*Russian Journal of Genetics, Vol. 39, No. 2, 2003, pp. 216–221. Translated from Genetika, Vol. 39, No. 2, 2003, pp. 280–285.*


# Transcription Regulation of Human *oct-1* Gene Requires Involvement of Two Promoters

**S. V. Zhenilo, I. E. Deyev, S. M. Serov, and O. L. Polanovsky**

*Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia; fax: (095)135-14-05; e-mail: pol@eimb.ru*

Received September 16, 2002

**Abstract**—Transcription initiation of human Oct-1 transcription factor-encoding gene involves two promoters, 1U and 1L, located at a substantial distance (about 100 kb) apart. The structure of these promoters and the adjacent sequences is different. Specifically, the 1U sequence is GC-rich, while the 1L sequence is AT-rich. Correspondingly, more than 25 GC-rich Sp1 *cis*-elements were localized within the 1U region, while in the 1L sequence nearly equal amount of homeo-specific NTAATNN sites along with two ATGCAAAT octamers were found. Analysis of transfection of recombinant plasmids, carrying the promoter fragments with or without enhancer indicated that expression from the 1L promoter was tissue-specific. In nonlymphoid HEK293 cells efficiency of transcription from the 1U promoter was several times higher than that from the 1L promoter. Another expression pattern was observed at transfection of the same constructs into Raji lymphoid cells. In this case the level of transcription from the L promoter (fragment L2) at the presence of external enhancer was higher than that from the fragments containing the 1U promoter. It was shown that the distal regions of 1U and 1L were capable of silencing activity. In Raji cells enhancer completely overcomes the activity of U silencer, but only partly overcomes the activity of L silencer. Our data on the interaction of two promoters with the enhancer and silencer in different cell types point to fine tissue-specific regulation of the *oct-1* gene expression, especially in lymphatic cells.

## INTRODUCTION

Transcription factors Oct-1 and Oct-2 were initially identified, purified, and functionally characterized in the late 1980s [1–3]. In double-stranded DNA optimal target sequence of these proteins is the ATGCAAAT octamer, initially revealed in the immunoglobulin gene promoters [4]. Later, this sequence was found in the promoters and enhancers of many other genes [5–7]. While Oct-1 protein is expressed in all eukaryotic cells and is regarded as ubiquitous, the activity of Oct-2 protein is mostly restricted to the immune and nervous systems. The Oct proteins are the members of a large family of proteins containing the DNA-binding POU domain, initially identified in the transcription factors Pit-1, Oct-1, and Oct-2, and also in Unc-86 [8]. Oct-1 can activate transcription through different mechanisms, involving either other transcription factors [5–7] or coactivators [5]. Recently, it was shown that POU domain of the Oct-1 protein could form direct contacts with the TBP (TATA box-binding protein), which was found to be essential for the interaction with a distant enhancer [9].

Multiple functions of the Oct-1 can result from different reasons, including the context of the DNA sequence recognized, the contacts with other proteins, and also the specific features of its own structure, which, along with the POU domain, contains some other functionally important regions. For example, N-terminal part of the molecule contains a glutamine-rich region, essential for activation of the promoters for the genes other than those encoding snRNA [10]. The C-terminal part of the molecule contains an alanine-rich domain, which can be involved in the Oct-1 silencing activity [11]. Earlier we described tissue-specific isoforms of Oct-1, differing from one another in the structure of 5'-terminal exon, termed exon L [12–14]. The Oct-1L isoform was found in murine lymphoid cells, and also in the murine cell lines, representing different stages of B- and T-lymphocyte differentiation. This isoform was not detected in the embryonal and nonlymphoid tissues and cell lines examined to that time.

It was suggested that L-containing isoforms were formed as a result of alternative transcription initiation and/or alternative splicing [13, 14]. This work was focused on the studies and comparison of the structure and specific features of tissue-specific transcription of the 5'-terminal fragments located upstream of exons 1U and 1L in lymphoid and nonlymphoid cells.

## MATERIALS AND METHODS

*Cell lines.* The HEK293 cells (embryonal renal carcinoma) were maintained in DMEM medium containing 10% fetal calf serum and 40 μg/ml kanamycin. Raji (B-lymphocytes) and Namalwa (B-lymphocytes) cells were maintained in RPMI-1640 medium containing 10% fetal calf serum and 40 μg/ml kanamycin.

1022-7954/03/3902-0216$25.00 © 2003 MAIK “Nauka /Interperiodica”

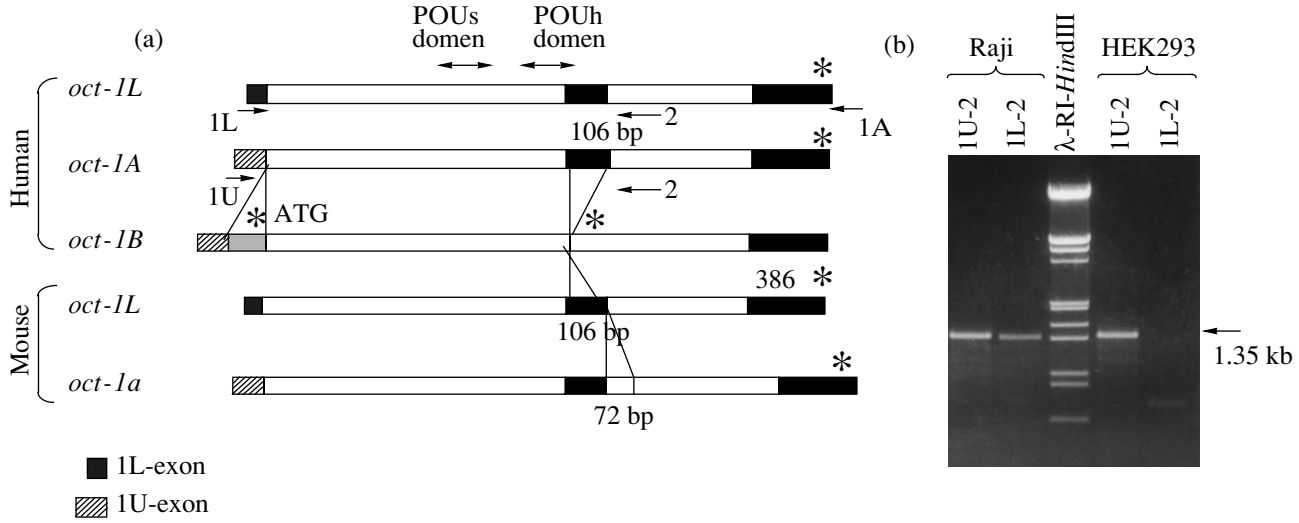


**Fig. 1.** (a) Comparison of *oct-1* isoforms from human (*oct-1L, A, B*) and mouse (*oct-1L, a*) (shown are POU homeodomain (POUh) and POU-specific domain (POUs). *, stop codone. (b) splicing of the *oct-1* RNA 5'-terminal exons (the products of reverse transcription and PCR; PCR between primer 2 and the primers to exon 1L, or exon 1U).

*cDNA.* RNA was isolated from the cell lines using guanidine thiocyonate as described [15]. Full-length cDNA was obtained as follows. RNA samples were reverse transcribed using the primer T15 (5'-gggaggcccttttttttttttttt-3'). Next, 1/5 of the cDNA sample was amplified with Hi-Fi polymerase (Sileks M, Russia) using the primers 1L(5'-gccatgctggacgcagtgac-3') and 1A(5'-actggatccagtccaatcacactgcagag-3'). The PCR program was as follows: 2 min at 95°C, then 5 cycles of 40 s at 95°C; 40 s at 59°C; and 3.5 min at 72°C; and 30 cycles of 40 s at 95°C; 40 s at 61°C; and 3.5 min at 72°C. The reaction products were cloned into pUC18 plasmid and sequenced. The 5'-fragments of cDNA were obtained in PCR reaction with *Taq* polymerase and the primers 1L-2(5'-aggtgagctgctggtccac-3') and 1U(5'-gagcagcgagtcaagatgag-3')-2. The reaction program was as follows: 2 min at 95°C and 35 cycles of 40 s at 95°C; 40 s at 57°C; and 2 min at 72°C.

*Plasmid constructs.* Fragments of the 5'-terminal regions of exons 1L and 1U were generated by use of PCR with Hi-Fi polymerase (Sileks M, Russia) and using human genomic DNA as a template. Construct L-1 was generated with the primers 3L(5'-ggtaccgttaggtccaagaagccagg-3')-2L(5'-gaacaaacctgagagacgag-3'); construct L-2, with the primers 4L(5'-ggtacctatttatgcaacgatttaaatt-3')-2L; construct L-3, with the primers 5L(5'-ggtaccaagcttcattatggaaactgcc-3')-2L; construct U-1, with the primers 4U(5'-ggtacctttccccaataccat-3')-2U (5'-gaatatttaaccaaaatcgccc-3'); and construct U-2, with the primers 3U(5'-ggtacctgaagaaagtgcggagcaca-3')-2U. The reaction products were digested with the *Kpn*I restriction endonuclease and cloned into the pGL3-enhancer and pGL3-basic vectors digested with the *Kpn*I and *Sma*I enzymes.

*Transient transfection.* Transfection of the HEK293 cells was carried out by calcium phosphate method using 0.5 μg reporter plasmid and 0.3 μg plasmid containing *β-gal* gene under the control of CMV promoter and enhancer. The Raji cells were transfected using electroporation with 1.5 μg reporter plasmid and 1 μg plasmid, containing *β-gal* gene: 1.1 kV, 40 μs; and dual pulse with the interval of 2 min ($2 \times 10^6$ cells; 0.4-mm cuvette; gap volume, 135 μl). The cells were harvested after 48 h. The luciferase activity for each point was normalized for β-galactosidase activity.

## RESULTS

### *Cloning of the Full-Length Isoform of Human oct-1 mRNA*

The full-length *oct-1* cDNA, termed *oct-1L,* was isolated from human Raji cells using reverse transcription and PCR techniques. It was homologous to the murine *oct-1L* cDNA cloned by us earlier from myeloma cells [13]. The 5'-terminal region of human *oct-1L* was identical to the murine *oct-1R* isoform [16], while the rest of the molecule was identical to human *oct-1A* [17] and murine *oct-1b* [18] cDNAs. The 5'-terminal exon of human *oct-1L,* along with murine *oct-1L* and *oct-1R* cDNAs encode 10 N-terminal amino acid residues, while the 5'-terminal exons of human isoforms *oct-1A* and *oct-1B,* and murine isoforms *oct-1a, b,* and *c* encode 21 amino acid residues of the protein (Fig. 1a). These exons were termed 1L and 1U (L, lymphocyte; U, ubiquitous), respectively.

It was shown that *oct-1* mRNA with the long 5'-terminal exon 1U is expressed both in nonlymphoid (HEK293 kidney cells) and lymphoid human cells, while mRNA with short 5'-terminal exon 1L is found only in lymphoid Raji (Fig. 1b) and Namalwa cells (data not shown).

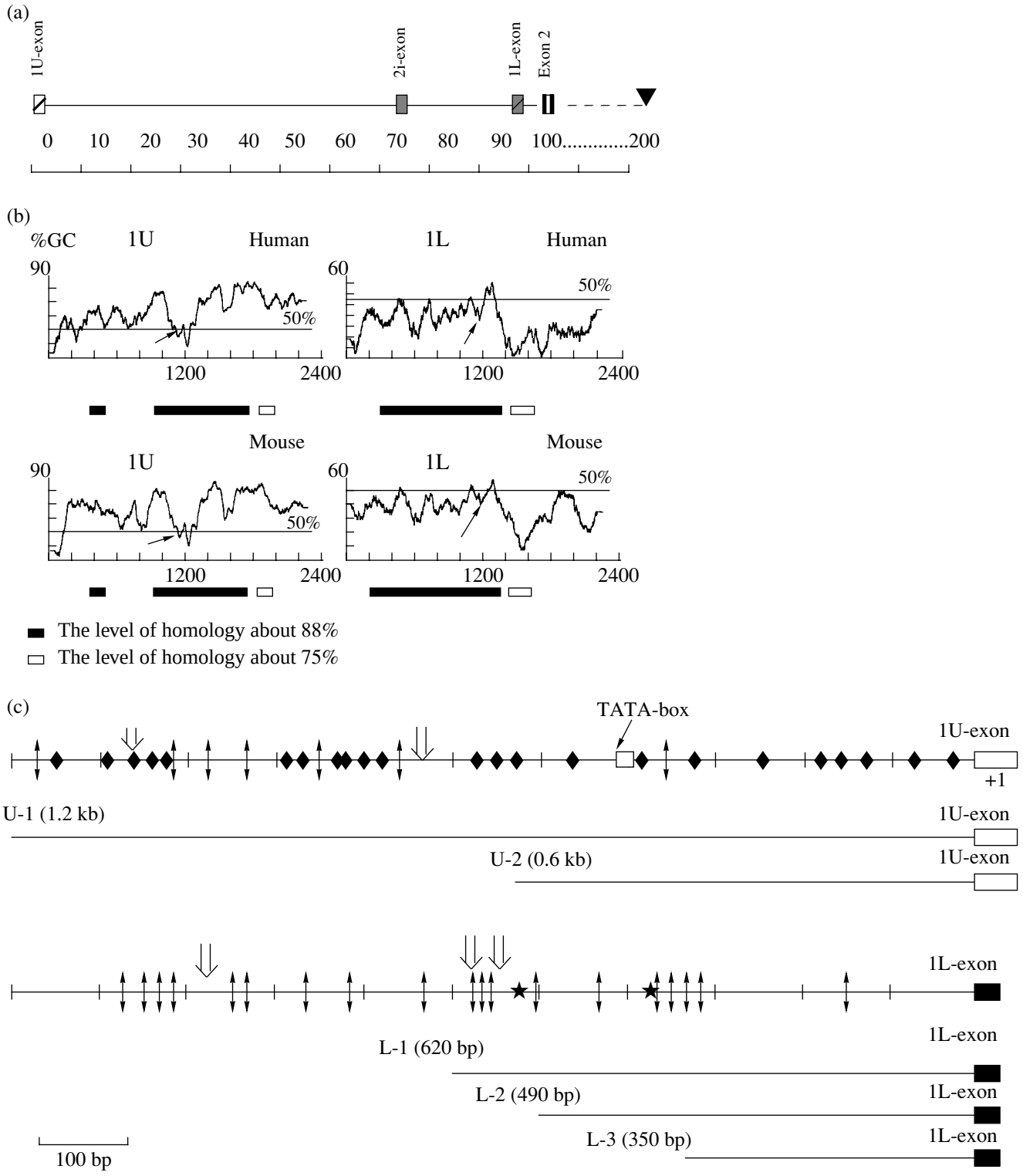


**Fig. 2.** Organization of the *oct-1* 5'-region. (a) 5'-region of the *oct-1;* (b) analysis of the regions around exons 1U and 1L in mouse and human: the level of GC-content and the level of human/mouse homology. Locations of exons 1U and 1L are indicated by asterisks (the level of GC-content was estimated by use of Gene Runner software with the step of 100 bp); (c) schematic representation of the 5'-nontranslated regions upstream of human *oct-1* exons 1U and 1L. Also shown are the regions U-1, U-2, L-1, L-2, and L-3, whose role in the transcription initiation was examined. ⇓, CCAAT binding site; ◆, Sp1 binding site; ↕, TAAT binding site (for homeo proteins); ★, oct site; ▼, stop codon.

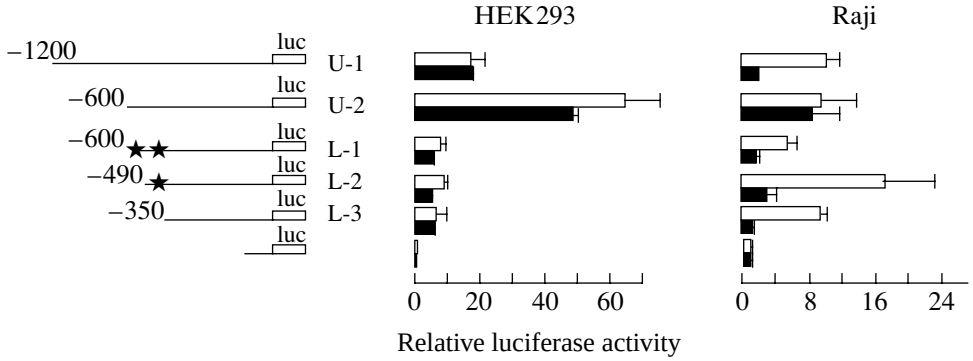


**Fig. 3.** Promoter activity of the regions upstream of human exons 1U and 1L as a result of transient transfection into HEK293 and Raji cells determined by use of luciferase analysis. □, promoter activity of the constructs cloned into the pGL3-enhancer vector; ■, constructs in the pGL3-basic vector; ★, oct site. The constructs demonstrated in Fig. 2 are shown at the left.

*Genomic Organization of the oct-1 5'-Terminal Region*

Human *oct-1* locus spans the region of about 200 kb, and furthermore, the distance between exons 1U and 1L constitutes 100 kb, a half of the whole locus (Fig. 2a). Comparison of exon 1U and 1L flanking regions from man and mouse showed that these sequences were highly homologous (Fig. 2b). The region around exon 1U in both organisms appeared to be extremely GC-rich. The GC-content usually constituted 60 to 70%, and even 90% in some 100-bp tracks. This area was also rich for the sites for the Sp1 factor. This pattern is typical to the promoters of the constitutive genes, i.e., the genes expressed in all cell types.

As regards the region around exon 1L, it appeared to be AT-rich. Moreover, it included numerous sites for homeoproteins, CCAAAT-binding proteins, and also two sites for Oct-1. Almost all of these sites can be found both in man and mouse, pointing to their possible involvement in regulation of transcription.

*Promorer Activity of the Regions Upstream Human Exons 1U and 1L*

To examine promoter activity, five constructs (see Fig. 2c), containing 5'-nontranslated regions upstream of exons 1U and 1L, within the pGL3-enhancer (containing the SV40 virus enhancer) and pGL3-basic vectors were generated. Among these there were two constructs containing the regions upstream of exon 1U subsequently truncated from the 5'-end, and three constructs, where the regions upstream of exon 1L were truncated in such a manner that in each subsequent construct at first one, and then another, Oct-1 binding site was excised. Promoter activity of the regions upstream of exon 1U and 1L was studied using HEK292 and Raji cell lines.

In nonlymphoid HEK293 cells maximum activity was typical of the region upstream of exon 1U. The activity of the U-2 construct was 50 times higher than that of the vector lacking promoter insert. Elongation of this sequence (construct U-1) resulted in the 2.5–3 times decrease of the promoter activity. Promoter activity of the regions upstream of exon 1L appeared to be only five times higher than that of the empty vector. The Oct-1 binding sites located in this region had no influence on promoter activity. The presence or absence of the enhancer in the reporter vector had very slight influence of the promoter activity. These findings indicate that in nonlymphoid cells activity of the regions examined is only weakly affected by the enhancer.

At transfection of lymphoid Raji cells another pattern of activity was observed. At the absence of the enhancer the level of the U-1 promoter activity is only slightly higher than that of the empty vector, while inclusion of the enhancer results in the tenfold increase of the promoter activity. The activity of smaller U-2 construct in both vectors was similar, i.e., the presence of the enhancer had no effect on its activity. Promoter activity of the regions L-1, L-2, and L-3 at the lack of the enhancer was 1.2 to 3 times higher than that of the empty vector, while the presence of the enhancer resulted in a remarkable increase of the activity, up to 3 to 8 times. Thus, in Raji cells the presence of the enhancer is essential for the activity of both promoters.

## DISCUSSION

The genes involved in complex cellular processes are often regulated at different levels of expression. One such regulatory mechanism is alternative usage of two or more promoters. A large number of such genes have been described (see [19]). Alternative promoters provide fine transient or tissue-specific regulation of gene expression. Usually, one of the promoters is a constitutive one. It is responsible for the gene expression in many tissues and cell types, providing the synthesis of housekeeping or ubiquitous proteins. Another promoter is characterized by certain tissue-specificity and is responsible for the synthesis of the proteins expressed only in several tissues during specific stages of ontogeny.

The synthesis of nitric oxide is a vivid example of gene expression with the involvement of nine enhanc-

ers [20]. In the cell, transcription factors play diverse roles. Some of the comprehensively studied genes encoding these proteins have two or even more promoters. For example, the gene encoding transcription factor Pax-6 is regulated by two promoters, 1A and 1B [21]. It was shown that the DNA regions upstream of the promoters contain activating elements and silencers. The 1B promoter is autoregulated.

Earlier it was shown that transcription factor Oct-1 was expressed as a set of isoforms, some of which were found in all cell types examined while the others were tissue-specific [12–14]. These two groups differed from one another in the structure of the first exon. Ubiquitous factors contained exon 1U, while tissue-specific factors carried exon 1L. Exons 1U and 1L in human and mouse genomes are located 70 and 100 kb apart, respectively. It was suggested that tissue-specific and ubiquitous isoforms of Oct-1 were formed as a result of alternative splicing and/or alternative transcription initiation.

In this study we examined 5'-terminal sequences upstream of exons 1U and 1L. It was found that these regions remarkably differed from one another in the content of GC- and AT-pairs. An extended nucleotide sequence upstream of exon 1U was GC-rich. The content of GC-pairs in some 100-bp fragments of this sequence was higher than 80% (Fig. 2b). Analysis of *cis*-elements revealed the presence of a great number of Sp1 binding sites.

Alternatively, the sequence upstream of exon 1L was AT-rich and contained a great number of homeospecific sites along with two octameric sites. Substantial differences in the structures of these regulatory regions suggested the existence of functional differences.

Our experiments were focused on detection of tissue-specific regulation of expression. Our data indicated that transcription from the L promoter was tissue-specific and dependent of external enhancer (Fig. 3). With respect to transcription initiation efficiency in lymphoid Raji cells L-constructs can be roughly arranged in the following order: L-1 < L-3 < L-2. This can be explained as follows: though L-1 construct contains promoter and is activated by external enhancer, this activation is not complete (only a threefold increase), which may be due to the presence of the silencer in the 5'-terminal region. At the presence of an enhancer, the highest level of activity was observed for the L-2 construct. This construct contains a proximal Oct-1 site, which can be involved into the formation of a contact with TBP protein, which is, in turn, essential for the interaction with the enhancer proteins [9].

Especially intriguing and important result is that activity of L-promoter in lymphoid cells at the presence of an enhancer is comparable with the activity of U-promoter responsible for the synthesis of the ubiquitous Oct-1 isoform. In nonlymphoid HEK293 cells the 1U-promoter is stronger than the L-promoter. An interesting feature of the 1U region is that it probably contains a silencer within a distance between 600 and 1200 bp. In nonlymphoid cells the activity of this silencer cannot be overcome by the inclusion of external enhancer, while in Raji cells its activity is completely overcome by the enhancer. Thus, expression of the gene for ubiquitous transcription factor Oct-1 is realized at least at two levels: alternative transcription initiation, described in the present study, and alternative splicing reported in our previous works [12–14].

## ACKNOWLEDGMENTS

We thank E.E. Egorov for providing the cuvettes for electroporation.

This work was supported by the Russian Foundation for Basic Research (grants nos. 02-04-49129 and 02-04-06752) and the Russian State Program “Studies and Projects in Priority Directions in Science and Technology.”

## REFERENCES

1. Fletcher, C., Heintz, N., and Roeder, R.G., Purification and Characterization of OTF-1, a Transcription Factor Regulating Cell Cycle Expression of a Human Histone *H2b* Gene, *Cell* (Cambridge, Mass.), 1987, vol. 51, pp. 773–781.
2. Scheidereit, C., Heguy, A., and Roeder, R.G., Identification and Purification of a Human Lymphoid-Specific Octamer-Binding Protein (OTF-2) That Activates an Immunoglobulin Promoter in Vivo, *Cell* (Cambridge, Mass.), 1987, vol. 51, pp. 783–793.
3. Stepchenko, A.G., Luchina, N.N., and Polanovsky, O.L., A Protein Factor Interacting with the Promoter Region of the Immunoglobulin κ Genes, *Dokl. Akad. Nauk SSSR*, 1988, vol. 303, pp. 1008–1010.
4. Falkner, F.C. and Zachau, H.G., Correct Transcription of the Immunoglobulin κ Gene Requires an Upstream Fragment Containing Conserved Sequence Elements, *Nature*, 1984, vol. 310, pp. 71–74.
5. Herr, W. and Cleary, M., The POU Domain: Versatility in Transcriptional Regulation by a Flexible Two-in-One DNA-Binding Domain, *Genes Dev.*, 1995, vol. 9, pp. 1679–1693.
6. Stepchenko, A.G. and Polanovsky, O.L., Interaction of Oct Proteins with DNA, *Mol. Biol.* (Moscow), 1996, vol. 30, pp. 503–513.
7. Hyillips, K. and Luisi, B., The Virtuoso Versatility: POU Proteins That Flex to Fit, *J. Mol. Biol.*, 2000, vol. 302, pp. 1023–1039.
8. Herr, W., Sturm, R.A., Clerc, R.G., *et al.*, The POU Domain: A Large Conserved Region in the Mammalian *pit-1*, *oct-1*, *oct-2*, and *C. elegans unc-86* Gene Products, *Genes Dev.*, 1988, vol. 2, pp. 1513–1516.
9. Bertolino, E. and Singh, H., POU/TBP Cooperativity: A Mechanism for Enhancer Action from a Distance, *Mol. Cell*, 2002, vol. 10, pp. 397–407.
10. Tanaka, M., Lai, J.-S., and Herr, W., Promoter-Selective Activation Domains in Oct-1 and Oct-2: Direct Differential Activation of an snRNA and mRNA Promoter, *Cell* (Cambridge, Mass.), 1992, vol. 68, pp. 755–767.

11. Kim, M., Lesoon-Wood, L.A., Weintraub, B.D., and Chung, J.H., A Soluble Transcription Factor, Oct-1, Is Also Found in the Insoluble Nuclear Matrix and Possesses Silencing Activity in Its Alanine-Rich Domain, *Mol. Cell. Biol.*, 1996, vol. 16, pp. 4366–4377.
12. Pankratova, E.V., Deyev, I.E., and Polanovsky, O.L., Tissue-Specific Splicing of the 5' Exons of the Gene for Transcription Factor Oct-1, *Mol. Biol.* (Moscow), 2001, vol. 35, pp. 34–41.
13. Pankratova, E.V., Deyev, I.E., Zhenilo, S.V., and Polanovsky, O.L., Tissue-Specific Isoforms of the Ubiquitous Transcription Factor Oct-1, *Mol. Gen. Genet.*, 2001, vol. 266, pp. 239–245.
14. Pankratova, E.V., Deyev, I.E., Zhenilo, S.V., *et al.*, Isoforms of Transcription Factor Oct-1 Produced in Lymphocytes, *Mol. Biol.* (Moscow), 2001, vol. 35, pp. 816–823.
15. Chomczinski, P. and Sacchi, N., Single-Step Method of RNA Isolation by Acid Guanidinium–Phenol–Chloroform Extraction, *Anal. Biochem.*, 1987, vol. 162, pp. 156–159.
16. Stepchenko, A.G., The Nucleotide Sequence of Mouse Oct-1 cDNA, *Nucleic Acids Res*, 1992, vol. 20, p. 1419.
17. Das, G. and Herr, W., Enhanced Activation of the Human Histone H2B Promoter by an Oct-1 Variant Generated by Alternative Splicing, *J. Biol. Chem.*, 1993, vol. 268, pp. 25 026–25 032.
18. Suzuki, N., Peter, W., Ciesiolka, T., *et al.*, Mouse Oct-1 Contains a Composite Homeodomain of Human Oct-1 and Oct-2, *Nucleic Acids Res.*, 1993, vol. 21, pp. 245–252.
19. Rubtsov, P.M., Alternative Promoters and RNA Processing in Expression of the Eukaryotic Genome, *Mol. Biol.* (Moscow), 2000, vol. 34, pp. 626–634.
20. Wang, Y., Newton, D.C., Robb, G.B., *et al.*, RNA Diversity Has Profound Effects on the Translation of Neuronal Oxide Synthase, *Proc. Natl. Acad. Sci. USA*, 1999, vol. 96, pp. 12 150–12 155.
21. Orladnova, O., Syagailo, Y.V., Mossner, R., *et al.*, Regulation of *pax-6* Gene Transcription: Alternative Promoter Usage in Human Brain, *Brain Res. Mol. Res.*, 1998, vol. 60, pp. 177–192.