Vol. 154, No. 1, 1988 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 15, 1988 Pages 91-100

ISOLATION OF PUTATIVE PROMOTER REGION FOR HUMAN TERMINAL DEOXYNUCLEOTIDYLTRANSFERASE GENE

Osamu Koiwai and *Akimichi Morita

Institute for Developmental Research, Aichi Prefectural Colony, Kasugai, Aichi, 480-03,

and

*School of Medicine, Nagoya City University, Nagoya, Japan

Received May 24, 1988

SUMMARY We have isolated the genomic clone, which contained 5'-flanking region for human terminal deoxynucleotidyltransferase (TdT) gene, by screening Charon 4A library and determined its nucleotide sequence. Two transcription initiation sites (cap sites) were localized at 81 and 106 bp upstream from the third ATG codon which is translation initiation site, by primer extension analysis. Although neither typical TATA nor CAAT sequences were detected within 100 bp upstream region of the 5' cap site, a palindromic structure comprised of 28 bp was formed just upstream of the 5' cap site and double repeats of 5'-CTGGC sequence were found in its palindrom. The typical octamer of 5'-ATGCAAAT and three octamer-like sequences, which are involved in the transcripton of lymphoid-specific genes, were detected at 212, 293, 347 and 770 bp upstream from the first cap site.

Terminal deoxynucleotidyltransferase (TdT; EC2.7.7.31) is a DNA polymerase which is normally detected in thymocyte and bone marrow cells in higher vertibrates. High level of TdT is also found in acute lymphoblastic leukemia (ALL) cells or chlonic myelogenious leukemia (CML) cells at blastic crisis. TdT catalyzes the polymerization of deoxyribonucleotides to the 3'-hydroxyl end of the oligonucleotide primer in the absence of DNA template (1). Based on its *in vitro* function and tissue specificity, TdT has been considered to be closely related to B- and T-cell development in thymus and

0006-291X/88 $1.50

bone marrow. Several lines of evidence suggest that TdT promotes the insertion of the extra nucleotides at the joining regions (N region) of the V and D or D and J gene segments in an immunoglobulin (Ig) gene (2,3). Landau *et al*. constructed the retrovirus vector (TdT-vector) which can express human TdT in mammalian cells and transfected the vector into the PD31 cells, which have no detectable TdT activity (3). In the course of development, the exta nucleotides were synthesyzed at the N region of superinfected artificial Ig gene rearrangement substrate in the cells expressing TdT, in a fivefold greater frequency than the cells which did not express TdT.

Human, mouse and bovine TdT cDNA clones have been reported (4-7). We isolated nearly full-length TdT cDNA clone from human T-cell leukemia MOLT3 cDNA library to investigate the regulation of TdT gene expression in normal and malignant leukemia cells. By using this human TdT cDNA as a probe, we tried to isolate the promoter region of normal human genomic TdT gene from the Charon 4A genomic library. In this paper, we report the nucleotide sequence of the 5'-flanking region for TdT gene and the transcription initiation sites of TdT mRNA by primer extension analysis.

MATERIALS AND METHODS

Isolation of human TdT cDNA. Human TdT cDNA clone (H14 clone) was isolated by screening the pcD library which was constructed from MOLT3 cells. The bovine TdT cDNA was used as a probe. The procedures for construction of the library and the screening condition were same as reported previously (7).

Screening of human genomic library. Charon 4A genomic library constructed from human sperm was kindly gifted by Dr. J. Nathans. 70,000 plaques were screened by using human TdT cDNA (mixture of Pst1-Eco R1 (740 bp; PE probe) and Eco R1 fragment (1,100 bp; E probe) of H14 clone) as a probe. The

plaques were transfered to nitrocellulose filters and the filters were hybridized with the probe at 45° C in 6xSSPE containing 20 % formamide (8). The filters were washed in 2xSSPE at 45° C. Nucleotide sequences were determined by the M13 dideoxy chain termination method (9).

5' end mapping of the TdT transcript. The oligonucleotide primer (50-mer) complimentary to the underlined region (Fig 2; from the position of 52 to101) was chemically synthesized. The 5' end of the primer (1 μg) was labeled with [γ-^{32}P] ATP with T4 polynucleotide kinase (8). 20 μg of poly A-RNA was mixed with 300 ng of the labeled primer and coprecipitated in ethanol at -20° C. The precipitant was suspended with 10 μl of the buffer A (0.1 M NaCl, 20mM Tris-HCl pH 8.0, 0.1 mM EDTA), annealed at 100° C for 2 min, then 60 C for 2 hs and mixed with 10 μl of the buffer B (80 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 20 mM 2-mercaptoethanol, 400 μM dNTP, 400 units of avian myeloblastosis virus reverse transcriptase) for 60 min and 120 min at 37° C. The mixture was treated with phenol and precipitated in ethanol at -20° C for 12 hs. After centrifugation, the precipitant was suspended with 20 μl of Tris-EDTA buffer (pH 8.0) and 1 μg was loaded onto a polyacrylamide gel. The concentration of the gel was 8 %.

RESULTS

Isolation of human TdT cDNA clone. Human TdT cDNA clone (H14 clone) longer than H10 clone was isolated from the pcD library constracted from MOLT3 cells (7). The insert of H14 clone contains 1,966 bp with an open reading frame comprising 1,524 bp. There are three ATG codons at the 5'end region of H14 clone (Fig 2). The first and second ATG are not in frame, the third ATG codon is in frame. The molecular weight deduced from this H14 clone is 58,308 daltons which is identical to the value reported by Peterson *et al* (4).

Isolation of human genomic TdT gene. With this H14 clone as a probe (mixture of PE and E probes), we isolated the genomic clone HG3-1,HG2-3 and HG25-2 by screening the Charon 4A library constructed from human sperm. The HG25-2 clone which contained about 35,000 bp of insert, encoded the

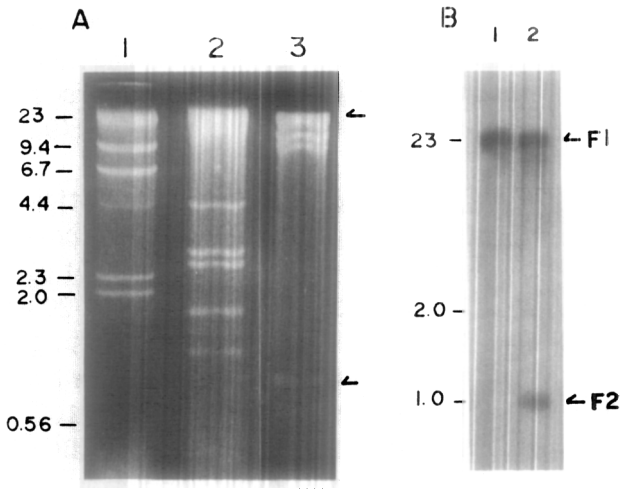

Fig 1. Southern blot hybridization analysis of genomic clone. A, Genomic clone of HG25-2 was digested with EcoR1 (lane 2) and Bam H1(lane 3) enzymes and analyzed on an agarose gel of 1 %. Lane 1, size marker (kilobases). Arrows indicate the hybridyzed Bam H1 fragments with ^{32}P labeled probe. B, Southern blot analysis of Eco R1 (lane 1) and Bam H1 (lane 2) digests. PE probe was used.

5'-flanking region for TdT gene. HG25-2 was digested with the restriction enzymes of Eco R1 and Bam H1. Figure 1 shows the electrophoretic pattern by agarose gelelectrophoresis and Southern hybridization pattern of the Eco R1 and Bam H1 digests with ^{32}P labeled PE probe. One Eco R1 fragment of 25,000 bp (Fig 1; lane 1), and two Bam H1 fragments of 23,000 bp (F1) and 1,000 bp (F2) (lane 2) were hybridized with this probe. We determined the entire DNA sequence (999 bp) of F2 for Bam H1 digests and 5' end region of F1 (Fig 2). F1 contains the 5'-flanking region of TdT gene and, F2, the first exon and first intron of TdT gene. The first exon starts from the positon of +1 (the cap site) to +309. The first intron starts from the consensus sequence of 5'-GT and extends at least longer than 23,000 bp.

5' end mapping of the TdT transcript. To identify the precise 5' termini of TdT mRNA, we synthesized the oligo-nucleotide primer (50-mer) from the

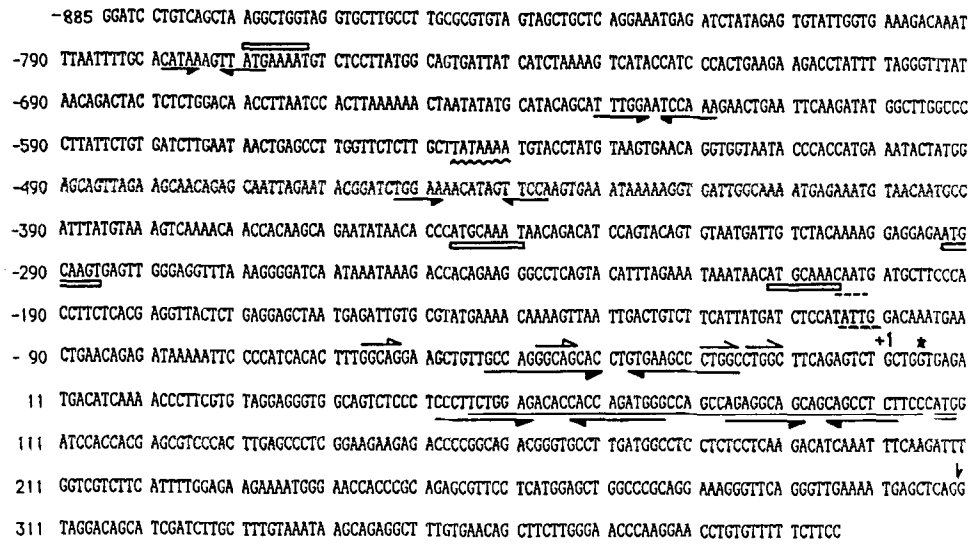

Fig 2. DNA sequence of 5'-flanking region for TdT gene. The Bam H1 fragment of HG25-2 was sequenced by chain termination method. ---- and ~~~~ indicate the typical CAAT and TATA sequences, respectively. ——, synthesyzed oligonucleotide (50-mer) for primer extention analysis. ⟶ and ⟶ indicate the direct repeats. ⟶ ⟵, palindromic structure. ═, translation initiation site. ↓, consencus site for exon-intron boundary, +1 and • denote the transcription initiation sites. ▭, consensus octamer found in enhancer and promoter regions of Ig gene. *, 5' end of TdT cDNA (H14 clone).

position of +56 to +105 which is complimentary to the underlined region (Fig 2). Figure 3 shows the pattern of the autoradiography for the primer extension analysis. One major and one minor bands were detected at the position of +1 and +26. The +1 cap site was 4 bp upstream from the 5' end of H14 clone.

DISCUSSION

Three clones (HG3-1, HG2-3 and HG25-2) were isolated by screening the Charon 4A human genomic library. We analyzed the HG25-2 clone which contained the 5'-flanking region of TdT gene, where the promoter of the eukaryotic gene is normally located. The extensive homology was observed

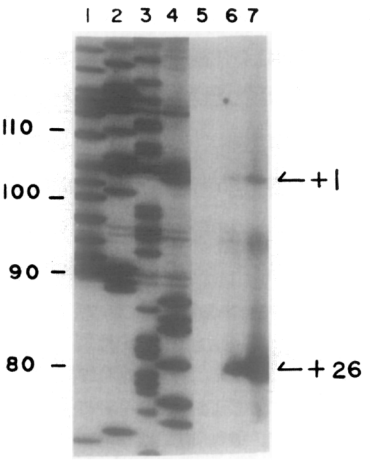

Fig 3. Mapping of transcription initiation sites by primer extention analysis. Lane 1-4, DNA sequencing pattern used for size marker (M13 phage DNA which inserted Pst1-EcoR1 fragment of H14 clone). Lane1, 2, 3 and 4 indicate the nucleotides for A, G, C and T, respectively. Lane 5-7, synthetic oligonucleotide of 50-mer was annealed with MOLT3 mRNA and the mixture was incubated with reverse transcriptase for 0 (lane 5), 60 min (lane 6) and 120 min (lane 7). The products were analyzed on an 8 % polyacrylamide gel. Numbers (base) on the left shows the position from the 3' end of 50-mer. +1 and +26 denote the transcription initiation sites.

among the sequences of this 5'-flanking region of TdT gene and the promoter or enhancer regions of Ig genes (10). Especially, the sequence from -357 to -377 in TdT gene bears a high similarity with the core domain of enhancer region in Ig gene, which encode a lymphoid-specific octamer of 5'-ATGCAAAT (Fig.4) (11,12). Besides this typical octamer of 5 '-ATGCAAAT, the octamer-like sequences of 5'-ATGAAAAT, 5'-ATGCAAAC and 5'-ATGCAAGT were also located at -770, -293 and -212, respectively. Even though this octamer was found in the promoter regions of other prokaryotic and eukaryotic genes (13), the region which contained this octamer might be involved in the regulation of lymphoid-specific gene expression because this sequence was detected in the promoter region of TdT gene, specifically

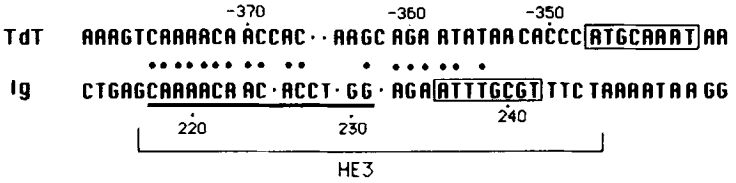

Fig 4. Putative lymphoid-specific regulatory region for transcription. The sequences enclosed by boxes indicate the consensus octamer found in enhancer and promoter regions of Ig genes. The underline shows the binding site (E4) for a nucler protein proposed in mouse Ig heavy-chain (IgH) enhancer by Ephrussi *et al.* (11). HE3 region shows the protected region in human Ig gene enhancer by DNase I footprint analysis (12). The numbers on the TdT sequence are the positions from the first cap site of TdT gene. The numbers below the sequence of human Ig enhancer region are according to Hayday *et al.* (19).

expressed in B- and T-cells. The transcription of TdT and Ig genes might be controlled by same lymphoid-specific regulatory factors in a nuclei.

Although neither typical TATA nor CAAT sequences were found within 100 bp from the cap sites, we can form the palindromic structure composed of 28 bp from the position of -43 to -16 just upstream of the cap site, where the eukaryotic promoter is canonically present (Fig 5). The double repeats of 5'-CTGGC sequence which its biological meaning is not clear, were found at this palindrom (Fig 2). We suspect that the regions of the consensus octamer and the palindromic structure in the 5'-flanking region for TdT gene is crucial to the regulation of TdT gene expression. We are now underway to clarify the functional significance of these structures, namely, whether this structure functions as the promoter for TdT gene by CAT assay (14). We have also analyzed the HG3-1 and HG2-3 clones. The human genomic TdT gene is composed of at least 10 exons (data not shown).

The primer extention analysis showed two transcription initiation sites for TdT mRNA, 9 bp upstream (+1 cap site) and 14 bp downstream from the

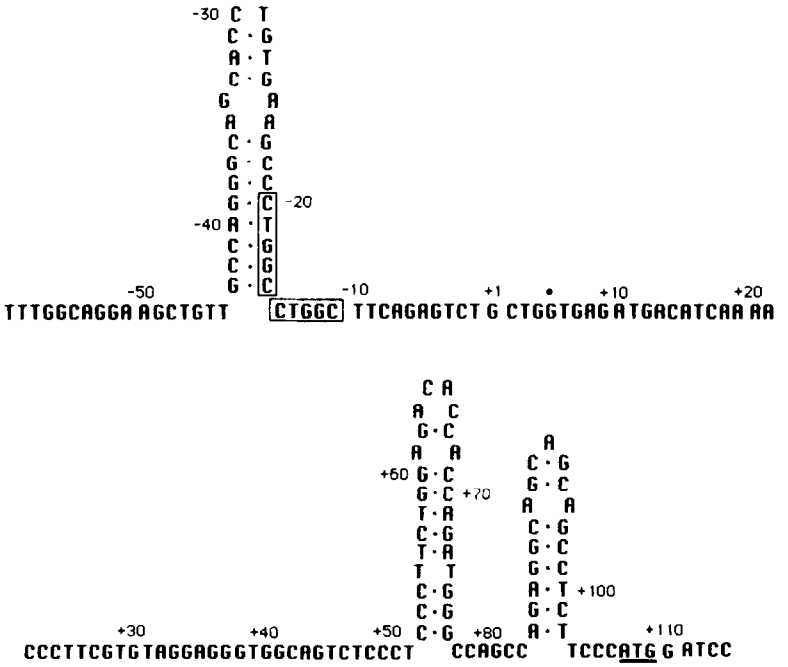

Fig 5. Structure of 5'-flanking region from transcription initiation site. The palindromic structure can be formed at 5' non-coding region of TdT gene. Double repeats of 5'-CTGGC are boxed. Underlined ATG is the translation initiation site. +1 shows the first cap site for TdT gene. • is 5' end of TdT cDNA (H14 clone).

first ATG codon of H14 clone. Although the band of +1 cap site is weaker than the second, the first might be the major transcription initiation site. The reaction of reverse transcriptase might be trapped at the second cap site under the condition of primer extension analysis, because more than 10 TdT cDNA clones except H10 clone, which we sequenced, extended longer than the second (data not shown).

This H14 clone has three ATG codons at the 5' end region of TdT cDNA. The third ATG is in frame. This is inconsistent with the H10 clone whose translation initiation site is the second ATG (7). The structure of the promoter region for TdT gene including the translation initiation site might be altered in malignant leukemic cells such as MOLT3. We can not entirely

rule out the possibility that TdT gene has two copies on the chromosome 10 where TdT gene is localyzed (15). We are isolating the promoter region of TdT gene for malignant leukemia cells to compare with the structure from the normal cells.

Recently, ALL patients that undergo the translocation between chromosome 10 and 14 (T-cell recepter α gene is located) and that are expressing high TdT activities were reported (16,17). TdT activity is also detected in adult T-cell leukemia cells (18). The isolation of regulatory element for TdT gene will facilitate the investigation of the regulation of the TdT gene expression for malignant leukemia cells.

ACKNOWLEDGMENT: We thank Dr. H. Sato and Dr. T. Kishi for helpful discussions. This research was supported by a grant-in-aid from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Bollum, F.J. (1974) The Enzymes, (Academic Press, New York), 10, 145-171
2. Alt, F.W., and Baltimore, D. (1982) Proc. Natl. Acad. Sci. USA, 79, 4118-4122
3. Landau, R.N., Schats, G.D., Rosa, M., and Baltimore, D. (1987) Mol. Cell. Biol., 7, 3237-3243
4. Peterson, R.C., Cheung, L.C., Mathaliano, R.J., Chang, L.M.S., and Bollum, E.J. (1985) J. Biol. Chem., 260, 10495-10502
5. Landau, R.N., John, P.St.T., Weissman, L.I., Wolf, C.S., Silverstone, E.A., and Baltimore, D. (1984) Proc. Natl. Acad. Sci. USA, 81, 5836-5840
6. Koiwai, O., Yokota, T., Kageyama, T., Hirose, T., Yoshida, S., and Arai, K. (1986) Nucleic Acid Research, 14, 5777-5792
7. Koiwai, O., Kaneda, T., and Morishita, R. (1987) Biochem. Biophys. Res. Commun. 144, 185-190
8. Maniatis, T., Fritsh, E.F., and Sambrook, J. (1982) Molecular Cloning: a Laboratory Manual, (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY)
9. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA. 74, 5463-5467

10. Singh, H., Sen, R., Baltimore, D., Sharp, P.A. (1986) Nature 319, 154-158
11. Ephrussi, A., Church, G.M., Tonegawa, S., and Gilbert, W. (1985) Science 227, 134-140
12. Maeda, H., Araki, K., Kitamura, D., Wang, J., and Watanabe, T. (1987) Nucleic Acid Res. 15, 2851-2869
13. Mocikat, R., Falkner, F.G., Mertz, R., and Zachau, H.G. (1986) Nucleic Acids Res. 14, 8829-8844
14. Gorman, C.M., Moffat, L.F., and Howard, B.H. (1982) Mol. Cell. Biol. 2, 1044-1051
15. Isobe, M., Huebner, K., Erikson, J., Peterson, R., Bollum, F.J., Chang, L.M.S., and Croce, C.M. (1985) Proc. Natl. Acad. Sci. USA, 82, 5836-5840
16. Dube, L.D., Raimondi, S.C., Pi, D., and Kalousek, D.K. (1986) Blood, 67, 1181-1184
17. Kagan, J., Finan, J., Letofsky, J., B, E.C., Nowell, P.C., and Croce, C.M. (1987) Proc. Natl. Acad. Sci. USA, 84, 4543-4546
18. Hinuma, Y., Nagata, K., Hanaoka, M., Nakai, M., Matsumoto, T., Kinoshita, K., Shirakawa, S., and Miyoshi, I. (1981) Proc. Natl. Acad. Sci. USA, 78, 6476-6480
19. Hayday, A.C., Gillies, S.D., Saito, H., Wood, C., Wiman, K., Hayward, W.S., and Tonegawa, S. (1984) Nature 307, 334-340