Volume 15 Number 13 1987 Nucleic Acids Research

Transcription of a human transposon-like sequence is usually directed by other promoters

K.Eric Paulson, A.Gregory Matera, Niren Deka and Carl W.Schmid

Department of Chemistry, University of California, Davis, CA 95616, USA

Received March 17, 1987; Accepted June 1, 1987

ABSTRACT

The transcriptional activity of a human transposon-like family of repeats, called the THE-1 family, has been studied in cell culture and in human tissue. Both strands of THE-1 are present in several discrete length poly A plus RNAs. Primer extension studies and the structures of cDNA clones show that these THE-1 transcripts are usually the product of other transcription units. The THE-1 LTR provides the polyadenylation processing site for two transcripts, which result from upstream non THE-1 promoters. Yet another transcript, containing an internal THE-1 element in the probable sense orientation, is greatly enriched in a polysomal size fraction.

INTRODUCTION

Transposable genetic elements, such as yeast Ty elements, Drosophila copia-like elements and structurally related mammalian proretroviruses usually code for several gene products including activities that are required for their own mobility (1-4). Transcriptional control signals contained within the LTRs insure that upon insertion at a new genomic position, transposons retain their transcriptional competence (4,5). Although the two LTRs are nearly identical sequences, the 5' LTR typically provides enhancer and promoter functions whereas the 3' LTR supplies 3' processing signals (5,6). Given these transcriptional control elements, the insertion of a transposon often affects the expression of nearby genes. For example, the insertion of the yeast Ty element at some loci causes constituitive overproduction of adjacent gene products and at other loci silences gene expression (6,7,8). Proretroviral LTRs can promote downstream transcription, enhance upstream transcription and cause termination of transcription (5).

The distinction between mammalian proretroviruses and lower eukaryotic transposons, such as copia and Ty, is somewhat arbitrary (9). These two classes of sequences have very similar structures, code for similar protein activities and utilize common pathways of transposition (1,2,3,4,9,10,12). Transposons in lower eukaryotes are often present as middle repetitive sequences and usually lack an identified extracellular viral form. Such transposons may be described as defective retroviruses.

The human THE-1 family of repeated sequences consists of 10,000 members which share a 2.3 kb consensus sequence (13). The consensus structure resembles that of lower eukaryotic transposons and mammalian proretroviruses in several details. However, the THE-1 sequence is not related to that of known retroviruses nor does it have an identified extracellular form. Furthermore, transposition of its member sequences has not yet been observed. THE-1 repeats might encode their own products or affect the expression of other genes. Here we examine the transcription of THE-1 in various human cell lines and tissues.

We find that both orientations of THE-1 are transcribed. Hence, the major transcript does not necessarily identify the sense strand. To avoid prejudging this question, we arbitrarily denote the transcriptional orientation of Paulson et al. (13) as the left strand and the other orientation as the right strand. The polyadenylation of two distinct THE-1 containing transcripts suggests that the left strand is indeed the sense strand for putative THE-1 gene products. Interestingly, another transcript containing this orientation of THE-1 as a readthrough transcript is greatly enriched in polysomal size particles.

MATERIALS AND METHODS

Maintenance of Cultured Cells:

HeLa cells were maintained in suspension in Joklik's modified minimal essential media and 10% calf serum (GibCo) at a cell density of 2-8 x 10^5 cells/ml.

Isolation of RNA:

HeLa cytoplasmic RNA was isolated using a modification of the method of Weiner (14). HeLa cells were washed twice in an

equal volume of ice-cold RSB (150 mM NaCl, 1.5 mM $MgCl_2$, 1 mM KH_2PO_4, 10 mM Tris-HCl pH 7.6). The cytoplasmic fraction was separated from nuclei by resuspending the final cell pellet at a density of 2 x 10^7 cell/ml in ice-cold RSB+NP-40 (RSB, 0.65% NP-40), followed by incubating on ice for 15 min. Nuclei were pelleted at 6000 x g, 5 min., 4°C. The cytoplasmic supernatant was removed and extracted with an equal volume of 0.2 M Tris-HCl pH 7.6, 5 mM EDTA, 2% SDS, followed by sequential extractions with 1 volume phenol, phenol/chloroform, and chloroform. Nucleic acids were ethanol precipitated from the aqueous phase, then RNA was purified by pelleting in 5.7 M CsCl at 100,000 x g, 20°C, 24 hr.

Total RNA from frozen tissue was isolated using a modification of the method described by Maniatis et al. (15). Frozen tissue (5g at -70°C) was crushed to powder, then added to 20 ml 4 M guanidinium thiocyanate (15) and dispersed with a dounce homogenizer. The suspension was heated to 60°C and DNA sheared using an 18 gauge needle. The solution was extracted with 25 ml phenol (60°C) while continuing to shear DNA. 25 ml 0.1 M NaAc, 10 mM Tris-HCl pH 7.6, 1 mM EDTA was added to the emulsion followed by 25 ml chloroform. The aqueous supernatant was extracted sequentially with one volume phenol/chloroform, then one volume chloroform. Nucleic acids were ethanol precipitated and RNA was purified by pelleting in 5.7 M CsCl at 100,000 x g, 20°C, 24 hr. and was fractionated on oligo (dT) cellulose (15).

Preparation of Polysomal RNA:

200 ml HeLa cells at 6 x 10^5 cell/ml were chilled by pouring over 100 ml Earle's Salts on ice containing 10 µg/ml cycloheximide. Cells were pelleted by centrifugation and washed three times in ice-cold Earle's Salts with cycloheximide. The cell pellet was resuspended in 4 ml lysis buffer (100 mM KCl, 1 mM DTT, 3 mM $MgCl_2$, 10 mM HEPES pH 6.8, 10 µg/ml cycloheximide) then lysed by adding 200 µl 10% NP-40 followed by a dounce homogenization. Nuclei were pelleted at 7000 x g for 10 min. at 4°C. The supernatant was removed and 1 ml layered over a 12.5 ml 18-42% sucrose gradient in lysis buffer in a Beckman SW 40 rotor. Polysomes were fractionated at 30,000 rpm for 2 hr at 4°C. The

sucrose gradient was divided into the following three fractions using a 280 nm flow cell: 1-2 ribosomes, 3-7 ribosomes, +7 ribosomes. Proteins were removed by consecutive extractions with 1 volume phenol, phenol/chloroform, and chloroform. RNA was then ethanol precipitated.

Primer extension Analysis of RNA:

A 16 base oligonucleotide complementary to the sequence of cDNA α (data not shown) and an 18 base oligonucleotide complementary to cDNA 7 (sequence not shown) was chemically synthesized using a Systec automated DNA synthesizer. The sequence of the oligonucleotides are: cDNA α complement; 5' GGCCAGGGTAGAATGA 3', cDNA 7 complement; 5' GTATCTTTTCAGCAACGC 3'. The cDNA α complement differs by one nucleotide from the THE-1A sequence (13). Primer extension was performed on both methyl mercury (5 mM) treated and untreated RNAs with no difference in results. 0.5 pmole of T4 polynucleotide kinase labelled primer (15) was hybridized to 3 µg poly A+ RNA at either 45°C or 48°C in 300 mM NaCl, 10 mM Tris-HCl pH 7.5, 1 mM EDTA for 1.5 hrs. The RNA was precipitated with 2.5 volumes ethanol then resuspended in 25 µl reverse transcriptase buffer (60 mM KCl, 25 mM Tris-HCl pH 8.3, 10 mM $MgCl_2$, 30 mM 3-mercaptoethanol, 1 mM deoxynucleotide triphosphates) containing 40 units RNasin (Promega Biotec). The reaction was initiated with 15 units AMV reverse transcriptase (Molecular Genetics Resources) and incubated at 42°C for 1 hr. The reaction was stopped by the addition of 1 volume phenol/chloroform. The aqueous phase was precipitated with ethanol, resuspended in formamide-dye mix and run on a polyacrylamide/7M urea sequencing gel.

Gel Electrophoresis and Blot Hybridization:

RNA was fractionated in a 1% agarose gel containing 6% formaldehyde in 20 mM $NaHPO_4$ pH 7.0 and transferred to nitrocellulose (15). Filters were prehybridized in 3xSSC, 5 x Denhardt's, 0.1% SDS, 200 µg/ml denatured salmon sperm DNA and 500 µg/ml yeast RNA at 60°C. Hybridization was performed with a nick translated restriction fragment (16) or with an RNA probe prepared using the SP6 transcription system (17). Filters were typically washed in 1 X SSC, in 0.1% SDS at 60°C.

Selection of cDNA Clones:

Size selected subfractions of the cDNA library prepared by Okayama and Berg (18) were transformed into E. Coli strain HB 101 or JM83. Plate lifts were prepared as described by Grunstein and Hogness (19) and screened for THE-1 containing sequences using a nick-translated restriction fragment of THE-1A.

Preparation of Probes for Northern Analysis:

Single copy sequences flanking the THE-1 element in each of the cDNAs were subcloned into SP6 transcription vectors such that the transcribed strand was complementary to the RNA in question. Regarding cDNA α, the probe was derived from a 0.4 kb Pst/Hae fragment on the 5' end of the genomic locus of the transcription unit (unpublished data) and subcloned into pSP64 (Promega Biotec). For cDNA γ, a 0.8 kb Bam/Pvu fragment from the 5' end was subcloned into pUC and then liberated as a Hind/Eco fragment for further subcloning into pSP65. The probe for cDNA δ was derived from a subclone of a 1.45 kb Dra fragment from the 3' flank. An internal 0.6 kb Bam fragment was subcloned into pUC and then mapped to determine its orientation. This sequence was then liberated as a 0.6 kb Hind/Eco fragment and subcloned into pSP65.

RESULTS

1) Both Strands of THE-1 Are Represented in Discrete Length RNAs

The two transcriptional orientations of THE-1 were subcloned into the SP6 transcription system to generate hybridization probes (17). These probes were used in Northern blot hybridization studies of poly A plus RNAs from several cell lines and tissues. The left strand of THE-1, i.e. the published orientation of Paulson et al. (13), corresponds to a major 5 kb transcript in all tissues and cell lines investigated, Fig. 1. We refer to these RNAs according to their length and orientation. In addition to the major 5 kb-L (left) transcript in HeLa cells, we detect at least two minor RNAs: 2 kb-L and 0.8 kb-L transcripts, Fig. 1. However, the complementary right strand is also represented in poly A plus RNAs, Fig. 1. There are two major right strand RNAs in HeLa, a 4 kb-R and a 1.9 kb-R transcript as

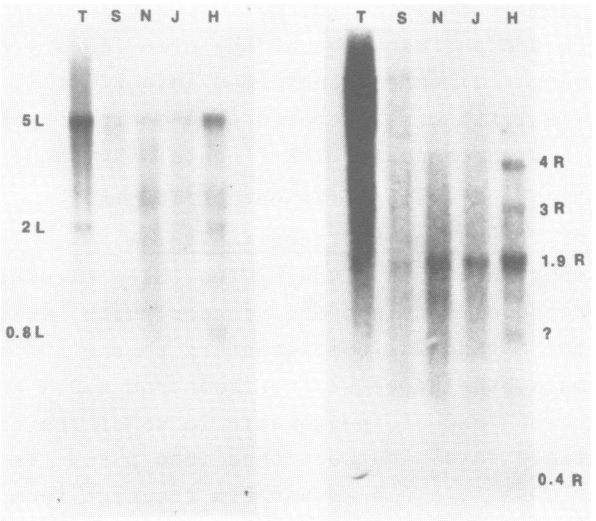


Fig. 1. Northern blots of human RNAs probed with both transcriptional orientations of THE-1. 3 μg poly A^+ RNA from human thyroid tissue (T), spleen tissue (S), neuroblastoma cell line (N), Jurkat lymphoblastoid cell line (J), and HeLa cells (H) were electrophoresed on denaturing agarose gels and transferred to nitrocellulose as described in Methods. Duplicate transfers were hybridized with two SP6 RNA probes corresponding to the left (L) or right (R) transcriptional orientation of the 0.51 kb Hae III fragment from THE-1A (13). The left (L) transcriptional orientation corresponds to the transcript sense reported by Paulson et al. (13). Numerical values identify the approximate lengths of these RNAs in kilobase units.

well as at least two minor right strand RNAs, 3 kb-R and 0.4 kb-R. Careful inspection of the RNA blots reveals additional minor transcripts (Fig.1).

The 4 kb-R transcript, present in HeLa cells, is not readily detectable in other tissues or cell types, Fig. 1. This transcript is barely detectable in cells grown in 5% serum but is more abundant in cells grown on 10% serum (data not shown). The abundance of this transcript does not depend on the density of the cell culture (1.2 to 8 x 10^5 cells/ml; data not shown). However, the detection of the various transcripts reported in Figure 1 does depend on hybridization conditions, in particular the use of high specific activity SP6 probes, Fig. 1. Differ-

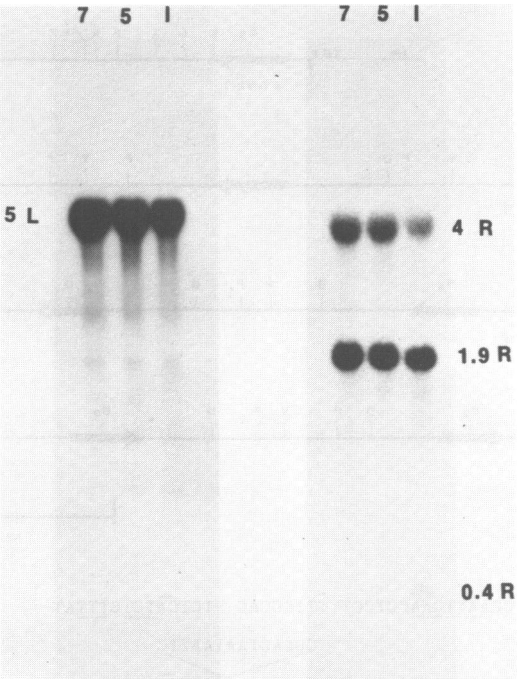
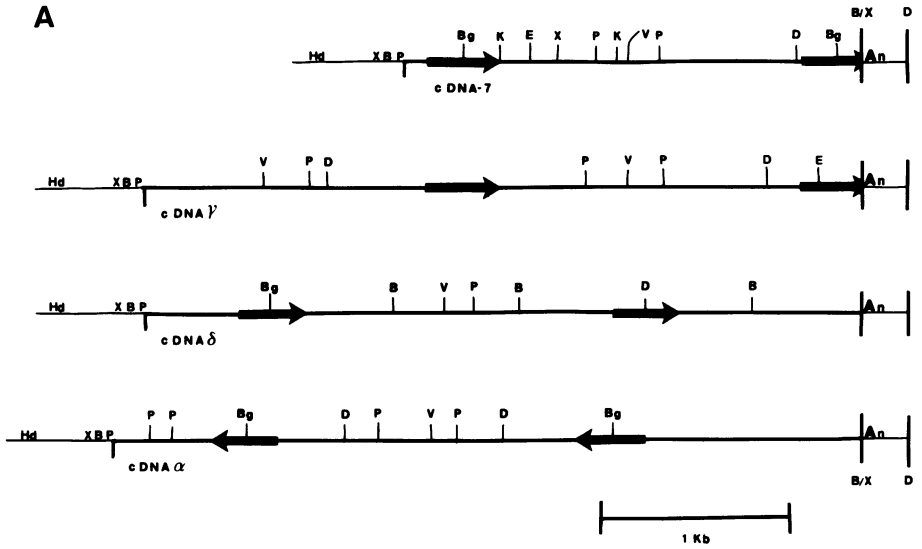


Fig. 2. Cosedimentation of THE-1 RNAs with HeLa polysomes. Northern blots of 10 µg HeLa polysomal RNA representing 1-2 ribosomes (1), 3-7 ribosomes (5), or greater than 7 ribosomes (7) were hybridized with the SP6 probes described in Figure 1. The transcriptional sense is defined in Figure 1.

ences in growth conditions and hybridization probes account for the additional transcripts detected in this study (13).

2) THE-1 Containing RNAs Cosediment with Polysomes

The presence of both strands of THE-1 sequences on different length RNAs raises the question of whether either strand or both strands encode a peptide. Except for the 2kb-L RNA, all major HeLa transcripts, 5 kb-L, 4 kb-R and 1.9 kb-R RNAs, cosediment with polysomes, Fig. 2. We have not determined whether the THE-1 RNAs are associated with polysomes or with other RNPs which merely cosediment with polysomes. None of the minor THE-1 RNAs are enriched in the polysomal fraction, whereas one minor RNA, the 0.4 kb-R transcript, is preferentially associated with monosome size particles, Fig. 2. cDNA sequence studies reported



B

5' AGTCCAATTGAACCTCCTTTTTCCCAG--TCTCATGTCTTTAT 3' THE 1-A

GGAACTAATAATTC

5' AAGCTAATTAAAC-TG-TTTTGTTGAAAGTCTCA$(A)_N$.... 3' THE 1-C

TAA

5' AGTCTAATTAAACCTCTTTTTATTGATAATCTCA$(A)_N$.... 3' cDNA 7

5' AGTCC-AATAAACTTCTTTCTTTTT------GTA$(A)_N$.... 3' cDNA Y

Fig. 3A. Orientation and position of THE-1 elements within cDNA clones. The 5' and 3'ends of cDNA inserts are assigned by restriction cleavage sites in the Okayama-Berg vector: Hd, X, B and P (Hind III, XhoI BamHI and PstI, respectively) map at the 5' positions indicated and B, X and D (DraI) map at at the 3' positions indicated. Vector sequences are shown by a thin line and insert sequences by a thicker line. The arrows denote positions and orientations of the THE-1 LTRs with respect to the transcriptional orientation reported by Paulson, et al. (13) "An" schematically depicts the 3' polyA tail. Subcloned regions of THE-1 were used to map the orientation and position of the THE-1 element. These subclones (13) are: an LTR specific probe, a 0.51 kb Hae III fragment mapping at the 3' end of THE-1, a 0.38 kb PstI fragment mapping in the middle of THE-1 and a 0.55 kb PvuII fragment mapping near the THE-1 5' end. All four hybridization probes were used in mapping cDNA 7 and cDNA α; the Hae III" subclone was not used in mapping cDNA γ and cDNA δ. The precise 5' end of cDNA 7 was determined by sequence analysis, as were the 3' ends of cDNA 7 and cDNA γ. Additional restriction enzymes used include BglII, (Bg), KpnI(K), EcoR1(E), PvuII(V).

B) Base sequence comparison of genomic clones THE-1A (13) and THE-1C (24) with cDNA 7 and cDNA γ. Under and overlining designate sequence elements that serve in identifying and processing the 3' ends of transcripts (20). Homology between the four sequences continues through their respective 2 kb THE-1 elements positioned 5' to the sequences depicted here.

below identify specific transcripts in these polysomal size RNA fractions.

3) Some THE-1 Transcripts Result from Other Transcription Units

THE-1 sequences hybridize to several RNAs which are longer than the 2.3 kb consensus sequences, Fig. 1. To identify the origin of these additional non THE-1 consensus sequences, cDNA clones containing THE-1 sequences were selected from a cDNA library of poly A plus RNA from SV40 transformed fibroblasts (18). Ten of thirteen selected cDNA clones were derived from a common transcript consisting of the right strand of THE-1 linked to an additional 1 kb of unique 3' flanking sequence (Fig. 2A). These ten cDNA clones have identical restriction patterns on their 3' ends, demonstrating that they are the product of a unique transcription unit. This interpretation is confirmed by genomic blot hybridization results (data not shown). cDNA-α, containing a 4 kb insert, is the longest representative of this series of ten clones, Fig. 3A.

The left strand of THE-1 is represented by three clones: cDNA 7, cDNA γ and cDNA δ, Fig. 3A. The THE-1 element is entirely contained within the cDNA δ sequence whereas the THE-1 element terminates the 3' ends of both cDNA 7 and cDNA γ, Fig. 2A. Of particular interest, the 3' LTRs in both cDNA 7 and cDNA γ served as polyadenylation sites, Fig. 3B. These THE-1 elements evidently inserted into other transcription units and now affect the processing and polyadenylation of the mutated transcript. The poly A addition site in cDNA 7 is the consensus site, CA, which is preceded by AATTAAA, consisting of two known variants of the polyadenylation signal, AATTAA and ATTAAA (Fig. 3B, 20). The polyadenylation signal AATAAA is positioned 16 nt away from the polyadenylation site in cDNA γ. A short run of T's as well as the sequence TGT are present downstream from the poly A addition site in some genomic LTR sequences (THE-1A, Fig. 3B). These

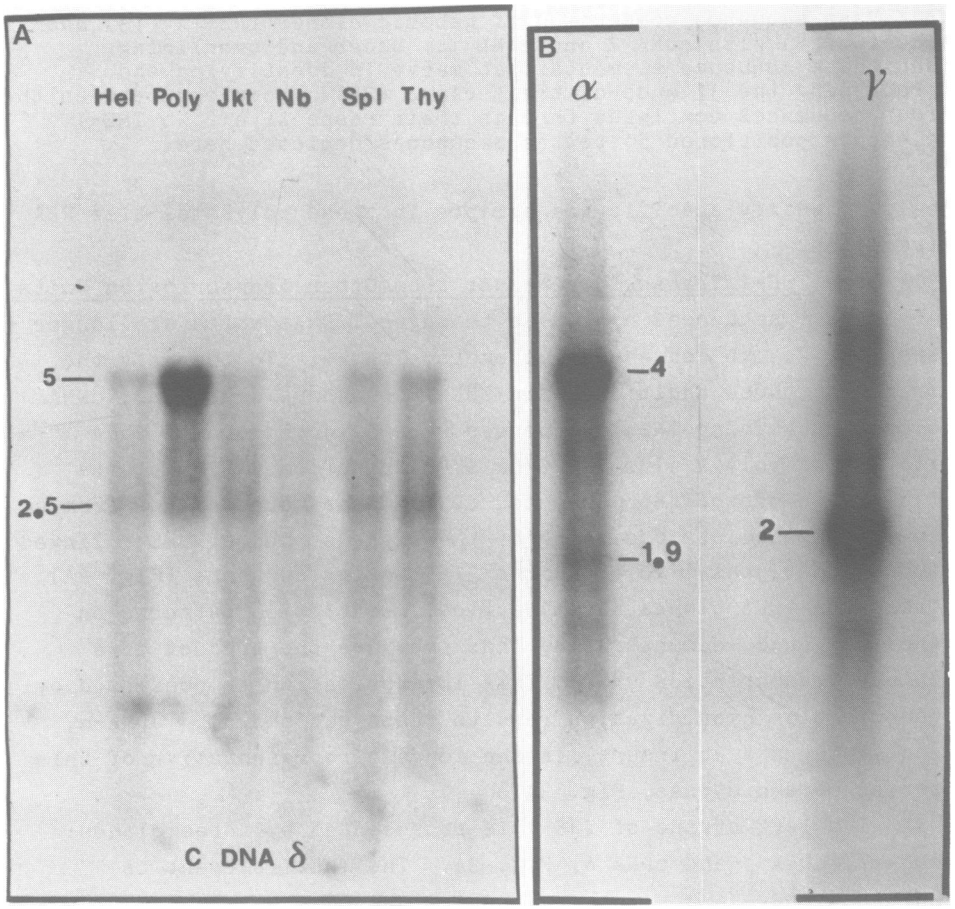


Fig. 4. Northern blot analysis of cDNA subclones. Unique sequences derived from the cDNAs of Figure 3A were mapped and used to construct SP6 subclones (Materials and Methods). A) The resulting SP6 cDNA δ transcript was hybridized to RNAs from HeLa (Hel), a polysomal size fraction between 3 and 7 ribosomes (Poly), Jurkat (Jkt), neuroblastoma (Nb), spleen (Spl) and thyroid carcinoma (Thy) cells and tissues (Materials and Methods). B) The resulting SP6 cDNA α and cDNA γ transcripts were as indicated hybridized to HeLa poly A+ RNA. Exposure of all of these blots is typically 24 hours.

short sequence elements are also associated with 3' RNA processing (20). Interestingly, the 3' LTR of another genomic sequence, THE-1C, terminates in a run of A's at the same position as the two cDNA clones (Fig. 3B). The 3' end structure of THE-

1C is typical of processed RNA pseudogenes (21) and indicates that the THE-1 polyadenylation signal is functional on three separate transcripts in two different cell types.

We have not investigated the transcription of THE-1 in the cell line used to construct the Okayama-Berg library. Consequently, these cDNA clones do not necessarily correspond to the RNA bands observed in Figure 1. Northern blot analysis using unique sequence subclones from the cDNAs was used to investigate this issue.

4) Several cDNA Clones Correspond to HeLa Transcripts

Blot hybridization to genomic DNA was employed to identify single copy regions in cDNAs α, γ and δ, each of which was found to be the product of a single transcription unit. An SP6 subclone derived from the 3' unique sequence in cDNA δ (Fig. 3A) hybridizes to both a 5 kb-L and 2.5 kb-L transcript in HeLa cells and several other human cell lines and tissues, Fig. 4A. As previously mentioned, cDNA δ is the product of a unique transcription unit, so that these two products must result from differences in processing. The longer transcript, 5 kb-L, approximates the length of cDNA δ and is presumed to have the same structure. In agreement with this interpretation, the 5 kb δ transcript is identical in length to the 5 kb-L transcript which is homlogous to THE-1 repeats in HeLa polysomal size fractions, Fig. 2. The 2.5 kb cDNA δ transcript does not correspond to any THE-1 containing transcripts (Figs. 1,2), and presumably results from the removal of THE-1 sequences by processing of the longer δ transcripts. It is noteworthy that compared to their cellular abundances, the 5 kb-L transcript is greatly enriched in polysomal RNA fractions relative to the 2.5 kb transcripts (Fig. 4A and unpublished). We speculate that the THE-1 element is responsible for the enrichment of this transcript in the polysomal size fractions (Discussion).

A hybridization probe derived from an SP6 subclone of the unique 5' flanking sequence of cDNA α hybridizes to a 4.3 kb-R transcript and less intensely to a 1.9 kb-R transcript in HeLa RNA, Fig. 4B. Additional blots show that the same 4.3 kb-R transcript as well as the 3 kb-R transcript are present in thyroid tissue. Both the 4.3 kb and 1.9 kb α transcripts are

also detected in HeLa polysomal size fractions. These lengths coincide exactly with the 4.3 kb-R and 1.9 kb-R THE-1 containing transcripts, which are present in both total HeLa RNA (Fig. 1) and polysomal size fractions (Fig. 2). The ~4 kb cDNA α is probably a full length cDNA to this 4.3 kb-R transcript as demonstrated by both sequence analysis (unpublished) and primer extension data, which is reported below. In agreement with this prediction, a hybridization probe derived from the 3' unique region of cDNA α hybridizes to the same 4.3 kb-R transcript as does the 5' α probe. The structure of the 1.9 kb-R transcript is not known, but, based on these assignments, it should include regions from both the 5' unique region of cDNA α and THE-1 sequences.

An SP6 hybridization probe derived from the unique 5' region of cDNA Υ (Fig. 2A) hybridizes primarily to a 1.9 kb transcript, Fig. 4B. Additional blot experiments also show a 4.4 kb transcript is present in HeLa polysomal size fractions. Although the length (1.9 kb) and orientation of the 1.9 kb Υ transcript resemble the THE-1 containing 2 kb-L transcript (Fig. 1, 2) we believe that they are different RNAs. The 1.9 kb Υ transcript is present in polysomal size fractions whereas the 2 kb-L THE-1 transcript (Fig. 1) is not detectable in polysomal size fractions (Fig. 2).

A 40 bp subcloned restriction fragment flanking the 5' end of THE-1 sequences in cDNA 7 (Fig. 3A) does not hybridize to any HeLa RNA bands (data not shown). Either the transcription unit containing cDNA 7 is inactive in HeLa cells or the unique cDNA 7 hybridization probe is too short to detect its complementary RNA.

In summary, the cDNA clones isolated from the Okayama-Berg library account for at least two and probably four of the THE-1 containing transcripts detected in HeLa RNA: the 5 kb-L, 4 kb-R and possibly the 3 kb-R and 1.9 kb-R transcripts. These assignments raise the question of whether any THE-1 RNAs are independently promoted by the THE-1 LTR.

5) Lack of THE-1 Self-Promoted Transcripts in HeLa RNA

Primer extension by reverse transcriptase was employed to detect transcripts which might initiate within the THE-1 LTRs. Two primers were selected for this purpose: a 16 nt primer

Fig. 5. Mapping of THE-1 containing transcripts by primer extension. An oligonucleotide primer complementary to cDNA α was hybridized to HeLa poly A+ RNA and primer-extended products were synthesized and analyzed as described in Methods. The two lanes correspond to two hybridization temperatures 45°C and 48°C. Arrows at 98 nt and 670 nt indicate major primer-extended products. The gel was exposed to XAR 5 film for 24 hrs at -70°C with an intensifying screen. A lighter exposure reveals the 670 nt band as more intense than the other high molecular weight reverse transcripts. A similar experiment with a primer complementing the cDNA 7, Y and δ orientation (Materials and Methods) did not yield discrete length reverse transcripts.

complementing the sequence of cDNA α near the junction of the 5' LTR and internal sequence and an 18 nt primer complementing the THE-1 consensus left strand near the junction of the 5' LTR and THE-1 internal sequence (Materials and Methods). Extension of the 16 nt primer results in a minor 98 nt band and several prominant bands clustered about 670 nt. (Fig. 5) The longer reverse transcripts correspond to the distance from the priming site to the 5' end of cDNA α (Fig. 3A, sequence data not shown). The minor 98 nt band might result from right strand transcripts,

which initiate within the LTR, or might be incomplete reverse transcript of longer RNAs containing THE-1 elements.

Polyadenylation of THE-1 containing transcripts (Fig. 2B) suggests that the left strand is more likely to encode THE-1 gene products. However, extension of the 18 nt consensus sequence primer by reverse transcriptase does not yield any discrete length reverse transcripts (data not shown). Because the 5' ends of cDNA γ, δ are far from the priming site (>1 kb), it is not surprising that their reverse transcripts are undetected in this experiment. The absence of a reverse transcript corresponding to transcriptional initiation within the LTR is a negative result and therefore provides inconclusive evidence that the LTR does not promote transcription of the left THE-1 strand in HeLa cells. In agreement with this finding, we have not identified any cDNA clones corresponding to RNAs which initiated within the LTR. Most of the transcriptional activity of THE-1 containing sequences detected here results from other promoters.

DISCUSSION

Transposon-mediated rearrangements in lower eukaryotes can affect the transcription of genes either within the transposable element or adjacent to it (Introduction). This initial investigation reveals a number of transcriptional properties of THE-1 sequences. First, several different discrete length transcripts contain the left or right THE-1 sequence in human cell lines and tissues. Second, both transcriptional senses are transcribed by virtue of their insertion into other transcription units. Third, three major THE-1 transcripts, 5 kb-L, 4 kb-R and 1.9 kb-R RNA transcripts in HeLa cells cosediment with polysomes. Each of these transcripts include other non-THE-1 sequences. Fourth, the left strand of THE-1 contributes a polyadenylation signal to several upstream transcription units.

It is curious that both strands of THE-1 are represented on poly A plus transcripts and even more surprising that both transcriptional orientations may be associated with polysomes. Transposons and retroviral RNAs normally have only one sense strand and presumably this generalization applies to human THE-1

repeats (1,3,4). The antisense strand of Drosophila copia-like elements constitutes a minor transcription product giving some precedence for finding both transcriptional orientations of human THE-1 sequences (22).

We previously identified the left THE-1 strand as the major transcript (13). The 3' LTR of the left strand contains a well defined and functional polyadenylation site suggesting that it is indeed the sense strand. This strand contains several potential, albeit short, open reading frames (13,24). Because the human THE-1 family is polymorphic, base sequence determinations of randomly selected genomic clones do not necessarily identify the functional open reading frame(s) and the corresponding sense strand. Sequence analysis of the cDNA clones reported here will reveal whether open reading frames exist in these transcripts.

The enrichment of the 5 kb-L δ transcript relative to the 2.5 kb δ transcript in polysomal size fractions is noteworthy. The δ transcript results from an upstream promoter and presumably its unique sequences encode a protein, thus accounting for the presence of the 2.5 kb δ transcript in polysome size fractions. The 5 kb δ transcript might additionally code for THE-1 gene products and these products may enrich its polysomal abundance. Polycistronic eukaryotic RNAs offer the ribosome a choice of initiation codons to translate, resulting in either altered or entirely diffierent gene products (25). The presence of the δ as well as the α and γ transcripts in polysomal size fractions raises the possibility that other promoters might ultimately cause expression of THE-1 gene products.

The insertion of THE-1 repeats into other transcription units may affect transcriptional expression as discussed in the Introduction. In the case of both cDNA-7 and cDNA γ, THE-1 repeats affect RNA 3' end processing. Analogously, an inserted mouse B2 Alu short interspersed repeat supplies the polyadenylation signal to an upstream transcription unit (23).

The human THE-1 family of transposon-like sequences consists of about 10,000 members as well as an additional 10,000 solitary LTR's. The approach used here only detects transcription units containing THE-1 sequences and does not test the effects of THE-1

elements on adjacent transcription units. Accordingly, THE-1 repetitive sequences are likely to have other effects on transcription in addition to those readily observed here.

ACKNOWLEDGEMENTS

We thank Dr. Roger Duncan for generously assisting us with the polysome preparation, Drs. Mark Rush and Ravi Misra for providing RNAs from Jurkat and Neuroblastoma cells, Dr. Steve Hinricks for providing tissue samples, Dr. H. Okayama for providing cDNA libraries, and Ms. F. Wong and C. Willard for DNA sequence analysis. This research was supported by USPHS grant GM21346.

REFERENCES

1. Clare, J. and Farabaugh, P. (1985) Proc. Natl. Acad. Sci. 82, 2829-2833.
2. Saigo, K., Kugimiya, W., Matsuo, Y. Inouye, S., Yoshioka, K. and Yuki, S. (1984) Nature 312, 659-661.
3. Mount, S. and Rubin, G. (1985) Mol. Cell. Biol. 5, 1630-1638.
4. Varmus, H. E. (1983) Chapter 10 "Retroviruses" p. 411-503 in Mobile Genetic Elements, J. A. Shapiro, Editor, Academic Press, Orlando, FL.
5. Temin, H. M. (1982) Cell 28, 3-5.
6. Roeder, G. S. and Fink, G. R. (1983) Chap. 7 "Transposable Elements in Yeast," 299-328 in Mobile Genetic Elements J. A. Shapiro Ed., Academic Press, Orlando, FL.
7. Young, T., Williamson, V., Taguchi, A., Smith, M., Sledziewski, A., Russell, D., Osterman, J., Deris, C., Cox, D. and Beier, D. (1982) in Genetic Engineering of Microorganisms for Chemicals, Hollander et al., Editor, p. 335-361, Plenum, NY.
8. Rothstein, R. and Sherman, F. (1979) Genetics 94, 891-898.
9. Baltimore, D. (1985) Cell 40, 481-482.
10. Shiba, T. and Saigo, K. (1983) Nature 302, 119-124.
11. Arkhipova, I. R., Mazo, A. M., Cherkasova, V. A., Gorelova, T. V., Schuppe, N. G. and Ilyin, Y. V. (1986) Cell 44, 555-563.
12. Boeke, J. D., Garfinkel, D. J., Styles, C. A. and Fink, G. R. (1985) Cell 40, 491-500.
13. Paulson, E. K., Deka, N., Schmid, C. W., Misra, R., Schindler, C. W., Rush, M. G., Kadyk, L. and Leinwand, L. (1985) Nature 316, 359-361.
14. Weiner, A. (1980) Cell 22, 209-218.
15. Maniatis, T., Fritsch, E. F. and Sambrook, J. in 'Molecular Cloning: A Laboratory Manual" (1982) Cold Spring Harbor, N.Y.
16. Rigby, P. W. J., Dieckmann, M. Rhodes, C. and Berg, P. (1977) J. Mol. Biol. 113, 237-251.
17. Melton, D. A., Kreig, P. A.,v Rebagliati, M. R., Maniatis, T., Zinn, K. and Green, M. R. (1984) Nucl. Acids Res. 12, 7035- 7056.
18. Okayama, H. and Berg, P. (1983) Mol. Cell. Biol. 3, 280-289.

19. Grunstein, M. and Hogness, D. S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961-3966.
20. Birnstiel, M. L., Busslinger, M. and Strub, K. (1985) Cell 41, 349-359.
21. Sharp, P. A. (1983) Nature 301, 471-472.
22. Georgiev, G. P., Ilyian, Y. V., Chemliavskaite, V. G., Ryskov, A. P., Kramerov, D. A., Skryabin, K. G., Krayev, A. S. Lukanidin, E. M. and Grigoryan, M. S. (1981) Cold Spring Harbor Symp. Quant. Biol. 45, 641-654.
23. Kress, M., Barra, Y., Seidman, J. G., Khoury, G. and Jay, G. (1984) Science 226, 974-977.
24. Willard, C.W. (1987) Ph.D. Thesis, in preparation.
25. Kozak, M. (1986) Cell 47, 481-483.