Isolation and Characterization of a Novel Zinc Finger Gene, ZNF219, and Mapping to the Human Chromosome 14q11 Region

Tsuyoshi Sakai, Atsushi Toyoda, Katsuyuki Hashimoto, and Hidekatsu Maeda

Department of Bioengineering, Faculty of Engineering, Soka University, 1-236 Tangi, Hatouji, Tokyo 192-8977, Japan, Department of Science, Kitazato University, 1-15-1 Sagamihara, Kanagawa, Tokyo 228-8555, Japan, and Division of Genetic Resources, National Institute of Infectious Diseases, Shinjuku-ku, Tokyo 162-8640, Japan

(Received 13 January 2000)

Abstract

We have isolated and characterized a novel zinc finger gene by screening a human testis cDNA library. The isolated cDNA, termed ZNF219, contains an open reading frame of 2169 nucleotides encoding 723 amino acids. Sequence analysis revealed 9 sets of Kruppel-related zinc finger structures and proline-rich regions in several parts. ZNF219 exhibited ubiquitous expression in all fetal and adult tissues examined. The transcript size was 5.5 kb in adult tissues, while the main transcript size in the embryo stage was 3.5 kb. The transcript size is developmentally regulated. When the plasmid cloned with green fluorescent protein (GFP)-tagged ZNF219 was expressed in HeLa and COS7 cells, strong fluorescence intensities were observed only in the nucleus of both cells by fluorescence confocal microscopy. These data suggest that ZNF219 may be related to the regulation of transcription and developmental regulation. Genomic structure analysis mapped ZNF219 to chromosome 14q11 between markers D14S72 and D14S990, because a sequence tagged site mapped to the locus was found in the intron region of the ZNF219 gene.

Key words: ZNF219, zinc finger protein 219, human chromosome 14q11

Zinc finger proteins comprise an important class of eukaryotic DNA binding proteins and contain an evolutionarily conserved structure that defines large families of transcription factors. The zinc finger structure is formed from invariant pairs of cysteines and histidines, which bind to a zinc ion. The most common zinc finger motif is the C2H2 type, which was initially found in the TFIIB transcription factor of Xenopus, and subsequently in the Drosophila segmentation gene, Kruppel. Since then, several hundred Kruppel-related zinc finger genes, in which the consensus zinc finger sequences CX2CX3FX5LX2HX5H are connected by an H/C linker, a characteristic 7-amino-acid domain, TGEKPYK, have been shown to be encoded by the human genome. We report here the isolation and characterization of a zinc finger gene, termed ZNF219, and its mapping to human chromosome 14q11.

We have attempted to isolate novel cDNA clones from Not I-linking cosmid clones mapped to the human chromosome Xq28 region by direct cDNA selection. When one of the cosmids clones, cKT263, was used as a probe, we isolated a 174-bp cDNA fragment. Furthermore, we performed plaque hybridization with a human testis cDNA library using the above 174-bp cDNA fragment as a probe, and succeeded in isolating five clones. The poly(A) tail and the same sequence were present in the 3′ region of all cDNA clones. The clone with the largest insert was designated as ZNF219. However, ZNF219 was not mapped to this cKT263 cosmid clone isolated from the Xq28 region. 0.83 kb of a part of cKT263 cosmid DNA was found to be 85% similar to the 3′ partial sequence of ZNF219 at the nucleotide level. In an attempt to determine the genomic structure of ZNF219, the BLAST search revealed that a sequence tagged site (STS) (Gene Bank Accession No. G27915) (Myers, R. M., unpublished data) was exhibited in the intron region of the ZNF219 gene, as shown in Fig. 1B. The UniGene collection revealed that the STS was mapped to the locus between markers D14S72 and D14S990 of the chromosome 14q11 region; therefore, we could map the ZNF219 gene to this region.

Communicated by Yoshiyuki Sakaki

* To whom correspondence should be addressed. Fax. +81-426-91-9312, E-mail: sakai@t.soka.ac.jp
† Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. AB015427.
Figure 1. A: The nucleic acid and deduced amino acid sequences of ZNF219 cDNA. The zinc finger motifs in the protein are underlined, and the cysteine and histidine residues of the nine consensus zinc finger motifs are circled and numbered. The nearest in-frame stop codon prior to the putative translation start site is indicated by an asterisk. The poly(A) signal (AATAAA) is doubly underlined. The vertical bars and the figures are the boundary lines between introns and exons, and the exon numbers, respectively. Dotted lines denote proline-rich (P) or glutamic acid-rich (E) regions. B: genomic structure of the ZNF219 cDNA. The restriction map is shown by Not I (N), Xba I (X) and EcoR I (E). The exon-intron structures of ZNF219 cDNA and other cDNAs located upstream of ZNF219 cDNA are shown. The black boxes show exons of 3.0-kb cDNA encoding ZNF219. The open boxes show exons of 1.4-kb EST clones. The hatched boxes show exons of 1.3-kb cDNA. The horizontal boxes show exons of 0.3-kb cDNA. The sequence tagged site denoted by STS is located in the intron region of the ZNF219 gene.
Figure 2. Northern analysis of adult tissues (A) and fetal tissues (B) using ZNF219 cDNA as a probe. The same membranes probed with GAPDH are shown beneath. The lanes are numbered according to the tissues: pancreas (1), kidney (2), skeletal muscle (3), liver (4), lung (5), placenta (6), brain (7), and heart (8). The sizes of the transcripts are shown on the right. Each lane contains equal amounts of mRNA (2 \mu g).

ZNF219 cDNA spans 3055 nucleotides, with an open reading frame of 2169 nucleotides encoding 723 amino acids. The nucleotide and predicted amino acid sequences of ZNF219 are shown in Fig. 1A. The nucleotide sequence around the methionine codon at the translation start site was confirmed to be the consensus start codon as described by Kozak, and is likely to be the site of translation initiation. In addition, a TGA stop codon exists in front of the presumed translation start site. The predicted peptide sequence indicated that the protein contains nine sets of C2H2 type zinc finger motifs. All these zinc finger motifs are related to the Kruppel-type motif.

ZNF219 contains proline-rich regions in several parts; 14.4 percent of the total ZNF219 amino acids comprises proline. It has been found that 78% of proteins containing proline-rich regions are transcription factors on database analysis. Proline-rich regions are seen in several other transcription factors, including proto-oncogenes such as Jun, AP2 and Oct-2, CTF-1, a ubiquitous transcription factor, and Krox-20, which is expressed in a time- and site-specific manner in the developing murine nervous system. These proline-rich regions may play a critical role in orienting DNA-binding domains and in activating RNA polymerase.

To examine expression of the ZNF219 gene, we used a Northern blot system for multiple human tissues (Clonetech). Northern blot analysis with 32P-labeled ZNF219 cDNA probe revealed that both the 3.5-kb and 5.5-kb transcripts steadily increase in expression with increasing embryo stage, although the level of the 3.5-kb transcript was significantly higher than the 5.5-kb transcript in all embryo stages examined. These data suggest that the size of transcript is developmentally regulated.

We analyzed the chromosome organization of ZNF219 by isolating nine overlapping lambda clones spanning 25 kbp from a lambda DASH II genome phage library (Stratagene). To determine the exon-intron structure of ZNF219, we subcloned and sequenced the corresponding regions in lambda clones. The results are summarized in Fig. 1B. The open reading frame of 2169 nucleotides was composed of 8 exons which spanned within 9 kb.

In an attempt to sequence the genomic region 5' upstream from ZNF219, 1.4 kb EST cDNA registered in a database (Gene Bank Accession No. AA284798) was found to be located upstream of our ZNF219 gene (Fig. 1B). Furthermore, we isolated a 1.3-kb cDNA by screening a human testis cDNA library using this EST clone as a probe. In addition, we attempted to confirm the sequence between EST cDNA and ZNF219 cDNA by RT-PCR of human brain total RNA (Clonetech) using one primer derived from the EST cDNA and another primer derived from the 5' region of ZNF219 cDNA. The isolated 0.3-kb cDNA covered the gap between the EST and ZNF219 cDNAs. We sequenced these cDNAs and determined their exon-intron structures (Fig. 1B). As a result, at least three types of transcripts were obtained on alternative splicing, and the alternative splicing was shown to occur in the 5' non-coding region. Alternative splicing in the 5' non-coding region of Kruppel-related zinc finger genes has been reported on several occasions.
The genetic meaning of the long 5' non-coding sequence is not known clearly at present. However, unusually long 5' non-coding sequences (i.e., >100–150 bp) are frequently detrimental to translation, because long 5' non-coding sequences probably form stable RNA secondary structures and such structures impede ribosome scanning. It is likely that the long 5' non-coding sequence of the 5.5-kb transcript expressed in adult tissues could reduce the efficiency of translation of the ZNF219 protein, and that the ZNF219 protein could be translated efficiently only in the fetal stage. These data suggest that ZNF219 protein may be related to developmental regulation.

Next, we expressed the full-sized ZNF219 protein fused to GFP and monitored the cellular localization of fusion products. GFP is an autofluorescent protein that can be used to study subcellular localization.

The coding region of ZNF219 was amplified by Pyrobost DNA Polymerase (Takara) using the primers 5'-TTGAATTCAGCCACCATGGAGGGCTCACGT-3’ and 5'-GCGGATCCAACCGTTCTTGCCC-3’. The amplified fragment was cloned as an EcoRI-BamHI fragment into pEGFP-N1 vector (CLONTECH) and GFP was fused to the C terminus of ZNF219. The amplified DNA were verified by DNA sequencing later. The constructed plasmid was introduced into HeLa cells by using Effectene transfection reagent (QIAGEN). Transient expression of the ZNF219-GFP fusion protein was examined by fluorescence confocal microscopy.

Many of the known zinc finger genes have also been shown to be evolutionary conserved. The nucleotide sequence was subjected to a BLAST search, 13 ESTs of mouse were found to exhibit 78–98% homology to ZNF219 cDNA. It was suggested that ZNF219 was conserved among mammalian genomes.

We found that another zinc finger protein, KIAA0390, registered in the Gene bank (Gene Bank Accession No. AB002388) is similar to ZNF219. The ZNF219 cDNA exhibited 51% identity at the nucleotide level and 26% similarity at the amino acid level with KIAA0390. KIAA0390 encodes ten sets of Kruppel-related zinc finger structures. Every zinc finger domain of ZNF219 exhibited high homology with those of KIAA0390 in the same arrangement (Fig. 5).

Figure 4. Detection of GFP-tagged ZNF219 protein in transfected HeLa cells by using fluorescence confocal microscopy. Light of microscopy is on (A) and off (B).

Figure 5. Comparison of the zinc finger regions between ZNF219 and KIAA0390 at the amino acid level.
of KIAA0390 was absent in ZNF219. These data suggest that these proteins may have the same evolutionary origin. The DNA binding sequence and genetic role of ZNF219 may be similar to that of KIAA0390. However, the biological function of the KIAA0390 gene remains unknown. In order to determine the precise role of ZNF219, it is necessary to identify the DNA binding sequence and the gene regulated by ZNF219.

References
