Analysis of a novel defective HTLV-I provirus and detection of a new HTLV-I-induced cellular transcript

Satoshi Kubota**, Rika A. Furuta, Haruhiko Siomi***, Masatoshi Maki, Masakazu Hatanaka*

Laboratory of Human Tumer Viruses, Institute for Virus Research, Kyoto University, 53 Shogoin-Kawaharachou, Sakyo-ku, Kyoto 606-01, Japan

Received 29 August 1995; revised version received 26 September 1995

Abstract HTLV-I generally integrates at least one full-length copy in adult T-cell leukemia (ATL) cells. A group of patients without full-length provirus have a unique conserved truncation of the provirus which retains *env-pX*-3′LTR. Tumor cells of a patient from this group were genetically analyzed. Analysis of the 5′ and 3′ cellular flanking region adjacent to the provirus suggest that the defective provirus was integrated immediately downstream of a promoter of an unknown cellular gene. The activity of the promoter was weak but was responsive to Tax-like HTLV-I LTR. The provirus may have utilized it as a substitute for the 5′LTR and thus 3′LTR may have become an alternative promoter for the cellular gene, which may give similar viral–cellular interactions to that of general cases with full-length proviruses. Surprisingly, the 3′ cellular flanking region which is thought to be controlled originally by the promoter is constitutively expressed specifically in an HTLV-I producing ATL cell line HUT1O2G, in which the corresponding region is not modified by provirus. The detection of this HTLV-I-induced transcript provides a probe to find an HTLV-I inducible unknown cellular gene that may be related to the pathogenesis of ATL.

Key words: Adult T-cell leukemia; HTLV-I; Tax; Oncogenesis; Provirus

1. Introduction

Evidence suggests that human T-cell leukemia virus type I (HTLV-I) must be the etiological agent of adult T-cell leukemia (ATL) [1,2]. These observations are supported by recent molecular and cellular biological studies that revealed the involvement of the viral *trans*-activator, Tax protein of HTLV-I, in its oncogenic potential. Tax protein can immortalize human T-lymphocyte [3] and transform Rat-1 cells in vitro [4], and is required for the maintenance of the malignant phenotype of Tax-transformed Rat-1 cells [5]. The oncogenic properties of Tax are thought to be based on interactions of Tax with cellular factors, such as the NF-kB family [6], Ets1 [7] and $p67^{SRF}$ [6,8], which eventually cause the transactivation of various host genes including IL-2 [9], IL-2 receptor a chain [10,11], GM-CSF [12,13], MHC class I [14], c-*fos* oncogene [15,16] and so on [17,18,19]. In spite of this evidence, however, no one has succeeded to transform normal T-cells to malignancy via direct infection with HTLV-I particles so far. Therefore the pathogenic pathway from virus entry to ATL may include numerous complex interactions between cellular and viral factors, especially Tax.

While analyzing many cases of ATL, we found a group that contains a defective HTLV-I provirus (without any full-length provirus) composed of *env-pX*-LTR [20]. One of them was genetically analyzed and extraordinarily long direct repeats were identified at the integration sites as previously reported [21]. We have further characterized this isolate and obtained data suggesting the *env-pX*-LTR provirus may be able to cause typical ATL through a novel *cis*-interaction with an HTLV-I-induced cellular gene to simulate a full-length HTLV-I provirus in the cell.

2. Materials and methods

2.1. Molecular clones

The molecular clone pOHS13 which contains the defective *env-pX*-LTR provirus of HTLV-I with cellular flanking sequence at both ends was obtained as previously described from an ATL patient R [11,20,21]. The clinical findings of patient R are summarized in Table 1. The other clone named pBS(5 + 3)A includes the fragment which corresponds to the sequence around the integration point derived from the genomic DNA of MOLT-4. As previously characterized, there was no major rearrangement among various human cell lines including HUT1O2G in the subcloned area [21]. Nucleotide sequences was determined by an estublished chain-termination sequencing method.

2.2. Cells

COS7 cells (a monkey kidney cell line) were maintained in Dulbecco's modified Eagle minimum essential medium (D-MEM) supplemented with 10% fetal calf serum (FCS). An HTLV-I producing human T-cell line HUT1O2G [1], HTLV-I-free human lymphoid cell lines including MOLT-4 and Jurkat were grown in RPMI 1640 medium supplemented with 10% FCS.

2.3. Plasmids for CAT assay

Every plasmid for CAT assay possesses a part of the region around the 5′ viral–cellular junction of pOHS13 upstream of bacterial chloramphenicol acetyl-transferase (CAT) gene. The plasmid pF1CAT was constructed by inserting a *Pst*I-*Hin*dIII fragment of pOHS13 into the *Hin*dIII site of pSVOCAT [22]. Similarly, an *Eco*RI-*Hin*cII fragment was subcloned into pSVOCAT to be pF2CAT. In this construct, a *Hin*dIII site was conserved on the HincII-end of the insert. The *Pst*I-*Bam*HI fragment of pF2CAT which contains the common region of pF1CAT and pF2CAT from pOHS13 fused to CAT gene, SV40 t-splicer and polyadenylation signal was subcloned at the corresponding sites of pUc119. The resultant plasmid was designated as pUcF3CAT. To construct a deletion mutant pUcF4CAT, the polymerase chain-reaction (PCR) product from pUcF3CAT was substituted for the *Pst*I-

*Corresponding author. Fax: (81) (75) 751-3998.

** *Present address:* The Dorrance H. Hamilton Laboratories, Division of Infectious Diseases, Department of Medicine, Thomas Jefferson University, Philadelphia, PA 19107-6799, USA.

*** *Present address:* The Howard Hughes Medical Institute Research Laboratories, University of Pennsylvania, School of Medicine, Philadelphia, PA 19104-6148, USA.

SSDI 0014-5793(95)01166-8

*Hin*dIII fragment of pUcF3CAT. Primers were synthesized by an automated DNA synthesizer. Cloning diagram is shown in Fig. 2A.

2.4. DNA transfection and CAT assay

Cesium chloride-purified plasmids were introduced into COS7 cells by either calcium-phosphate co-precipitation method or DEAE-dextran-mediated method following established protocols as previously described [23,24]. Twenty-four hours prior to transfection, 2×10^5 cells were seeded in a 6 cm tissue-culture dish. For calcium-phosphate transfections, 5 μg of pF1CAT, pF2CAT, pUcF3CAT or pUcF4CAT and 10 mg of pKCR40M (a Tax expressor) or pKCRH2 (the parental plasmid of pKCR40M without exogenous cDNA) were subjected to a single experiment. For DEAE-dextran transfections, we used 10 μg of pF1CAT, pF2CAT, pUcF3CAT or pUcF4CAT. Seventy-two hours after transfection, cells were harvested and their lysate was prepared by 3 cycles of freeze-thawing followed by centrifugation to remove debris. The CAT activity was estimated by a standard protocol as previously shown [22,25].

2.5. RNA preparation and RNase protection analysis

Total RNA extracts were made from HUT102G, MOLT-4 and Jurkat cells by an acid guanidinium thiocyanate (GTC)/phenol/chlorophorm (AGPC) method as previously described [26]. Riboprobes were transcribed in vitro by bacteriophage T3 RNA polymerase using *Bam*HI or *Spe*I-digested pBS(5 + 3)A as a template (for probe strand) or by bacteriophage T7 RNA polymerase using *Hin*dIII-digested pBS(5+3)A as a template (for sense strand) in the presence of [α-^{32}P]UTP. The transcription reaction was performed following the manufacturer's protocol with a commercial kit (Stratagene, CA). Diagrams of the resultant probes are shown in Fig. 4A. Ten micrograms of each total RNA was hybridized with 5×10^5 cpm of ^{32}P-labeled probe in a hybridization buffer containing 80% formamide at 45°C overnight and digested by RNase A and RNase T1 under the condition described previously [27]. After digestion protected probe fragments were analyzed under denaturing conditions (Fig. 4B).

2.6. GenBank/EMBL/DDBJ accession numbers

Sequence data appeared in this study are partially available under accession number X75259 and fully available under accession number X84992 for gene ARP in Fig. 1 and X84993 for gene ART in Fig. 3.

3. Results

3.1. Tax-responsive promoter activity in the 5′-flanking sequence

Since the defective provirus was lacking its 5′ LTR, we suspected the existence of a functional substitute of the 5′ LTR in the 5′ cellular flanking region. Nucleotide sequence analysis of this region revealed typical eukaryotic TATA-box unit and its homologue. In addition, several enhancer-like sequences [28,29] and an element homologous to the NF-κB binding consensus [30] were also identified (Fig. 1), which suggest that it may contain a functional promoter. Subfragments derived from the 5′-flanking region were isolated and subcloned into pSVO-CAT to determine their activity as promoters, but we could detect no CAT activity in COS7 cells transfected with these plasmids alone. However, when co-transfected with a Tax-expression plasmid, 3 out of 4 plasmids displayed detectable CAT activity (Fig. 2). As the three plasmids (pF1CAT, pF2CAT and pUcF3CAT) share the common *Pst*I-*Hin*cII subfragment, the Tax-responsive weak promoter activity comes from this region. This fragment includes the putative TATA-box and the enhancer-like elements, however, pUcF4CAT has them as well, but has shown no detectable promoter activity even in the presence of Tax. All the CAT assays were repeated at least three times and gave similar results (data not shown). The nucleotide sequence was subjected to the homology-search of GenBank and EMBL. No significant homologues were found in reported sequences.

3.2. Nucleotide sequence of the 3′-flanking corresponding region in MOLT-4 cells

Since the putative TATA box was located near the integration site, the putative gene which is thought to have been driven by the promoter should exist beyond the defective provirus in

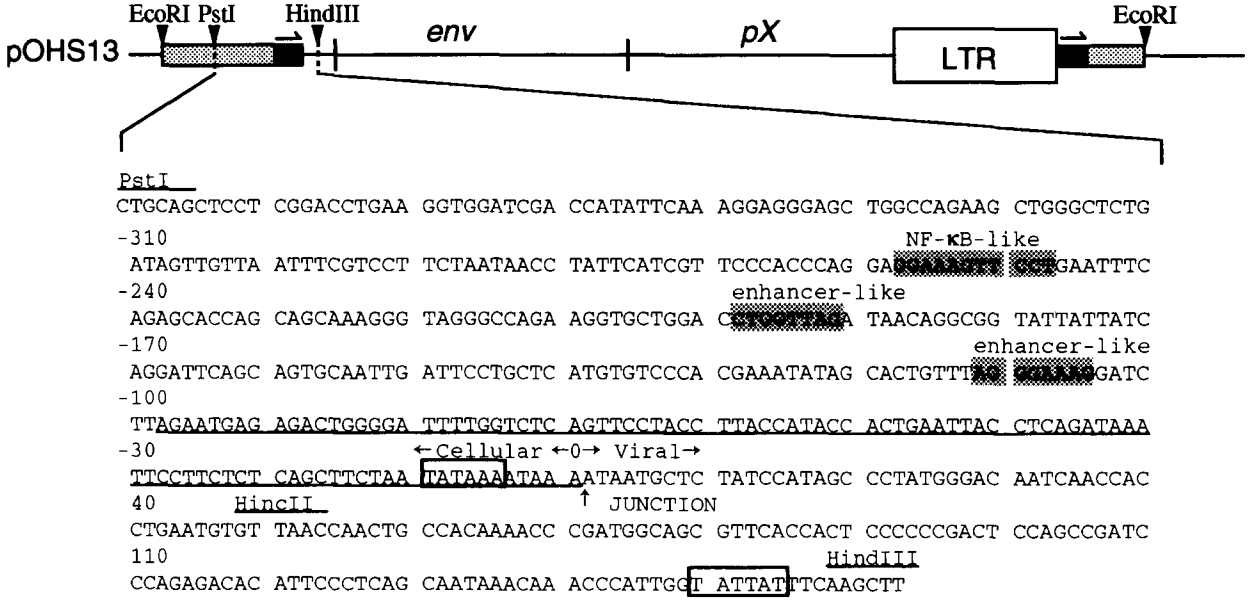

Fig. 1. Structure of the defective HTLV-I provirus from patient R and the nucleotide sequence around the 5′ integration site. Small boxes in the scheme represent 5′ and 3′ cellular flanking regions. Within these areas, the novel direct repeats reported previously are shown in solid black and marked with arrows over the boxes. The repeated region is underlined in the nucleotide sequence. TATA box-like elements are boxed. Shaded regions indicate NF-κB-like binding site and enhancer core-like sequences. Several recognition sites for restriction endonucleases are shown.

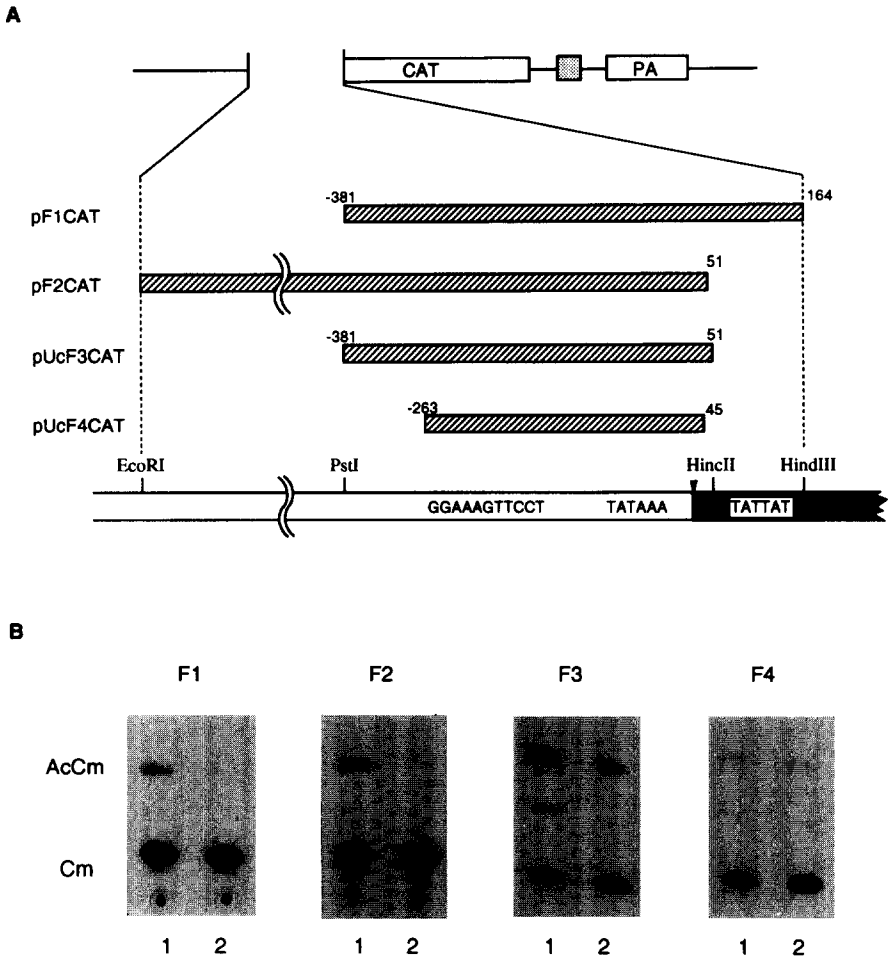

Fig. 2. Tax-responsive promoter activity in the 5′-flanking region. (A) Construction of the 5′-flanking-CAT plasmids. The regions subcloned upstream of bacterial CAT gene are shown as hatched boxes. Sequence numbers displayed here correspond to those in Fig. 1. Schematic representation of the structure around the 5′ viral–cellular junction is shown below for comparison. The NF-κB-like binding site and TATA box-like sequences are also included. (B) CAT activities from COS7 cells transfected with these plasmids. Lane 1 = co-transfected with pKCR40M (a Tax expresser); lane 2 = co-transfected with pKCRH2 (parental vector of pKCR40M). F1, F2, F3 and F4 denote pF1CAT, pF2CAT, pUcF3CAT and pUcF4CAT, respectively.

the case of patient R. Therefore, the 3′-flanking region of pOHS13 may contain a fragment of the putative gene. First, we determined the nucleotide sequence of the 3′-flanking corresponding region of the MOLT-4 genome; however, there were no significant coding frames long enough to encode a protein in the sequenced area (Fig. 3B). In our previous study, we reported that in the case of patient R, both 5′-3′-flanking sequences contain 99 bp-long direct repeats at the viral integration sites, whereas the 5′-3′-fused correspondent in MOLT-4 genome contains a single sequence. We found no difference between patient R's 3′ flanking sequence and its corresponding sequence of MOLT-4 except the repeats. Also the electromobility of PCR-derived 5′-3′-fused correspondent from HUT102G was exactly the same as MOLT-4 [21]. Fig. 3A summarizes these previous findings. The nucleotide sequence has shown a unique *Spe*I site at the middle area of the 3′-flanking corresponding region. No significant homologues were found by searching the data bases EMBL and GenBank.

Table 1
Clinical observations in patient R. The values of physical and hematologic findings of 35 cases [32] are summerized

		Reported cases	Patient R
Age (Y)		27–73 (median = 51)	51
Symptoms	Enlarged lymphonode	86%	+
	Hepatomegary	77%	+
	Splenomegary	51%	−
	Skin lesions	49%	++
Peripheral W.B.C. (/ cm^2)		9700–499000	198000
Survival perod	~6M	68%	6M
	6M–12M	26%	6M
	12M ~	6%	6M

3.3. *Detection of specific transcripts within the 3′-flanking corresponding region in HUT102G*

Utilizing the internal *Spe*I site and vector-derived *Bam*HI site, we could prepare two riboprobes for detection of transcripts using pBS(5+3)A in which the 5′-3′-fused correspondent from MOLT-4 is subcloned. RNase protection analysis with these radiolabeled probes has shown several distinct protected fragments when hybridized with the total RNA from

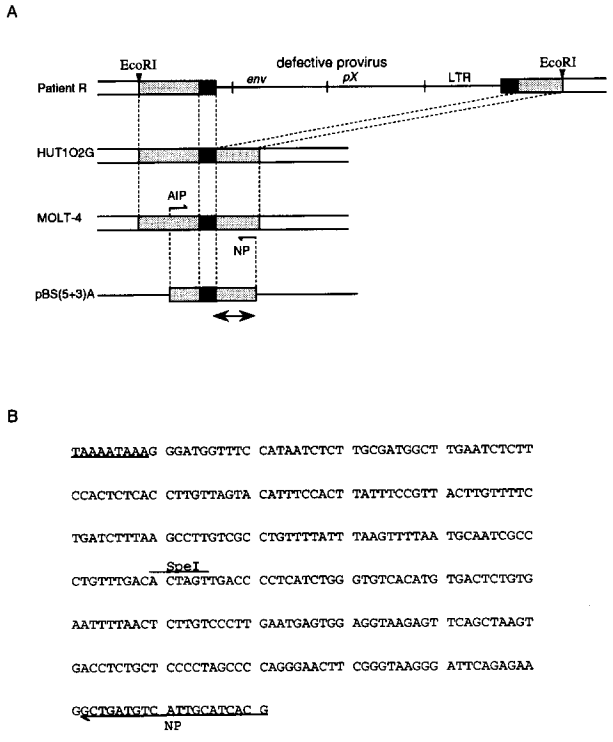

Fig. 3. (A) Structural comparison of cellular flanking regions and their correspondents. The subcloned area of pOHS13 is the *Eco*RI fragment of patient R's genome. Closed boxes represents the subcloned flanking regions and the correspondents. Hatched areas denote sequences without repeat. Solid areas represent the sequences repeated in R's genome. Primers (AIP, NP) used for isolating and subcloning the corresponding part of MOLT-4 genomic DNA to pBS(5+3)A are also illustrated. The large arrow below pBS(5+3)A indicates the sequenced region displayed in panel B. (B) Nucleotide sequence of the 3′-flanking corresponding region of MOLT-4. Most of the repeated sequence appeared in Fig. 1 is omitted and its 3′ end is involved here and underlined. The internal *Spe*I recognition site and the binding site of a primer are indicated. ←

HUT1O2G (Fig. 4), whereas no such signal occurred in HTLV-I-free T-cell lines, Jurkat and MOLT-4. Similar experiments under a radiolabeled sense-strand RNA from pBS(5+3)A revealed no specific band in every cell line tested (data not shown). Thus, this region is actively expressed specifically in HUT1O2G sharing the same sense-strand with the defective HTLV-I in R's genome. The longer probe contains most area of the 5′ flanking as well as the 3′-flanking region. The shorter probe covers only a part of the 3′-flanking region of approximately 160 bases. As shown in Fig. 4B, there was no difference in size of the protected probe RNAs between these two probes. This finding proves that the HUT1O2G-specific transcripts originate in the corresponding region of the 3′-flanking region. Although we attempted to determine the transcription initiation site(s) of these transcripts by primer-extension analysis, it has not been successful, due to no information about which part of the region covered by the shorter probe would be transcribed. Accordingly it is likely that we have not prepared an appropriate primer.

4. Discussion

Our study has pointed out two interesting aspects of ATL. One is the detection of an unknown transcript expressed specifically in an ATL cell line that possesses full-length HTLV-I provirus and actively produces virus particles. The induction of the expression should be undertaken in *trans* possibly by viral Tax protein, since there is no proviral integration around the transcribed region, which was previously shown by PCR [21] and genomic Southern blot hybridization (data not shown),

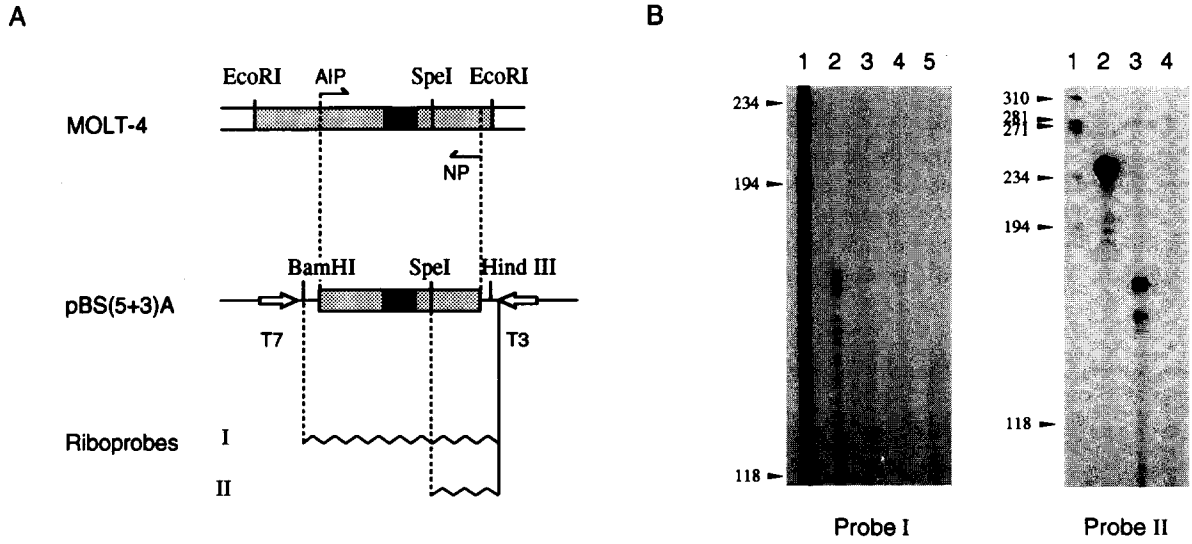

Fig. 4. (A) Riboprobes obtained from pBS(5+3)A. T3 and T7 represent specific promoters for in vitro transcription with bacteriophage T3 and T7 RNA polymerase, respectively. *Bam*HI and *Hind*III recognition sites originate in the multiple cloning site of the parental plasmid, while *Spe*I is the internal one shown in Fig. 3B. Other information is displayed in Fig. 3. (B) Detection of specific transcripts in HUT1O2G. Left panel: RNase protection analysis with radiolabeled probe I. Lane 1 = DNA size marker; lane 2 = HUT1O2G; lane 3 = MOLT-4; lane 4 = Jurkat; lane 5 = tRNA. Right panel: RNase protection analysis with radiolabeled probe II. Lane 1 = DNA size marker; lane 2 = undigested probe; lane 3 = HUT1O2G; lane 4 = MOLT-4. Sizes of DNA markers are shown in numbers of bases with arrowheads. The undigested probe looks longer than 160 bases due to the difference of electromobility between RNA and DNA and addition of an extra vector-derived sequence.

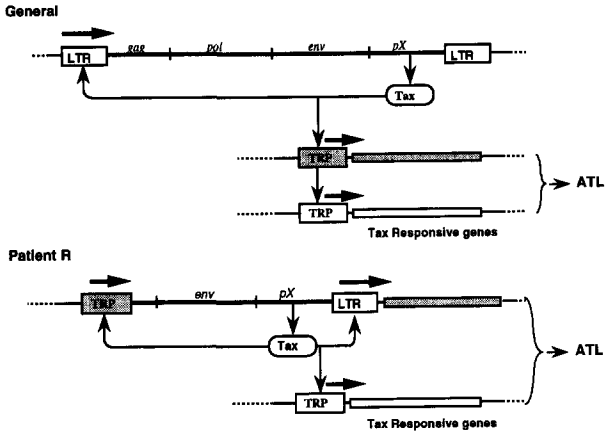

Fig. 5. Schematic comparison of the proviral–host interactions of a general ATL case versus patient R. 'TRP' is an abbreviation of Tax-responsive promoter. Wide boxes with 'TRP' followed by skinny long boxes denote Tax-responsive genes in general. Stippled one represents the gene analyzed in this study.

and a Tax-responsive promoter upstream of the transcribed area is strongly suggested. Although we have not strictly specified which element was responsible for the Tax dependent promoter activity, we found a 118 bp area that is crucial for it. In this case, something more than the NF-kB binding site-like element was required for its Tax-responsiveness. It is likely that unknown cellular factor(s) may mediate this regulation like several cellular genes, which should be clarified in the near future. Then interestingly, in the ATL case we analyzed, the same region was modified by a defective proviral integration at a point between the putative promoter and the transcribed area, then eventually the disrupted gene could revive under the control of 3'LTR. In the area we sequenced, we could not specify any supposed protein encoded. However, it is likely for detected RNA fragments to be connected with other exon(s) downstream of the sequenced area by splicing. In fact, we detected at least three different fragments in the same area, suggesting the involvement of alternative splicing. Although we have not yet succeeded to isolate the entire gene at present, we are currently trying to obtain a better probe by exon-trapping method in order to pick up the whole cDNA.

The other interesting point is the novel mode of the integration of patient R's provirus. The role of this kind of defective provirus in ATL has been controversial, since it lacks the 5'LTR for transcription of viral gene products. It might look like a fossil without transcriptional machinery. However, it had been reported that Tax is essential for maintenance of malignant phenotypes in transformed cells in vitro [5]. If so, even such a defective provirus may have expressed Tax to maintain malignancy in vivo.

In the case of patient R, the defective provirus was the only HTLV-1 provirus in the host genome and could simulate the intact provirus by integrating in a novel manner (Fig. 5). By integrating between the promoter and the transcribed region of a certain gene which may be related to the progression and maintenance of ATL, the defective provirus may have regained the regulatory machinery and could provide biological effects on host cells similar to that of a full-length provirus integrated anywhere in the host cell genome. Unfortunately, since neither RNA nor cell extract from this patient is in our hands, we are not able to verify the hypothesis further. However, the defective provirus could act like a full-length HTLV-I provirus without *gag/pol* genes that may be dispensable for ATL-leukomogenesis. It is noteworthy that this patient R displayed typical symptoms and findings of ATL throughout his illness (Table 1), despite having the defective provirus.

Recently it has been reported that the *pol-env* region of HTLV-I provirus contains a weak Tax-independent internal promoter which may act as an early promoter for Tax-expression [31]. It is a most interesting discovery that gave insights in the complex regulation of gene expression of HTLV-I. However, the promoter does not necessarily guarantee the oncogenic potential of such an *env-pX*-LTR provirus. If the promoter is strong enough to cause and maintain the malignant phenotype of ATL cells, the clinical feature of patient R should have been far different from typical ATL. On the contrary, if the promoter is too weak to cause it, an additional factor to trigger ATL should be required. In the case of patient R, the additional factor could be the novel *cis* interaction with the host genome. In this point of view, it is also interesting to analyze the clinical features of ATL cases that have only one *env-pX*-LTR defective provirus in relation to the mode of *cis* interaction of the provirus and the host genome; for the clinical features of these cases may be strongly influenced by such additional factors provided in *cis* in the absence of 5'LTR. Finally, they may give information that help us to explain the long pathway from HTLV-I infection to ATL.

Acknowledgements: We thank Deanna Gotte for critically reading the manuscript, Masumi Fujiki and Risa Kubota for the preparation of the manuscript.

References

[1] Poiesz, B.J., Ruscetti, F.W., Gazdar, A.F., Bunn, P.A., Minna, J.D. and Gallo, R.C. (1980) Proc. Natl. Acad. Sci. USA 77, 7415–7419.
[2] Hinuma, Y., Nagata, K., Hanaoka, M., Nakai, M., Matsumoto, T., Kinoshita, K., Shirakawa, S. and Miyoshi, I. (1981) Proc. Natl. Acad. Sci. USA 78, 6476–6480.
[3] Grassmann, R., Berchtold, S., Radant, I., Alt, M, Fleckenstein, B., Sodroski, J.G., Haseltine, W.A. and Ramstedt, U. (1992) J. Virol. 66, 4570–4575.
[4] Tanaka, A., Takahashi, C., Yamaoka, S., Nosaka, T., Maki, M. and Hatanaka, M. (1990) Proc. Natl. Acad. Sci. USA 87, 1071–1075.
[5] Yamaoka, S., Tobe, T. and Hatanaka, M. (1992) Oncogene 7, 433–437.
[6] Suzuki, T., Hirai, H., Fujisawa, J., Fujita, T. and Yoshida, M. (1993) Oncogene 8, 2391–2397.
[7] Dittmer,J., Gitlin, S.D., Reid, R.L. and Brady, J.N. (1993) J. Virol. 67, 6087–6095.
[8] Fujii, M., Tsuchiya, H., Chuhjo, T., Akizawa, T. and Seiki, M. (1992) Genes Dev. 6, 2066–2076.
[9] Hoyos, B., Ballard, D.W., Böhnlein, E., Siekevitz, M. and Greene, W.C. (1989) Science 244, 457–460.
[10] Ruben, S., Poteat, H., Tan, T.-H., Kawakami, K., Roeder, R., Haseltine, W. and Rosen, C.A. (1988) Science 241, 89–91.
[11] Sabe, H., Tanaka, A., Siomi, H., Koyasu, S., Yonehara, S., Yahara, I., Tagaya, Y., Sugie, K., Teshigawara, K., Yodoi, J. and Hatanaka, M. (1988) Int. J. Cancer 41, 880–885.
[12] Miyatake, S., Seiki, M., Yoshida, M. and Arai, K. (1988) Mol. Cell. Biol. 8, 5581–5587.
[13] Schreck, R. and Baeuerle, P.A. (1990) Mol. Cell. Biol. 10, 1281–1286.

[14] Sawada, M., Suzumura, A., Yoshida, M. and Marunouchi, T. (1990) J. Virol. 64, 4002–4006.
[15] Fujii, M., Sassone-Corsi, P. and Verma, I.M. (1988) Proc. Natl. Acad. Sci. USA 85, 8526–8530.
[16] Nagata, K., Ohtani, K., Nakamura, M. and Sugamura, K. (1989) J. Virol. 63, 3220–3226.
[17] Li, C-C.H., Ruscetti, F.W., Rice, N.R., Chen, E., Yang, N.-S., Mikovits, J. and Longo, D.L. (1993) J. Virol. 67, 4205–4213.
[18] Li, M. and Siekevitz, M. (1993) Mol. Cell. Biol. 13, 6490–6500.
[19] Miura, S., Ohtani, K., Numata, N., Niki, M., Ohbo, K., Ina, Y., Gojobori, T. Tanaka, Y., Tozawa, H., Nakamura, M. and Sugamura, K. (1991) Mol. Cell. Biol. 11, 1313–1325.
[20] Konishi, H., Kobayashi, N. and Hatanaka, M. (1984) Mol. Biol. Med. 2, 273–283.
[21] Kubota, S., Furuta, R., Maki, M., Siomi, H. and Hatanaka, M. (1993) Oncogene 8, 2873–2877.
[22] Gorman, C.M., Moffat, L.F. and Howard, B.H. (1982) Mol. Cell. Biol. 2, 1044–1051.
[23] Graham, F.L. and Van der Eb, A.J. (1973) Virology 52, 456–467.
[24] Cullen, B.R. (1987) Methods Enzymol. 152, 684–704.
[25] Mori, K., Sabe, H., Siomi, H., Iino, T., Tanaka, A., Takeuchi, K., Hirayoshi, K. and Hatanaka, M. (1987) J. Gen. Virol. 68, 499–506.
[26] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[27] Gilman, M. (1987) in: Current Protocols in Molecular Biology (Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. eds.) Wiley, New York, pp. 4.7.1.–4.7.8.
[28] Lenz, J., Celandor, D., Crowther, R.L., Patarca, R., Perkins, D.W. and Haseltine, W.A. (1984) Nature 308, 467–470.
[29] Golemis, E.A., Speck, N.A. and Hopkins, N. (1990) J. Virol. 64, 534–542.
[30] Fujisawa, J., Toita, M. and Yoshida, M. (1989) J. Virol. 63, 3234–3239.
[31] Nosaka, T., Ariumi, Y., Sakurai, M., Takeuchi, R. and Hatanaka, M. (1993) Nucleic Acids. Res. 21, 5124–5129.
[32] Takatsuki, K., Uchiyama, T., Ueshima, Y and Hattori, T. (1979) Jpn. J. Clin. Oncol. 9, 317–324.