Molecular and Cellular Probes (1996) **10,** 443–451

Identification of a human endogenous LTR-like sequence using HIV-1 LTR specific primers

Chiara Menin,[1,2]* Stefano Indraccolo,[1,2] Marco Montagna,[1] Barbara Corneo,[1] Laura Bonaldi,[1] Christine Leib-Mösch,[3,4] Luigi Chieco-Bianchi[1] and Emma D'Andrea[1]

[1]*Istituto di Oncologia,* [2]*IST-Biotechnology Section, University of Padova, via Gattamelata 64, 35128 Padova, Italy,* [3]*III Medizinishcen Klinikum Mannheim der Universität Heidelberg, Mannheim D-68305, Germany and* [4]*GSF-Forschungszentrum, Institut für Molekulare Virologie, D-85764 Oberschleißheim, Germany*

(Received 23 May 1996, Accepted 23 August 1996)

Using 'consensus' primers derived from the LTR region of 15 HIV-1 isolates, a fragment of 583 bp was amplified from human DNA. Even though specificity was confirmed by Southern blot analysis with a conserved LTR oligonucleotide probe, no significant homologies were detected to either retroviral regions or human or non-human published sequences. Nevertheless, when used as a probe, the 583-bp fragment identified a unique DNA sequence in the human genome on chromosome 1, and cross-reactive sequences in monkey, but not mouse, DNA. This novel, unique and conserved sequence of 583 bp was used to isolate a human HS-1 clone in which the structural property of a viral LTR could be identified. © 1996 Academic Press Limited

KEYWORDS: HIV-1, LTR consensus primers, unique LTR-like sequence, conserved human DNA.

INTRODUCTION

The human genome contains a complex variety of endogenous retroviral sequences (HERSs).[1,2] Genetic events, such as chromosome rearrangements, transposition, and amplification can affect the genomic organization of HERSs, thereby generating either full-length or truncated proviruses; such proviral sequences make up at least 1% of the human genome. As retrotransposons, these sequences might make important contributions in shaping the human genome.[3]

Most of these sequences were isolated based on their relationship to mammalian endogenous and exogenous retroviruses using several different approaches, the most common of which is low-stringency hybridization to regions of exogenous and endogenous retroviruses.[4–7] Other retroviral sequences were identified by analysis of the flanking regions of cellular genes,[8] and by polymerase chain reaction (PCR) amplification of conserved retroviral segments.[9–11]

The discovery of HIV-1-related human DNA sequences designated EHS-1 and EHS-2[12] suggests the presence of a family of HIV-1-related sequences in primate genomes. Since all the HERS families iden-

* Author to whom all correspondence should be addressed

The nucleotide sequence data reported in this paper (HS-1) have been submitted to GenBank and assigned accession number U29896.

 © 1996 Academic Press Limited

tified thus far conserved at least one long terminal repeat (LTR), the possible existence of human sequences with homology to the LTRs of HIV-1 was investigated using PCR amplification.

The regulation of HIV-1 gene expression is modulated by complex mechanisms that involve transcriptional as well as post-transcriptional events.[13] The genetic elements which control expression of the integrated provirus have been identified within the 5′ LTR, which begins with the U3 region, followed by R and U5. From a functional point of view, the LTR can be divided into three main domains: the basal promoter region, which encompasses the transcriptional start site and exerts a positive effect on transcription; the enhancer region, which increases the effect of the downstream promoter; and the negative regulatory element (NRE), upstream of the enhancer site, which contains several binding sites for factors activating or repressing viral transcription.[14,15]

The NRE region appears to be specific for the HIV-1 LTR. In particular, this domain contains the following components: (1) two sequences containing AP-1 binding sites for Fos and Fos-related antigens;[16] (2) two protein binding sites (A and B) recognized by human T-cell factors, one of which is a member of the steroid/thyroid hormone receptor superfamily;[17] (3) the binding site for NFAT-1, a nuclear factor involved in the activation of LTRs by mitogens;[18] and (4) the binding site for a nuclear factor termed MLTF/USF, which down-regulates transcription.[15]

Using a PCR approach specific for the NRE region, a human DNA sequence with an HIV-1 LTR-related organization was identified.

MATERIALS AND METHODS

DNA sources

High-molecular weight DNA was prepared by cell lysis, proteinase K digestion, extraction with phenol/chloroform, and precipitation with ethanol.[19] Genomic DNA was isolated from human peripheral blood lymphocytes and placental tissue from normal subjects and from the following cell lines: erythroleukaemic K562 cells; T-lymphoid Jurkat, HUT78 and JM cells; and monocytoid U937 cells. Mouse DNA was extracted from the tail of a C57BL mouse, monkey DNA was extracted from the brain of a Macaca Fascicularis (Instituto Superiore di Sanità, Rome, Italy).

Synthetic oligonucleotides

Forward primer NR 1 (5′- GACAAGATATCCTTGATCTGTG-3′) spans nucleotide positions –426 to −405, while reverse primer NR 2 (5′-TCCGGATGCATCTCTCGG-3′) spans nucleotide positions −158 to −141 in the LTR of HIV-1[14] and they are highly conserved between 15 HIV-1 prototypes. The oligonucleotide probe NR 3 (5′-GCAGAACTACACACCAGGGC-3′) is also well conserved among different HIV-1 isolates and spans nucleotide position −371 to −352 of the LTR of HIV-1.

PCR and hybridization conditions

One μg of DNA was amplified in 100 μl of reaction mixture containing 2·5 U of AmpliTaq DNA polymerase (Perkin-Elmer-Cetus, Norwalk, CT, USA), 0·2 mM of each dNTP, 30 pmols of each primer, 2 mM $MgCl_2$, 10 mM Tris-HCl (pH 8·3), and 50 mM KCl. PCR conditions were as follows: denaturation at 94°C for 3 min; 30 cycles of 95°C for 1 min 30 s, 56°C for 1 min 30 s and 72°C for 2 min; and a final extension at 72°C for 7 min. Forty-five μl of the PCR products were electrophoresed on 1·5% agarose gel with 1 × TAE (40 mM Tris-acetate, pH 8; 2 mM EDTA). The gel was then soaked in 0·5 M NaOH and 1·5 M NaCl for 30 min and neutralized for 30 min in a solution containing 1·5 M NaCl, 0·5 M Tris-HCl (pH 7·4), and 1 mM EDTA, prior to Southern blotting onto nylon membrane with 20 × SSC (3M NaCl, 0·3 M sodium-citrate). The membrane was hybridized to the NR 3 oligonucleotide probe in 30% formamide, 6 × SSC, 1% sodium-dodecylsulphate (SDS), 1 mg salmon-sperm DNA, at 42°C. The probe was labelled with [γ^{32}P]ATP using T4 polynucleotide kinase. The blot was washed twice in 6 × SSC containing 0·1% SDS for 15 min at room temperature, and then twice for 30 min at 55°C in 2 × SSC containing 0·1% SDS. The filter was then exposed to an X-ray film for 15–36 h.

For nucleotide sequencing, the size-selected amplification product was eluted from the agarose and ligated into plasmid pCR using the TA Cloning Kit (Invitrogen BV-Leek, The Netherlands), according to the manufacturer's instructions.

Genomic DNA analysis

Human, monkey and mouse DNAs (10 μg) were digested with different restriction enzymes, fractionated by 0·8% agarose gel electrophoresis, and transferred to nylon membrane according to Southern. Filters were hybridized to the 583-bp human amplified product, under standard high-stringency conditions. The amplified product probe was isolated from plasmid pCR by *Eco*RI digestion and radioactively labelled

with [32P]dCTP by the random-priming method (Amersham International, England). Blots were washed twice in a solution of 6 × SSC, 0·1% SDS for 30 min at room temperature and twice in 0·1 × SSC, 0·1% SDS for 1 h at 65°C, and then exposed to X-ray film.

Chromosome 1 mapping was carried out using human × rodent somatic cell hybrid DNA (mapping panel No. 2, NIGMS Human Gentic Mutant Cell Repository, Camden, NJ, USA).

Preparation and screening of a human genomic DNA library

DNA was extracted from normal human placenta, partially digested with *Sau*3A and ligated to *Bam*HI-digested EMBL 3 phage arms using T4 DNA ligase, as described by Sambrook.[19] Ligated DNA was packaged using the Packagene System (Promega, Madison, WI) and recombinant phages were plated on *Escherichia coli* K802. Positive clones were identified by hybridization with the 583-bp human amplified product, under the same high-stringency conditions described for Southern blot analysis of genomic DNAs.

Sequence analysis

All sequencing reactions were performed using the Fmol Sequencing System Kit (Promega, Madison, WI). The 583-bp amplified product was sequenced both directly and after cloning in plasmid pCR. M13 forward and reverse primers were used for sequencing of pCR. Sequence comparisons were made using the PC-GENE software package (IntelliGenetics).

RESULTS

Identification of a human sequence by PCR with HIV-1 LTR specific primers

The PCR technique was used to identify human DNA sequences with homology to HIV-1. This approach was chosen on the basis of its past use in detecting short HERSs with homology to exogenous retroviruses.[9–11] Primers NR 1 and NR 2 were chosen from a consensus sequence among LTRs of 15 HIV-1 isolates (Los Alamos). These primers comprise an HIV-1 region of 286 bp (from position −426 to position −141) that contains the negative regulatory element (NRE). In addition, the NR 3 oligonucleotide probe used to assess specificity of the amplified product is well conserved among the 15 HIV-1 isolates. Using high-stringency PCR conditions, NR 1 and NR 2 specifically amplified a segment of 583 bp from genomic DNA isolated from all human cell lines tested, which included K562 (erythroleukaemic); Jurkat, HUT78, and JM (T-lymphoid); and U937 (monocytoid). The same PCR product was amplified from normal human placenta and peripheral blood lymphocytes (PBL) (Fig. 1a). Moreover, Southern blot analysis demonstrated that, under high-stringency conditions, the human 583-bp fragment was recognized by the LTR-specific probe NR 3 (Fig. 1b). Sequence comparison of the amplified product with the NR 3 oligonucleotide probe revealed two regions showing 60 and 55% homology. Computer-assisted comparison of the human 583-bp sequence to GenBank and EMBL databases did not detect significant homologies to retroviral regions nor to other human and non-human sequences.

This novel human DNA sequence appeared to be present as a single copy per human genome, as shown by Southern blot hybridization (Fig. 2). Indeed, using high-stringency conditions, the 583-bp hybridized with unique sequences in human and monkey DNA, but not with sequences in mouse DNA. The different fragment sizes of monkey and human DNA digested by *Eco*RI, *Pvu*II and *Msp*I, and the lower intensity of the monkey bands, likely reflect partial sequence homology of the two primate genomic regions. Thus, the 583-bp fragment appeared to be a unique, well-conserved sequence in both primate species suggesting that it might be endowed with a biological function.

Identification of flanking regions of the amplified product

To characterize the structure of the regions flanking the 583-bp fragment, recombinant lambda clones were isolated by screening a human genomic DNA library under high-stringency conditions using the amplified 583-bp product as a probe. Two positive clones (lambda HS-1 and lambda HS-2) were sequenced in both directions and shown to contain the 583-bp fragment; both clones disclosed the same sequence with few mismatches (5 out of 2400 sequenced bases: position 505, A→G; 570, A→C; 988, G→A; 1229, C→T; 1689, A→C) (Fig. 3). To confirm that the lambda HS-1 clone contained the genomic flanking regions of the 583-bp fragment, we isolated a 1·6 kb *Pst*I fragment, which hybridized to the same restriction fragments detected by the 583-bp probe in Figure 2, and mapped to human chromosome 1 using a human × rodent somatic cell hybrid mapping panel (data not shown).

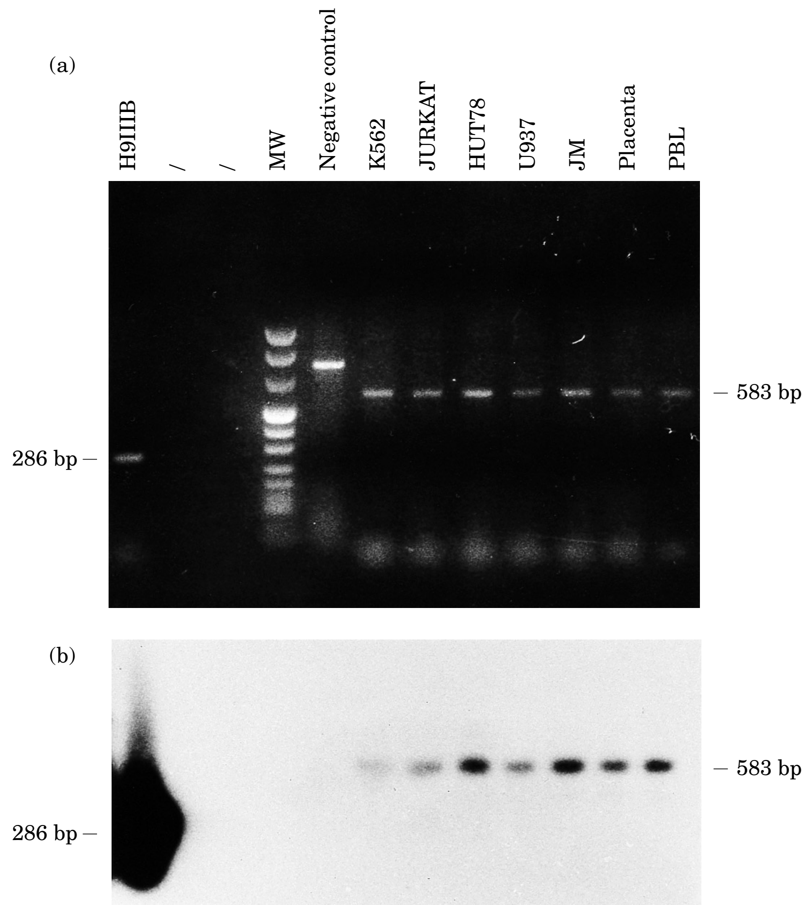

Fig. 1. Amplification of a human DNA fragment using HIV-1 LTR consensus primers. (a) Primers NR 1 and NR 2 amplified a 286-bp fragment in H9 IIIB (HIV-1 infected) cell DNA (positive control); the same primers amplify a 583-bp fragment from human DNA of cell lines K562, Jurkat, HUT78, U937, and JM and from placenta and peripheral blood lymphocytes (PBL). H9 IIIB cell DNA amplified with primers specific for the HIV-1 envelope region was used as a control (negative control) for any cross-hybridization to the LTR-specific probe. Fragments were separated by electrophoresis on agarose gel and stained with EtrBr. MW = molecular weight DNA; (b) Southern blot of the same experiment demonstrating that an oligonucleotide probe specific for HIV-1 LTR recognizes the 286-bp and 583-bp PCR products detected in HIV-infected and non-infected cells, respectively.

Sequence analysis

Sequence analysis of lambda HS-1 identified two regions, up- and downstream of the 583-bp amplified product, with 70 and 83% homology to NR 1 and NR 2 primers, respectively (Fig. 3). Examination of the sequence disclosed several characteristic features of LTR sequences found in mammalian retroviruses[20,21] (Fig. 4). These include a 5-base inverted repeat bordering the LTR (TGCTG and CAGCA) and a putative TATA box (TATATA) followed by a transcriptional start site which begins with GC, and defines the R region containing a putative polyadenylation signal AGAAAA.

Like the U3 region of HIV-1 LTR (Fig. 4a), the upstream region of lambda HS-1 containing the TATA box disclosed some genetic elements which could control expression of adjacent genes (Fig. 4b). In particular, at 10 bases downstream from the 5′ inverted repeat, there was a 7-bp region with 71% homology to the binding site for the inducible transcription factor AP-1 found in U3 of HIV-1[16] (Table 1). This AP-1 like sequence in lambda HS-1 is followed by a sequence of 18 bp showing 72% homology to the protein binding site A of HIV-1. This sequence has been reported to interact with the protein binding site B, which is related to steroid/thyroid hormone response elements, and has a negative effect on HIV-1 LTR-driven gene expression.[17] HIV-1 protein binding site B contains another binding site for the inducible transcription factor AP-1. Furthermore, this second AP-1 site of HIV-1 presents a 86% homology to the 7-bp region of lambda HS-1. Between these two AP-1-related sequences, there are two other putative protein binding sites with a good homology (86 and 92%) to the corresponding sites in the U3-LTR of the

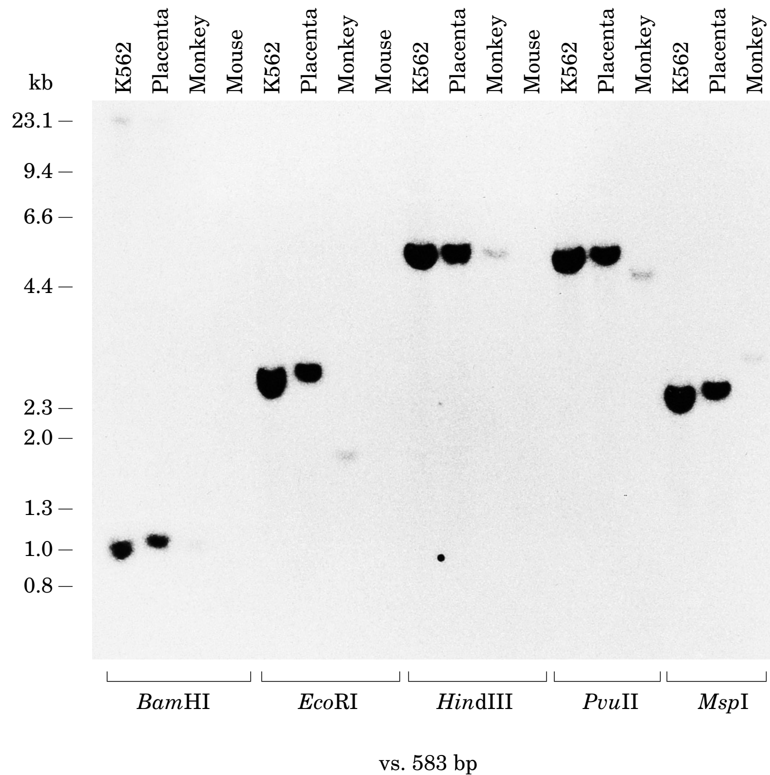

Fig. 2. Southern blot analysis of human, monkey, and mouse DNA using the human 583-bp amplified product as probe. DNA from the specified sources was digested with the indicated restriction enzymes, separated by agarose gel electrophoresis, blotted, and hybridized to the P^{32}-labelled 583-bp fragment. Single bands corresponding to single copy genes could be detected in human and monkey DNA, but not in mouse DNA.

HIV-1 ELI isolate. One of these sites is homologous to the target element for nuclear factor USF/MLTF, and the second element is related to the binding site for the inducible transcription factor NF-kB, which increases the rate of viral gene transcription in response to mitogenic signals.[22]

Therefore, based on sequence analysis, a region (from nucleotide 689 to nucleotide 1073) could be identified in lambda HS-1 with the genomic organization of the HIV U3-LTR region. In addition to a U3-like region, lambda HS-1 also contained a sequence from nucleotide 1074 to nucleotide 1324 resembling elements specific to the R region of an LTR, which could be involved as a transcription start site and as a polyadenylation signal. Furthermore, the region from nucleotide 1325 to nucleotide 1509 of lambda HS-1 could represent a U5-like sequence. However, lambda HS-1 did not appear to contain an open reading frame; indeed, no transcripts were detected in Northern blot experiments using a 5-kb *Hin*dIII fragment of HS-1 clone, including the entire LTR-like fragment, as hybridization probe on RNA from human placenta, freshly isolated PBMC, and different hematopoietic cell lines including the Jurkat, K562, Raji and U937 cell lines (data not shown).

DISCUSSION

The human genome carries multiple copies of sequences related to endogenous retroviruses, including LTRs and genes; in many cases, LTR-like sequences contain potential transcriptional regulatory sequences which may control the expression of linked cellular genes.[23,24] In this study we investigated whether the human genome contains sequences related to the HIV-1 LTR by means of PCR, which allows the extension of primer pairs annealed to non-target but homologous nucleic acid sequences (mis-priming).[25] HIV-1 LTR consensus primers amplified a human 583-bp fragment which enabled the identification of a new sequence, designated HS-1, from a human genomic DNA library. This sequence did not show significant homology to the HIV-1 LTR but did possess an LTR-like organization. In particular, the presence of human elements related to protein binding sites present in

```
                10          20          30          40          50          60          70
1     AAGTTTCCTGGATCCAGGTCTCACTCGCTGGGTTCCCTGCTCTTCTTCGCCATTGTAGGCTCCCTTTCCTCCCACCCTCC   80
81    CACCTGGCATCCTAGCTGGAACTCACAGTCCCAGAGCCCAACCAATTTGGCCAAGCCTCTTGGCTAAAAGAAGTAAAGCA  160
161   ACGAATGTGCACCACTGGGAGTACCTCCTAAGCAGCTCCAGCACCTTAATGGAGGATGTGGGGGACCTTCTTTATGATGC  240
241   TTTATCAAGCAGCCTTTTCTGTTTTGTTCTTTCATTGGGAACCTGGGTAGCAAGTCCAGCACCTGTGTCTTTCCTAACTC  320
321   TTACAGCCCCTCTGCCCTCGCTGCTGTAATTGTTAGTACCCATGGCCTGTCTCCCTGACTAGTCTGCAGGCTTCATGAGG  400
401   TCAAGTAAAGCCTTCATGAGGTCAAGTATCACATTCCAACAAGGGTCAGCGAAGCTTAACCCAGTGTCTGGTATGATAGA  480
481   TTCCATAGATTATAGAAATAAAATATCAAGAACAGCTAATGTTCATGTAACACATTTAAACTTGTTAAATTTACAAATCA  560
561   TTTTCATGTAAATTATTTCATTTATAGTTTCAAAAGCCCTGCGAGATTCAAAAAGATTACTTGCCCAAGATAATCCAATT  640
                                                               IR
641   AGCAAGAGGCAGAGACAAATTCAAACTAATGTCTTCTTCCAGGTTATGTGCTGTTCCACAATGCCTTAGCTGTCTTCCCT  720

721   GCTATCTAGAAGTCAACATTTAAAATCTACAAGAGGGATTTCATTTCAGCCCAAGGTTGAGGAAAATAAGGATATTAAAA  800
                                                                       70% NR1
801   CAGATGGGTTTTGTTTTTCTTTAAGGGACTATCCTTTTCATTGAGACAGTAgCATaAaAatccAttgCtctgtgTTGGGT  880

881   TATTAAAGACCTAGAATAAAATTCATAGCGTTAGTCAAGCTAGTGATATACTGCTTGATTTTGTGTCCATTTATTCATAC  960

961   ACATATTGCATTCTGGAGAGAACAGGAGCAAATAAGATTCAGCTTCATAAACCTTAGATTTAACATGCCCATCAGGCTCA 1040
          TATA-box                 60% NR3
1041  TGGGTATATATGTGGCcagaCctGcGAaAcaAAgcATACTCCTTTAACTGGTCCCTCTCTCCTCTGAGGATCTGAGTACC 1120

1121  CTGTCCTCTTCACAGAGGATCCTGGGGTGATTCGGATGATGATGCATATTACGGTTATTTAGTTTACCAGATGACTGCTG 1200
                          55% NR3
1201  AATTCCCCCACTGAAAGgCaGCTcaTaGcagggcACTATGAATATCTTAATTTATCCAGTTCCCTTGGAGCAAGGCATAG 1280
      poly-A
1281  AAAATGTAAAAGGGGGAATCTGGCCTATAGAAGAAGAGTGTTCATATTCCCTAAGGCCCCTTGGCAACAGCCTGTGGCCT 1360
                                                                          83% NR2
1361  AAGGTAAGAGGATAGGTAATTGGCTCCTTTTCATTTGTTCTATCCAGGGTTTCTACAccgagagatgcaAcAgAaATCAA 1440
                                                                            IR
1441  GAAGGGATGCATCTAGATTTCAGAATTCCAAACAAGGGGAACAAAAACAAATCCCCATGTGATGCAGCAGAAACACCACC 1520
1521  ATTCTATTTTCAGAGCTGGGGATCTAGGTATCCTCTGCTGCTTACAAACTGGACAGTCACCTTGCTGTCTCAATTTCTTC 1600
1601  ATCAGTAAGAGTTGACATTAGGACCCACATAAACCCTAGGCTCATAACCAGAGGCAGGGATAGATCTGGTTGACTCTGAG 1680
1681  AGTCCTGAATAGATTCCCATAAGAGATTGTACAGTGACACAAATCCAAGTCGTGTTTCATTCCCGATACTTCTCCTTACC 1760
1761  CAATTCTGTCATCCTTAATTTTAGGTTGCTCCTCCAGCCCCAAACCTGGCTCCTCTGAGGGAGGATAACCTGTAGGTCCC 1840
1841  CTCACATACGGCTCTTGCATACAGGGAAGTAAGGTGAATTGGAGTTAAGCATGGATGAGCCTCTACTCTTAGATACTACA 1920
1921  TTGATGCAAGGTTTGGGTTTCTAAATTTCAGTCATGGAAACTAAGAACAAGGTGGATCTCTTTGCCCTGCATGGTTGAGC 2000
2001  ACAGTTTTCTTTTATCCTGGTTACTACATTTGTATTTGTGACTGCTTTTTTTCTGCAGCAGTTTACATGAGAAATATAAA 2080
2081  TTGTAAAACTCAACTATTCATACACAGCTAATTCTATATAATTAATCATTAAGATGAAGTGGAAGAATTACATAATCTCA 2160
2161  TACAAGGTAGAATGCCCAAAACATTTATCATGTGTGCCTGGAGGGATGAGGGGAGAGAAGACAGAGACTTCTTTGAAATC 2240
2241  ATTTTAGGTTATTATTAAAATTAGCATTTTATGAATGTATGAACCTATAAGGAGCCACTGCTACAGACAAATCACAAAGT 2320
2321  GGCTAATTTCTGCCTTAATGAAGTAACTGGTTTTAGCCCATAGGCTGATACTAGGAGGGTCTTCGGTCATTCACATTTCT 2400
```

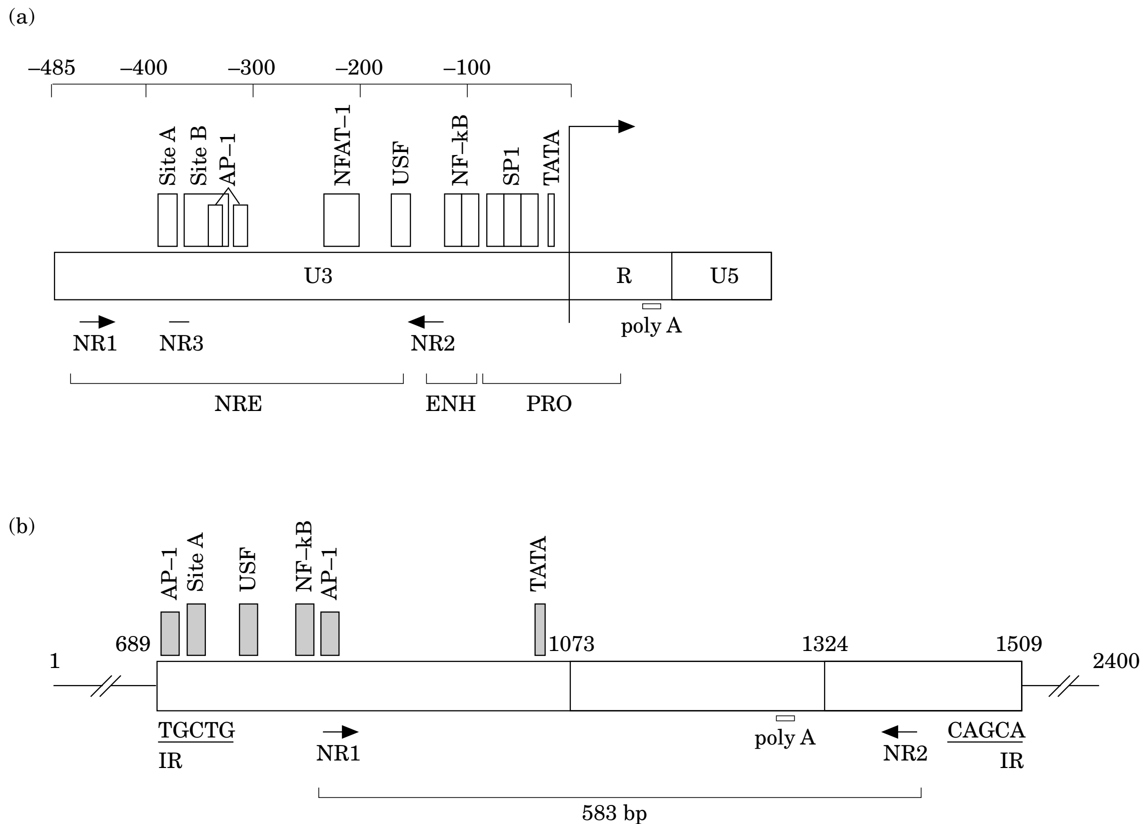

Fig. 4 Schematic representation of the functional features of HIV-1 LTR and HS-1 clone. (a) HIV-1 LTR with binding sites for: T-cell site A and B proteins; activator protein (AP) -1; nuclear factor of activated T cells (NFAT) -1; upstream stimulatory factor (USF); nuclear factor (NF) -kB; SP1; and TATA binding protein. Transcriptional start site is indicated by an arrow at the U3-R boundary. Upper negative numbers indicate nucleotide positions upstream of the transcription start site. The negative regulatory element (NRE), enhancer (ENH) and promoter (PRO) regions are indicated at the bottom. Primers NR 1 and NR 2 and the NR 3 probe are indicated. (b) HS-1 clone with HIV-1 homologous binding sites. The amplified 583-bp product is indicated. IR = inverted repeats.

the U3 segment of the HIV-1 LTR indicates that this sequence might regulate gene expression. Indeed, an endogenous LTR can affect neighbouring gene expression in several ways (i.e. by promoter insertion, enhancer or repressor insertion, or differential polyadenylation). In the case of salivary amylase genes, integration of an endogenous sequence (HERV-E element) upstream of the gene resulted in a parotid-specific enhancer effect.[26,27] Furthermore, transcripts from two other cellular genes have been found to be driven by endogenous retroviral promoters: H-plk, a Kruppel-related gene encoding a zinc-finger protein, is driven by an HERV-R promoter in placental tissue,[28] while PLA2L, a phospholipase A2-related gene expressed in teratocarcinoma cell lines, can be driven by a HERV-H promoter.[29] Finally, endogenous retroviral LTRs might also provide poly(A) signals for cellular genes, as shown in the case of PLT, a new gene which can be polyadenylated by a HERV-H LTR.[30]

On the other hand, a full proviral organization is not always required for the regulation of surrounding genes, as observed for the human cytochrome c1 gene, which carries regulatory sequences homologous only to the HERV-I LTR.[31] This might also be the case for HS-1, which is present as an isolated LTR-like sequence on human chromosome 1. As isolated LTRs are usually present as multicopies in the host genome,[32] the finding that HS-1 is a single copy sequence does not support its retroviral origin;

Fig. 3. Nucleotide sequence of the human HS-1 clone. The underlined regions indicate: the inverted repeats (IR) which defined the putative LTR organization; the TATA-box and poly-A sequences; and the regions with their percentage homology to the NR1 and NR2 primers, and NR3 probe. Lower-case letters indicate identical nucleotides between HS-1 and synthetic oligonucleotides.

Table 1. Sequence homology of the binding site present in U3-HIV-1 and HS-1 clone

	Sequence	Binding site	Homology
		AP-1	
U3-HIV-1	−336 CACTGAC −330		5/7
HS-1 clone	699 CAATGCC 705		71%
	* *		
		Site A protein	
U3-HIV-1	−383 CTTCCCTGATTAGCAGAA −366		13/18
HS-1 clone	714 CTTCCCTGCTATCTAGAA 731		72%
	* ****		
		USF	
U3-HIV-1	−174 CATTTCATCACGTGG −160		11/15[a]
HS-1 clone	762 CATTTCAGCCCAAGG 776		73%
	* * **		
		NF-kB	
U3-HIV-1	−106 AAGGGACTTTCC − 95		11/12
HS-1 clone	823 AAGGGACTATCC 834		92%
	*		
		AP-1	
U3-HIV-1	−348 GAGTCAG −342		6/7
HS-1 clone	843 GAGACAG 849		86%
	*		

[a] 13/15 (86%) between HS-1 clone and ELI isolate.
USF = upstream stimulatory factor, NF = nuclear factor.

hence, at present we favour the alternative explanation, that HS-1 is a novel cellular sequence with some homology to the HIV-1 LTR.

ACKNOWLEDGEMENTS

This work was supported in part by grants from the Italian Association for Research on Cancer (AIRC); National Research Council (target projects ACRO, BTBS, and FATMA); ISS, AIDS Project; MURST 40% and 60%. We are grateful to D. D'Agostino and A. De Rossi for helpful suggestions and critical comments and we thank M. Quaggio and D. Zullato for excellent technical assistance, P. Gallo for artwork, and P. Segato for help in preparing the manuscript.

REFERENCES

1. Leib-Mösch, C., Brack-Werner, R., Werner, T. *et al.* (1990). Endogenous retroviral elements in human DNA. *Cancer Research* **50**, 5636s–42s.
2. Wilkinson, D. A., Mager, D. L. & Long, J-A. C. (1994). Endogenous human retroviruses. In *The Retroviridae*. (Levy, J. A., ed.) Pp. 465–535. New York: Plenum Press.
3. Finnegan, D. J. (1989). Eukariotic transposable elements and genome evolution. *Trends in Genetics* **5**, 103–7.
4. Perl, A., Rosenblatt, J. D., Chen, I. S. Y. *et al.* (1989). Detection and cloning of new HTLV-related endogenous sequences in man. *Nucleic Acids Research* **17**, 6841–54.
5. Kröger, B. & Horak, I. (1987). Isolation of novel human retrovirus-related sequences by hybridization to synthetic oligonucleotides complementary to the $tRNA^{Pro}$ primer-binding site. *Journal of Virology* **61**, 2071–75.
6. Ono, M. (1986). Molecular cloning and long terminal repeat sequences of human endogenous retrovirus genes related to types A and B retrovirus genes. *Journal of Virology* **58**, 937–44.
7. Leib-Mösch, C., Brack, R., Werner, T., Erfle, V. & Hehlmann, R. (1986). Isolation of an SSAV-related endogenous sequence from DNA. *Virology* **155**, 666–77.
8. Mager, D. L. & Henthorn, P. S. (1984). Identification of a retrovirus-like repetitive element in human DNA. *Proceedings of the National Academy of Sciences, USA* **81**, 7510–4.
9. Shih, A., Misra, R. & Rush, M. G. (1989). Detection of multiple, novel reverse transcriptase coding sequences in human nucleic acids: relation to primate retroviruses. *Journal of Virology* **63**, 64–75.
10. Cordonnier, A., Casella, J. F. & Heidmann, T. (1995). Isolation of novel human endogenous retrovirus-like elements with foamy virus-related pol sequence. *Journal of Virology* **69**, 5890–7.
11. Indraccolo, S., Gunzburg, W. H., Leib-Mösch, C., Erfle, V. & Salmons, B. (1995). Identification of three human sequences with viral superantigen-specific primers. *Mammalian Genome* **6**, 339–44.
12. Horwitz, M. S., Boyce-Jacino, M. T. & Faras, A. J. (1992). Novel human endogenous sequences related to human immunodeficiency virus type 1. *Journal of Virology* **66**, 2170–9.
13. Garcia, J. A. & Gaynor, R. B. (1994). Regulatory mechanisms involved in the control of HIV-1 gene expression. *AIDS* **8** (Suppl. 1), S3–S17.
14. Gaynor, R. (1992). Cellular transcription factors involved in the regulation of HIV-1 gene expression. *AIDS* **6**, 347–63.
15. Giacca, M., Gutierrez, M. I., Menzo, S., D'Adda di Fagagna, F. & Falaschi, A. (1992). A human binding site for transcription factor USF/MLTF mimics the negative

regulatory element of human immunodeficiency virus type 1. *Virology* **186**, 133–47.
16. Franza, B. R., Rauscher, F. J., Josephs, S. F. & Curran, T. (1988). The fos complex and fos-related antigens recognize sequence elements that contain AP-1 binding sites. *Science* **239**, 1150–3.
17. Orchard, K., Perkins, N., Chapman, C. *et al.* (1990). A novel T-cell protein which recognizes a palindromic sequence in the negative regulatory element of the human immunodeficiency virus long terminal repeat. *Journal of Virology* **64**, 3234–9.
18. Shaw, J. P., Utz, P. J., Durand, D. B., Toole, J. J., Emmel, E. A. & Crabtree, G. R. (1988). Identification of a putative regulator of early T cell activation genes. *Science* **241**, 202–5.
19. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*. 2nd edn. New York: Cold Spring Harbor Laboratory Press.
20. Temin, H. M. (1981). Structure, variation and synthesis of retrovirus long terminal repeat. *Cell* **27**, 1–3.
21. Coffin, J. (1984). Endogenous viruses. In *RNA Tumor Viruses*. (Weiss, R., Teich, N., Varmus, H. & Coffin, J., eds) 2nd edn, Pp. 1109–1203. New York: Cold Spring Harbor Laboratory Press.
22. Nabel, G. & Baltimore, D. (1987). An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature* **326**, 711–3.
23. Feuchter, A. & Mager, D. (1990). Functional heterogeneity of a large family of human LTR-like promoters and enhancers. *Nucleic Acids Research* **18**, 1261–70.
24. Feuchter, A. E., Freeman, J. D. & Mager, D. L. (1992). Strategy for detecting cellular transcripts promoted by human endogenous long terminal repeats: identification of a novel gene (CDC4L) with homology to yeast CDC4. *Genomics* **13**, 1237–46.
25. Chou, Q., Russel, M., Birch, D. E., Raymond, J. & Bloch, W. (1992). Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplification. *Nucleic Acids Research* **20**, 1717–23.
26. Samuelson, L., Wiebauer, K., Snow, C. & Meisler, M. (1990). Retroviral and pseudogene insertion sites reveal the lineage of human salivary and pancreatic amylase genes from a single gene during primate evolution. *Molecular and Cellular Biology* **10**, 2513–20.
27. Ting, C-N., Rosenberg, M. P., Snow, C. M., Samuelson, L. C. & Meisler, M. H. (1992). Endogenous retroviral sequences are required for tissue specific expression of a human salivary amylase gene. *Genes & Development* **6**, 1457–65.
28. Kato, N., Shimotohno, K., VanLeeuwen, D. & Cohen, M. (1990). Human proviral mRNAs down regulated in choriocarcinoma encode a zinc finger protein related to Kruppel. *Molecular and Cellular Biology* **10**, 4401–5.
29. Feuchther-Murthy, A. E., Freeman, J. D. & Mager, D. L. (1993). Splicing of a human endogenous retrovirus to a novel phospholipase A2 related gene. *Nucleic Acids Research* **21**, 135–43.
30. Goodchild, N., Wilkinson, D. & Mager, D. (1992). A human endogenous long terminal repeat polyadenylates a novel alternatively spliced transcript in normal placenta. *Gene* **121**, 287–94.
31. Suzuki, H., Hosokawa, Y., Toda, H., Nishikimi, M. & Azawa, T. (1990). Common protein-binding sites in the 5′-flanking regions of human genes for cytochrome c1 and ubiquinone-binding protein. *Journal of Biological Chemistry* **265**, 8159–63.
32. Leib-Mösch, C., Haltmeier, M., Werner, T. *et al.* (1993). Genomic distribution and transcription of solitary HERV-K LTRs. *Genomics* **18**, 261–9.