α region. This insertion consists of a 7-bp insertion (AAGGACT) followed by a 22-bp direct repeat followed by a 3-bp insertion (TAC) which generates another 16-bp TCF-1α-like site. The AAGGACT sequence is itself a direct repeat of the first 7 bp of the 22-bp repeat. Clones 3A-7 and 3A-8 have an 11-bp insertion in the Sp1 site cluster which creates a fourth Sp1 site (designated Sp1 site IV).

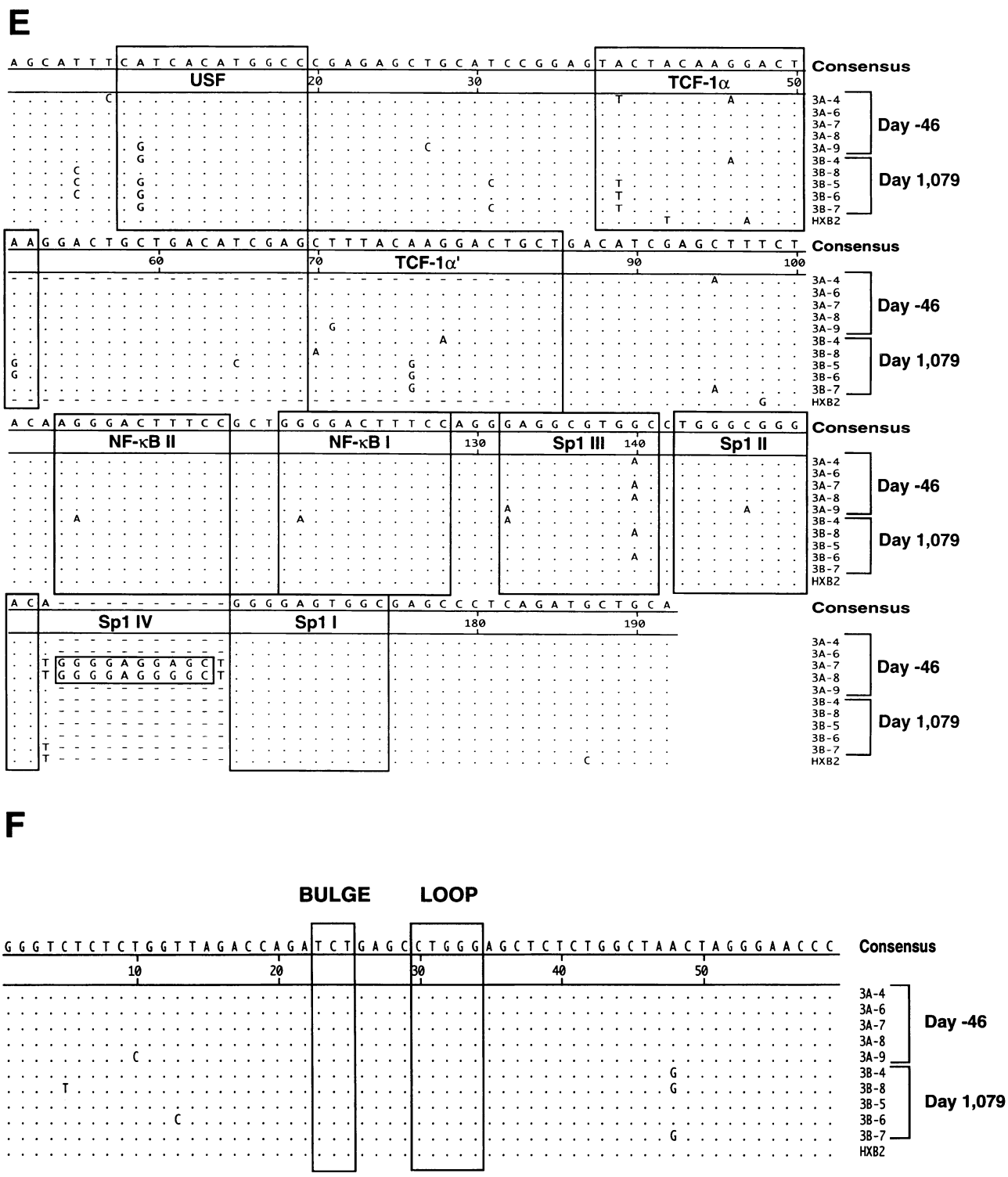
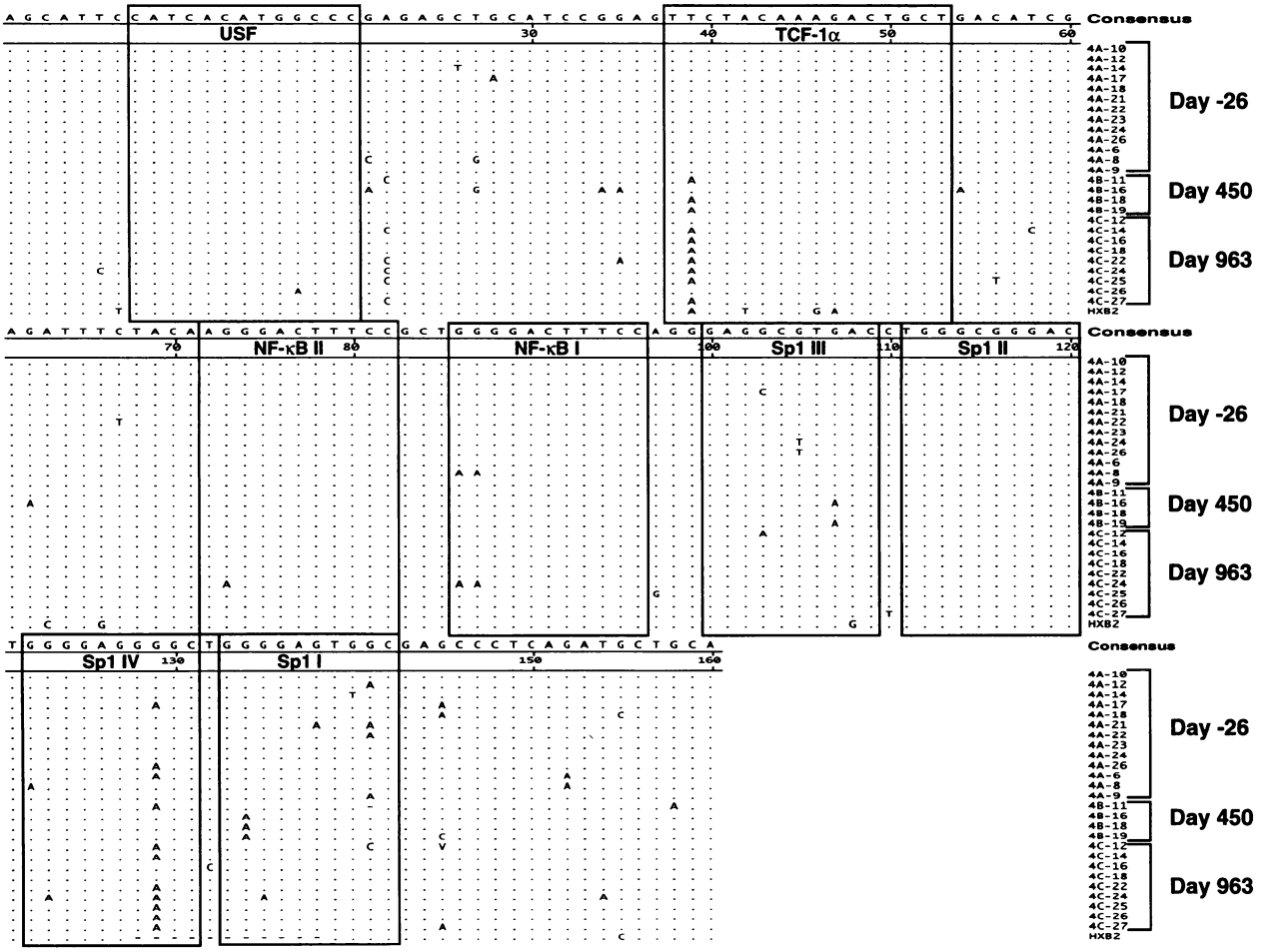


FIG. 2—*Continued.*

This Sp1 site is identical to Sp1 site I except for a T-to-G change at position 7 in both clones and a G-to-A transition at position 8 in clone 3A-7. Sp1 site III lacks the single-base pair insertion seen in the LTR clones from patients 21.1.1 and 21.1.2. The TAR, USF, TATA box, primer binding site, and major 5′ splice donor sequences from this patient do not diverge significantly from those of HXB2. As shown in Fig. 3, however, all clones from this patient show either insertions (3A-7, 3A-8, 3B-4, 3B-5, and 3B-7) or deletions (3A-4, 3A-6, 3A-9, 3B-6, and 3B-8) of adenylate residues in the ψ site (1, 28) similar to the single clade B sequences HIVSF2 and HIVYU2 and the clade O sequences HIVMAL and CPZGAB in the Los Alamos data base (39). The importance of distal sequences in the function of the ψ site is indicated by the dashed arrow in



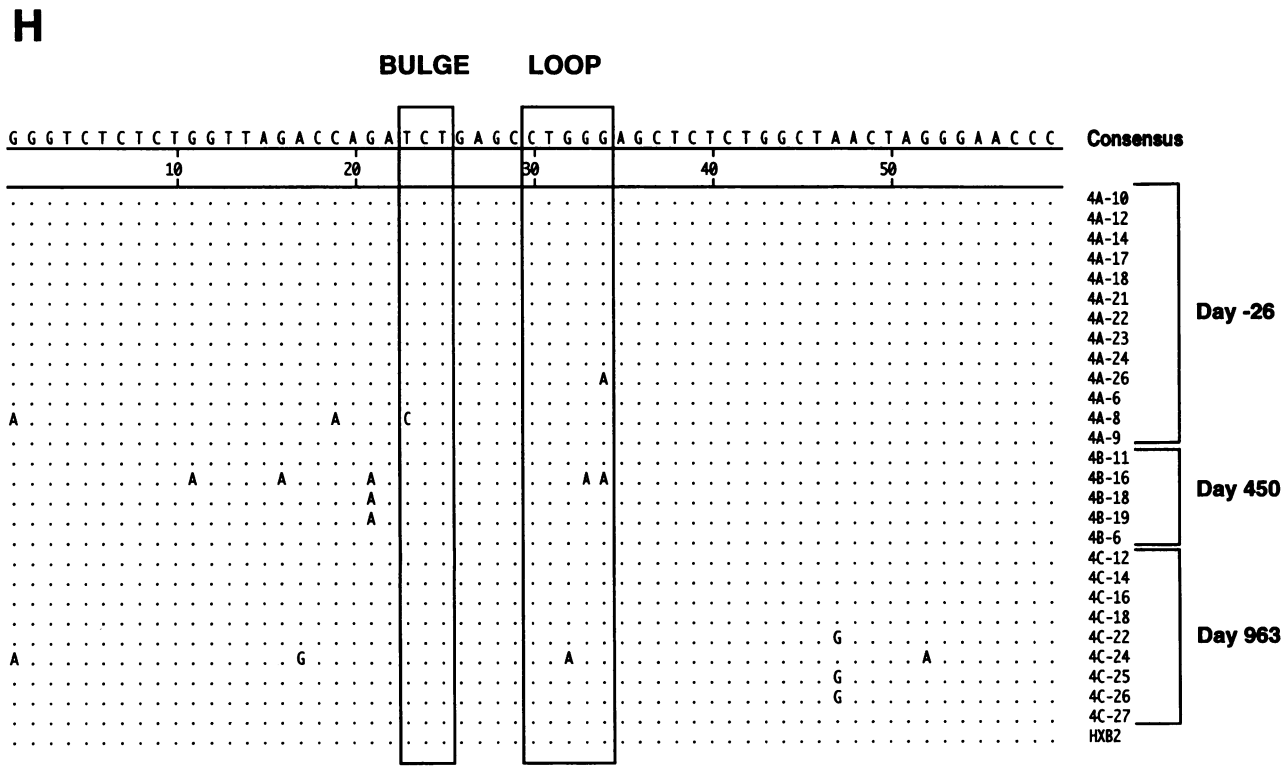


FIG. 2—*Continued.*

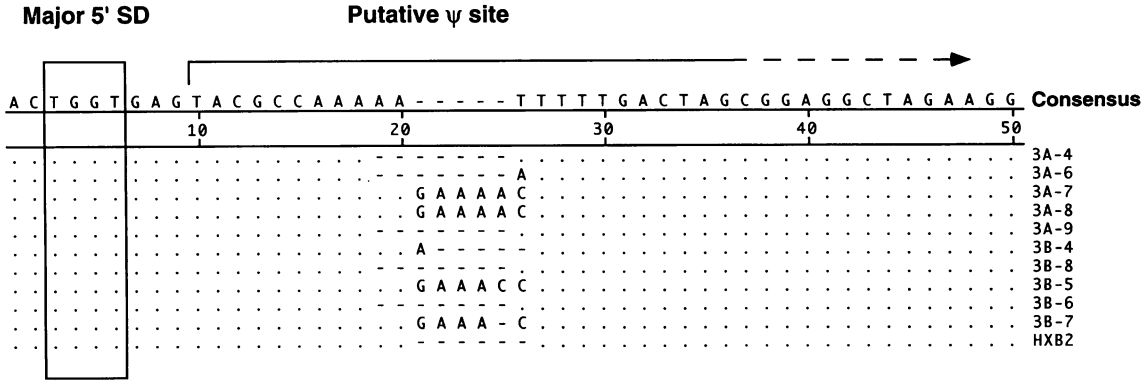


FIG. 3. Length polymorphism of ψ site sequences from patient 21.1.3. The length polymorphisms observed are shown by alignment to the cognate region of HXB2 as given in Fig. 2. The positions of the major 5′ splice donor and ψ site (1, 28) are indicated by the box and angled arrow, respectively.

Fig. 3. The insertion of the sequence $GA_{3-4}C$ within the ψ site seen in clones 3A-7, 3A-8, 3B-5, and 3B-7 has not previously been reported in the HIV-1 data base (39).

Figures 2G and H contain the enhancer/promoter and TAR sequences, respectively, from days −26 (clones 4A), 450 (clones 4B), and 963 (clones 4C) from patient 21.1.4. All clones from all time points for this patient have an insertion that creates a fourth Sp1 site identical to Sp1 site I except for a T-to-G transversion at position 7. This is precisely the insertion seen in patient 21.1.3 clones 3A-7 and 3A-8 (Fig. 2E). Sp1 site III lacks the single-base pair insertion seen in patients 21.1.1 and 21.1.2. There is a polymorphism at position 8 in this site with a tendency for a G-to-A transition in the day 963 clones. With time, the TCF-1α site displays a T-to-A transversion in position 2. Clone 4A-8 has a GG-to-AA dinucleotide transition in NF-κB site I, and clone 4C-24 contains this change as well as a G-to-A transition in the second position of NF-κB site II. The presence of a G in the second position of the NF-κB sites is highly conserved in reported HIV-1 sequences (39). Three clones from this patient, 4A-26, 4B-16, and 4C-24, have mutations in the highly conserved TAR loop heptanucleotide CTGGG, and an additional clone, 4A-8, has a T-to-C transition in the conserved TAR bulge trinucleotide TCT (Fig. 2H). As seen for some of the clones from patient 21.1.3, all clones except 4B-18 and 4B-19 contain an insertion of the sequence $GA_{3-4}C$ within the ψ site (data not shown). The USF site, primer binding site, and major 5′ splice donor sequences do not deviate substantially from previously reported HIV-1 sequences.

Transcriptional activity of LTR genotypes by reporter gene analysis. Portions (409 bp) of the recovered LTR genotypes were subcloned upstream of the CAT gene and the simian virus 40 intron-poly(A) sequence to experimentally assess the effect of the observed mutations in transcriptional control regions. These subcloned sequences begin at the 5′ end of the original amplicon (position −347 relative to the cap site) and extend through the TAR element to position +81 just downstream of the poly(A) signal (Fig. 1). The results of transient transfection of these reporter gene constructs into SupT1 cells are shown in Fig. 4. It should be noted that these assays can be compared for each patient but not between patients, as studies for each patient were performed on different days with different amounts of lysate. Three clones from patient 21.1.1 on day 975, 1E-6, 1E-10, and 1E-17, were selected for reporter gene analysis. Clones 1E-6 and 1E-10 had mutations within NF-κB sites I and II, Sp1 site II, and the TAR loop (see Fig. 2A and B). Clone 1E-10 also contained mutations in Sp1 site I. Clone 1E-17 appeared to be the wild type, as assessed strictly by sequence analysis. The relative basal-level transcriptional activities of these clones are shown in Fig. 4A. Whereas the putative wild-type clone 1E-17 displayed a typical dose response in this assay, the mutant clones 1E-6 and 1E-10 were essentially devoid of transcriptional activity. These mutant clones, containing mutations in the highly conserved CTGGG motif of the TAR loop, were also insensitive to transactivation by Tat in comparison with the wild-type clone 1E-17 (Fig. 4B). Clone 1E-17 was fully capable of Tat transactivation despite the presence of an atypical TAR bulge sequence, TCC versus TCT, that was present in all of the genotypes recovered from patient 21.1.1. In contrast to the range of transcriptional activities of the genotypes recovered from patient 21.1.1, the clones from patient 21.1.2 displayed only a twofold range of basal transcriptional activities (Fig. 4C). These clones differed from those of patient 21.1.1 and from most clade B U3 region sequences reported in the Los Alamos data base (39) in that they contained a 29-bp insertion upstream of NF-κB site II that imperfectly duplicated the TCF-1α site. The transcriptional activity levels of these representative clones from patient 21.1.2 were significantly lower than that of the control plasmid pU3R-III that lacked this 29-bp insertion (Fig. 4C).

In contrast, the basal-level transcriptional activities of most of the genotypes recovered from patient 21.1.3 compare favorably to that of pU3R-III (Fig. 4D). All of these genotypes, except for clone 3A-4, contain a 32-bp insertion in the TCF-1α region. Clone 3A-4 was the most transcriptionally active in the reporter gene analysis shown in Fig. 4D. Clone 3A-7 differed from clone 3A-6 by the presence of an additional Sp1 site (site IV) between Sp1 sites II and I, although the addition of this Sp1 site had little effect on the basal-level activity of this clone compared with that of clone 3A-6, which contained three Sp1 sites. Clone 3A-9, the least transcriptionally active of the genotypes tested from patient 21.1.3, contained a C-to-A transversion in the highly conserved central core of Sp1 site II and a G-to-A transition in the conserved first position of Sp1 site III.

The basal-level activities of two selected genotypes from patient 21.1.4 are shown in Fig. 4E. All but a single genotype recovered from this patient displayed an extra Sp1 site inserted between Sp1 sites II and I. Clone 4A-23 had a higher basal-level transcriptional activity than the wild-type clone 1E-17, which contains three Sp1 sites. Clone 4A-8, which contains tandem G-to-A transitions in the first two positions of NF-κB site I (Fig. 2G), had markedly impaired basal-level transcription. This same clone also contains a T-to-C transition in the first position of the TAR bulge (Fig. 2H). Despite these cumulative mutations, clone 4A-8 is capable of transactivation

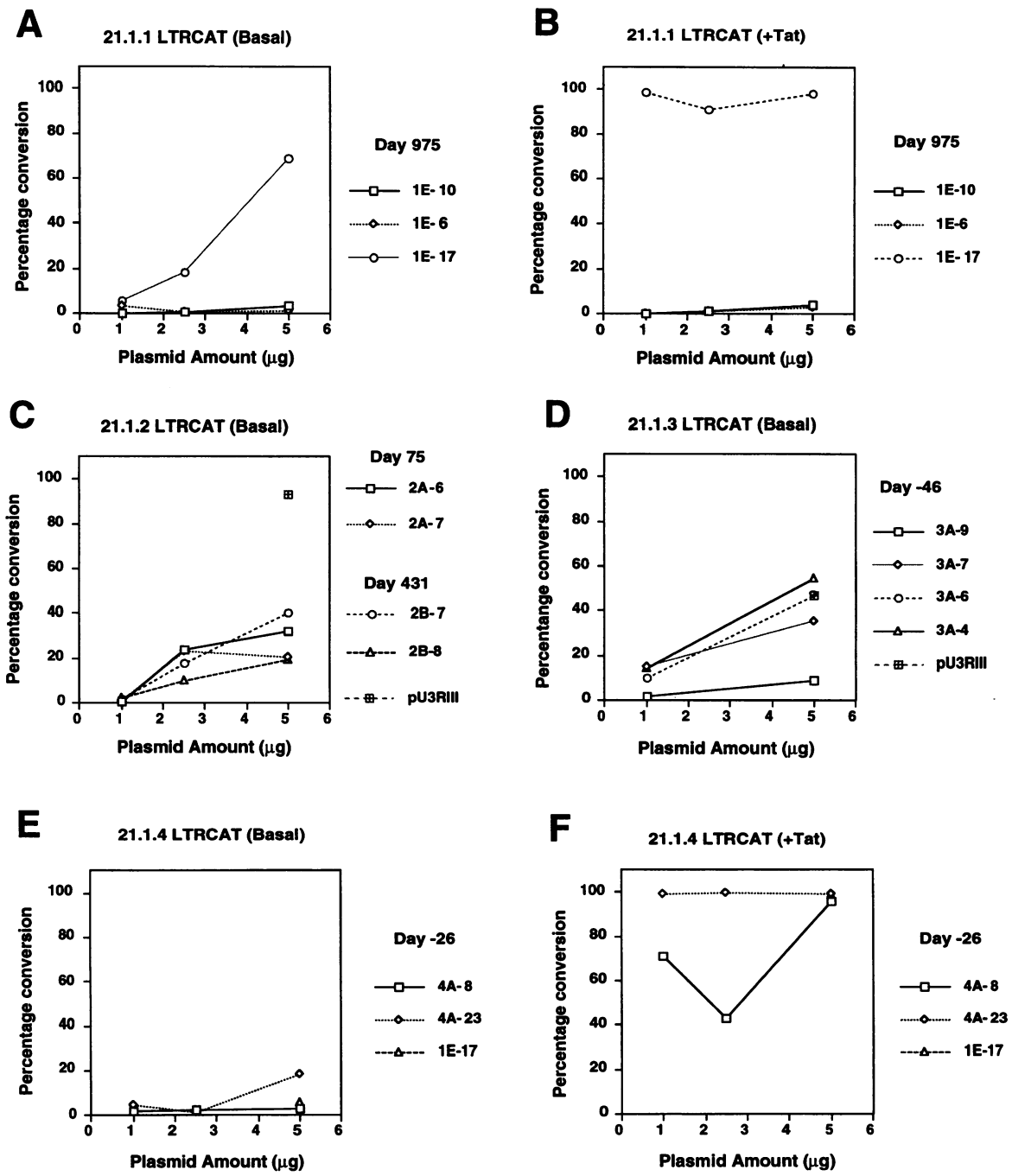


FIG. 4. Reporter gene analyses of LTR mutations. The results of CAT assays of recovered LTR sequences cloned upstream of the CAT gene are shown in graphic form. Plasmid DNA constructions (LTRCAT), in amounts indicated on the horizontal axis for each graph, were transfected into SupT1 cells by electroporation, and CAT activity was determined 48 h later. CAT activity, expressed as the percentage of chloramphenicol converted to acetylated forms, is given on the vertical axis of each graph. The identities of clones tested are indicated to the right of each graph. Basal, transfections with LTRCAT constructions alone. Transactivated cotransfections (+Tat) of LTRCAT constructions were done with 1.0 μg of the Tat expression plasmid pgTat. CAT assays can be compared for each patient but not between patients, as studies for each patient were performed on different days. (A) Patient 21.1.1, basal. (B) Patient 21.1.1, transactivated. (C) Patient 21.1.2, basal. (D) Patient 21.1.3, basal. (E) Patient 21.1.4, basal. (F) Patient 21.1.4, transactivated (note that the data for clones 4A-23 and 1E-17 overlap at 100% conversion).

by Tat (Fig. 4F [note that the data for clones 4A-23 and 1E-17 overlap at 100% conversion]) unlike clone 1E-6 in Fig. 4B, which contained both NF-κB and Sp1 site mutations and a T-to-C transition in the third position of the TAR bulge. Previous work using engineered mutations of laboratory-adapted strains of HIV-1 suggests that Sp1 site mutations impaired both basal-level and transactivation-dependent transcriptions (6). The ability of clone 4A-8, but not clone 1E-6, to be Tat transactivated may reflect either the role of Sp1 site mutations in transactivation or a more severe consequence of

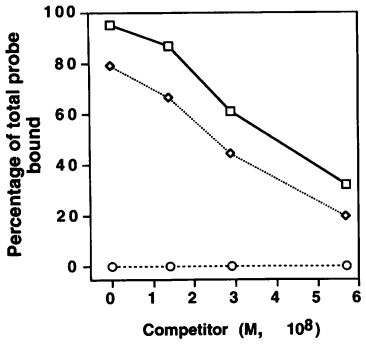


FIG. 5. Electrophoretic mobility shift assays with purified Sp1. ^{32}P end-labeled double-stranded oligodeoxyribonucleotide probes representing clones from patient 21.1.1 1E-6 (◇), 1E-10 (○), and 1E-17 (□) on day 975 were mixed with a fixed amount of purified human Sp1 protein in the presence of increasing amounts of unlabeled probe prior to electrophoresis through a nondenaturing polyacrylamide gel and storage phosphor imaging. The percentage of labeled probe bound was then determined by dividing the amount of slowly migrating signal by that for free probe.

changes in the third versus the first position of the TAR bulge on transactivation. Although not directly tested, a single clone from patient 21.1.4 in each time point containing mutations in the TAR loop would be expected to be transactivation defective on the basis of data presented in this paper and data previously reported (12).

In order to interpret these data in light of the potential for experimentally induced sequence variation, we performed the following control experiment. A defined template was amplified by the 54-cycle strategy described above. Twelve independent molecular clones were subsequently recovered, and their inserts were completely sequenced. The amplicon was 657 bp long but 42 nucleotides of this amplicon were represented by the amplifying primers and, thus, not subject to PCR-generated misincorporation error. The remaining 615-bp region contained 10 mutations distributed among the 12 independent clones. This resulted in a misincorporation rate calculated as follows: 10 mutations/(615 bp)(12 clones)(54 cycles) = 2.5×10^{-5} PCR/cloning mutations per base pair per cycle. This represents the background variation introduced into these sets of sequence data strictly on the basis of the experimental recovery of the genotypes analyzed. No length polymorphisms were observed in these control experiments. Overall, the amount of sequence variation in our patient LTR data sets exceeded that of the control template by a factor of 10. Thus, we believe that the sequence data presented in this paper represent an accurate reflection of the sequence variation that existed in these patients in vivo.

Correspondence between Sp1 site mutations and Sp1-DNA binding activities. The mutations in Sp1 sites observed in three genotypes recovered from patient 21.1.1, assessed in terms of transcriptional activity by reporter gene analysis, were also studied by using a DNA binding assay and highly purified Sp1 protein (Fig. 5). Three sets of double-stranded oligodeoxyribonucleotides spanning the entire Sp1 region (HXB2 sequence positions 370 to 411) were synthesized to represent the wild-type genotype 1E-17 and mutant genotypes 1E-6 and 1E-10. The DNA probe from clone 1E-6, containing wild-type Sp1 sites I and III and a mutant Sp1 site II, bound somewhat less purified Sp1 protein than the fully wild-type probe. In contrast, the probe based on clone 1E-10, with mutant Sp1 sites I and II and wild-type Sp1 site III, completely failed to bind Sp1 protein. This suggests that Sp1 site III, by itself, is incapable of binding purified Sp1 in this assay. The fact that the clone 1E-6 probe, representing wild-type Sp1 sites I and III, bound Sp1 only slightly less well than the fully wild-type Sp1 site probe suggests that Sp1 site III is capable of binding Sp1 protein in the presence of Sp1 site I. Despite the near wild-type level of Sp1 binding to the Sp1 sites representative of clone 1E-6, this clone was incompetent for either basal-level or Tat-induced transactivation in reporter gene analyses (Fig. 4A and B). This suggests that other mutations in clone 1E-6, perhaps the accompanying NF-κB site mutations, were important in their transcriptionally impaired phenotypes.

DISCUSSION

The serial analysis of LTR/*gag* leader sequences amplified directly from the PBMCs of these four HIV-1-infected patients revealed a remarkable degree of intra- and interpatient diversity associated with a wide range of basal and Tat-induced transcriptional activity. This diversity included both point mutations and extensive length polymorphisms. The majority of point mutations were G-to-A transitions. This is consistent with previous analyses of HIV-1 quasispecies diversity (33). The sequences from three of these patients, 21.1.2, 21.1.3, and 21.1.4, diverged significantly from the published group of clade B sequences primarily on the basis of length polymorphisms (39). All three of these patients displayed length polymorphisms associated with duplication of regulatory sequence motifs. Previous reports have described duplications of Sp1, NF-κB, and TCF-1α sequences (10, 18, 26). Koken and colleagues found that the LTR sequences of 1 of 17 patients studied had a perfect duplication of Sp1 site III inserted between site III and II (26). Four other patients in this study had partial duplications of the TCF-1α site. In agreement with the reporter gene data presented by us for the four Sp1 site LTRs recovered from patient 21.1.4, the four-Sp1 site variant studied by Koken had a slightly higher basal transcriptional rate than the three-Sp1 site variant. A distinct growth advantage of the four-Sp1 site variant was demonstrated by Koken by the cocultivation of equal amounts of isogenic recombinant proviruses containing three and four Sp1 sites. This suggests that variants containing an additional Sp1 site may represent a naturally occurring genotype of HIV-1 with a replicative advantage over variants containing three Sp1 sites. However, the variant from patient 21.1.3 that contained a direct repeat in the TCF-1α region as well as four Sp1 sites (clone 3A-7) was less transcriptionally active in the reporter gene assay than variants with three Sp1 sites (Fig. 4D). This was true for three-Sp1 site variants that either contained (clones 3A-4 and 3A-6) or lacked (pU3R-III) a duplication of the TCF-1α site. This suggests that the increased transcriptional advantage of four-Sp1 site variants may only be realized in viruses that lack direct repeats in TCF-1α sequences.

In contrast to the enhanced transcriptional effect of Sp1 site duplications, the imperfect duplications of the TCF-1α sites observed for patients 21.1.2 and 21.1.3 revealed a dichotomous effect on basal-level transcription. The transcriptional activity levels of representative LTR genotypes from patient 21.1.2, which contained a 29-bp insertion imperfectly duplicating the TCF-1α site, were all significantly lower than that of pU3R-III. Conversely, the basal-level promoter strengths of representative genotypes from patient 21.1.3, which contained a 32-bp insertion in the TCF-1α region, were similar to that of pU3R-III. The variants reported by Koken containing partial TCF-1α site duplications were reported to have slightly less transcriptional activity than those without these duplications

(26), while a combined TCF-1α site duplication and point mutation between NF-κB sites I and II in a variant reported by Golub and colleagues showed increased promoter activity and replication kinetics (18). Careful analysis of the duplications described by both our work and that of others reveals that none of the TCF-1α duplications are exactly alike. Thus, the differential effects on transcriptional activity of the TCF-1α variants are complex and remain incompletely defined. We observed no NF-κB site polymorphisms that either duplicated or deleted enhancer elements as reported in a study of cultivated viral variants (10).

The results presented in this work also point to the role of point mutations in the transcriptional activity of LTR genotypes. Point mutations in the conserved domains of Sp1 sites, especially site II, and the NF-κB sites showed effects on transcriptional activities of their cognate LTRs. Sp1 site mutations impaired basal transcription, in agreement with mutational analysis of laboratory strains of HIV-1 (6). However, our data from Sp1 site mutations for patient 21.1.3 are not consistent with a previous report that all three Sp1 sites need be mutated to see transcriptional impairment in reporter gene analyses (23). This may be due to the use of HeLa cells in this earlier work as opposed to the use of T-lymphoblastoid lines in our experiments (23). Point mutations in the TAR loop from patient 21.1.1 were found to be inactivating for Tat-induced transactivation as suggested by previous work using laboratory-generated TAR mutants (12). Although not directly tested, clones 4A-26, 4B-16, and 4C-24 would also be presumed to be similarly transactivation impaired given their TAR loop mutations. Interestingly, clone 1E-17 from patient 21.1.1 was fully capable of Tat-induced transactivation despite the presence of an atypical TAR bulge sequence, TCC versus TCT, that was present in all of the genotypes recovered from patient 21.1.1. These data suggest a greater degree of functional tolerance to TAR bulge region mutations than would be inferred by inspection of published HIV-1 TAR sequences (39).

Given that all four of the patients described in this report were undergoing open-label therapy with recombinant gp160 during the course of this study, it is possible that some of the mutations observed from nonbaseline time points were related to vaccination. One patient in this study, 21.1.1, showed a trend toward the accumulation of transcriptional control region mutations over time. We do not know whether this trend was the consequence of vaccination or the natural history of this individual's disease. Further, the functional significance of this observation to the clinical course of this individual remains unclear. Although this report of four patients over 3 years, and a previous report of a single patient over 4 years (9), did not demonstrate a uniform temporal variation in basal or Tat-induced transcriptional activity, we feel that this question has yet to be definitively answered by such relatively small studies. The growing body of data showing the common occurrence of regulatory region duplications in naturally occurring LTR genotypes, versus genotypes derived from cocultivated isolates, suggests that this question warrants attention in a larger number of HIV-1-infected individuals.

This report highlights the need to focus on the range of naturally occurring mutations that have an impact upon the transcriptional control of HIV-1 RNA expression. We have detected a large number of point mutations and length polymorphisms that affect the basal and Tat-induced levels of HIV-1 transcription. Further work is required to evaluate the functional consequences of mutations described in this work that map outside of the U3/R regions, such as the primer binding site, major 5′ splice donor, and ψ site. We do feel that dissection of the structural and functional features of HIV-1 LTR/*gag* leader genotypes may be useful in the molecular genetic analysis of patients undergoing therapeutic intervention in HIV-1 disease. This would represent a deviation away from a strictly quantitative analysis of HIV-1 viral load and toward a more functional approach to the assessment of in vivo response to therapy.

ACKNOWLEDGMENTS

We thank the Infectious Disease Service at the Walter Reed Army Medical Center, especially C. Oster, N. Aronson, S. Johnson, and J. Lennox, and the ward 11 nursing staff of the Henry M. Jackson Foundation, especially M. Swatson and L. Pascal, for their care of the patients involved in these studies. We thank M. Vahey for her contributions to the early part of this work and review of the manuscript. We thank J. Kim, D. Birx, D. Burke, R. Carroll, J. Mosca, and D. St. Louis for helpful discussions and critical review of the manuscript. B. Cullen provided the pgTat construction.

REFERENCES

1. **Aldovini, A., and R. A. Young.** 1990. Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. J. Virol. **64:**1920–1926.
2. **Arrigo, S. J., S. Weitsman, J. A. Zack, and I. S. Chen.** 1990. Characterization and expression of novel singly spliced RNA species of human immunodeficiency virus type 1. J. Virol. **64:**4585–4588.
3. **Arya, S. K., C. Guo, S. F. Josephs, and F. Wong-Staal.** 1985. Trans-activator gene of human T-lymphotropic virus type III (HTLV-III). Science **229:**69–73.
4. **Bagasra, O., S. P. Hauptman, H. W. Lischner, M. Sachs, and R. J. Pomerantz.** 1992. Detection of human immunodeficiency virus type 1 provirus in mononuclear cells by in situ polymerase chain reaction. N. Engl. J. Med. **326:**1385–1391.
5. **Bagnarelli, P., S. Menzo, A. Valenza, A. Manzin, M. Giacca, F. Ancarani, G. Scalise, P. E. Varaldo, and M. Clementi.** 1992. Molecular profile of human immunodeficiency virus type 1 infection in symptomless patients and in patients with AIDS. J. Virol. **66:**7328–7335.
6. **Berkhout, B., and K. T. Jeang.** 1992. Functional roles for the TATA promoter and enhancers in basal and Tat-induced expression of the human immunodeficiency virus type 1 long terminal repeat. J. Virol. **66:**139–149.
7. **Cullen, B. R., and W. C. Greene.** 1989. Regulatory pathways governing HIV-1 replication. Cell **58:**423–426.
8. **Delassus, S., R. Cheynier, and S. Wain-Hobson.** 1991. Evolution of human immunodeficiency virus type 1 *nef* and long terminal repeat sequences over 4 years in vivo and in vitro. J. Virol. **65:**225–231.
9. **Delassus, S., A. Meyerhans, R. Cheynier, and S. Wain-Hobson.** 1992. Absence of selection of HIV-1 variants in vivo based on transcription/transactivation during progression to AIDS. Virology **188:**811–818.
10. **Englund, G., M. D. Hoggan, T. S. Theodore, and M. A. Martin.** 1991. A novel HIV-1 isolate containing alterations affecting the NF-kappa B element. Virology **181:**150–157.
11. **Feinberg, M. B., D. Baltimore, and A. D. Frankel.** 1991. The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation. Proc. Natl. Acad. Sci. USA **88:**4045–4049.
12. **Feng, S., and E. C. Holland.** 1988. HIV-1 tat trans-activation requires the loop sequence within TAR. Nature (London) **334:** 165–167.
13. **Folks, T. M., J. Justement, A. Kinter, S. Schnittman, J. Orenstein, G. Poli, and A. S. Fauci.** 1988. Characterization of a promonocyte clone chronically infected with HIV and inducible by 13-phorbol-12-myristate acetate. J. Immunol. **140:**1117–1122.
14. **Franza, B. J., F. E. Rauscher, S. F. Josephs, and T. Curran.** 1988. The Fos complex and Fos-related antigens recognize sequence elements that contain AP-1 binding sites. Science **239:**1150–1153.
15. **Garcia, J. A., F. K. Wu, R. Mitsuyasu, and R. B. Gaynor.** 1987. Interactions of cellular proteins involved in the transcriptional

regulation of the human immunodeficiency virus. EMBO J. **6**:3761–3770.
16. **Gaynor, R.** 1992. Cellular transcription factors involved in the regulation of HIV-1 gene expression. AIDS **6**:347–363.
17. **Giacca, M., M. I. Gutierrez, S. Menzo, F. F. Di, and A. Falaschi.** 1992. A human binding site for transcription factor USF/MLTF mimics the negative regulatory element of human immunodeficiency virus type 1. Virology **186**:133–147.
18. **Golub, E. I., G. G. Li, and D. J. Volsky.** 1990. Differences in the basal activity of the long terminal repeat determine different replicative capacities of two closely related human immunodeficiency virus type 1 isolates. J. Virol. **64**:3654–3660.
19. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2**:1044–1051.
20. **Greene, W. C.** 1990. Regulation of HIV-1 gene expression. Annu. Rev. Immunol. **8**:453–475.
21. **Guatelli, J. C., T. R. Gingeras, and D. D. Richman.** 1990. Alternative splice acceptor utilization during human immunodeficiency virus type 1 infection of cultured cells. J. Virol. **64**:4093–4098.
22. **Hahn, B. H., G. M. Shaw, S. K. Arya, M. Popovic, R. C. Gallo, and F. Wong-Staal.** 1984. Molecular cloning and characterization of the HTLV-III virus associated with AIDS. Nature (London) **312**:166–169.
23. **Harrich, D., J. Garcia, F. Wu, R. Mitsuyasu, J. Gonzalez, and R. Gaynor.** 1989. Role of SP1-binding domains in in vivo transcriptional regulation of the human immunodeficiency virus type 1 long terminal repeat. J. Virol. **63**:2585–2591.
24. **Jones, K. A., J. T. Kadonaga, P. A. Luciw, and R. Tjian.** 1986. Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1. Science **232**:755–759.
25. **Kao, S. Y., A. F. Calman, P. A. Luciw, and B. M. Peterlin.** 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature (London) **330**:489–493.
26. **Koken, S. E., J. L. B. Van Wamel, J. Goudsmit, B. Berkhout, and J. L. Geelen.** 1992. Natural variants of the HIV-1 long terminal repeat: analysis of promoters with duplicated DNA regulatory motifs. Virology **191**:968–972.
27. **Laspia, M. F., A. P. Rice, and M. B. Mathews.** 1989. HIV-1 Tat protein increases transcriptional initiation and stabilizes elongation. Cell **59**:283–292.
28. **Lever, A., H. Gottlinger, W. Haseltine, and J. Sodroski.** 1989. Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. J. Virol. **63**:4085–4087.
29. **Lu, X., T. M. Welsh, and B. M. Peterlin.** 1993. The human immunodeficiency virus type 1 long terminal repeat specifies two different transcription complexes, only one of which is regulated by Tat. J. Virol. **67**:1752–1760.
30. **Lu, Y., M. Stenzel, J. G. Sodroski, and W. A. Haseltine.** 1989. Effects of long terminal repeat mutations on human immunodeficiency virus type 1 replication. J. Virol. **63**:4115–4119.
31. **Luo, Y., and B. M. Peterlin.** 1993. Juxtaposition between activation and basic domains of human immunodeficiency virus type 1 Tat is required for optimal interactions between Tat and TAR. J. Virol. **67**:3441–3445.
32. **Malim, M. H., J. Hauber, R. Fenrick, and B. R. Cullen.** 1988. Immunodeficiency virus rev trans-activator modulates the expression of the viral regulatory genes. Nature (London) **335**:181–183.
33. **Meyerhans, A., R. Cheynier, J. Albert, M. Seth, S. Kwok, J. Sninsky, M. L. Morfeldt, B. Asjo, and S. Wain-Hobson.** 1989. Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations. Cell **58**:901–910.
34. **Michael, N. L., G. C. Chang, P. K. Ehrenberg, M. T. Vahey, and R. R. Redfield.** 1993. HIV-1 proviral genotypes from the peripheral blood mononuclear cells of an infected patient are differentially represented in expressed sequences. J. Acquired Immune Defic. Syndr. **6**:1073–1085.
35. **Michael, N. L., M. Vahey, D. S. Burke, and R. R. Redfield.** 1992. Viral DNA and mRNA expression correlate with the stage of human immunodeficiency virus (HIV) type 1 infection in humans: evidence for viral replication in all stages of HIV disease. J. Virol. **66**:310–316.
36. **Michael, N. L., M. Vahey, L. d'Arcy, P. K. Ehrenberg, and R. R. Redfield.** 1994. Negative-strand RNA transcripts are produced in human immunodeficiency virus type 1-infected cells and patients by a novel promoter downregulated by Tat. J. Virol. **68**:979–987.
37. **Muesing, M. A., D. H. Smith, C. D. Cabradilla, C. V. Benton, L. A. Lasky, and D. J. Capon.** 1985. Nucleic acid structure and expression of the human AIDS/lymphadenopathy retrovirus. Nature (London) **313**:450–458.
38. **Muesing, M. A., D. H. Smith, and D. J. Capon.** 1987. Regulation of mRNA accumulation by a human immunodeficiency virus trans-activator protein. Cell **48**:691–701.
39. **Myers, G., B. Korber, S. Wain-Hobson, R. F. Smith, and G. N. Pavlakis.** 1993. Human retroviruses and AIDS 1993. Los Alamos National Laboratory, Los Alamos, N.Mex.
40. **Nabel, G., and D. Baltimore.** 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature (London) **326**:711–713. (Erratum, **344**:178, 1990.)
41. **Piatak, M. J., M. S. Saag, L. C. Yang, S. J. Clark, J. C. Kappes, K. C. Luk, B. H. Hahn, G. M. Shaw, and J. D. Lifson.** 1993. High levels of HIV-1 in plasma during all stages of infection determined by competitive PCR. Science **259**:1749–1754.
42. **Pomerantz, R. J., M. B. Feinberg, R. Andino, and D. Baltimore.** 1991. The long terminal repeat is not a major determinant of the cellular tropism of human immunodeficiency virus type 1. J. Virol. **65**:1041–1045.
43. **Purcell, D. F. J., and M. A. Martin.** 1993. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. J. Virol. **67**:6365–6378.
44. **Ratner, L., W. Haseltine, R. Patarca, K. J. Livak, B. Starcich, S. F. Josephs, E. R. Doran, J. A. Rafalski, E. A. Whitehorn, K. Baumeister, L. Ivanoff, S. R. Petteway, M. L. Pearson, J. A. Lautenberger, T. S. Papas, J. Ghrayeb, N. T. Change, R. C. Gallo, and F. Wong-Staal.** 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature (London) **313**:277–284.
45. **Redfield, R. R., D. L. Birx, N. Ketter, E. Tramont, V. Polonis, C. Davis, J. F. Brundage, G. Smith, S. Johnson, A. Fowler, T. Wierzba, A. Shafferman, F. Volvowitz, C. Oster, D. S. Burke, and The Military Medical Consortium for Applied Retroviral Research.** 1991. A phase I evaluation of the safety and immunogenicity of vaccination with recombinant gp160 in patients with early human immunodeficiency virus infection. N. Engl. J. Med. **324**:1677–1684.
46. **Rosen, C. A., J. G. Sodroski, and W. A. Haseltine.** 1985. The location of cis-acting regulatory sequences in the human T cell lymphotropic virus type III (HTLV-III/LAV) long terminal repeat. Cell **41**:813–823.
47. **Schnittman, S. M., J. J. Greenhouse, H. C. Lane, P. F. Pierce, and A. S. Fauci.** 1991. Frequent detection of HIV-1-specific mRNAs in infected individuals suggests ongoing active viral expression in all stages of disease. AIDS Res. Hum. Retroviruses **7**:361–367.
48. **Schnittman, S. M., M. C. Psallidopoulos, H. C. Lane, L. Thompson, M. Baseler, F. Massari, C. H. Fox, N. P. Salzman, and A. S. Fauci.** 1989. The reservoir for HIV-1 in human peripheral blood is a T cell that maintains expression of CD4. Science **245**:305–308.
49. **Schwartz, S., B. K. Felber, D. M. Benko, E. M. Fenyo, and G. N. Pavlakis.** 1990. Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1. J. Virol. **64**:2519–2529.
50. **Selby, M. J., E. S. Bain, P. A. Luciw, and B. M. Peterlin.** 1989. Structure, sequence, and position of the stem-loop in tar determine transcriptional elongation by tat through the HIV-1 long terminal repeat. Genes Dev. **3**:547–558.
51. **Siekevitz, M., S. F. Josephs, M. Dukovich, N. Peffer, F. Wong-Staal, and W. C. Greene.** 1987. Activation of the HIV-1 LTR by T cell mitogens and the trans-activator protein of HTLV-1. Science **238**:1575–1578.
52. **Sodroski, J., R. Patarca, C. Rosen, F. Wong-Staal, and W. Haseltine.** 1985. Location of the trans-activating region on the genome of human T-cell lymphotropic virus type III. Science **229**:74–77.
53. **Waterman, M. L., W. H. Fischer, and K. A. Jones.** 1991. A thymus-specific member of the HMG protein family regulates the human T cell receptor C alpha enhancer. Genes Dev. **5**:656–669.