Binding of the ubiquitous nuclear transcription factor YY1 to a *cis* regulatory sequence in the human LINE-1 transposable element

Kevin G.Becker, Gary D.Swergold[1,+], Keiko Ozato and Ronald E.Thayer[1,*]

Laboratory of Molecular Growth Regulation, National Institute of Child Health and Human Development and [1]Laboratory of Biochemistry, National Cancer Institute, NIH, Bethesda, MD 20892, USA

The first step of the currently favored model for the mechanism of transposition of the human LINE-1 element involves the synthesis of full length LINE-1 mRNA. Previous work demonstrated that the 5'-terminal 100 base pairs of the human LINE-1 element (L1Hs) has an important role in regulating it's expression. Here we report further deletion analysis revealing the presence of a *cis* regulatory element overlapping the region between base pairs +12 and +18. Oligonucleotides containing this sequence form a specific complex with a nuclear protein extracted from NTera2D1 and Jurkat cells, and with recombinant YY1 produced in *E. coli*. The complex is competed by YY1 binding sites found in other genes, and is ablated by anti-YY1 serum. These results suggest that YY1 is involved in the regulation of L1Hs transcription and therefore transposition.

INTRODUCTION

The haploid human genome is estimated to contain 4,000 near-full length LINE-1 (L1Hs) elements (1, 2). An as yet undefined number of these elements are capable of transposition. New insertions of L1Hs elements into expressed regions of the genome may result in mutational events. Examples of both somatic (3, 4) and presumed germ-line (5, 6) insertional mutagenesis events have been reported. The rate of L1Hs induced insertional mutagenesis, as well as the overall rate of L1Hs transposition, have not yet been determined but presumably depend on the regulation of expression of the elements.

The consensus human L1Hs (Fig. 1A) is a 6.2 kb sequence that contains two open reading frames of approximately 1 and 4 kbp on one strand, but no long terminal repeats (7). The currently favored model for the mechanism of L1Hs transposition involves (i) synthesis of full-length polyadenylated L1Hs RNA directed by regulatory sequences located within the element, (ii) translation of this RNA into proteins, one or more of which are required for transposition, (iii) reverse transcription of the RNA by an L1Hs-encoded enzyme, and (iv) insertion of the cDNA into staggered chromosomal breaks.

Evidence in support of this model has been accumulating over the past several years. Skowronski and co-workers characterized full-length, polyadenylated L1Hs RNA in the human terato-carcinoma cell line NTera2D1 (8, 9). Swergold (10) identified regulatory sequences located within the L1Hs 5' untranslatable region capable of promoting the transcription of abundant L1Hs RNA in a cell-specific manner. Dombroski *et al.* reported the

isolation of a full-length genomic L1Hs element (L1.2B) likely to be the source of a truncated element newly inserted into the factor VIII gene of a patient with hemophilia A (11). L1.2B was reported to be a member of a small class of L1Hs elements, several of which contain 2 fully open long reading frames (ORFs) in contradistinction with most genomic L1Hs elements in which the ORFs are disrupted by rearrangement, random stop codons, and other mutations. L1.2A, an allele of L1.2B, was found to differ from L1.2B by only 3 base pairs, all within the second ORF.

Polypeptides translated from L1Hs RNAs have recently been identified. p40, a leucine-zipper-motif containing phosphoprotein, encoded by the L1Hs first ORF, has been reported in human teratocarcinoma and several other tumor cell lines (12, 13)(8, 9, Thayer *et al.* in preparation). Mathias *et al.* fused a short fragment of the second ORF of the yeast Ty transposable element to the second ORF of L1.2A (14). The L1Hs sequence contained in this construct directed the expression of reverse transcriptase activity upon introduction into *Saccharomyces cerevisiae*.

Initial characterization of the L1Hs transcription regulatory sequences revealed that the first hundred bases were most critical for the expression of a L1Hs-*lacZ* construct in NTera2D1 cells (10). Deletion of these bases resulted in an approximately 300-fold reduction in the expression of the beta-galactosidase reporter gene. Mathias reported the ability of the first 31 base pairs of L1.2A to bind proteins in NTera2D1 and HeLa cell nuclear extracts (15). Minakami *et. al.* (16) showed that a widely
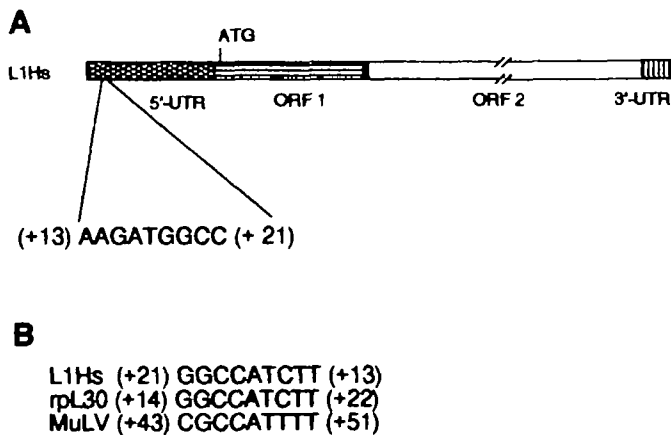
Figure 1. Diagram of the LINE-1 element (L1Hs) and sequence of the YY1 binding site. (A) Schematic diagram of L1Hs showing the 5' (wavy lines) and 3' (vertical lines) untranslated regions (UTR) as well as the ORF1 (horizontal lines), ORF2 (open) and inter-ORF region (black). The position of the YY1 binding site is shown. (B) Comparison of the YY1 site found in the L1Hs 5'UTR with the sites from the mouse ribosomal protein, rpL30 (40), and murine Moloney Leukemia Virus (MULV) (41). The nucleotide numbers of The L1Hs and rpL30 sites are numbered relative to the RNA transcription initiation site; the MuLV site is numbered relative to the 5' end of the LTR.

distributed nuclear protein protected the sequence between +3 and +26 of the L1Hs element pCGL1-1 in DNAse I footprinting experiments. Mutation of base pairs +10 to +17 reduced the activity of a pCGL1-1 (+1 to +155) fusion gene 25 fold.

In this paper we report on the further characterization of a regulatory sequence within the first 30 bases of the L1Hs sequence. The recently identified transcription factor YY1, also called UCRBP, δ, and NF-E1, binds to cis regulatory elements that differ substantially in sequence (17, 18, 19, 20). We show that the L1Hs sequence between base pairs 13 and 21 specifically binds YY1 protein, and that deletion of this sequence decreases the expression of L1Hs 5-fold in NTera2D1 cells.

RESULTS

Deletion analysis of the L1Hs promoter

Plasmid p1LZ (fig. 2A) was previously described (10). It consists of (i) a full length L1Hs 5' UTR constructed mostly from cDNA sequences, (ii) 15 codons of L1Hs ORF1, (iii) the bacterial *lacZ* reporter gene fused in frame to the L1Hs ORF1, and (iv) an L1Hs 3' UTR. The sequence of the first 100 base pairs of p1LZ is identical to the sequence of L1.2 except at positions 3 (C instead of G on the coding strand) and 73 (G-C base pair not present in L1.2). Previous experiments established that deleting p1LZ base pairs −17 to +101 resulted in a dramatic reduction in the expression of β-galactosidase upon transfection into NTera2D1 cells (10).

To further delineate the specific sequence motifs contributing to this activity, a series of small deletions was made in p1LZ starting in the Bluescript plasmid at base pair −32 and extending into the L1Hs sequences (Materials and Methods). The resulting plasmids were tested for their ability to express β-galactosidase in NTera2D1 cells by a transient transfection protocol as previously described (10, 21). As shown in Figure 2, deletion of the upstream plasmid sequences from −32 to −4, or L1Hs sequences from base pairs +1 to +11, had no significant effect
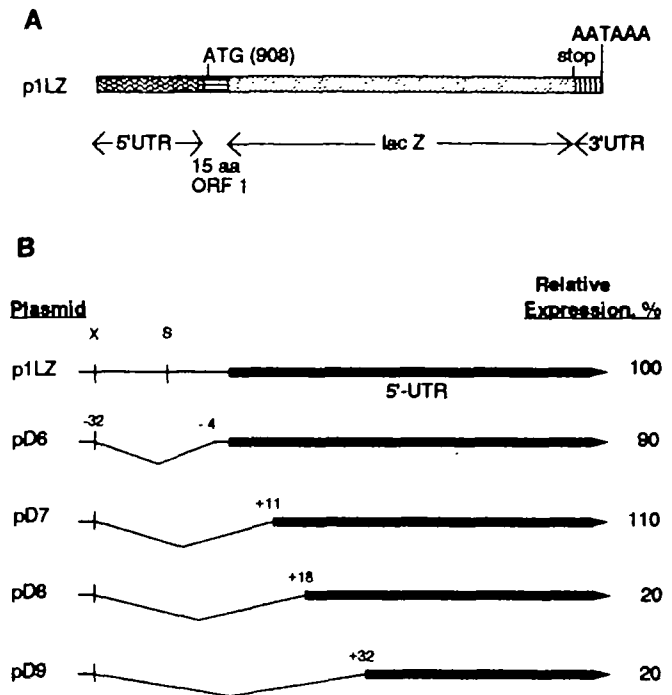
Figure 2. (A) The structure of p1LZ, a LacZ reporter construct. The relevant areas are shown by the arrows below. (B) 5' L1Hs UTR deletion constructs and their expression of β-galactosidase in NTera2D1 cells relative to expression by the undeleted construct, p1LZ. Vector segment is shown by thin lines; deletions limits by nucleotide numbers; 5' UTR by thick lines. The ORF1, lacZ and remaining vector segments are not shown. X, Xba1 site; S, Sma1 site. The number to the right of each construct is the average relative β-galactosidase activity found for each deletion, based on spectrophotometric analysis of three individual transfection assays and expressed as a percentage of the expression of p1LZ.

on the expression of the reporter gene. In contrast, when the deleted sequences extended to base pair +18, a five-fold reduction in expression was noted. Further deleting the sequence to base pair +32 resulted in no further reduction in expression. These results suggested that an important regulatory sequence overlaps the region between base pairs +12 and +18.

Recognition of the L1Hs binding site by recombinant YY1

We next investigated whether this region had sequence similarity to the DNA binding sites for known transcription factors. YY1, a protein factor capable of exerting both activating and repressing effects on transcription, and reportedly present in many or all mammalian cell types, binds to DNA sequences with significant similarity to L1Hs +13 to + 21 (17, 18, 19, 20). Binding sites for YY1 have been found both upstream and downstream of transcription start sites. The mouse ribosomal protein L30 gene contains a binding site for YY1 identical in sequence to L1Hs +13 to +21, and nearly identical in location relative to the initiation site although in the opposite orientation (Fig. 1b). The YY1 binding site in the 5' LTR of the murine leukemia virus is also highly similar to the L1Hs sequence.

Initial studies were performed with a Jurkat cell nuclear extract previously shown to contain the YY1 protein (Becker,K.G., Jedlicka,P.Templeton,N.S., Liotta,L., and Ozato,K., manuscript in preparation), and with double-stranded oligonucleotide probes corresponding to (i) the first 40 base pairs of the actively transposing L1Hs element L1.2B (L1), and to (ii) bases +34
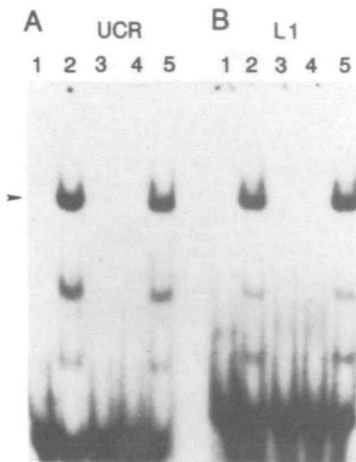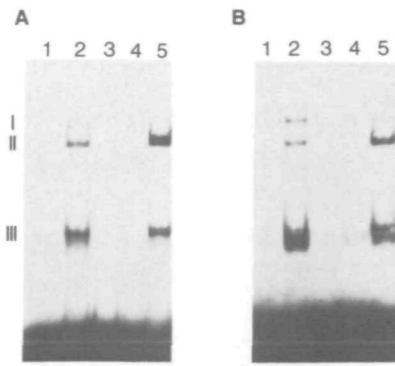
Figure 3. EMSA analysis with recombinant YY1 protein and ^{32}P oligonucleotides. (A) ^{32}P-UCR; (B) ^{32}P-L1 oligonucleotide. Lanes in both panels are: 1, free probe; 2, probe plus YY1 protein; 3 to 5, competition with 100 fold excess of unlabeled L1, UCR and control oligonucleotide, respectively. The arrow identifies the major YY1-oligonucleotide complex.

Figure 4. EMSA analysis with nuclear extract prepared from NTera2D1 cells. (A) EMSA analysis with ^{32}P-L1. (B) EMSA analysis with ^{32}P-UCR. Competitions were done with 100 fold molar excess of unlabeled oligonucleotides. Lanes: 1, free probe; 2, NTera2D1 nuclear extract; 3, competition with L1; 4, competition with UCR; 5, competition with control oligonucleotide.

to +63 of the MuLV 5'LTR (UCR), previously shown to contain a site (bp +43 to +51) bound by YY1 (17). These experiments indicated that both L1 and UCR formed specific competable complexes with factors in the Jurkat extract that migrated with equal mobilities in nondenaturing gels (data not shown), and suggested that L1 is capable of binding YY1.

To test directly the ability of YY1 to bind to the LINE-1 sequence, EMSAs were performed using recombinant YY1 (Materials and Methods). Both UCR (Fig. 3A) and L1 (Fig.3B) bound to recombinant YY1 protein and the band shift patterns produced were identical (lanes 2). The most abundant complex (indicated by the arrow) comigrated with the complexes formed with the Jurkat cell nuclear extract (data not shown). Both the UCR and the L1 complexes could be completely competed by 100-fold molar excess of non-labeled UCR or L1 oligonucleotides (lanes 3 and 4). A control double stranded oligonucleotide, which does not contain a YY1 binding site, was ineffective in competing for the UCR or L1 complexes (lanes 5). In addition to the most abundant complexes, two less abundant and faster migrating complexes were noted (lanes 2). These most likely represent complexes formed between the probes and partially degraded recombinant YY1 as determined by (i) the absence of any bands in EMSA using extracts prepared from control bacteria, and (ii) a decrease in the intensity of the upper band and concomitant increase in the intensity of the lower bands after the extract was subjected to multiple cycles of freezing and thawing (not shown). These results indicate that L1 is capable of binding YY1.

NTera2D1 cells express the YY1 protein

High level transcription of L1Hs has been reported only in a narrow range of cell types including NTera2D1 (8, 9, 10). p40, the protein encoded by the L1Hs first open reading frame, has been identified primarily in those cell types known to be permissive for L1Hs transcription (12, 22). p40 has not been detected in Jurkat cells (R.T. unpublished observation) suggesting that L1Hs transcription does not occur in Jurkat cells. In order for YY1 to positively modulate the expression of L1Hs it must be present in cells actively transcribing these elements. To

determine whether NTera2D1 cells express the YY1 protein, a NTera2D1 nuclear extract was prepared (Materials and Methods) and substituted for the recombinant YY1 in EMSAs.

As shown in figure 4, both probes bound factors present in the NTera2D1 nuclear extract forming three complexes each, and the resulting complexes migrated with similar mobilities (lanes 2). Once again, the complexes formed by both the UCR and the L1 probes were specifically competed by both UCR and L1 but not the control oligonucleotide (lanes 3−5). One of these complexes (II) migrated identically to the YY1 complex formed in the Jurkat extract (see fig. 5) and therefore most likely represents the binding of the oligos to YY1 protein present in the NTera2D1 extract. Complexes III are distinct from any formed in the Jurkat extract but are recognized by a YY1 antiserum (see below). Complexes I varied in intensity between individual experiments and were frequently absent, and therefore their significance is unknown.

Analysis of the complex formed between the L1 oligonucleotide and the NTera2D1 cell nuclear extract with an antibody to the YY1 protein

To determine whether the specific complexes formed in NTera2D1 cell nuclear extract resulted from the binding by YY1 or antigenically related protein to both UCR and L1, we performed EMSA in the presence or absence of a specific and high titer antiserum raised in mice against human YY1 synthesized in *E.coli* (Methods and Materials). As seen in figure 5, the specific binding complexes evident in the absence (lanes 3 and 7) or presence (lanes 5 and 9) of non-immune mouse serum were ablated by the addition of the anti-YY1 serum (lanes 4 and 8). Complexes II (see fig. 4) were completely ablated (fig 5, lanes 4 and 8) while complexes III were greatly reduced; complexes III could be completely ablated by the addition of greater amounts of the anti-YY1 serum (data not shown). The apparent disruption of YY1-DNA complexes by anti-YY1 serum, as opposed to the expected shift in the mobility of these complexes, has been noted previously (23). These results establish that YY1, or an antigenically related protein, is responsible for the UCR and L1 complexes formed in NTera2D1 cell nuclear extract. In addition, these results suggest that several forms of YY1, or proteins related to YY1, may exist in NTera2D1 cells.
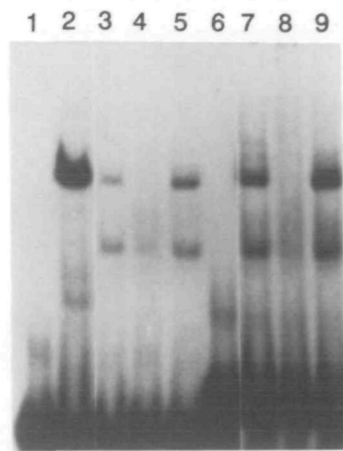
Figure 5. Effect of anti-YY1 serum on DNA-protein complex formation in NTera2D1 nuclear extract. ^{32}P-UCR, lanes 1 to 5; ^{32}P-L1 lanes 6 to 9. Lanes: 1, free probe; 2, Jurkat nuclear extract; 3, NTera2D1 nuclear extract; NTera2D1 nuclear extract with: 4, anti-YY1 serum and 5, non-immune serum; 6, free probe; 7, NTera2D1 nuclear extract; NTera2D1 nuclear extract with: 8, anti-YY1 serum; 9, non-immune serum.
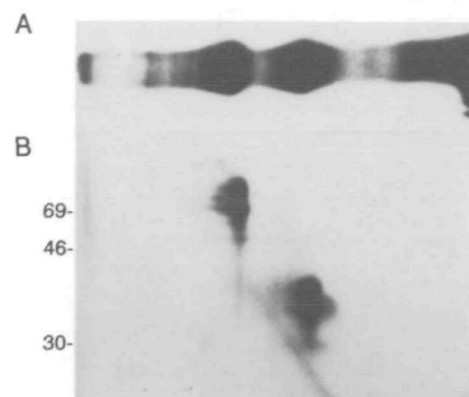

Figure 6. UV cross-linking and 2-dimensional analysis of EMSA reaction using NTera2D1 extract and L1 probe. A. One dimensional EMSA depicted horizontally. B. Separation of the cross-linked EMSA products in the second dimension on a denaturing SDS-polyacrylamide gel. The position of the molecular weight markers (in kd) are shown. Panels A and B derive from separate binding reactions.

UV crosslinking and 2D analysis of L1 complexes formed in NTera2D1 nuclear extracts

To confirm that the specific complexes formed by L1 in NTera2D1 cell nuclear extract resulted from the binding of YY1, we subjected the EMSA reactions, after resolution in the first dimension, to UV crosslinking followed by separation in the second dimension by denaturing SDS-PAGE (Materials and Methods). Two L1-protein complexes (corresponding to complexes II and III of fig. 4) with apparent molecular weights of approximately 69 and 40 kd, were successfully cross-linked and resolved (fig. 6). Although the sequence of the human YY1 protein predicts a molecular weight of only 44.6 kd, several studies have noted that it migrates with an apparent molecular weight of approximately 68 kd on SDS-PAGE (17, 18, 19, 20). These results establish that YY1 is the DNA binding protein responsible for the L1 complex II formed in NTera2D1 cell nuclear extract.

DISCUSSION

LINE-1 elements constitute several percent of the human genome and transpositions of these elements contribute both to the evolution of the genome and its accumulation of mutations. Current models suggest that L1Hs transposition is preceeded by, and at least in part is regulated by, the transcription of full-length elements. Efforts to understand the overal impact of L1Hs transposition on the human genome require therefore, a detailed description of the mechanisms by which L1Hs transcription is regulated.

Several independent studies have suggested that the region near the 5′ terminus of L1Hs contains one or more *cis* elements capable of regulating transcription. Swergold (10) fused the entire L1Hs 5′ UTR to the *lacZ* reporter gene. Expression of this gene was efficient in NTera2D1 cells but inefficient in HeLa cells; deletion of base pairs −17 to +101 essentially abolished the expression of this gene. Mathias (15) demonstrated that nuclear extracts prepared from both NTera2D1 and HeLa cells were

capable of forming specific complexes with a double-stranded oligonucleotide representing the sequence from +1 to +31 of L1.2A. DNAse I footprints obtained with HeLa and NTera2D1 extracts were identical and indicated the protection of the sequence between base pairs −6 to +22. In contrast, when EMSAs were performed using these same extracts, the patterns of bands seen were overlapping but clearly distinct. The identity of the factor(s) binding to this region was not determined, although the possibility that YY1 or a similar protein might be involved was suggested. (16) fused the sequence from bp +1 to +155 of pCGL1-1 to the chloramphenicol acetyl transferase reporter gene. Mutation of base pairs +10 to +17 resulted in a significant decrease in expression of the reporter gene. In addition, the region between base pairs +3 and +26 of pCGL1-1 was protected in DNAse I footprinting experiments by a nuclear factor present in several different cell types. A double stranded oligonucleotide with the sequence of pCGL1-1 from −1 to +25 formed a single specific complex in EMSAs using HeLa cell nuclear extracts. A similar complex was reportedly seen when EMSA was done with extracts from a variety of other cell types. The protein binding to this region was termed TFL1-A but its identity was not established.

In this report, we confirm the existence of a *cis* regulatory element between base pairs +12 and +18 of L1Hs (figs. 1 and 2). Deletion of the first 11 base pairs does not impair the expression of p1LZ while deletion of the first 18 base pairs results in a 5-fold reduction in expression. In addition, we establish that the ubiquitous nuclear transcription factor YY1 binds specifically to this region (figs. 3−6). Thus, nuclear extracts from Jurkat and NTera2D1 cells contain a factor that binds specifically to a double-stranded oligonucleotide with a sequence matching the first base pairs of L1.2A. The band representing this complex is ablated upon the addition of anti-YY1 serum. Finally, recombinant YY1 protein partially purified from bacteria also binds specifically to the L1 probe. In addition, we show that YY1, or an antigenically related protein with the same electrophoretic mobility on denaturing SDS gels, is present in NTera2D1 cells (fig. 4−6), and therefore is likely to be involved in the regulation of L1Hs transcription.

Several characteristics of YY1 make it an excellent candidate for the regulation of L1Hs transcription. The sequences necessary

and sufficient for the regulation of L1Hs transcription appear to be located entirely within the transcription unit (10). YY1 binding sites are frequently found downstream of transcription initiation sites (18, 19, 20, 24). Transcription of L1Hs is highly restricted; to date, only a few tumor cell lines have been found to express either L1Hs RNA (8, 9, 10) or proteins (12, 13, 22, 25, 26). Because a large region of DNA appears to be necessary for high level expression of L1Hs in NTera2D1 cells, it is likely that many different transcription factors, both positive and negative, are involved in its regulation. YY1 is capable of both enhancing and suppressing transcription, in part depending on the presence or absence of ancillary proteins (18). Although YY1 is apparently a ubiquitous protein, its dual function suggests that it could be involved in both the activation of L1Hs transcription in permissive cells, and its suppression in non-permissive cells.

It is interesting to note that YY1 can act as an initiator protein; on its own, YY1 is capable of directing the initiation of RNA synthesis at its binding site (23). In L1Hs and mouse ribosomal protein L30 however, the synthesis of RNA begins not at the YY1 binding site, but at a position 13 or 14 base pairs, respectively, upstream of it. L1Hs does not contain a TATA sequence. It will be interesting to determine whether YY1 is acting as an initiator for the transcription of L1Hs (and ribosomal protein L30) and if so, how it is capable of selecting the appropriate transcription start site.

Both L1Hs and MuLV encode enzymes with reverse transcriptase activity (14, 27), and both are capable of transposition (3, 4, 5, 28). Despite this, the LINE-1 family of elements is clearly distinct from the retroviruses (29). It is notable therefore, that 90% of the C-type retroviruses, and the LINE-1 elements found in humans, a prosimian (30), and in rabbits (31), all contain potential binding sites for YY1 in the same relative positions. No binding sites for YY1 are identifiable in published mouse or rat LINE-1 sequences (K. B., unpublished observation), indicating that the regulation of human and murine LINE-1 elements differ significantly in molecular detail (32, 33, 34, 35). It will be interesting to determine whether the binding of a YY1-like activity is a feature of other mammalian LINE-1 elements, and what the role of YY1 is in the regulation and start site positioning of retrovirus and LINE-1 transcription.

MATERIALS AND METHODS

Cell culture and nuclear extracts

Ntera2D1 (10) and Jurkat cells (36) were grown as previously described. Nuclear extracts from eukaryotic cells were prepared by the method of Dignam (37), with minor modifications.

Construction and assay of p1LZ deletions

Overlapping deletions of the region from −32 to +32 of plasmid p1LZ (10) were made by the method of Henikoff (38) using the double stranded nested deletion kit from Pharmacia. Briefly, plasmid DNA was first digested with *Xba I* and the overhang was filled in with deoxythionucleotides. The DNA was further digested with *SmaI*; exonuclease III digestion was carried out in 75 mM NaCl at 15°C for either 5 or 10 minutes. The sequence surrounding the deleted region of individual clones was determined by dideoxy sequencing. Plasmids were purified over Qiagen maxi columns. The expression of β-galactosidase by individual clones was determined by the spectrophotometric assay of whole cell extracts from transiently transfected NTera2D1 cells as previously reported (10, 21).

Oligonucleotides

All oligonucleotides used in this study were synthesized on an Applied Biosystems 380B DNA synthesizer. The UCR oligonucleotide (UCR) was previously described (17). Its sequence (5′CTGCAGTAACGCCATTTTGCAAGGCATGAA-3′) (24) corresponds to bases 34 to 63 of the MuLV 5′ LTR (based on sequence of MuLV 5′ LTR in Genbank accession no. X05157) and includes the previously reported

YY1 binding site (17,24). The L1Hs oligonucleotide (L1) corresponds to bp 1 to 40 of the L1Hs element L1.2 (Genbank accession no. M80343). Its sequence is 5′-GGGGGAGGAGCCAAGATGGCCGAATAGGAACAGCTCCGGT-3′. The sequence of the control oligonucleotide (C), which does not contain a YY1 binding site, is: 5′-CTCAACAGAGGGGACTTTCCGAGAGGCCAT-3.

Production of a polyclonal antibody specific for recombinant YY1

The production and purification of recombinant YY1 and the production of a polyclonal anti-YY1 serum are described elsewhere (Becker, K.G., Jedlicka, P. Templeton, N.S., Liotta, L., and Ozato, K. manuscript in preparation). Briefly, the protein was expressed in *E. coli* from a plasmid that contained a cDNA fragment which included the entire human YY1 coding sequence. The protein was purified to >90% purity by passage over heparin-sepharose and agarose-Ni^{++} columns. A specific, high-titer antiserum against this protein was prepared by repeated injections into BALB/C mice.

Electrophoretic mobility shift assay (EMSA)

EMSA was performed as previously described (17). Briefly, double stranded, complementary oligonucleotide probes were end labeled with [^{32}P-γ-]ATP using polynucleotide kinase. Labeled probe (5×10^4 cpm) was incubated with either 4μg of nuclear progein extracts or 1μg of a bacterial extract prepared from *E.coli* expressing recombinant human YY1 (17) and 2μg of poly dI-dC (Pharmacia) in 40μl of binding buffer (50mM NaCl, 20mM Tris pH 7.6, 1mM MgCl$_2$, 0.2mM EDTA, 5% glycerol and 5mM DTT) for 40 min. at 4°C. Binding reactions were analyzed on 6% non-denaturing polyacrylamide gels and subjected to autoradiography. Competitor oligonucleotides were added at a 100 fold molar excess 5 min prior to the addition of radiolabeled probe and incubated at 4°C. When the EMSA was done in the presence of antibody, 2 μl of either non-immune serum or anti-YY1 antiserum was added after the initial reaction, and the incubation was extended for an additional 30 min at 4°C.

UV-crosslinking EMSA

UV-crosslinking was performed as described (39) with minor modifications. Nuclear extracts from NTera2D1 cells were incubated with a ^{32}P-labeled UCR substituted on the strand shown above with 5-bromo-2′-deoxyuridine. The DNA-protein complexes were resolved in the first dimension on a 6% non-denaturing polyacrylamide gel as above. The gel was irradiated by placing it directly on a UV transilluminator (305nm, Fotodyne) for 20 min. The entire lane containing the cross-linked EMSA products was then excised from the gel, equilibrated for 15 min. in loading buffer (30% glycerol (w/v), 15% 2-mercaptoethanol, 7% SDS, 0.2 M Tris HCl pH 6.8), placed on top of a 10.5% SDS-polyacrylamide gel and run in the second dimension.

ACKNOWLEDGEMENTS

REFERENCES

1. Adams, J. W., R. E. Kaufman, P. J. Kretschmer, M. Harrison and A. W. Nienhuis. 1980. Nucleic Acids Res. 8: 6113−6128.
2. Grimaldi, G., J. Skowronski and M. F. Singer. 1984. EMBO J. 3: 1753−1759.
3. Miki, Y., I. Nishisho, A. Horii, Y. Miyoshi, J. Utsunomia, K. W. Kinzler, B. Vogelstein and Y. Nakamura. 1992. Canc. Res. 52: 643−645.
4. Morse, B., P. G. Rothberg, V. J. South, J. M. Spandorfer and S. M. Astrin. 1988. Nature. 333: 87−90.
5. Kazazian, H. H. J., C. Wong, H. Youssoufian, A. F. Scott, D. G. Phillips and S. Antonarakis. 1988. Nature (London). 332: 164−166.
6. Narita, N., H. Nishio, Y. Kitoh, Y. Ishikawa, Y. Ishikawa, R. Minami, H. Nakamura and M. Matsuo. 1993. Journal of Clinical Invest.. 9: 1862−1867.
7. Scott, A. F., B. J. Schmeckpeper, M. Abdelrazik, C. T. Comey, B. O'Hara, J. P. Rossiter, T. Cooley, P. Heath, K. D. Smith and L. Margolet. 1987.
8. Skowronski, J. and M. F. Singer. 1985. Proc. Natl. Acad. Sci., USA. 82: 6050−6054.
9. Skowronski, J., T. G. Fanning and M. F. Singer. 1988. Mol. Cell. Biol. 8: 1385−1397.
10. Swergold, G. D. 1990. Mol. Cell Biol. 10: 6718−6729.
11. Dombroski, B. A., S. L. Mathias, E. Nanthakumar, A. Scott F. and H. H. J. Kazazian. 1991. Science. 254: 1805−1808.

12. Leibold, D. M., G. D. Swergold, M. F. Singer, R. E. Thayer, B. A. Dombroski and T. G. Fanning. 1990. Proc. Natl. Acad. Sci. USA. 87: 6990−6994.
13. Bratthauer, G. L. and T. G. Fanning. 1992. Oncogene. 7: 507−510.
14. Mathias, S. L., A. F. Scott, H. H. J. Kazazian, J. D. Boeke and A. Gabriel. 1991. Science. 254: 1808−1810.
15. Mathias, S. L. 1992. Functional Studies of a Human Transposable Element. Thesis, The Johns Hopkins University.
16. Minakami, R., K. Kurose, K. Etoh, Y. Furuhata, M. Hattori and Y. Sakaki. 1992. Nucl. Acids Res. 20: 3139−3145.
17. Flanagan, J. R., K. G. Becker, D. L. Ennist, S. L. Gleason, P. H. Driggers, B.-Z. Levi, E. Appela and K. Ozato. 1992. Mol. Cell. Biol. 12: 38−44.
18. Shi, Y., E. Seto, L.-S. Chang and T. Shenk. 1991. Cell. 67: 377−388.
19. Hariharan, N., D. Kelley E. and R. P. Perry. 1991. Proc. Natl. Acad. Sci. USA. 88: 9799−9803.
20. Park, K. and M. L. Atchison. 1991. Proc. Natl. Acad. Sci. USA. 88: 9804−9808.
21. Simon, J. A. and J. T. Lis. 1987. Nucleic Acids Res. 15: 2971−2988.
22. Holmes, S. E., M. F. Singer and G. D. Swergold. 1992. Jour. Biol. Chem. in Press: .
23. Seto, E., Y. Shi and T. Shenk. 1991. Nature (London). 354: 241−245.
24. Flanagan, J. R., A. M. Krieg, E. E. Max and A. S. Khan. 1989. Mol. Cell. Biol. 8: 739−746.
25. Brattauer, G. L. and T. G. Fanning. In preparation.
26. Brattauer, G. L., R. D. Cardiff and T. G. Fanning. In preparation.
27. Rothenberg, E. and D. Baltimore. 1976. J. Virol. 17: 168−174.
28. Varmus, H. and P. Brown. 1989. Retroviruses, 53−108. *In* D. Berg E. and M. Howe M. (ed.), Mobile DNA. Vol. American Society for Microbiology, Washington, DC.
29. Hutchison III, C. A., S. C. Hardies, D., Loeb, W. R. Shehee and M. H. Edgell. 1989. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. Vol. American Society for Microbiology, Washington, DC. 593−617.
30. Sakaki, Y., M. Hattaro, A. Fujita, K. Yoshioka, S. Kuhara and O. Takenaka. 1986. Cold Spring Harbor Symp. Quant. Biol. 51: 465−469.
31. Price, D. K., J. A. Ayres, D. Pasqualone, C. H. Cabell, W. Miller and R. C. Hardison. 1992. Genomics. 14: 320−331.
32. Schichman, S. A., D. M. Severynse, M. H. Edgell and C. A. Hutchison III. 1992. J. Mol. Biol. 224: 559−574.
33. Severynse, D. M., C. A. Hutchison III and M. H. Edgell. 1992. Mamm. Genome. 2: 41−50.
34. Martin, S., L. 1991. Mol. Cell. Biol. 11: 4804−4807.
35. Nur, I., E. Pascale and A. V. Furano. 1988. Nucleic Acids Res. 16: 9233−9251.
36. Weiss, A., R. L. Wiskocil and J. D. Stobo. 1984. J. Imunol. 133: 123−128.
37. Dignam, J. D., R. M. Lebovitz and R. G. Roeder. 1983. Nucl. Acid. Res. 11: 1475−1489.
38. Henikoff, S. 1984. Gene. 28: 351−359.
39. Ballard, D. W., E. Bohnlein, J. A. Hoffman and H. P. Bogerd. 1989. New Biol. 1: 83−92.
40. Andrews, P. W., P. N. Goodfellow, L. H. Shevinsky, D. L. Bronson and B. B. Knowles. 1982. Int. J. Cancer. 29: 523−531.
41. Fawcett, D. H., C. K. Lister, E. Kellet and D. J. Finnegan. 1986. Cell. 47: 1007−1015.