Molecular Biology, Vol. 35, No. 1, 2001, pp. 28–34. Translated from Molekulyarnaya Biologiya, Vol. 35, No. 1, 2001, pp. 34–41.

GENE MOLECULAR BIOLOGY

UDC 577.21

The Tissue-Specific Splicing of 5'-Terminal Exons of the Oct-1 Gene

E. V. Pankratova, I. E. Deev, and O. L. Polanovsky

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 117984 Russia

Received April 18, 2000

Abstract—It has been shown that the pre-mRNA of transcription factor Oct-1 undergoes tissue-specific splicing. The Oct-1R isoform is synthesized in mouse lymphoid myeloma cells NS/0 of B series and is not found in other somatic and embryonal cells. The initiation sites of *oct-1R* mRNA transcription are located at positions –159 and –307 from the AUG codon of the 1L exon. In both cases, no TATA box was found in the region preceding these sites (25–30 bp). A different form of *oct-1* mRNA (containing the 1U exon) was also found in NS/0 cells, which is probably synthesized in all cell lines. This isoform differs from *oct-1R*, in particular, by the structure of 5'-terminal exon, which is the result of alternative splicing.

Key words: transcription factor Oct-1, *oct-1*pre-mRNA, tissue-specific splicing, transcription initiation, eukaryotes

INTRODUCTION

Transcription factor Oct-1 was found in all types of dividing eukaryotic cells [1, 2]. Oct-1 is involved in regulation of expression of many genes, including the housekeeping genes such as those for snRNA [3] and histone H2B [4]. This transcription factor participates in cell proliferation and differentiation, the knock-out of *oct-1* gene is lethal [2]. Oct-1 is also involved in the tissue-specific expression of genes, in particular, of immunoglobulin genes [5–8].

Oct-proteins (Oct-1, Oct-2, Oct-3, and Oct-6) belong to the POU family of transcription factors. All members of this family contain POU domain that with a high affinity recognizes an "octamer" ATGCAAAT and/or homeospecific sites containing the core sequence TAAT [9, 10] and sites localized in promoters and enhancers of many genes. Other transcription factors or tissue-specific coactivators also take part in their recognition by Oct-proteins [10, 11]. One of such coactivators OCA-B protein (synonyms OBF-1 and Bob-1) is synthesized in B cells and takes part in their differentiation and regulation of transcription of immunoglobulin genes [12–15]. This protein forms a triple complex with Oct-1 and Oct-3 proteins and with octamer ATGCAAAT present in promoters and enhancers of immunoglobulin genes. The A-T mutation in position 5 of the octamer prevents formation of the triple complex with the target DNA sequence. Thus, the ATGCTAAT sequence interacting with Oct-1 protein is not recognized by the Oct-1/OCA-B complex.

Along with POU domains [16], formation of the Oct-protein complexes with coactivators and transcription factors requires other domains of Oct-proteins [17, 18]. Thus, N-terminal part of Oct-1 and Oct-2 contains a Gln-rich region that is involved in activation of promoters of genes encoding mRNA rather than snRNA [17]. Oct-1 also contains domains responsible for activation of the snRNA gene promoters [18]. Two isoforms, Oct-1A and Oct-1B, which are products of alternative splicing, were found in human cells [19]. They are synthesized in all cell lines. It has been shown that Oct-1A activates mainly snRNA genes, whereas Oct-1B (with an insert in its N-terminal region and with reduced C end) activates the histone H2B gene promoter [19]. Three forms of oct-1 mRNA (oct-1a, b, c) with marked distinctions in 3'-terminal region were isolated from a cell line of mouse teratocarcinoma F9 [20]. Multiple isoforms of Oct-2 with different functions were identified [21].

Structural analysis of Oct-1 isoforms and organization of the OTF-1 locus encoding Oct-1 [22] reveals several variants of pre-mRNA splicing, which may result in synthesis of functionally different forms of this protein. We have found tissue-specific splicing of 5'-terminal exons of Oct-1. The NS/0 lymphoid myeloma cells of B series synthesize the Oct-1R isoform, not found in other somatic and embryonal cells. It has been also shown that the *oct-1R* mRNA tran-

0026-8933/01/3501-0028$25.00

scription initiation sites are localized in positions –159 and –307 bp from the AUG codon of the 1L exon of locus OTF.

EXPERIMENTAL

Cell lines. The following mouse cell lines were used in this work: myeloma NS/0, neuroblastoma NB41A3, teratocarcinoma F9, and fibroblasts 10(0). Cells were grown in DMEM medium supplemented with 10% bovine serum and 40 μg/ml Gentamycin.

RNA isolation. RNA was isolated from cells using guanidine thiocyanate as described in [23].

Oligonucleotides. The following synthetic oligonucleotides were used: oligo 1: 5'-GTCACTGCAGTCCAGCATGGC-3' corresponding to exon 1L of *oct-1L* cDNA; oligo 2: 5'-GGTTTCTGATGGATTATTCATTC-3' corresponding to the second exon; oligo 3: 5'-GCCATGCTGGACTGCAGTGAC-3' corresponding to exon 1L; oligo 4: 5'-GAGCAGCGAGTCAAGATGAG-3' corresponding to exon 1U, and oligo 5: 5'-ACTAAGCTTCTCTTCCCACCCTTGTT-3' corresponding to 5' nontranslated region before exon 1L of *oct-1L* cDNA. Oligos 1 and 2 were used in reverse transcription, oligos 2–5 were used in PCR.

Reaction of reverse transcription was carried out using the M-MLV reverse transcriptase (BION M). RNA (3-5 μg) was mixed with a specific primer in 15 μl water, heated for 5 min at 70°C and cooled in ice. 5 μl of five-fold buffer (250 mM Tris-HCl, pH 8.3, 375 mM KCl, 15 mM $MgCl_2$, 50 mM dithiothreitol, 5 μl of 2.5 mM solution of deoxynucleoside triphosphates, 25 units of ribonuclease inhibitor, 200 units of reverse transcriptase, deionized water up to 25 μl were mixed and added to the RNA mixture with primer, mixed and incubated for 60 min at 37°C. Then, the mixture was treated in succession with phenol, phenol-chloroform mixture (1:1), and chloroform and precipitated with ethanol. Reaction products were analyzed in 6% PAGE in the presence of 8 M urea. To determine the transcription start, oligo 1 corresponding to the first exon of the *oct-1L* cDNA were used. Products of alternative splicing were determined with oligo 2.

Polymerase chain reaction (PCR). Single-stranded cDNA obtained in the reverse transcription reaction (1/5–1/10 part) was used as the template for PCR. Oligo 2 was labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase (Fermentas, Lithuania) as described by the supplier. PCR was carried out with Hi-Fi- polymerase (Silex, Moscow) as described by the supplier. Reaction conditions were 95°C, 2 min; (95°C, 40s; 47°C, 1 min; 72°C, 3 min) for 5 cycles; (95°C, 40s; 57°C, 1 min; 72°C, 2 min) for 30 cycles. The PCR products were analyzed in denaturing 6% PAGE. Products of known sequence (5'-untranslated region of *oct-1*, primer, oligo 1) obtained by the sequencing reaction were used as markers.

Sequencing. Nucleotide sequence was determined by PCR technique using a sequencing kit of Fermentas (Lithuania) according to the supplier's instruction.

RESULTS

Determination of *oct-1* transcription initiation site. The *oct-1* cDNA from mouse myeloma NS/0 cells was cloned earlier in our laboratory (*oct-1L*, EMBL X56230). This cDNA and the cDNA of *oct-1* obtained by Suzuki *et al.* [20] differ by their 5' regions. The first one encodes the N-terminal sequence of nine amino acid residues, which differs from the N-terminal sequence of 20 residues of Oct-1a, b, c isoforms. Reverse transcription from a primer complementary to the 5'-terminal exon of *oct-1R* has shown that in all cell cultures studied the initiation of *oct-1* transcription begins from two different sites. The proximal point is adenine at –159, the transcription initiation site is 5'-TAGTTT-3', and the homologous distal initiation site 5'-TTGTTT-3' with thymine in the start position –306 relatively to the AUG codon of 1L exon (Fig. 1). The presence of two transcription initiation sites in *oct-1* gene at a distance of 148 bp from each other indicates that probably different *cis* and *trans* elements of transcription may be involved in regulation of *oct-1* expression. The comparison and analysis of base sequences preceding each of these sites [25, 26] show that the distal site of transcription initiation is preceded by two canonic octamers ATGCAAAT in the direct and reverse orientation, and protein Oct-1 exhibits the highest affinity just to these sequences. Here there are also five clusters containing homeospecific sites interacting with POU and homeoproteins [25, 26]. No specific sites for binding POU proteins were found ahead of proximal site of transcription initiation.

No canonical TATA box was found in the region of –25–30 bp of both sites of transcription initiation. This shows that promoter of *oct-1L* belongs to those free of TATA box.

Alternative splicing of *oct-1* 5'-terminal exons. This work deals with the study of alternative splicing of the first 5'-terminal exons of the transcription factor Oct-1 in the cell lines of various tissue origin. Reverse transcription and PCR have shown that the neuroblastoma, teratocarcinoma, and myeloma cells and fibroblasts contain the *oct-1U* mRNA having a long 5'-terminal exon (universal exon 1U). This mRNA was identified in all cell types, whereas mRNA with a short 5'-terminal exon 1L is expressed only in myeloma cells NS/0 and is a product of tissue-specific alternative splicing. To reveal the *oct-1* mRNA containing exon 1L, two pairs of oligonucleotides were used. The first contained a common for all PCR ^{32}P-labeled oligonu-

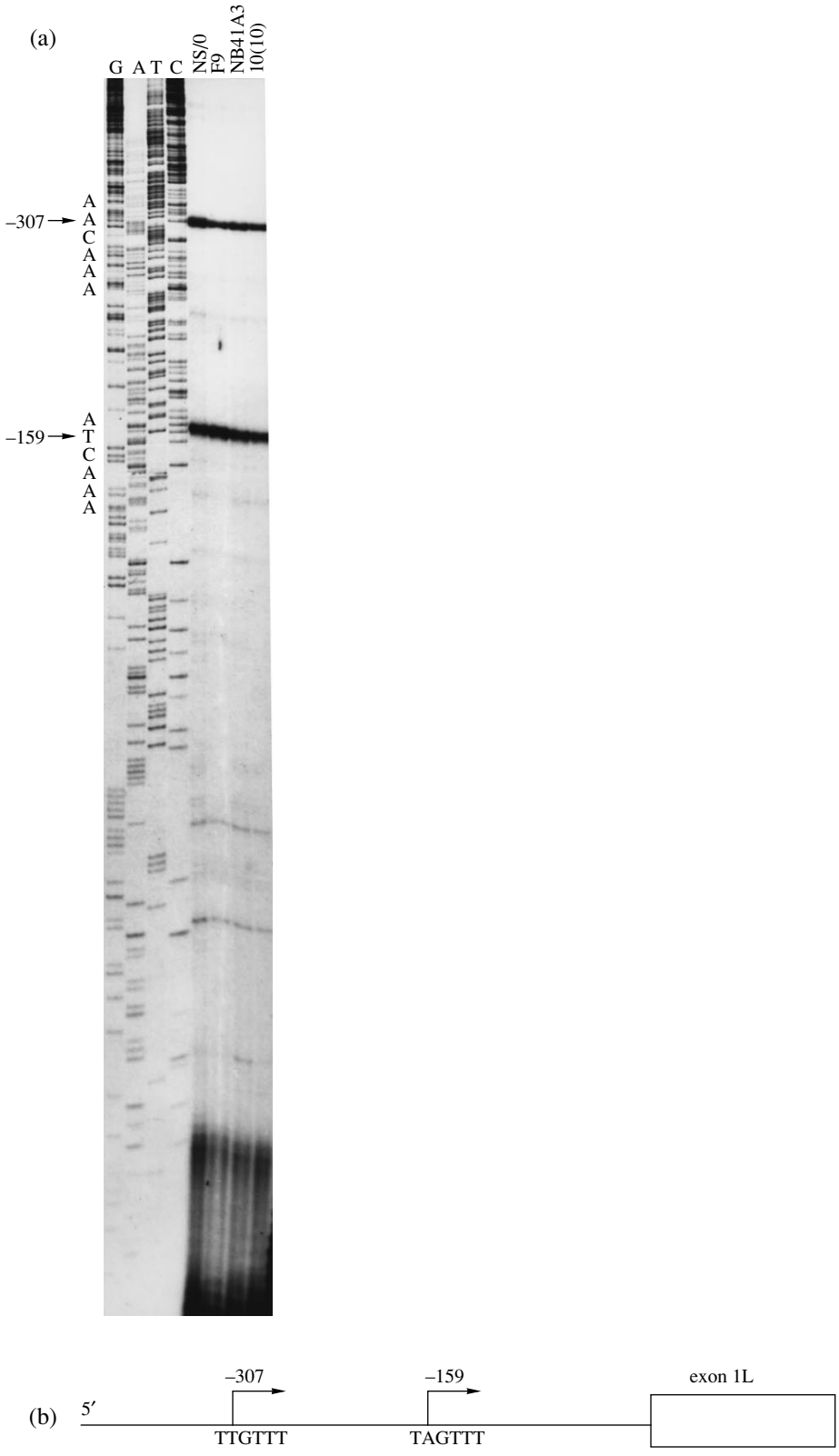

Fig. 1. Transcription initiation sites of *oct-1R* mRNA. (a) G, A, T, C are products of sequencing of *oct-1* cDNA with oligo 1 to the *oct-1L* exon. Next lanes show electrophoretic separation of products of reverse transcription from an oligonucleotide complementary to the 5'-terminal exon of *oct-1L*. RNA from cell lines NS/0, F9, NB41A3, 10(10). (b) Schematic arrangement of transcription initiation sites about exon 1L.

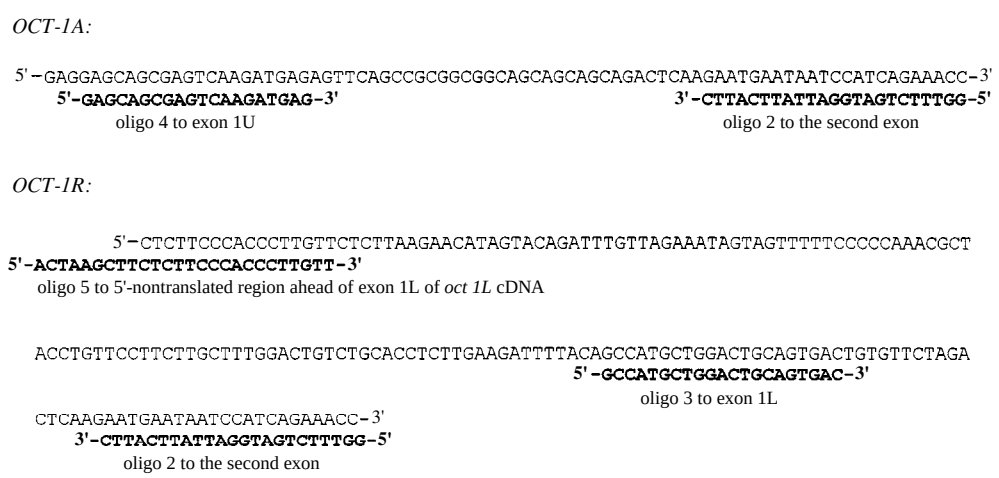

Fig. 2. mRNA and primers used for identifying mRNA of *oct-1R* and ubiquitous forms of *oct-1*. Primers are designated as described in Experimental.

cleotide corresponding to the second exon (oligo 2 and oligo 3 to exon 1L), whereas the second pair along with the labeled oligo 2 contained oligo 5 to 5' nontranslated region of *oct-1R* mRNA (Fig. 2). In both cases, the 60 and 195 bp products were synthesized by PCR on cDNA from the myeloma NS/0 cells, which correlates with the distance between primers in mature mRNA of *oct-1* (Fig. 3a, b). These PCR products were not found with RNA of other cell lines (Fig. 3). To reveal mRNA containing exon 1U, primers corresponding to exon 1U and exon 2 (oligo 4 and oligo 2, respectively) were used. The 77 bp PCR product corresponding to 5' end of mature *oct 1U* was found in all cell lines studied in this work.

Earlier, we cloned the genomic DNA fragment containing exon 1L with the promoter region and the intron region [25, 26]. Sequence analysis of cDNA of *oct-1L* and 5' region of *oct-1* gene shows that exon 1L consists of 28bp and has at the 3' end an acceptor splice site (Fig. 4). In human *oct-1*, exon 1U, encoding a long 5'-terminal domain, contains also the splicing acceptor site and binds to the following exon with formation of the same GAC triplet. Splicing produces the *oct-1* mRNA forms with a short 1L or a long 1U 5'-terminal exon. No *oct-1* forms with both exons were found.

DISCUSSION

The *oct-1* mRNA, whose 5'-terminal sequence differs from that of corresponding mRNA from nonlymphoid cells, was isolated from the myeloma cell line NS/0. The table shows N-terminal amino acid sequences of isoforms of Oct-1 protein deduced from the primary structure of *oct-1* mRNA obtained from different cell lines. The results correlate with the data listed in the Table. The *oct-1R* form, so far detected only in myeloma cells of B series, is absent from nonlymphoid cells. The mutual arrangement of the 1U and 1L exons in the OTF-1 locus is not yet established. No simultaneous presence of these two 5' exons in *oct-1* mRNAs has been detected by us as well as by others. This fact can be explained by alternative splicing, which is probably associated with initiation of transcription from different promoters localized in the 5' regions of preceding exons.

The exon splicing is carried out with involvement of donor and acceptor sequences localized at the flanks of introns and exons. Fig. 4 shows sequences that are involved in splicing of 5'-terminal exons of *oct-1* pre-mRNA. The donor site of exon 1U and acceptor site of exon 2 of human *oct-1* were established earlier [22]. As seen in Fig. 3, the 5' donor sequence of exon 1L well correlates with consensus [29, 30], and the exon 1L joining to exon 2 can be efficient.

Synthesis of the Oct-1R isoform in lymphoid cells of B series suggests its participation in regulation of immunoglobulin gene expression. Molecular mechanisms of tissue-specific regulation of expression of these genes are still not fully elucidated. More than ten years ago Schaffner put forward a hypothesis [31] that transcription factor Oct-2, synthesized mainly in B cells and neurons, tissue-specifically activates gene promoters of light and heavy immunoglobulin chains. This hypothesis was transformed into a postulate that was included in many reviews and manuals, although it was soon shown that the "knock-out" of the *oct-2* locus in somatic B cells does not cause a decrease in the immunoglobulin gene transcription [6, 7]. In accordance with a new model that was confirmed in a number of works [13–15, 32], Oct-1 is also involved in synthesis of immunoglobulins along with Oct-2. These proteins are not completely interchangeable: in mice with oct-2–/– genotype, synthesis of IgM, IgG1,

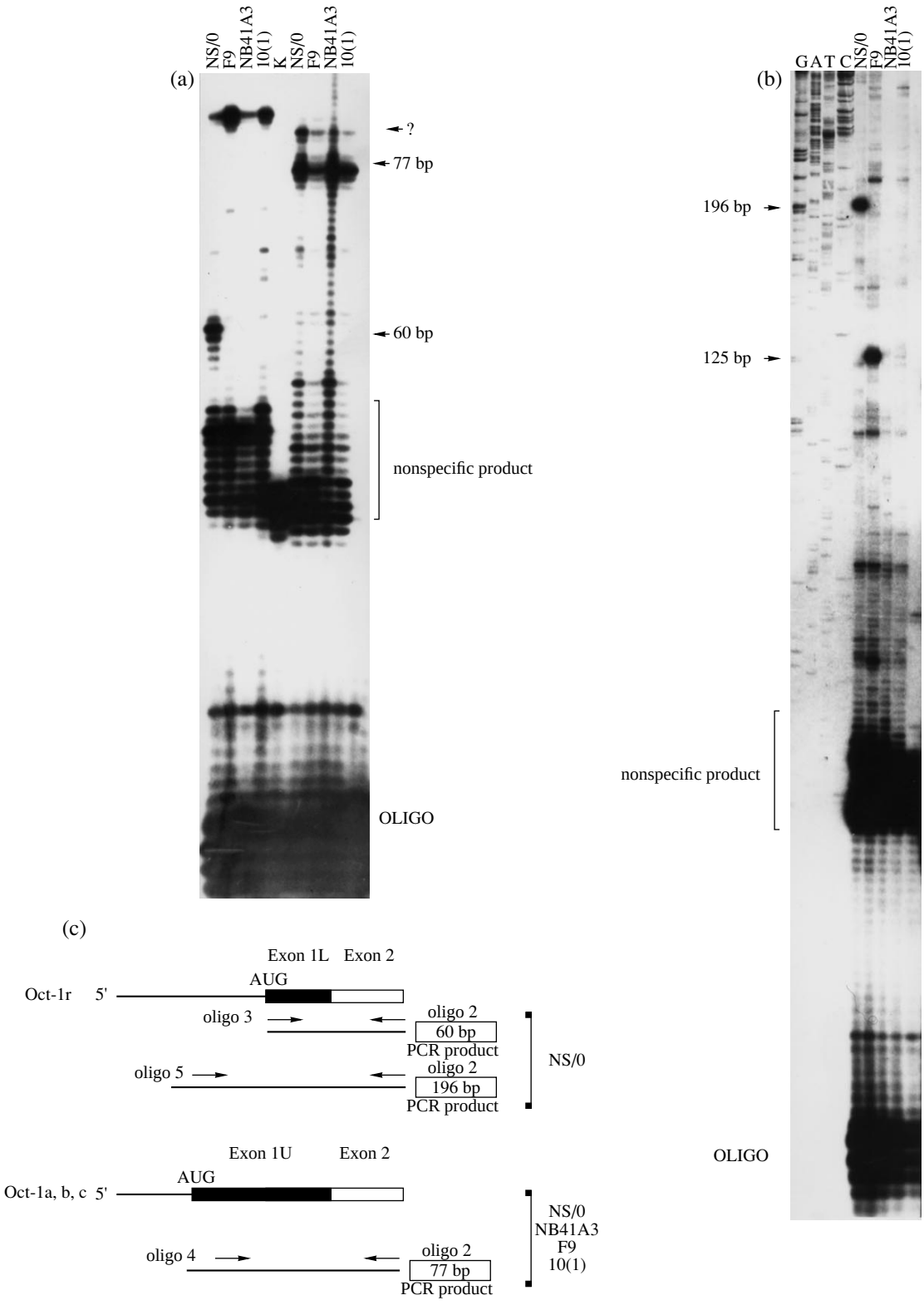

Fig. 3. Identification of PCR products. (a) Electrophoretic separation of PCR products obtained on single-stranded cDNA in the reaction of reverse transcription from oligo 2. Lanes 1–4, PCR products obtained using oligos 2 and 3; 5, control reaction without products of reverse transcription; the rest lanes, PCR products obtained using oligos 4 and 2. (b) G, A, T, and C, products of sequencing reactions. The rest lanes - PCR products obtained with oligos 5 and 2. Arrows mark specific reaction products. Cell lines (sources of mRNA) are indicated over corresponding lanes. (c) A scheme of an experiment supporting the existence of alternative tissue-specific splicing of *oct-1R* mRNA.

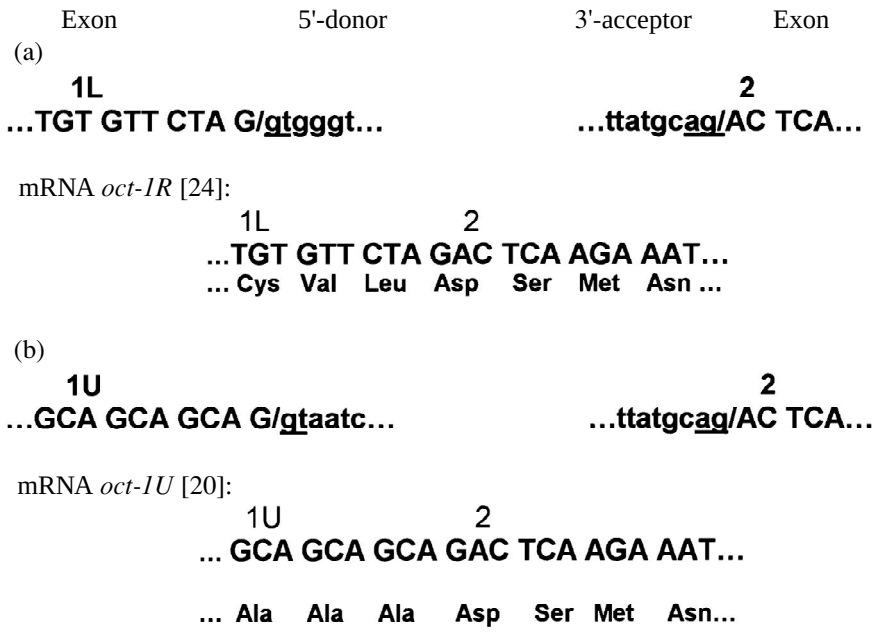

Fig. 4. Base sequences involved in splicing of 5' exons of *oct-1* pre-mRNA. (a) Exons 1L and 2; (b) exons 1U and 2. Consensus of donor A/CAGgtaagt and acceptor $(Py)_nNpyag$ sequences, respectively [29, 30].

IgG3, and IgG2B isotypes is altered [33]. This is supposed to be due to structural differences between remote enhancers localized in the 3' region of the heavy-chain genes.

An obligatory component of activation of the immunoglobulin gene transcription is a co-activator OCA-B. This protein synthesized in B cells only [13–15] establishes contacts with POUs and POUh domains [15] as well as with A5 of an octamer ATGCAAAT [34]. Synthesis of Oct-1R in lymphoid cells may be indicative of additional possibilities for tissue-specific regulation of immunoglobulin genes. In addition to structural differences in 5' region of Oct-1R and other known forms of Oct-1, it contains no exons 4, 15, and 16 [24]. It was shown previously [19] that alternative splicing of human *oct-1* pre-mRNA results in formation of an isoform Oct-1B that activates the histone H2B transcription markedly more efficiently than Oct-1A. Recently, different variants of alternative splicing of the *oct-1* 3' region have been detected [35]. However, only in the case of Oct-1B, a definite function characteristic of this isoform was found. Oct-1 is known as a ubiquitous transcription factor regulating expression of the cell housekeeping genes. This work shows for the first time the tissue-specific synthesis of a definite Oct-1 isoform in myeloma cells. Activation of *oct-1* takes place in all cells, but

N-Terminal amino acid sequences of transcription factor Oct-1 deduced from the primary structure of 5'-terminal sequences of different isoforms of *oct-1* mRNA

Factor	Amino acid sequence	Reference
mOct-1r	MLDCSDCVLDSRMNNPSETNKSSMES…	[24]
mOct-1, exon 1L	MLDCSDCVLD	[25, 26]
mOct-1	MNNPSETNKSSMES…	[20]
mOct-1	NKSSMES…	[27]
hOct-1	MADGGAASQDESSAAAAAAADSRMNNPSETSKPSMES…	[22]
hOct-1, exon 1U	MADGGAASQDESSAAAAAAAD	[22]
hOct-1, exon 2	SRMNNPSETSKPSMES…	[22]
hOct-1	MNNPSET…	[27]
hOct-1	MNNPSET…	[28]

Note: mOct-1 and hOct-1, mouse and human Oct-1, respectively.

specific interaction of the products of *oct-1* alternative splicing with the coactivator proteins (OCA-B, VP-16) or universal transcription factors and formation of oligomeric protein complexes can be a key mechanism that activates promoters or enhancers of certain genes, thus regulating their expression.

ACKNOWLEDGMENTS

Authors are grateful to G.F. Yudina for help in the manuscript preparation. The work was supported by the State program Frontiers in Genetics, the Russian Foundation for Basic Research (project no. 99-04-49180), the State Human Genome program, and a grant for young scientists from the Presidium of the Russian Academy of Sciences.

REFERENCES

1. Verrijzer, C.P. and van der Vilet, P.C., *Biochim. Biophys. Acta*, 1993, vol. 1173, pp. 1–21.
2. Ryan, A.K. and Rosenfeld, M.G., *Genes Dev.*, 1997, vol. 11, pp. 1207–1225.
3. Mittal, V., Cleary, M.A., Herr, W., and Hernandez, N., *Mol. Cell. Biol.*, 1996, vol. 16, pp. 1955–1965.
4. Sturm, R.A., Das, G., and Herr, W., *Genes. Dev.*, 1988, vol. 2, pp. 1582–1599.
5. Shah, P.C., Bertolino, E., and Singh, H., *EMBO J.*, 1997, vol. 16, pp. 7105–7117.
6. Corcoran, L.M., Karvelas, M., Nossal, G.J., Ye, Z.S., Jacks, T., and Baltimore, D., *Genes Dev.*, 1993, vol. 7, pp. 570–582.
7. Feldhaus, A.L., Klug, C.A., Arvin, K.L., and Singh, H., *EMBO J.*, 1993, vol. 12, pp. 2763–2772.
8. Schoeler, H.R., *Trends Genet.*, 1991, vol. 7, pp. 323–329.
9. Stepchenko, A.G., *FEBS Lett.*, 1994, vol. 337, pp. 175–178.
10. Herr, W. and Cleary, M.A., *Genes Dev.*, 1995, vol. 9, pp. 1679–1693.
11. Roberts, S.G. and Green, M.R., *Nature*, 1995, vol. 375, pp. 105–106.
12. Luo, Y., Fujii, T., Gerster, T., and Roeder, R.G., *Cell*, 1992, vol. 71, pp. 231–241.
13. Gstaiger, M., Knoepfel, O., Georgiev, O., Schaffner, W., and Hovens, C.M., *Nature*, 1995, vol. 373, pp. 360–362.
14. Luo, Y. and Roeder, R.G., *Mol. Cell. Biol.*, 1995, vol. 15, pp. 4115–4124.
15. Strubin, M., Newell, J.W., and Matthias, P., *Cell*, 1995, vol. 80, pp. 497–506.
16. Tanaka, M. and Herr, W., *Cell*, 1990, vol. 60, pp. 375–386.
17. Tanaka, M. and Lai, J.-S., *Cell*, 1992, vol. 68, pp. 755–767.
18. Zurlling, S., Annweiler, A., and Wirth, T., *Nucleic Acids Res.*, 1994, vol. 22, pp. 1655–1662.
19. Das, G. and Herr, W., *J. Biol. Chem.*, 1993, vol. 268, pp. 25026–25032.
20. Suzuki, N., Werner, P., Cieselka, T., Gruss, P., and Schoeler, H.R., *Nucleic Acids Res.*, 1993, vol. 21, pp. 245–252.
21. Wirth, T., Priess, A., Annweiler, A., Zurlling, S., and Oeler, B., *Nucleic Acids Res.*, 1991, vol. 19, pp. 43–51.
22. Sturm, R.A., Cassady, J.L., Das, G., Romo, A., and Evans, G.A., *Genomics*, 1993, vol. 16, pp. 333–341.
23. Chomczynski, P. and Sacchi, N., *Anal. Biochem.*, 1987, vol. 162, pp. 156–159.
24. Stepchenko, A.G., *Nucleic Acids Res.*, 1992, vol. 20, p. 1419.
25. Pankratova, E.V. and Polanovsky, O.L., *FEBS Lett.*, 1998, vol. 426, pp. 81–85.
26. Pankratova, E.V. and Polanovsky, O.L., *Mol. Biol.*, 1998, vol. 32, pp. 970–975.
27. Lerner, A.D., D'Adamio, L., Diener, A.C., Clayton, L.K., and Renherz, E.L., *J. Immunol.*, 1993, vol. 151, pp. 3152–3162.
28. Sturm, R.A., Das, G., and Herr, W., *Genes Dev.*, 1988, vol. 2, pp. 1582–1599.
29. Senepathy, P., *Proc. Natl. Acad. Sci. USA*, 1988, vol. 85, pp. 1129–1133.
30. Green, M.E., *Annu. Rev. Cell Biol.*, 1991, vol. 7, pp. 559–599.
31. Schaffner, W., *Trends Genet.*, 1989, vol. 5, pp. 37–39.
32. Shah, P.C., Bertolino, E., and Singh, H., *EMBO J.*, 1997, vol. 16, pp. 7105–7117.
33. Corcoran, L.M., and Karvelas, M., *Immunity*, 1994, vol. 1, pp. 635–645.
34. Chasman, D., Cepek, K., Sharp, P.A., and Pabo, C.O., *Genes Dev.*, 1999, vol. 13, pp. 2650–2657.
35. Riss, J. and Laskov, R., *Biochim. Biophys. Acta*, 1999, vol. 1944, pp. 295–308.