Isolation of three kinds of human endogenous retrovirus-like sequences using tRNA$^{\text{Pro}}$ as a probe

Fumio Harada, Naoko Tsukada and Nobuyuki Kato

Biophysics Division, Cancer Research Institute, Kanazawa University, 13-1, Takaramachi, Kanazawa 920 and 'Virology Division, National Cancer Center Research Institute, 5-1-1, Tsukiji, Chuo-ku, Tokyo 104, Japan

Received September 25, 1987; Accepted October 23, 1987 Accession no. Y00491

ABSTRACT

Three kinds of human endogenous retrovirus-like sequences (HuERS-P1, 2 and 3) were isolated from a HeLa cell genomic library using the 3'-half fragment of proline tRNA as a hybridization probe. These elements contained putative primer binding sites complementary to the 3'-terminus of proline tRNA and long terminal repeats (LTRs) characteristic of retrovirus provirus. The LTR sequence of HuERS-P1 consisted of about 690 nucleotides and contained a CAT box, a TATA box and a polyadenylation signal. A complete unit of an Alu family sequence was inserted into the 5'-LTR of one of the clones. HuERS-P2 also contained a TATA box and a polyadenylation signal in its LTR (about 840 nucleotides long), but the LTR sequence of this element was quite different from that of HuERS-P1. Although clone HuERS-P3 contained only the 5'-LTR region, this LTR sequence contained a CAT box, a TATA box and a polyadenylation signal and was quite similar to the LTR sequence of the recently isolated human retrovirus-related sequence HuRRS-P (Kröger, B. and Horak, I. (1987) J. Virol., 61, 2071-2075). Human and simian DNAs contain 10 to 40 copies of these elements, but mouse DNA does not contain these elements.

INTRODUCTION

RNA directed DNA synthesis by reverse transcriptase of retroviral genomic RNA is initiated from a specific primer tRNA, tRNA$^{\text{Trp}}$ for Rous sarcoma virus (1-3), tRNA$^{\text{Pro}}$ for Moloney murine leukemia and avian reticuloendotheliosis viruses (4-6) and tRNA$^{\text{Lys}}$ for mouse mammary tumor virus (7) are used as primers. Nucleotide sequence analysis of DNA clones of each retrovirus showed the presence of a 18 nucleotide sequence (primer binding site, PBS), which is complementary to the 3'-end of primer tRNA on the immediate 3'-side of the 5'-long terminal repeat (LTR)(8-11). The PBS has been found downstream of the 5'-LTR in the sequences of various other retrovirus proviruses and all mammalian type C retroviruses whose sequences have been reported to possess a PBS corresponding to that of tRNA$^{\text{Pro}}$ (12).

Recently, several endogenous retrovirus-like sequences have been isolated from the human genome and their LTR sequences have been determined. These sequences were isolated by low stringency hybridization of various
known retroviral sequences (13-17) or were accidentally found in flanking regions of other genes (18-19). They have PBS sequences complementary to tRNA\(^{\text{Gln}}\) (20, 21), tRNA\(^{\text{Arg}}\) (22), tRNA\(^{\text{His}}\) (18), tRNA\(^{\text{Ile}}\) (19) or tRNA\(^{\text{Lys}}\) (17, 23) but no sequence corresponding to tRNA\(^{\text{Pro}}\) has yet been reported.

In the present study, we isolated three families of human endogenous retrovirus-like sequences (HuERS-P1, 2 and 3) from a human genomic library using tRNA\(^{\text{Pro}}\) as a hybridization probe. Quite recently, Kröger and Horak isolated a human retrovirus-related sequence corresponding to tRNA\(^{\text{Pro}}\) (HuRRS-P) (24). They used a synthetic oligonucleotide homologous to the PBS for tRNA\(^{\text{Pro}}\). Their HuRRS-P belongs to the same family as our HuERS-P3.

**MATERIALS AND METHODS**

**Preparation of tRNA\(^{\text{Pro}}\) probe.**

\(\text{tRNA}^{\text{Pro}}\) was purified from mouse liver by successive column chromatographies as described previously (5). The nucleotide sequence of human tRNA\(^{\text{Pro}}\) is the same as that of mouse tRNA\(^{\text{Pro}}\) (F. Harada, unpublished result). Purified tRNA\(^{\text{Pro}}\) was cleaved at m\(^{\text{7G}}\) by the procedure of Wintermeyer and Zachau (25) and the fragments were separated by electrophoresis in 10\% polyacrylamide gel. The 3'-fragment (30 nucleotides long) was eluted from the gel, labeled with [5'-\(^{32}\)P]pCp (Amersham) at the 3'-end (26) and used as a probe in genomic cloning.

**Screening of phage clones.**

A HeLa cell genomic library cloned in Charon 4A vector was prepared as described by Maniatis et al. (27). Hybridizations with the \(^{32}\)P-labeled 3' fragment of tRNA\(^{\text{Pro}}\) were carried out in 50\% formamide, 5X SSC (1X SSC is 0.15M NaCl plus 0.015M sodium citrate), 5X Denhardt's solution (28), 0.2\% SDS and 10\% dextran sulfate sodium salt for 18 hours at 37\(^{\circ}\)C. The hybridized filters were washed with two changes of 3X SSC containing 0.1\% SDS for 20 min each time at room temperature and then with 0.1X SSC containing 0.1\% SDS at 45\(^{\circ}\)C for one hour. The filters were autoradiographed for 18 hours and then treated with 2 \(\mu\)g/ml of RNase A (29). After inactivation of RNase A with iodoacetate (30), the filters were washed in 0.1X SSC containing 0.1\% SDS for one hour at 45\(^{\circ}\)C and autoradiographed for several days. Restriction fragments from purified phage clones were subcloned in pKH47 vectors (31).

**DNA sequencing.**

For sequencing of HuERS-P1 and 2, DNA fragments were labeled either by end-filling with [\(\alpha-\(^{32}\)P]dNTP and the Klenow fragment of DNA polymerase I or by kination with [\(\gamma-\(^{32}\)P]ATP and T4 polynucleotide kinase. Strands of labeled
Figure 1. Restriction maps of the inserts of phage clones. Black boxes are LTRs. The white box in the 5'-LTR of HuERS-P1-2 is an Alu family sequence. The bars under the 5'-LTRs are DNA fragments for the probes in Southern hybridization experiments in Fig. 2. Upper bars are DNA fragments used as probes for the experiments in Fig. 6. Restriction sites shown are R, Eco RI; H, Hind III; B, Bam HI.

fragments were separated and sequenced by the method of Maxam and Gilbert (32). For sequencing of HuERS-P3, DNA fragments were subcloned in M13mp10 or M13mp11 vectors (33) and sequenced as described by Sanger et al. (34). All sequences shown were determined in both DNA strands.

Southern hybridization of total DNA.

High molecular weight DNAs of human (HeLa, Jurkat and Raji), African green monkey (CV-1) and Friend erythroleukemia (FVTCT) cell lines were prepared as described by Maniatis et al. (27). Eco RI digests of these DNAs were subjected to electrophoresis in 0.6% agarose gel and were blotted onto a nitrocellulose filter (35). Hybridization was carried out in 6X SSC, 5X Denhardt's solution, 5 mM EDTA and 0.5% SDS for 18 hours at 65°C. After hybridization, the filters were washed with two changes of 2X SSC containing 0.1% SDS for 15 min each time at 55°C and then with 0.1X SSC containing 0.1% SDS at 55°C (low stringency condition) or 65°C (high stringency condition) for one hour.

RESULTS

Isolation of recombinant phage clones hybridizing with tRNA<sup>Pro</sup> probe.

We used the 3'-fragment of tRNA<sup>Pro</sup> 3'-end labeled with [5'-<sup>32</sup>P]pCp as a probe for cloning endogenous retrovirus sequences. This probe hybridized to not only the PBS of endogenous retroviral sequences, but also with the genomes of tRNA<sup>Pro</sup>. However, the genome of tRNA<sup>Pro</sup> does not have the 3'-CCA sequence of the mature tRNA molecule, whereas PBS contains this sequence.
Figure 2. Southern hybridization analysis of cloned DNAs probed with fragments containing PBS. Lane 1, HuERS-P3; lane 2, HuERS-P2; lane 3, HuERS-P1-1; lane 4, HuERS-P1-2. Each cloned DNA was digested with the same enzymes as used for preparation of probes. The probes used were A, the Pvu II-Eco RI fragment (0.4 kbp) of HuERS-P3; B, the Kpn I-Tthlll I fragment (0.65 kbp) of HuERS-P2; C, the Pvu II-Bam HI fragment (0.4 kbp) of HuERS-P1-1; D, the Stu I fragment (1.4 kbp) of HuERS-P1-2. Low stringency conditions were used.

Therefore, if the hybrid is treated with RNase A, the radioactivity of the probe hybridized to the tRNA$_\text{Pro}$ genome should be eliminated and only the endogenous retroviral sequence should be selected.

About $3 \times 10^5$ phage plaques were screened by this method and four positive clones were isolated. Phage DNA of each clone was digested with Eco RI, Hind III, Bam HI or combinations of these enzymes and the restriction maps were constructed (Fig. 1). The maps were all different, indicating that these fragments were from different loci in the genome.

Since the PBS is located immediately after the 5'-LTR in the provirus, the DNA fragment that hybridizes to primer tRNA should contain a part or all the 5'-LTR. We prepared a small fragment containing the sequence homologous to the 3'-terminus of tRNA$_\text{Pro}$ from each subclone. These fragments were
Figure 3. Nucleotide sequences of LTRs and their flanking regions of HuERS-Pl-1 and Pl-2. Only the altered nucleotides are shown for the 3'-LTR of Pl-1 and the LTRs of Pl-2. Dashes indicate the absence of nucleotides. (A) Comparison of LTR sequences. Arrows indicate a terminal inverted repeat and direct repeat. The TPA-responsive element, CAT box, TATA box and polyadenylation signal are boxed. The triangle indicates the position of insertion of an Alu sequence. (B) Flanking regions of LTRs. Bars above sequences indicate cellular direct repeats. The PBS and polypurine tract are boxed. The sequence complementary to the 3'-terminus of tRNA is shown. (C) Alu sequence inserted in the 5'-LTR of HuERS-Pl-2. Bars indicate cellular direct repeats.

nick-translation and used as probes in searching for the 3'-LTR in each clone and/or analyzing homology between each clone. As shown in Fig. 2, the HuERS-P3 probe hybridized to only one fragment of the HuERS-P3 clone. Therefore, this clone does not contain the 3'-LTR and is not homologous to 9157.
other clones. Since the HuERS-P2 probe hybridized to two fragments of the HuERS-P2 clone, this clone contains the 3'-LTR but is not related to other clones. The HuERS-Pl-1 probe hybridized not only to two fragments of the HuERS-Pl-1 clone but also to two fragments of HuERS-Pl-2. Therefore, HuERS-Pl-1 and HuERS-Pl-2 are homologous and both clones contain the 3'-LTR. The HuERS-P2 probe contained a repetitive sequence since it hybridized to various fragments of clones.

**Nucleotide sequences of LTRs and their flanking regions.**

The nucleotide sequences of regions corresponding to the 5'- and 3'-LTRs of each clone were determined. The PBS sequences of these clones shared 17 of 18 nucleotides with the complementary sequence of the 3'-terminus of tRNA Pro (Fig. 3B, Fig. 4B and Fig. 5). HuERS-Pl-1, P1-2 and P2 were flanked by cellular direct repeats of 4 base pairs (bp) (Fig. 3B and Fig. 4B) and each LTR started with the dinucleotide TG and ended with CA. Polypurine tracts were observed on the immediate 5'-side of the 3'-LTRs.

The LTR sequences of HuERS-Pl-1 and P1-2 were very similar each other (Fig. 3A). These sequences (about 690 bp) contained a CAT box (CCAAAT), a
Figure 5. Comparison of the 5'-LTR sequences of HuRRS-P (24) and HuERS-P3. The CAT box, TATA box, polyadenylation signal and PBS are boxed.

TATA box (TTAAAA) and a polyadenylation signal (AATAAA). These sequences also possessed a TPA-responsive element TGACTCAG (36, 37) 50 bp upstream of the CAT box. The short sequence TAGTTTA is repeated three times (Pl-1) or twice (Pl-2) in tandem near the 5'-terminus. In the 5'-LTR of HuERS-Pl-2, a complete unit of the Alu family sequence is inserted in the opposite direction to the LTR (Fig. 3A and C).

The LTR sequence of HuERS-P2 was about 150 bp longer than that of HuERS-Pl. This sequence (Fig. 4A) also contained a TATA box (TATAAA) and a polyadenylation signal (AATAAA), but was quite different from the LTR sequence of HuERS-Pl. The distance between the 5'- and 3'-LTRs was only 3.2 kbp.

The phage clone of HuERS-P3 did not contain the 3'-LTR region. However, the upstream sequence of PBS had the characteristic features of an LTR, including a CAT box (CCAAAT), a TATA box (TATAAA) and a polyadenylation signal (ATTAAA). Moreover the sequence of this region resembled that of the LTR of HuRRS-P (24) (Fig. 5).

Copy numbers of HuERS-Pl, P2 and P3 in human, simian and murine genomes.

For examination of the copy numbers of HuERS-Pl, P2 and P3 in human genomes and the appearance of these sequences in other species, small fragments of these elements (Fig. 1) were hybridized to Eco RI digestes of human, simian and murine DNAs (Fig. 6). All Huers-P probes hybridized with human and monkey DNAs, but not with mouse DNA. The hybridization patterns of human and monkey DNAs were similar. Internal probes of HuERS-Pl and P2 gave about ten and twenty bands, respectively, whereas the LTR probe of HuERS-P3
Figure 6. Southern hybridization analysis of HuERS-P1, P2 and P3 genes in human (HeLa, Jurkat and Raji), monkey (CV-1), and mouse (FVTCT) genomes. Samples of 5 µg of human and mouse DNAs and 3 µg of monkey DNA were digested with Eco RI and analyzed. Lane 1, HeLa DNA; lane 2, Jurkat DNA; lane 3, Raji DNA; lane 4, CV-1 DNA; lane 5, FVTCT DNA. The probes used were A, the Eco RI-Hae III fragment (1.1 kbp) of HuERS-P1-1; B, the Bam HI fragment (0.4 kbp) of HuERS-P2 and C, the Eco RV-Eco RI fragment (0.9 kbp) of HuERS-P3. Low stringency conditions were used.

gave about 40 bands. Since overlapping and closely adjacent bands were not counted, these numbers may be under-estimations. Therefore, there may be 10 to 20 copies of HuERS-P1, 20 to 30 copies of HuERS-P2 and 30-40 copies of HuERS-P3 in the human genome.

DISCUSSION

Using trNA Pro as a hybridization probe, we obtained three kinds of human endogenous retrovirus-like sequences. These elements contained sequences characteristic of the LTR structure of the retrovirus provirus. During preparation of this manuscript, a report by Kröger and Horak of a similar screening method using synthetic oligonucleotide complementary to trNA Pro was published (24). They obtained one clone, HuRRS-P, which belongs to the same
family as HuERS-P3 in our study. These methods are very useful to obtain new elements that are not homologous to other retroviruses.

The LTR sequences of these three elements do not resemble those of any retroviruses or endogenous retroviruses whose sequences have been reported. Since each LTR sequence contains base exchanges, deletions and insertions, these elements may not be expressed as virus particles. However, the promoter sequences were well preserved, the viral sequence and/or downstream sequence of the 3'-LTR might be transcribed in the cells. The possibility that transcription of HuERS-P1 is induced by TPA, UV-light or mitomycin C is of especial interest, because of the presence of a TPA-responsive element (36, 37) upstream of the promoter region of the LTR. The internal sequence of HuERS-P2 is very short, and we are now examining whether this length (3.2 kbp) is natural or whether this clone is a deletion mutant.

These elements exist not only in the human genome but also in the simian genome. Probably germ cells of the common ancestor of humans and monkeys were infected by retroviruses and these sequences still exist in the genomes of both species. By using these screening methods with other tRNAs, many other endogenous retrovirus-like sequences might be obtained.

ACKNOWLEDGMENTS

We thank Drs. Y. Ikawa, K. Shimotohno and S. Murakami for helpful discussions. We also thank Mses. K. Terada and T. Nishikawa for technical assistance.

This work was partly supported by a Grant-in-Aid from the Ministry of Health and Welfare for a Comprehensive 10-Year Strategy for Cancer Control, Japan and by a Grant-in-Aid for Cancer Research from the Ministry of Education, Science and Culture of Japan.

REFERENCES