Transcription Regulatory *cis*-Element $(GCC)_8$ in the 5'-Untranslated Region of the Gene for Human Very-Low-Density Lipoprotein Receptors

K. B. Kuteikin-Teplyakov, S. V. Orlov, E. B. Dizhe, and A. P. Perevozchikov

Translated from *Byulleten' Eksperimental'noi Biologii i Meditsiny*, Vol. 135, No. 5, pp. 552-557, May, 2003
Original article submitted September 3, 2001

The 5'-untranslated region of very-low-density lipoprotein (VLDL) receptor gene includes 2 groups of triplet repeats $(GCC)_n$. Four repeats are localized near the promoter region in position 15...23 from the transcription initiation site. Eight repeats were detected in position 573...597. Sequence $(GCC)_8$ in VLDL receptor gene forms specific complexes with nuclear proteins of HepG2 cells, the formation of these complexes depended on Zn^{2+}. Superexpression of the CGGBP-20 protein interacting with long sequences $(GCC)_n$ suppressed transcriptional activity of VLDL receptor gene. Removal of fragment (397…616) containing *cis*-element $(GCC)_8$ from the 5'-untranslated region of VLDL receptor gene led to activation of the linked marker gene *cat* in Hutu80 cells, but did not abolish the repressor effect of CGGBP-20 protein. Our results suggest that $(GCC)_n$-binding proteins differing from CGGBP-20 regulate activity of the VLDL receptor gene via *cis*-element $(GCC)_8$.

Key Words: *gene for very-low-density lipoprotein receptor; triplet repeats GCC; cis-element; DNA-binding proteins*

Receptors for human VLDL belong to the receptor family that includes receptors for low-density lipoproteins (most similar to VLDL receptors), megalin receptors, apolipoprotein E-2 receptors, proteins similar to VLDL receptors (*e.g.*, LRP-5 and LRP-11), and proteins of other groups [8]. VLDL receptor consists of 5 functional domains: N-terminal ligand-binding domain containing 8 cysteine-rich repeats; domain mediating acid-dependent dissociation of the ligand-receptor complex (homologous to epidermal growth factor precursor); O-glycosylated domain; transmembrane domain; and C-terminal cytoplasmic domain containing internalization signal sequence [9].

VLDL receptor gene is localized on the short arm of chromosome 9 and is expressed primarily in adipose tissue, muscles, small intestine, and brain. Small amounts of the corresponding RNA are expressed in liver cells [12,14].

The 5'-untranslated region of VLDL receptor gene contains 2 groups of triplet repeats $(GCC)_n$. Four repeats are localized near the promoter region in position 15...23 from the transcription initiation site. Eight repeats occur in position 573...597 [12]. The role of these sequences in the regulation of transcription remains unknown.

The functional role of triplet repeats $(GCC)_n$ (n=3-4) present in the regulatory regions of various mammalian genes was illustrated by the example of the regulatory region in the gene for mouse ribosomal protein L32 (*rpL32*) [3-5]. We found that nuclear proteins of mammalian cells (*e.g.*, human cells) can specifically interact with nucleotide sequence $GC(GCC)_4$, a component of a more complex synergic *cis*-element localized in the rpL32 gene promoter in position -19...+10 from the transcription initiation site [4].

Institute of Experimental Medicine, Russian Academy of Medical Sciences, St. Petersburg. ***Address for correspondence:*** app@iem.sp.ru. Kuteikin-Teplyakov K. B.

0007-4888/03/1355-0470$25.00 © 2003 Plenum Publishing Corporation

Much attention is given to triplet repeats $(GCC)_n$ (n>30) in human chromosomal DNA. Expansion of 50-200 GCC triplet repeats in the 5'-untranslated region of the *FMR-1* gene is associated with fragile X syndrome and other serious disease. Similar blocks of repeats were found in other genes of the FMR family [11]. Probably, the critical event in expansion of GCC triplets is their impaired interaction with specific $(GCC)_n$-binding nuclear proteins. One of these proteins (CGGBP-20) was identified and described [7].

Here we studied structural and functional properties of the potential *cis*-acting GC-rich $(GCC)_8$ element present in the 5'-untranslated region of VLDL receptor gene in position 573...597 from the transcription initiation site.

MATERIALS AND METHODS

We used reagents from Sigma, Serva, Pharmacia, Boehringer Mannheim, and Russian companies and enzymes for genetic engineering studies (Fermentas and SibEnzim).

Human HepG2 (hepatoma) and Hutu80 cells (duodenal adenocarcinoma) were obtained from the American Collection of Cell Cultures (ATCC, National Institute of Health, USA) and maintained at the Department of Cell Cultures (Institute of Cytology, Russian Academy of Sciences). VLDL receptor gene (*vldl-r*) was kindly provided by L. Chan (Bailor Medical College). The *rpL32* gene was gifted by N. V. Tomilina (Institute of Cytology, Russian Academy of Sciences). The cDNA gene for CGGBP-20 proteins (*cggbp-20*) was obtained from I.M.A.G.E. Consorcium (clone 269133) and Research Genetics Company.

Experiments were performed with the following synthetic double-stranded oligonucleotides: fragment of the rpL32 gene promoter (-24…+11) *L32-oligo* 5'-CTTGCGCGCCGCCGCCGCCTCTTCCTTCTTCCTCG-3', $(GCC)_9$ element 5'-GCCGCCGCCGCCGCCGCCG

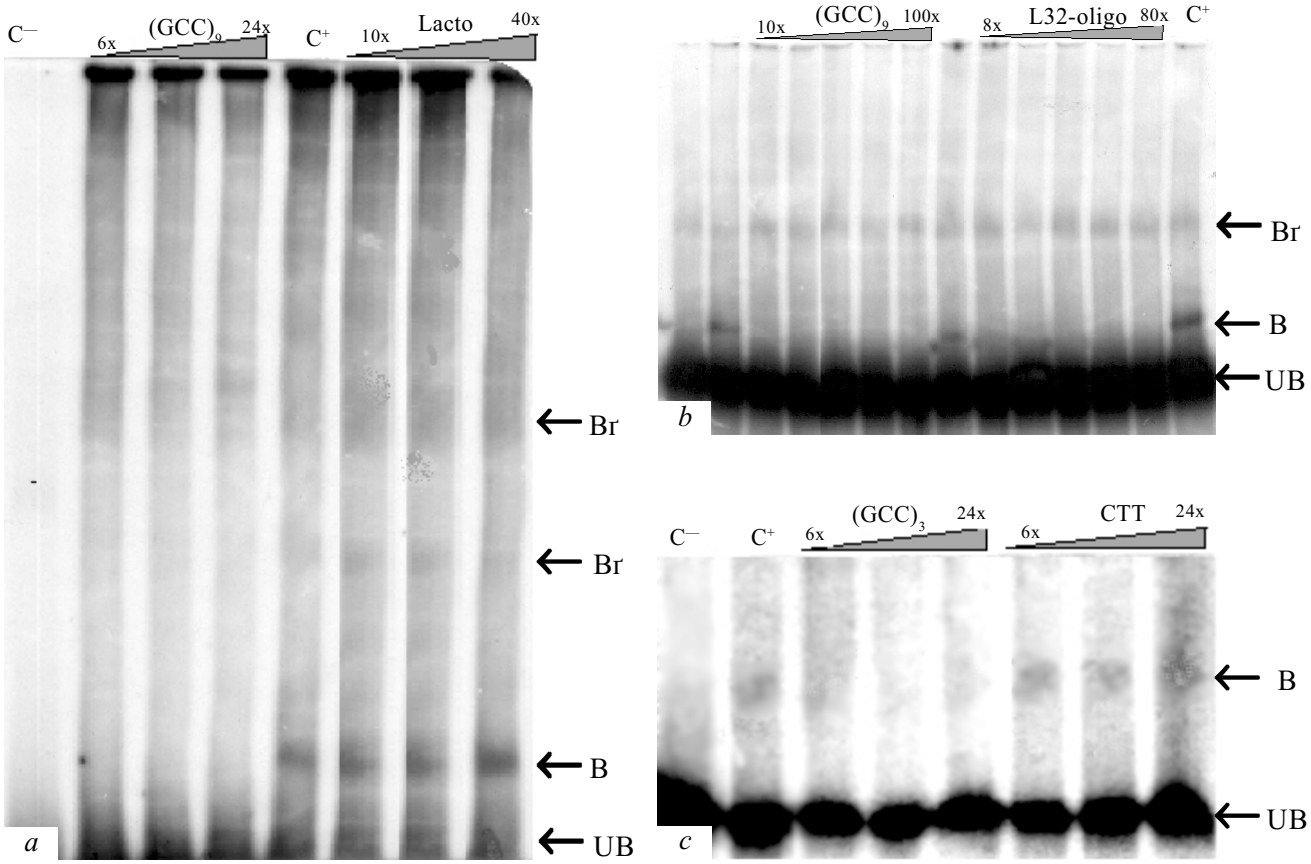

Fig. 1. Formation of DNA-protein complexes containing fragment 397…616 of the 5'-untranslated region from VLDL receptor gene with 8 GCC repeats and nuclear extracts of human hepatoma HepG2 cells. Competition with the copy of the $(GCC)_8$ element and nonspecific *Lacto* oligonucleotide (*a*); competition between the $(GCC)_8$ element of VLDL receptor gene and fragment of the *rpL32* gene promoter (*L32-oligo*) containing 4 GCC repeats for binding to proteins (*b*); competition between the polypyrimidine element and GCC element in the composite *cis*-element of the *rpL32* gene promoter (*c*). C^-, negative control (without proteins); C^+, positive control (without DNA-competing agent). Competitive DNA and their molecular excess are shown above tracks. Here and in Fig. 2: B, specific complex; B', nonspecific complexes; UB, unbound DNA.

CCGCCGCC-3', $(GCC)_4$ element 5'-TGCCGCCGC CGCCGCTC-3', polypyrimidine element *CTT-oligo* 5'-CTCTTCCTTCTTCCTCG-3', and lactoferrin-binding element *Lacto-oligo* 5'-CTAGTGCAAGTGCCA-3'. Genetic constructions were obtained by routine techniques of genetic engineering [2].

Synthetic oligonucleotides and VLDL receptor gene fragment (397...616) were radiolabeled with Klenow fragment of *E. coli* DNA polymerase I (3'-end) and phage T4 polynucleotide kinase (5'-end) using [α-^{32}P]-dCTP and [γ-^{32}P] ATP, respectively (NPO GIPKh, St. Petersburg).

Gel retardation assay was performed as described elsewhere [13]. The reaction medium contained 100 mM NaCl, 10.5 mM HEPES (pH 7.9), 0.8 mM Na_2-EDTA, 0.15 mM Na_2-EGTA, 0.8 mM DTT, 6 mM $MgCl_2$, 0.15 mM PMSF, 1.5 μg fragmented DNA from herring sperm, 8000-80,000 cpm labeled DNA (2-10 ng), and 2-20 μg nuclear proteins. In competition studies we added unlabeled oligonucleotides. Binding of DNA to proteins was performed at room temperature (20 min) or at 4°C (1 h). Reaction products were separated by electrophoresis in 5% polyacrylamide gel. The gel was fixed in 10% acetic acid and 10% ethanol, dried, and exposed with Hyperfilm-βmax X-ray film (Amersham) for 1-14 days depending on activity of samples.

The cells were cultured in DMEM medium containing 10% fetal bovine serum (Flow). Nuclear extracts were obtained from cultured cells as described previously [6].

Transient transfection of mammalian cells with plasmid DNA was performed by the method of calcium phosphate precipitation. The cells were cotransfected with the pCMVlac plasmid containing bacterial *lacZ* marker gene encoding β-galactosidase and controlled by the human cytomegalovirus immediate early promoter [1]. The samples were equalized depending on the efficiency of transfection. It was determined by measuring bacterial β-galactosidase activity in cell extracts with chromogenic substrate *o*-nitrophenyl-β-D-galactopyranoside [1]. CAT activity in cell extracts was measured by CAT-test using [^{14}C]-chloramphenicol (Amersham).

RESULTS

The pVLRcat plasmid was obtained by cloning 5'-untranslated region (-699...+616 from the transcription initiation site) of VLDL receptor gene and *cat* marker gene into commercial vector pUC18. To estimate functional activity of $(GCC)_8$ the pVLRcat/del$(GCC)_8$ plasmid was constructed by deletion of the 5'-untranslated region (384...615) from the pVLRcat plasmid by restriction sites *NcoI* and *PstI*. The pCMVcgp-1 plasmid is a eukaryotic vector of *cggbp-20* gene expression controlled by human cytomegalovirus immediate early promoter.

Gel retardation assay showed that the fragment of the 5'-untranslated region in VLDL receptor gene (397...616) can form several DNA-protein complexes with nuclear proteins of HepG2 cells (Fig. 1). Specificity of these interactions was determined by addition of excess unlabeled oligonucleotides [13]. The competition between nucleotide sequences and the test fragment of VLDL receptor gene for proteins is illustrated in Fig. 1, *a*. Addition of the excess of homologous $(GCC)_9$ oligonucleotide led to replacement of the initial radiolabeled DNA from the complexe, which was manifested in a reduction of the corresponding retardation band (Fig. 1, *a*, arrow B). However, the excess non-homologous oligonucleotide *Lacto* had no effect on this DNA-protein complex; the intensity of the corresponding band remained unchanged. These data indicate that the DNA-protein complex consisting of the 5'-untranslated region from VLDL receptor gene and nuclear proteins is formed by $(GCC)_8$ nucleotide sequence in position 573...597.

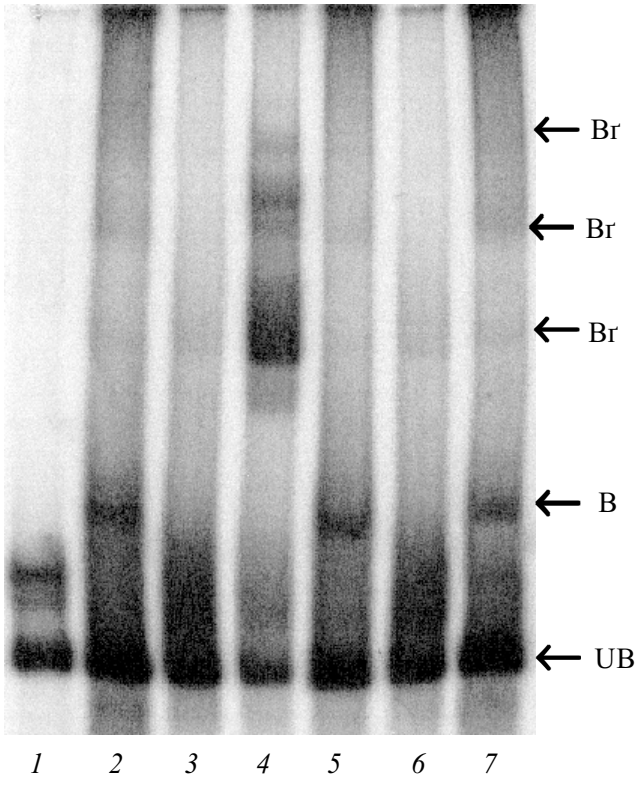

Fig. 2. Formation of DNA-protein complexes from fragment 397...616 of VLDL receptor gene and nuclear proteins of HepG2 cells depending on the presence of Zn^{2+}. Negative control (without nuclear extracts, *1*), positive control (without 1,10-phenanthroline, *2*, *7*), and samples with 1,10-phenanthroline in a concentration of 5 mM (*3*, *6*), 5 mM $ZnCl_2$ and 5 mM 1,10-phenanthroline (*4*), and 5 mM EGTA.

Competition assay with unlabeled oligonucleotide $(GCC)_9$ or fragment of the *rpL32* gene promoter (-24...+11) confirmed specificity of DNA-protein complexes containing the $(GCC)_8$ element of the 5'-untranslated region from VLDL receptor gene (Fig. 1, *b*). The fragment of the *rpL32* gene promoter containing sequence $GC(GCC)_4$ homologous to sequence $(GCC)_8$ of VLDL receptor gene fragment competed with this fragment for binding to nuclear proteins. Comparison of these oligonucleotides showed that the homologous nucleotide containing 9 GCC repeats more efficiently replaced the fragment of VLDL receptor gene from specific complexes (arrow B) than the fragment of the *rpL32* gene promoter containing only 4 GCC repeats.

Previous studies showed that the fragment of the *rpL32* gene promoter (-24...+11 from the transcription initiation site) includes the composite *cis*-element consisting of a GCC element (-19...-6) and polypyrimidine element (-5...+10). The fragment of the 5'-untranslated region from VLDL receptor gene (397...616) competes with this composite *cis*-element. It was interesting to evaluate which constituent of the composite element competes with the $(GCC)_8$ element of VLDL receptor gene for binding to proteins. Unlabeled oligonucleotide copies of the GCC element $(GCC)_4$ replace the labeled fragment of VLDL receptor gene from DNA-protein complexes (Fig. 1, *c*). However, another constituent of the composite *cis*-element in the *rpL32* gene promoter (polypyrimidine element) did not replace the gene fragment from complexes. These results attest to autonomous existence of GCC element (and its functional activity) at least in human VLDL receptor gene.

GC-rich *cis*-elements interact with transcription factors containing the Zn^{2+}-finger DNA-binding domain [10]. However, the only described DNA-binding CGGBP-20 protein specifically interacting with $(GCC)_n$ sequences does not contain the Zn^{2+}-finger domain and interacts with DNA independently on the presence of Zn^{2+} [7]. It was interesting to determine the role of Zn^{2+} in DNA-protein interactions of the $(GCC)_8$ sequence and nuclear proteins of HepG2 cells. Binding of the 5'-untranslated region from VLDL receptor gene to nuclear proteins of human hepatoma HepG2 cells in the presence of 1,10-phenanthroline inhibits the formation of DNA-protein complex (Fig. 2). Addition of EGTA binding bivalent cations produced a less pronounced effect, which is consistent with low Zn^{2+}-chelating capacity of EGTA. These data indicate that Zn^{2+}-dependent proteins (probably, Zn^{2+}-finger proteins) are involved in the formation of DNA-protein complexes specific for HepG2 cells. Destruction of these complexes with 1,10-phenanthroline is irreversible, since their formation was not restored after addition of excess Zn^{2+}. Our previous studies showed that the interaction of nuclear proteins from

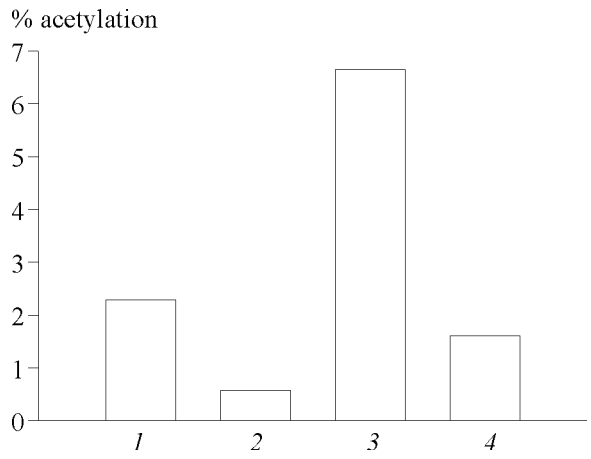

Fig. 3. Study of functional activity of the GCC element in the fragment of VLDL receptor gene by CAT-test. Ordinate: relative activity of the regulatory region in VLDL receptor gene expressed in percents of the content of acetylated chloramphenicol forms (% acetylation). pVLRcat (*1*), pVLRcat and pCMVcgp-1 (*2*), pVLRcat/del$(GCC)_8$ (*3*), and pVLRcat/del$(GCC)_8$ and pCMVcgp-1 (*4*).

HepG2 cells with the fragment of the *rpL32* gene promoter containing $GC(GCC)_4$ depends on the presence of Zn^{2+} [5]. Probably, DNA-binding proteins differing from CGGBP-20 are involved in DNA-protein interactions with GCC elements of VLDL receptor genes and *rpL32*.

The fragment of the 5'-untranslated region from VLDL receptor gene that included 2 types of repeats possessed protein-binding activity. It remained unclear whether these repeats play a role of *cis*-acting transcription elements in the regulation of gene expression. Functional activity of $(GCC)_8$ repeats in position 573...597 was evaluated. Hutu80 cells originating from cells of the small intestine and characterized by more intensive expression of VLDL receptor gene than hepatoma HepG2 cells were used to determine functional activity of $(GCC)_8$ [14]. In parallel experiments Hutu80 cells were transfected with pVLRcat plasmid containing bacterial reporter gene encoding CAT and controlled by 5'-untranslated region of VLDL receptor gene or pVLRcat/del$(GCC)_8$ plasmid with deleted segment (230 bp) of the 5'-untranslated region that included $(GCC)_8$. Experiments with cotransfection showed that superexpression of CGGBP-20 protein 4-fold reduces *cat* gene activity (Fig. 3). Deletion of the region (384...615) containing *cis*-element $(GCC)_8$ increased activity of this gene by 2.9 times, but did not block the repressive effect of CGGBP-20 (Fig. 3). It was probably related to removal of the $(GCC)_8$ element acting as a negative regulator of transcription. Functional activity of $(GCC)_8$ is realized via the interaction with DNA-binding proteins differing from CGGBP-20. These data are consistent with the results of gel retardation assay. CGGBP-20 protein can reduce activity of the *cat* marker gene independently on *cis*-element $(GCC)_8$, which is probably associated with

the interaction of this protein with other region of GCC triplet repeats in VLDL receptor gene (15...23).

Our results indicate that the nucleotide sequence containing triplet repeats $(GCC)_8$ and localized to the right from the transcription initiation site in the 5'-untranslated region of human VLDL receptor gene (position 573...597) can be considered as a potential *cis*-acting transcription element of this gene. In HepG2 cells this sequence interacts at least with one DNA-binding nuclear protein (probably, Zn^{2+}-finger protein). Deletion of a 230-bp segment containing triplet repeats $(GCC)_8$ in the 5'-untranslated region of VLDL receptor gene led to activation of the *cat* marker gene in Hutu80 cells linked to the 5'-regulatory region in this gene. However, exogenous CGGBP-20 protein does not interact with $(GCC)_8$ of VLDL receptor gene in Hutu80 cells. This protein affects expression of the *cat* gene, which is probably associated with other groups of GCC triplets in VLDL receptor.

These data show that various human cells (HeLa, HepG2, and Hutu80) include several nuclear GCC-binding proteins.

We are grateful to A. D. Denisenko for careful reading of the manuscript and helpful remarks.

This work was supported by the Russian Foundation for Basic Research (grants Nos. 00-04-49426, 01-04-06287, 01-04-06288).

REFERENCES

1. K. Gorman, *DNA Cloning* [in Russian], Ed. D. M. Glover, Moscow (1988), pp. 409-463.
2. T. Maniatis, E. Frich, and D. Sembruk, *Molecular Cloning. Methods of Genetic Engineering* [in Russian], Moscow (1984).
3. S. V. Orlov, E. B. Dizhe, K. B. Kuteikin, *et al.*, *Dokl. Ros. Akad. Nauk*, **366**, No. 2, 262-265 (1999).
4. S. V. Orlov, K. B. Kuteikin, E. B. Dizhe, *et al.*, *Biokhimiya*, **64**, No. 2, 255-261 (1999).
5. A. P. Perevozchikov, S. V. Orlov, and K. B. Kuteikin, *Dokl. Ros. Akad. Nauk*, **338**, No. 3, 411-415 (1994).
6. N. C. Andrews and D. V. Faller, *Nucl. Acids Res.*, **19**, 2499 (1991).
7. H. Deissler, M. Wilm, B. Genc, *et al.*, *J. Biol. Chem.*, **272**, 16,761-16,768 (1997).
8. M. M. Hussain, D. K. Strickland, and A. Bakillah, *Ann. Rev. Nutr.*, **19**, 141-172 (1999).
9. H. Jingami and T. Yamamoto, *Curr. Opin. Lipidol.*, **6**, 104-108 (1995).
10. P. J. Mitchell and R. Tjian, *Science*, **245**, 371-378 (1989).
11. J. E. Parrish, B. A. Oostra, A. J. Verkerk, and C. S. Richards, *Nat. Genet.*, **8**, 229-235 (1994).
12. J. Sakai, A. Hoshino, S. Takahashi, *et al.*, *J. Biol. Chem.*, **269**, No. 3, 2173-2182 (1994).
13. E. Schreiber, P. Matthias, M. M. Muller, and W. Schaffner, *Nucl. Acids Res.*, **17**, 6419 (1989).
14. J. C. Webb, D. D. Patel, M. D. Jones, *et al.*, *Hum. Mol. Genet.*, **3**, No. 4, 531-537 (1994).