Available online at www.sciencedirect.com

SCIENCE DIRECT

Gene 359 (2005) 127 – 137

GENE
SECTION
FUNCTIONAL GENOMICS

www.elsevier.com/locate/gene

Structural organization of the human complexin 2 gene (*CPLX2*) and aspects of its functional activity

Natalja M. Raevskaya [a], Lyudmila V. Dergunova [a,*], Irina P. Vladychenskaya [a], Vasily V. Stavchansky [a], Maria V. Oborina [a], Andrey B. Poltaraus [b], Svetlana A. Limborska [a]

[a] *Institute of Molecular Genetics, Russian Academy of Sciences, Moscow, 123182, Russia*
[b] *Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991, Russia*

Received 4 May 2005; received in revised form 7 July 2005; accepted 9 July 2005

Received by S.M. Mirkin

Abstract

We report here on the in vitro and in silico characterization of the organization of the human complexin 2 (*CPLX2*) gene. This encodes for a protein of 134 amino acid residues, contains five exons, is localized on human chromosome 5q35.3, and spans more than 87 kb. We performed in silico analysis of the *CPLX2* 5′untranslated region (UTR) and propose an alternative variant of the gene transcript. Compared to the mRNA reported earlier [McMahon, H.T., Missler, M., Li, C., Sudhof, T.C., 1995. Complexins: cytosolic proteins that regulate SNAP receptor function. Cell 83, 111–119.], this transcript bears a partly altered 5′-UTR associated with the same open reading frame. Both *CPLX2* transcripts share exons III–V; the alternative transcript is devoid of exons I and II, and includes exon A instead. Exon A is localized within *CPLX2* intron 2 about 7 kb upstream to exon III. Using reverse transcription polymerase chain reaction (RT-PCR) we detected both types of transcripts in human brain mRNA. In silico data suggest that two putative alternative TATA-less promoter regions separated by 74 kb govern the expression of two *CPLX2* transcripts. Several potential transcription start sites were detected by primer extension for each of two alternative *CPLX2* transcripts. The relative abundance of the alternative transcripts was investigated in human and rat forebrain, cerebellum, and hippocampus. Whereas both transcripts were detected in human and rat brain, their expression levels were found to vary significantly among the regions investigated. The organization of *CPLX2* transcripts is conserved in humans and rodents.

Keywords: Gene of neuron specific protein; Alternative transcripts; 5′- and 3′-untranslated regions; Promoter, in silico analysis

1. Introduction

Global sequencing of genomic and expressed sequences makes it possible to detail the structural organization of numerous known genes. Comparisons between genomic sequences and their transcripts reveal previously unknown exons coding for alternative transcripts and can help identify some features of mRNA untranslated regions (UTRs), whose functional elements often play critical roles in post-transcriptional regulation of gene expression. Specification of the genomic organization is especially important for those genes that have been identified by screening cDNA libraries with oligonucleotide probes corresponding to amino acid sequences of peptide products. The sequences and structures of the untranslated regions of the genes thus identified often remain uncovered. The gene for the cytoplasmic protein complexin 2 belongs to the group of genes identified by such an approach. Complexin 2, along with other proteins involved in synaptic transmission, was first found in rat brain homogenates in the

Abbreviations: EST, expressed sequence taq; bp, base pairs; kb, kilobase pairs; nt, nucleotide(s); mRNA, messenger RNA; cDNA, DNA complementary to RNA; DNase, deoxyribonuclease; NR, non-redundant; dATP, deoxyadenosine triphosphate; RNase, ribonuclease; RT, reverse transcription; PCR, polymerase chain reaction; UTR, untranslated region; uORF, upstream open reading frame; GAPDH, glyceraldehyde-3-phosphate-dehydrogenase.

* Corresponding author. Department of Human Molecular Genetics, Institute of Molecular Genetics RAS, Kurchatov sq., 2, 123182, Moscow, Russia. Tel.: +7 095 196 1858; fax: +7 095 196 0221.

E-mail address: dergunov@img.ras.ru (L.V. Dergunova).

0378-1119/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.gene.2005.07.005

mid-1990s (McMahon et al., 1995). As a result of screening cDNA libraries, the *CPLX2* mRNA sequence was established for mouse, rat, and human (McMahon et al., 1995; Takahashi et al., 1995). The genomic organization of coding exons corresponding to these transcripts was determined for the named organisms. However, the sizes and organization of noncoding regions, potential noncoding exons, and regulatory genomic sequences controlling the expression of *CPLX2* have not been described so far for any species.

We have previously isolated the brain-specific sequence Hfb1 (GenBank accession no. Y15161) and studied the organization of the corresponding genomic clone Ghfb (GenBank accession no. AF318943) obtained from a cosmid library of human chromosome 5 (Buiakova et al., 1992; Dergunova et al., 2001, 2003a). Analysing Ghfb we have found a new extended sequence corresponding to the 3′-UTR of the *CPLX2* mRNA. Here we have focused on the structural organization of the 5′-UTRs in human *CPLX2* transcripts. We have detected a novel *CPLX2* transcription variant that differs from the mRNA reported earlier by a portion of its 5′-UTR. This is encoded by an additional exon A located within intron II, 7 kb upstream to exon III. Using semiquantitative RT-PCR (sqRT-PCR) we have estimated the relative abundance of both alternative *CPLX2* transcripts in human and rat brain regions.

2. Materials and methods

2.1. RNA isolation and preparation of cDNA

The investigation was performed after the approval from the local Ethics Committee was obtained. Total RNA was isolated from human tissues (one individual, male, 60 years old, post mortem interval 4 h) and those of adult male Wistar rats (4 animals) using guanidine thiocyanate (Chomczynski and Sacchi, 1987). RNA integrity was assessed by analysing the ratio between rRNA bands after agarose gel electrophoresis under denaturing conditions. RNA samples were stored at −70 °C under ethanol. Residual genomic DNA was removed from the total RNA samples by treating each sample with RNase-free DNase I (MBI Fermentas, Vilnius, Lithuania) in accordance with the supplier's recommendations. RNA was then extracted with a 1 : 1 phenol:chloroform mix and precipitated with sodium acetate (3.0 M, pH 5.2). An RNA poly(A)$^+$ fraction was obtained by column affinity chromatography on microcrystalline oligo(dT) cellulose (Pharmacia, Rockville, MD, USA). The fraction was quantified by spectrophotometry.

2.2. Reverse transcriptase polymerase chain reaction (RT-PCR)

Aliquots of 5 μg of DNAse I-treated total RNA from human cerebellum were taken for the first strand cDNA synthesis. This was carried out with random primers using a RevertAid™ First Strand cDNA Synthesis Kit (MBI Fermentas) in accordance with the manufacturer's recommendations. First strand cDNA was used as a template for the PCR with gene-specific primers. All amplifications were performed in a Tercyc MC2 thermocycler (DNA Technology, Moscow, Russia). Primers specific to *CPLX2* were designed using OLIGO Primer Analysis software (Wojciech and Piotr Rychlik Copyright, version 6.31); the primer sequences and nucleotide coordinates are summarized in Table 1. PCR

Table 1
Nucleotide sequences and locations of the primers corresponding to human *CPLX2*, *GAPDH* and rat *cplx2*, *gapdh* transcripts

Primer	Nucleotide sequence, 5′-3′	Acc. no, location, *n*
Human *CPLX2*_v1		
		EMBL-DataBank BN000499
F1	GAAGAGGGGGAGGGAGAAACG	179–199
F2	AGCGACTGAAGGTGCCCG	295–312
F3	GCTAAGGCACGCTAACCAGA	398–417
F4	AGGCGGAGCGGGAGAAGGTCC	593–613
F5	CCCTCACACCTCCCTTCATC	984–1003
F6	CATTTGAAGGCTGCTGTGTG	2334–2353
R1	CCTTACGCTCCTCCTCCTGCTGCCGC	549–574
R2	GATGAAGGGAGGTGTGA	987–1003
R3	CCCAGCCTCCCACCCAGAT	1351–1369
R4	CATCCATGCACACAACTGTC	2374–2393
R5	CTCCCTTGTGTGTTGTGTG	3837–3855
R8	CCCCTCTTCCTTGACTCCTCCGCC	164–187
Human *CPLX2*_v2		
		EMBL-DataBank BN000500
F7	CAGTGGCTCAGACGGTTGC	192–210
F8	CATCTGGAAGGGGAGCGGT	15–33
R2	GATGAAGGGAGGTGTGA	871–887
R6	ACCTTCTCCCGCTCCGCCT	477–495
R7	CTCTCCTCTTCCTCCTCCTCCTCGT	630–654
R9	GGTGGTTCCTTCTGCTATTCC	147–167
Human *GAPDH*		
		GenBank NM_002046,
For	GGAAGGTGAAGGTCGGAGTC	80–99
Rev	GAGGGGGCAGAGATGATGA	431–449
Rat *cplx2*_v1		
		GenBank U35099,
F9	GGAGGAGTGAGGAGGACGGG	24–53
R10	CAGCAACCGTCTAAGCCACTG	135–155
Rat *cplx2*_v2		
		GenBank D70816,
F10	CTGGCTGAGGGGCGTGATT	25–43
R10	CAGCAACCGTCTAAGCCACTG	180–196
Rat *gapdh*		
		GenBank NM_017008,
For	TGCCATCAACGACCCCTTCA	936–955
Rev	ACTCAGCACCAGCATCACCC	1104–1123