Key words: HTLV-I/gag–pX fused protein/DNA sequence/peptide mapping

Structural Analysis of p28 Adult T-Cell Leukaemia-associated Antigen

By TAKASHI IINO, KAORU TAKEUCHI, SEOK HYUN NAM, HARUHIKO SIOMI, HISATAKA SABE, NOBUYUKI KOBAYASHI AND MASAKAZU HATANAKA*

Institute for Virus Research, Kyoto University, Kyoto 606, Japan

(Accepted 19 March 1986)

SUMMARY

The 28 000 mol. wt. polypeptide (p28) of adult T-cell leukaemia-associated antigen encoded by the 24S defective human T-cell leukaemia virus (HTLV-I) is associated with protein kinase activity. We have determined the nucleotide sequence of this defective HTLV-I provirus and found that it contains a portion of the gag gene (p19 and part of p24), the pX region, and two long terminal repeats, one at each end. The predicted p28 gag–pX fused protein consists of 190 amino acids and its mol. wt. was calculated as 21 055. The results of peptide mapping analysis showing that p28 contains p19 supported the nucleotide sequence data. That p28 was encoded by this defective provirus was also demonstrated by transient expression of p28 polypeptide in COS 7 cells transfected with a recombinant plasmid containing a simian virus 40 early promoter and the p28-coding region of the 24S HTLV-I.

INTRODUCTION

Adult T-cell leukaemia (ATL) is a new type of human T-cell malignancy, endemic in the south-west part of Japan (Takatsuki et al., 1977; Uchiyama et al., 1977) and also in the Caribbean area (Catovsky et al., 1982). A novel retrovirus associated with ATL, named human T-cell leukaemia virus (HTLV-I), was isolated from ATL cell lines and T-cell leukaemia (lymphoma) patients (Poiesz et al., 1980; Hinuma et al., 1981; Kalyanaraman et al., 1982; Yoshida et al., 1982). Recent aetiological, immunological and molecular biological studies strongly suggest that HTLV-I may be the causative agent of ATL (Yoshida et al., 1982; Sarin et al., 1983).

ATL patients and healthy HTLV-I carriers are known to have serum antibodies against antigens expressed in HTLV-I-bearing cell lines. These antigens, termed ATL-associated antigens (ATLA), have been found to include the HTLV-I core proteins p19 and p24, and the glycosylated envelope proteins gp61 and gp46 (Kalyanaraman et al., 1981; Robert-Guroff et al., 1981; Yamamoto & Hinuma, 1982; Yamamoto et al., 1982a; Manzari et al., 1983; Hattori et al., 1983; Kobayashi et al., 1984a).

Recently, HTLV-I has been molecularly cloned from the peripheral blood leukocytes of an ATL patient, and was shown to be composed of gag, pol, env and pX genes (Seiki et al., 1983). The genomic structure of HTLV-I is similar to the known leukaemic-type retroviruses in animals, except that HTLV-I contains the new pX region. The MT-2 cell line, which was established from normal human cord leukocytes by co-cultivation with leukaemic cells from an ATL patient (Miyoshi et al., 1981), expresses a unique 28 000 mol. wt. polypeptide (p28) in addition to the normally detected ATLA. These cells retain the ability to transform normal lymphocytes after co-cultivation (Yamamoto et al., 1982b). To understand this transforming function, we determined the genomic structure of HTLV-I in MT-2 cells and found that at least eight HTLV-I proviruses are integrated per genome (Kobayashi et al., 1984b). Southern blotting analysis of these proviruses revealed that one contains all the information necessary for virus replication, while the rest are defective. Among the latter, four proviruses showed the same structure, containing the 5' portion of the gag gene, a complete pX region, and two long terminal
repeats (LTRs), one at each end. We demonstrated that these amplified defective genomes could be transcribed into 24S HTLV-I mRNA which directs the synthesis of the p28 component of ATL A (Kobayashi et al., 1984a).

Recently, we found that p28 is associated with protein kinase activity specific for serine or threonine residues (Kobayashi et al., 1984c). From immunological analysis, p28 has been shown to share an antigenic determinant with p19 (Tanaka et al., 1983).

In this study, we have determined the nucleotide sequence of one of these defective HTLV-I proviruses, no. 42, analysed the structure of p28 compared with p19 by peptide mapping, and have proved that p28 is a gag–pX fused protein encoded by this defective genome.

**METHODS**

**Nucleotide sequence analysis of defective HTLV-I.** MT-2 DNA partially digested with Sau3AI was cloned into the vector lambda phage Charon 28 at the BamHI site. A complementary DNA clone of HTLV-I (Kobayashi et al., 1984b) was used for screening. The HTLV-I provirus clones were subcloned into pBR322. The nucleotide sequence of the appropriate fragments was routinely determined by the procedure of Maxam & Gilbert (1980), and partly by the procedure of Sanger et al. (1977).

**Cell culture and isotopic labelling of cells.** MT2 cells were maintained in RPMI 1640 medium supplemented with 10% foetal bovine serum, at 37 °C in humidified air with 5% CO2. The medium was changed every 3 to 4 days.

For tryptic peptide mapping of [3H]leucine-labelled proteins, 2 × 107 cells were incubated for 16 h in 12 ml of labelling medium (leucine-free RPMI 1640) containing 6 mCi [3H]leucine (120 to 190 Ci/mmol; Amersham). The cells were washed twice in phosphate-buffered saline, lysed, and immunoprecipitated with anti-p19 mouse monoclonal antibody GIN-7 (Tanaka et al., 1983).

For transfection of recombinant plasmids, COS 7 cells, which produce simian virus 40 (SV40) T-antigen and are permissive for the replication of SV40 replication origin-containing plasmids, were used. These cells were maintained in Dulbecco's modified MEM supplemented with 10% foetal bovine serum as described above.

**Immunoprecipitation.** Labelled cells were suspended in low salt extraction buffer [1 × 107 cells/ml, 10 mm-Tris–HCl pH 8.0, 0.14 M-NaCl, 1 mM-EDTA, 0.5% (w/v) sodium deoxycholate], and four times in low salt washing buffer [20 mM-Tris–HCl pH 7.6, 0.5 M-NaCl, 1 mM-EDTA, 0.5% (w/v) NP40, 1% (w/v) sodium deoxycholate], and four times in high salt washing buffer [20 mM-Tris–HCl pH 7.6, 0.5 M-NaCl, 1 mM-EDTA, 0.5% (w/v) NP40].

For tryptic peptide mapping of [3H]leucine-labelled proteins, immunoprecipitates were analysed by twodimensional gel electrophoresis according to the procedure of O'Farrell et al. (1977). To locate radiolabelled protein spots, fixed and Coomassie Brilliant Blue-stained gels were soaked in 1 M-sodium salicylate for 30 min. The washed gel slice was lyophilized to dryness and treated with 0.5 ml of cold performic acid (one part 30% hydrogen peroxide to 19 parts formic acid, preincubated for 1 h at room temperature) for 1 h on ice. The sample was then lyophilized three times to remove the remaining performic acid. To the oxidized gel slice was added 2 ml of 50 μg/ml trypsin–TPCK (Worthington) in 50 mM-ammonium bicarbonate, and the mixture was incubated at 37 °C for 8 h. An additional 1 ml of trypsin solution was added for 8 h, and a final 1 ml was added for 8 h. By this procedure, about 50 to 60% of the radioactivity in the gel slice was recovered in the trypsin solution. The sample was passed through a Millipore filter (0.22 μm) to remove debris, and lyophilized four times until all the ammonium bicarbonate was removed. The sample was dissolved in 10 μl of electrophoresis buffer (buffer I, acetic acid: formic acid: water, 15:5:80) and was applied to cellulose-coated glass plates (20 × 20 cm, 0.25 mm thick layer; Avicel, Funakoshi). [3H]Leucine-labelled tryptic fragments were resolved in two dimensions as described by Elder et al. (1977). Electrophoresis was performed in buffer I at 1000 V for 50 min in the cold-room. After air drying, the plates were developed in the second dimension by ascending chromatography at room temperature in buffer II (butanol: pyridine: acetic acid: water, 32.5:25:5:20). To locate the tryptic peptide fragments, the dried plates were dipped into molten 0.4% (w/v) diphenylazoazole in 2-methylnaphthalene as described by Bonner & Stedman (1978).

**Transfection of recombinant plasmids into COS 7 cells.** The expression vector pKCR H2, containing the SV40 early gene promoter, the donor and acceptor splice sites derived from the rabbit β-globin gene, and polyadenylation sites derived from both the rabbit β-globin gene and the SV40 early gene, was used (O'Hare et al.,
Structural analysis of p28 of ATLA

Fig. 1. Restriction map and sequencing strategy of the defective 24S HTLV-I provirus no. 42 from MT-2 cells. The provirus DNA is shown by the double line with a LTR (box) at each end. Flanking cellular sequences are shown by wavy lines. Arrows and their lengths indicate the direction of sequencing and sizes of regions sequenced.

1981). The SalI–SalI 3·0 kb DNA fragment which contains almost all the no. 42 provirus genome, or the RsaI–PstI 0·7 kb DNA fragment which contains just the p28 coding frame was inserted into pKCR H2 at the HindIII site located between the acceptor splice site and the polyadenylation site derived from the rabbit β-globin gene; these plasmids were named pKCR 42 and pKCR 42RP, respectively (Fig. 3a). Each recombinant plasmid was transfected into COS 7 cells by calcium phosphate precipitation (Graham & van der Eb, 1973). At 44 h after transfection, cells were labelled with [3H]leucine, immunoprecipitated with the anti-p19 mouse monoclonal antibody GIN-7, and analysed by SDS–polyacrylamide gel electrophoresis (12%, polyacrylamide) according to the procedure of Laemmli (1970).

RESULTS

Nucleotide sequence analysis of pMT-2-42

We previously reported the cloning of four defective gag–pX-type proviruses in the lambda Charon 28 phage vector (Kobayashi et al., 1984b). One of the clones, no. 42, corresponds to the provirus contained in a 7·0 kb EcoRI fragment. This fragment was subcloned into pBR322 at the EcoRI site. Fig. 1 shows the restriction endonuclease map of the resulting recombinant plasmid, pMT-2-42. The strategy of the sequence determination is also indicated.

Fig. 2 shows the nucleotide sequence of pMT-2-42. Clone 42 defective provirus consisted of 3706 bases with two LTRs of 755 bases. Compared with the sequence previously reported for HTLV-I (Seiki et al., 1983), the 570 base reading frame extended from ATG at position 802 to CATCC at position 1315 then jumped to CTG at 6640 (Seiki et al., 1983) and ended with TAA at 6696 in lambda ATK-1, indicating the deletion of entire pol and env genes. The reading frame of p28 therefore consists of the complete gag p19 gene, part of gag p24, and the region between env and pX-I of HTLV-I which we named pX-0. In this open reading frame, some nucleotides differed from the sequence of lambda ATK-1, resulting in changes of amino acids 58 and 59 (underlined in Fig. 2). The reading frame could encode a polypeptide of 190 amino acids, with a predicted mol. wt. of 21055 and having an unusually high content of proline (18%). The 5' and 3' LTR regions, intact TATA boxes (325 to 330 and 3276 to 3281) and putative enhancer sequences consisting of 21 bp repeats (Shimotohno et al., 1984) were detected.

Transient expression of p28 in COS 7 cells

To confirm that defective provirus no. 42 was functionally active, we transfected recombinant plasmids pKCR 42 or pKCR 42RP (Fig. 3a) into COS 7 cells and the transient expression of p28 was analysed (Fig. 3b). When the cells were transfected with the expression vector pKCR H2 no specific protein was detected, but after transfection with pKCR 42 or pKCR 42RP a 28000 mol. wt. polypeptide was formed. The amount of p28 expressed in COS 7 cells was greater following transfection with pKCR 42RP than with pKCR 42.
In this study, we determined the entire nucleotide sequence of a defective HTLV-I provirus (no. 42) from MT-2 cells. The sequence analysis demonstrated that p28 consists of the complete gag p19, part of gag p24, and part of pX-0. A 28000 mol. wt. polypeptide which was immunoprecipitated with an anti-p19 monoclonal antibody was expressed in COS 7 cells following transfection with the recombinant expression vectors pKCR 42 and pKCR 42RP. It was shown in Fig. 4, confirming that p28 includes p19.

**DISCUSSION**

Peptide mapping analysis of p28

Nucleotide sequence analysis revealed that p28 contains the p19 moiety. To confirm this at the protein level, tryptic peptide mapping analysis was performed on p28 and p19. The patterns obtained by peptide mapping of [3H]leucine-labelled p19 and p28 were remarkably similar as shown in Fig. 4, confirming that p28 includes p19.
### Structural analysis of p28 of ATLA

Fig. 3. (a) Schematic representation of recombinant plasmids. The expression vector pKCR H2 consists of the SV40 early gene promoter, a part of the rabbit β-globin gene containing the donor and acceptor splice sites, and polyadenylation sites derived from the rabbit β-globin gene and the SV40 early gene. The SacI–SacI 3·6 kb DNA fragment (which contains almost all of the p28 coding frame) was inserted into pKCR H2 at the HindIII site with HindIII linkers as indicated, and was named pKCR 42RP, respectively. (b) Forty μg of each recombinant plasmid was transfected into each dish of COS 7 cells by the calcium phosphate precipitation method. After 44 h, cells were labelled with [3H]leucine, immunoprecipitated with GIN-7, and analysed by SDS-PAGE. The plasmids used were pKCR H2 (lane 1), pKCR 42RP (lane 2) and pKCR 42 (lane 3).

Fig. 4. Tryptic peptide mapping patterns of [3H]leucine–labelled p19 and p28. The procedure of peptide mapping analysis was described under Methods. (a) p19; 36000 c.p.m. of radioactivity was loaded, run and exposed for 22 days. (b) p28; 31000 c.p.m., 29 days. (c) Diagrammatic representation of (a) and (b) superimposed. The spots common to p19 and p28 are indicated in black. The spots specific to p19 or p28 are indicated as open spots. The origin is indicated by a cross.
Thus, defective provirus no. 42 proved to be functionally active for synthesis of the p28 protein. The reason why the amount of p28 expressed in cells transfected with pKCR 42 is much less than that with pKCR 42RP is unclear. It is possible that the part of the LTR contained in pKCR 42 but not in pKCR 42RP may be inhibitory to transcription or translation. Although the predicted molecular weight, 21,000, appears different from that observed, 28,000, this may be due to the unusually high content of proline (18%) causing a slower mobility of the protein in SDS-containing gels (Zakut-Houri et al., 1983). Similar behaviour is observed for gag p19 whose molecular weight is estimated as 14,000 and contains 19% proline.

The sequence analysis shows that provirus no. 42 consists of two flanking LTRs, the 5' region of the gag gene and the entire pX region; the whole pol and env genes had been deleted. The homologous recombination event responsible for this must have occurred between the gag and pX regions, at the sequences CATCC in the gag p24 gene at position 1310 and in the end of the env gene at position 6636 of the complete HTLV-I sequence (Seiki et al., 1983). Homologous recombination has also been reported to take place in the case of Moloney murine sarcoma virus (Van Beveren et al., 1981).

Previously, we reported the isolation of four provirus clones of the same gag–pX defective type as no. 42 (Kobayashi et al., 1984b). We have determined the sequence of these clones in the region of the gag–pX junction, and have found them to be identical to no. 42 (data not shown). Therefore, the abundance of 24S RNA and its product p28 may be explained by a gene dosage effect of the gag–pX defective proviruses.

In this study, we have shown that the tryptic peptide mapping patterns of [3H]leucine-labelled p19 and p28 are remarkably similar. This result confirms that p28 contains p19, as was suggested by the nucleotide sequence data. Although a number of major spots coincided with those deduced from the DNA sequences of both p28 and p19, extra minor spots were also observed in the peptide mapping analysis (Fig. 4). It is considered that these resulted from unavoidable partial digestion by trypsin.

The physiological significance of p28 on leukaemogenesis is not clear yet. We have found that p28 is associated with protein kinase activity specific to serine or threonine residues (Kobayashi et al., 1984c), but it is necessary to investigate whether this activity is intrinsic to p28. Further studies are underway to elucidate the physiological function of p28.

This work was supported by Grants-in-Aid from the Ministry of Education, Science and Culture of Japan.

REFERENCES


(Received 10 September 1985)