Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1982), p. 121.
18. J. F. Lechner, A. Haugen, I. A. McClendon, E. W. Pettis, *In Vitro* **18**, 633 (1982).
19. A. E. Pegg, L. Wiest, R. S. Foote, S. Mitra, W. Perry, *J. Biol. Chem.* **258**, 2327 (1983).
20. Cytotoxicity of NMU, formaldehyde, or a combination of the two was measured by plating 150 to 3000 human fibroblasts per 35-mm dish, allowing 12 hours for attachment, and introducing the chemicals. The cells were refed with fresh culture medium weekly until clones developed to macroscopic size, at which time they were fixed, stained, and counted. Duplicates of four dishes were used per experimental datum point. The cloning efficiencies for the solvent-treated controls in four separate experiments ranged from 33 to 47 percent. The frequencies of 6-thioguanine–resistant mutations induced by formaldehyde or NMU or both were calculated from data obtained in three independent experiments. The mutation assays were conducted as described by L. L. Yang, V. M. Maher, and J. J. McCormick [*Mutat. Res.* **94**, 435 (1982)]. Exponentially growing cells (1×10^6 to 5×10^6) were seeded into 150-cm^2 flasks at a density calculated to give 2.5×10^5 surviving cells per flask (a total of at least 10^6 surviving cells per treatment). A growth medium consisting of Eagle's minimum essential medium and Earle's salts supplemented with 10 percent fetal calf serum (Sterile Systems) and gentamicin was used throughout the experiments. The cells were exposed to the test compounds within 12 hours of seeding. Just before exposure of the cells, the growth medium was replaced with serum-free medium. NMU (dissolved in dimethyl sulfoxide) or formaldehyde (diluted in sterile distilled water) or the combination was introduced by micropipette (final concentration of dimethyl sulfoxide, <0.5 percent). For experiments with NMU alone, the carcinogen-containing medium was removed after 1 hour and fresh serum-free medium was added for an additional 4 hours. For experiments with the combination of NMU and formaldehyde, both chemicals were present for 1 hour before the treatment medium was replaced with fresh serum-free medium containing various concentrations of formaldehyde. Incubation continued for another 4 hours, at the end of which the formaldehyde-containing medium was replaced with fresh growth medium. For experiments with formaldehyde alone, the procedure was identical except that no NMU was present during the first hour of exposure. The surviving cells were allowed to undergo six to eight population doublings (7 to 9 days after treatment) in order to fully express the mutant phenotype. Population doublings were estimated from electronic counts of cells that had been seeded into small flasks at identical densities and given identical treatments to those described above. The cells were kept in rapid growth during the expression period by replating into new flasks as necessary. For selection of mutants, cells from three or more flasks were pooled and 1×10^6 to 2×10^6 cells were plated for selection in 40 μM 6-thioguanine–containing medium at a density of 500 cells per square centimeter in 100-mm dishes. At the same time cells were seeded at clonal densities (150 cells per 35-mm dish) in growth medium to determine the replating efficiency. Incubation under the selective or nonselective conditions was continued for 14 days with one refeeding. Qualitatively identical results were seen with shorter (4- to 6-day) mutation expression periods. Reconstruction experiments with Lesch-Nyhan cells deficient in hypoxanthine-guanine phosphoribosyltransferase (HGPRT) were conducted for each treatment condition at each selection day to determine the efficiency of recovery of mutant cells in the presence of cells positive for HGPRT. The efficiency of recovery in these experiments was 60 to 100 percent. Mutation frequency was calculated by determining the frequency of mutant clones per number of cells seeded and correcting for cloning efficiency and efficiency of recovery. The spontaneous mutation frequency did not exceed 10×10^{-6} mutants per survivor for the mutagenesis experiments.
21. We appreciate the technical aid of W. Pettis and the secretarial assistance of N. Paige. Supported in part by Swedish Cancer Society grant 1623-B-84-03R to R.C.G.
* Present address: Department of Toxicology, Karolinska Institutet, S-10401 Stockholm, Sweden.
† To whom correspondence and requests for reprints should be addressed.

17 May 1984; accepted 17 December 1984

Thermosensitivity of a DNA Recognition Site: Activity of a Truncated *nut*L Antiterminator of Coliphage Lambda

Abstract. *Antitermination is an important transcriptional control. In bacteriophage lambda, the presence of the* nut *antiterminators between the promoters and terminators results in relatively unhindered transcription when the lambda* N *gene product and necessary host factors are supplied. This antitermination system has been rendered thermosensitive by modification of the* nut *site. A fragment of* λ *DNA [74 base pairs (bp) in length] that contained the 17-bp* nut*L core sequence, but lacked the 8-bp* box*A sequence, was cloned in a* p_p-N-t_{L1}-galK *plasmid between the* p_p *promoter and gene* N. *This fragment mediated antitermination of transcription at 30°C, as measured by assaying* galK *gene expression in* Escherichia coli. *At 42°C, however, antitermination at the* λ t_{L1} *terminator was abolished. Antitermination at 42°C was restored by replacing the 74-bp* nut*L fragment with longer sequences containing both* nut*L and* box*A or by cloning a synthetic* box*A sequence ahead of the 74-bp* nut*L fragment. Thus, efficient antitermination required both* box*A and the 17-bp* nut*L core, with the latter becoming conditionally defective when the* box*A sequence was deleted.*

Termination and antitermination of transcription are examples of negative and positive regulatory mechanisms that control gene expression. Terminator signals block the progress of transcription, while special recognition sites, such as *nut* (*1–5*), mark those operons that are subject to antitermination. In bacteriophage λ the *nut* sequences, which are located downstream of the promoters p_L and p_R and upstream of a series of terminators in the phage genome, are the recognition sites for the *N*-dependent transcriptional antitermination function. The *nut*L site (Fig. 1A) (*1*) functions as an antitermination site not only in λ but

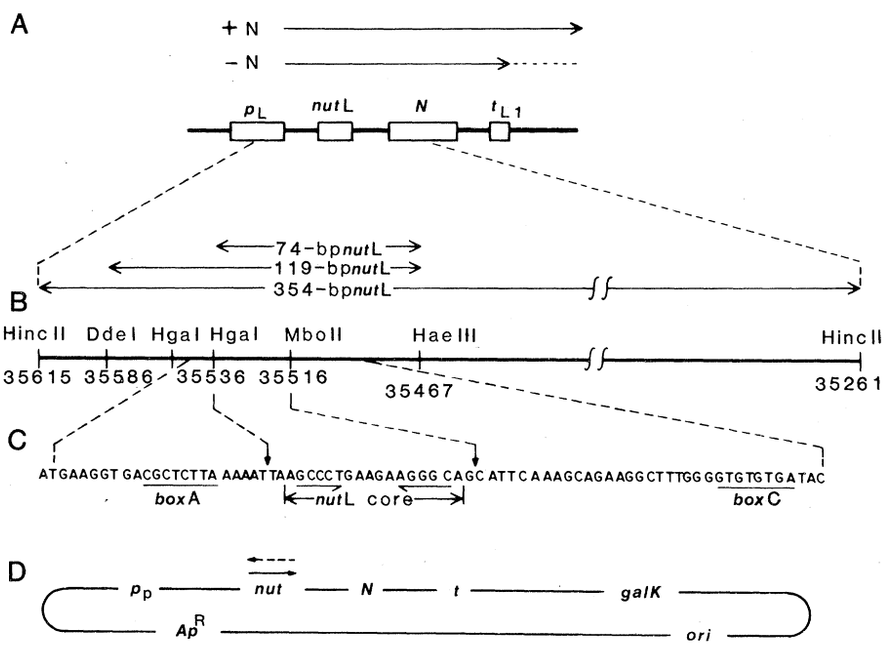

Fig. 1. Genetic and physical maps of the control regions of the leftward phage λ operon. The orientation of the map is opposite to that of the conventional λ map (*9*). (A) Diagrammatic representation of the p_L promoter, *nut*L antiterminator, gene *N*, and t_{L1} terminator in phage λ, with transcription originating at p_L and terminating at t_{L1} (lower arrow; $-N$). In the presence of gene *N* product, in conjunction with the *nut*L antiterminator, transcription proceeds across t_{L1} (upper arrow; $+N$). (B) Restriction map of the *nut*L region of λ. The locations of three λ DNA fragments are shown: (i) the 74-bp Hga I–Hae III fragment (*2*) containing the 17-bp *nut*L core (*10*) flanked by 7 bp upstream and 50 bp downstream, including the *box*C sequence (*4*); (ii) the 119-bp Dde I–Hae III fragment, containing the 17-bp *nut*L core flanked by 57 bp upstream [including the *box*A sequence (*4*)] and 50 bp downstream, including *box*C; (iii) the 354-bp Hinc II–Hinc II fragment. The coordinates of the restriction sites are specified (*9*). (C) Nucleotide sequence of the *nut*L region, including the 8-bp *box*A, the 17-bp *nut*L core, and the 8-bp *box*C. Nucleotide numbers (35536 and 35515) refer to the bases to the left of the restriction cut (*9*). (D) Schematic representation of plasmids pDX1, pDE3 (*2*), and p74nutL-1 derived from pKO3 (*11*), which carry the 74-bp *nut*L sequence. The *nut*L site is shown in its λ-like orientation (solid arrow) and in the reverse orientation (dashed arrow).

Table 1. Antitermination activity of cloned *nut* sites at 42°C with *N*-gene function supplied in *trans*. All inserts have the proper orientation with regard to the p_p promoter. The plasmids pD509 and pD571 have been previously described (*2*). The other plasmids were constructed as described (*12*). All plasmids were transformed into *Escherichia coli* lysogens N4830*galK*$^-$(λΔ-Bam N^+*c*Its857ΔH1) and N5260*galK*$^-$(λΔBam *N7N53c*Its857ΔH1), which were obtained as *galK*$^+$ [or *galE*$^-$(IS)] from S. Adhya (*5*) and converted to *galK*$^-$ by transduction with phage P1 grown on *galK*$^-$ cells and selection for the Gal$^-$ phenotype. Galactokinase was assayed 3 hours after transfer to 42°C (or at 30°C in controls) as described (*3, 11*).

Recombinant plasmid		Galactokinase units			
Code	Insert	N^+		N^-	
		30°C*	42°C	30°C	42°C
pD509	p_p-*galK*	76	84	72	81
pD571	p_p-t_{L1}-*galK*	6	9	8	11
pSPT8	p_p-119bp*nut*L-t_{L1}-*galK*	10	84	9	11
pSPT5	p_p-74bp*nut*L-t_{L1}-*galK*	6	15	7	12
pSP60	p_p-*nut*R-t_{R1}-t_{L1}-*galK*	5	116	7	12
pSPT3	p_p-74bp*nut*L-*nut*R-t_{R1}-*galK*	7	86	5	10

*The *N* gene is not expressed at 30°C.

also when cloned in expression plasmids as a part of the p_p-*nut*L-*N*-t_{L1}-*galK* transcriptional unit (Fig. 1D) (*2*).

In the control plasmid pD509, the *galK* gene is expressed from the p_p promoter at 30° and 42°C both in the presence or absence of the *N* gene product (Table 1, pD509). The t_{L1} terminator blocks transcription under all the above conditions (Table 1, pD571). Insertion of the 119-bp fragment carrying the entire *nut*L sequence (Fig. 1, B and C) results in readthrough across the t_{L1} terminator, but only when the *N* gene product is supplied (Table 1, pSPT8, N^+, 42°C). Under N^+ conditions, pSPT8 yields as much galactokinase activity as pD509.

The cloned 74-bp *nut*L fragment (Fig. 1B) exhibited no antitermination activity when *N* gene product was supplied in *trans* at 42°C (Table 1, pSPT5), which was surprising as the same fragment had been shown earlier to mediate antitermination at 30°C in several other constructs (*2*). One interpretation of this discrepancy was that the activity of the 74-bp *nut*L fragment was thermosensitive. To test this possibility, we studied the antitermination activity of the 74-bp *nut*L fragment at various temperatures. For that purpose, the *N* gene had to be constitutively expressed, which was achieved by placing it between *nut*L and t_{L1} in the p_p-*nut*L-t_{L1}-*galK* plasmid (*1*).

Two of the resulting p_p-74bp*nut*L-*N*-t_{L1}-*galK* plasmids, pDX1 and pDE3 (*2*), showed antitermination at 30° to 34°C, but not at 42°C (Fig. 2). The thermosensitivity of antitermination in the pDE3 and pDX1 plasmids must have been associated with the 74-bp *nut*L fragment, since (i) the p_p and *galK* modules are not thermosensitive (Fig. 2, pD509), (ii) the product of the *N* gene stimulates transcription at all temperatures tested (Fig. 2, pNP109 as compared to pD509), (iii) t_{L1} is subject to antitermination at 42°C by *nut*R (Table 1, pSP60), and (iv) the 74-bp *nut*L sequence does not inhibit transcription at 42°C (Fig. 2, pSP5) and does not interfere with antitermination by *nut*R (*6*) in the p_p-74bp*nut*L-*nut*R-t_{R1}-*galK* plasmid pSPT3 (Table 1, pSP60 as compared to pSPT3).

The *nut*L and *nut*R sites show a high degree of sequence homology at three elements: the 8-bp *box*A, the 17-bp *nut* core (or *box*B), and the 8-bp *box*C (Fig. 1C) (*4*); the main difference between the cloned *nut*R (Table 1) and the 74-bp *nut*L is the absence of *box*A in the latter. We therefore decided to test the *box*A-containing 119- and 354-bp *nut*L sequences (Fig. 1B), which were previously cloned in an analogous plasmid system under control of the p_p or p_{lac} promoter and found to be highly efficient in antitermination (*7*). The 119-bp (Table 1, pSPT8) and 354-bp *nut*L fragments (*7*) functioned as antiterminators at 42°C; the 74-bp *nut*L was inactive under the same conditions (Table 1, pSPT5).

To confirm that the absence of *box*A confers thermosensitivity on *nut*L, we cloned a synthetic *box*A sequence in front of the 74-bp *nut*L sequence (*8*). All *box*A constructs were sequenced. The presence of the *box*A sequence, CGCTCTTA (C, cytosine; G, guanine; T, thymine, A, adenine), in plasmid pBA74nutL-1 resulted in thermoresistance of the antitermination activity (Fig. 2, compare pBA74nutL-1 and p74nutL-1 at 38° and 42°C), but the *galK* activity was not as great as that observed with pNP109, an N^+ p_p-*N*-*galK* plasmid that lacks the terminator (Fig. 2). However, the distance between *box*A and the 17-bp *nut*L core is longer in pBA74nutL-1 (18

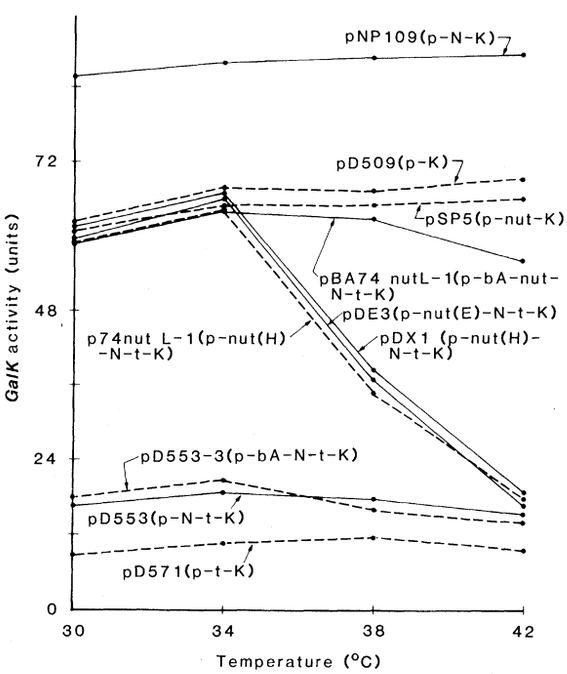

Fig. 2. Effect of temperature on the antitermination function mediated by cloned *nut*L sequences. The control plasmids (pNP109, pD509, pSP5, pD553, pD571) and *nut*L-carrying plasmids (pDX1 and pDE3) have been described (*2*) and are specified here by their code numbers and module arrangement (in parentheses). Abbreviations are p, p_p promoter; N, λ gene *N*; K, *E. coli* gene *galK*; nut, 74-bp *nut*L sequence (Fig. 1) cloned in the Hind III site (H) or Eco RI site (E) of pD553 (*2*); and t, rho-dependent t_{L1} terminator (*2*). The plasmid pBA74nutL-1 is analogous to pDX1 but contains a synthetic *box*A (bA; see Fig. 1C); it was constructed as described (*8*). The plasmid p74nutL-1 was constructed by excising the Hind III (*nut*L) fragment from pBA74nutL-1 and cloning it into pD553, yielding a plasmid with structure and thermal sensitivity analogous to pDE3 and pDX1. The *Escherichia coli* C600 *galK*$^-$ transformants carrying the specified plasmids were grown at the indicated temperatures to an optical density at 650 nm of 0.6 to 0.8 in Luria broth containing ampicillin (30 μg/ml), and then subjected to galactokinase assay (*2, 11*), which measured the transcriptional readthrough from the p_p promoter into the *galK* gene. The *N* gene carried by the plasmids shown here was expressed; this was measured by plaque formation by λ*N7N53* mutants on plasmid-carrying su° host strains, as previously described (*2*).

bp) (*8*) than the analogous distance in bacteriophage λ (7 bp; Fig. 1). Antitermination activity at 38° to 42°C must require *box*A, since a point mutation in the cloned *box*A sequence restored thermosensitivity to the level observed for the 74-bp *nut*L-mediated antitermination. Specifically, we constructed plasmid pmBAnutL-1, which was identical to pBA74nutL-1, except for a mutated *box*A (CG*G*TCTTA; where the italicized G represents mutation from C to G). The temperature sensitivity of *galK* expression for pmBAnutL-1 was indistinguishable from that for the *box*A-less plasmid p74nutL-1 (Fig. 2). The *box*A module alone had no antitermination activity (Fig. 2, pD553-3).

These results show that *box*A must be a component of the fully functional *nut*L site, but a truncated (*box*A-less) *nut*L sequence still retains a partial antitermination function at 30° to 34°C. At higher temperatures, however, the 74-bp *nut*L is thermosensitive, unlike the complete *nut*L site contained in the 119- or 354-bp sequences (*7*). This thermosensitivity might reflect the partial loss of contact points between the truncated *nut*L site and the proteins participating in the antitermination complex. The present results show that not only is the entire λ genome constructed from modular units, but even an individual recognition site, *nut*, is composed of separate subelements. While the *box*A-less *nut*L sequence displays an imperfect antitermination function, the addition of the *box*A subunit results in improved efficiency and thermal stability of the antiterminator.

STUART W. PELTZ
ALEXANDRA L. BROWN
NOAMAN HASAN
ANNA J. PODHAJSKA
WACLAW SZYBALSKI

McArdle Laboratory for Cancer Research, University of Wisconsin, Madison 53706

References and Notes

1. J. S. Salstrom and W. Szybalski, *J. Mol. Biol.* **124**, 195 (1978).
2. D. Drahos and W. Szybalski, *Gene* **16**, 261 (1981).
3. K.-C. Luk and W. Szybalski, *ibid.* **20**, 127 (1982).
4. E. R. Olson, E. L. Flamm, D. I. Friedman, *Cell* **31**, 61 (1982).
5. M. E. Gottesman, S. Adhya, A. Das, *J. Mol. Biol.* **140**, 57 (1980).
6. Antitermination by *nut*R in the plasmid system at 42°C was shown first by B. de Crombrugghe, M. Mudry, R. DiLauro, M. Gottesman [*Cell* **18**, 1145 (1979)].
7. N. Hasan and W. Szybalski, in preparation.
8. The plasmid pBA74nutL-1 was constructed by (i) chemical synthesis of the *box*A sequence,

5'-AGCTCGCTCTTACAA
GCGAGAATGTTTCGA

(ii) insertion of *box*A into the Hind III site of pD553 creating pD553-3, (iii) excision of the *nut*L sequence from pDX1 by AluI restriction; (iv) cloning of this *nut*L fragment into the Hind III site of pD553-3 employing Hind III linkers; and (v) sequencing of the resulting pBA74nutL-1 plasmid. The nucleotide sequence of the *box*A-*nut*L core junction in pBA74nutL-1 is

5'-AGCTCGCTCTTACAAAGCTTGCTAAAATT
TCGAGCGAGAATGTTTCGAACGATTTTAA
AAGCCCTGAAGAAGGGCA-3'
TTCGGGACTTCTTCCCGT

(the *box*A and *nut*L core sequences are underlined; see Fig. 1).
9. D. L. Daniels *et al.*, in *Lambda II*, R. W. Hendrix, J. W. Roberts, F. W. Stahl, R. A. Weisberg, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1983).
10. D. Drahos, G. R. Galluppi, M. Caruthers, W. Szybalski, *Gene* **18**, 343 (1982).
11. K. McKenney *et al.*, in *Gene Amplification and Analysis*, J. G. Chirikjian and T. S. Papas, Eds. (Elsevier/North Holland, New York, 1981), vol. 2, pp. 383–415.
12. The plasmid pSPT8 was constructed in two steps: (i) converting the Eco RI site into a Bam HI site by cutting pD571 (*2*) with Eco RI, filling in the ends, and ligating in a Bam HI linker, and (ii) cloning into this Bam HI site the 119-bp *nut*L sequence (Fig. 1) excised with Bam HI from plasmid pNH455 (see below). The plasmid pSPT5 was constructed by ligating the Pst I–Hind III (p_R-74bp*nut*L) fragment of pDE3 (*2*) to the Hind III–Pst I (t_{LI}-*galK*) fragment of pD571 (*2*). The plasmid pSP60 was constructed by insertion in proper orientation of the Hae III–Hinc II (*nut*R-t_{RI}) fragment excised from the λ genome [compare pD129 in (*3*)] into the Eco RI site of pD571 (*2*), with the use of Eco RI linkers. The plasmid pSPT3 was constructed by ligating the Pst I–Hind III fragment of pDE3 (see pSPT5 construction above) with the Hind III–Pst I (*nut*R-t_{RI}-*galK*) fragment of pD129 (*3*). Construction of plasmids pNH206 and pNH455, which contain the 354- and 119-bp *nut*L sequences, respectively (Fig. 1), cloned in a pKO3 derivative (*10, 11*) between Bam HI linkers, will be described elsewhere (*7*).
13. The skillful help of C. Seeliger is greatly appreciated. Supported by the NIH Program Project grant (5-PO1-CA-23076) (W.S.), core grant 5-P30-CA-07175, and NIH training grants CA-09135 (S.W.P.) and CA-09075 (A.L.B. and N.H.)

6 September 1984; accepted 3 December 1984

Expression in *Escherichia coli* of Open Reading Frame Gene Segments of HTLV-III

Abstract. *Human T-cell lymphotropic virus type III (HTLV-III), the causative agent of the acquired immune deficiency syndrome (AIDS), was recently isolated and its genomic structure analyzed by DNA cloning methods. In the studies reported here a combined cloning and expression system was used to identify HTLV-III encoded peptides that react immunologically with antibodies in sera from AIDS patients. Cloned HTLV-III DNA was sheared into approximately 500-base-pair fragments and inserted into an "open reading frame" expression vector, pMR100. The inserted DNA was expressed in* Escherichia coli *transformants as a polypeptide fused to the λCI protein at its amino terminus and to β-galactosidase at its carboxyl terminus. Sera from AIDS patients containing antibodies to HTLV-III were then used to screen for immunoreactive fusion proteins. Twenty clones, each specifying a fusion protein strongly reactive with AIDS serum, were identified. DNA sequence analysis indicated that the HTLV-III fragments were derived from the open reading frame DNA segments corresponding to the* gag *and* pol *gene coding regions and also the large open reading frame region* (env-lor) *located near the 3' end of the viral genome.*

The human T-cell lymphotropic virus type III (HTLV-III) has been routinely isolated from patients with the acquired immune deficiency syndrome (AIDS) (*1, 2*). More than 100 isolates have been obtained (*3–5*) and, although genetic variants with different restriction enzyme maps have been observed, serum samples from more than 90 percent of patients with AIDS and pre-AIDS contain antibodies reactive with the prototype HTLV-III isolate, H9/HTLV-III-B2, as demonstrated by Western blot analyses and solid phase immunoassays (*6, 7*). These and other results (*8–10*) suggest that the antibodies to HTLV-III are highly cross-reactive with the different genetic variants and that it may be possible to develop diagnostic assays and, perhaps, a preventive vaccine.

The helper T-cell tropism (*11*), the size of some of the viral proteins (*7, 12*), and serological (*6, 13, 14*) and DNA hybridization studies (*15*) indicate that HTLV-III is related to the leukemia-causing HTLV-I and HTLV-II (*16*). This information and sequence data (*17*) suggest that the HTLV-III RNA genome is similar to those of other retroviruses and contains at least (i) a *gag* gene that encodes the internal structural (nucleocapsid or core) proteins, (ii) a *pol* gene that encodes the reverse transcriptase, and (iii) an *env* gene that encodes the envelope glycoproteins of the virion. In addition, viruses in the HTLV family contain a pX region, which in HTLV-III consists of a coding sequence with one large open reading frame, designated *lor*, located between the *env* gene and the 3' end of the genome (*18–20*). This *lor* is continuous in the same frame with the *env* and, therefore, is referred to as *env-lor*. It has been suggested that the *lor* is responsible for the unique pathogenic properties of the viruses, namely, the