ELSEVIER

Biochimica et Biophysica Acta 1306 (1996) 127–132

BBA Biochimica et Biophysica Acta

Short sequence-paper

# A novel zinc finger gene preferentially expressed in the retina and the organ of Corti localizes to human chromosome 12q24.3 [1]

Marcelo N. Rivolta *, Clelia Negrini, Edward R. Wilcox

*Laboratory of Molecular Genetics, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Rockville, MD 20850, USA*

Received 1 June 1995; accepted 16 January 1996

**Abstract**

A cDNA encoding a novel member of the zinc finger gene family, designated *zfOC1*, has been cloned from the organ of Corti. This is the first transcriptional regulator cloned from this sensory epithelium. This transcript encodes a peculiar protein composed of 9 zinc finger domains and a few additional amino acids. The deduced polypeptide shares 66% amino acid similarity with MOK-2, another protein of only zinc finger motifs and preferentially expressed in transformed cell lines. Northern blot hybridization analysis reveals that *zfOC1* transcripts are predominantly expressed in the retina and the organ of Corti and at lower levels in the stria vascularis, auditory nerve, tongue, cerebellum, small intestine and kidney. The human gene was mapped, using a human × hamster somatic cell hybrid panel and fluorescent in situ hybridization, to chromosome 12q24.3. Because of its relative abundance in sensorineural structures (retina and organ of Corti), this regulatory gene should be considered a candidate for hereditary disorders involving hearing and visual impairments that link to 12q24.3.

*Keywords:* Zinc finger gene; Organ of Corti; Retina; Gene regulation

DNA binding proteins play a pivotal role in cellular homeostasis. One of the largest and most important groups of DNA binding proteins are the transcription factors [1,2]. These proteins are critical in controlling the events that lead a cell to its differentiated state and in preserving the complex and dynamic balance between expressed and silent genes once the differentiated phenotype has been reached. The positive and negative control mechanisms of gene expression are therefore critical for determining and maintaining the particular phenotype of a cell [3].

Transcription factors bind to specific DNA sequences and also interact with the other components of the transcription initiation complex. Several transcription factor families have been described, and homologous proteins are included in a family based on the structure of the domain involved in binding to DNA [2]. Examples of such domains are the homeodomain, the paired domain, the zinc finger domain, etc. The zinc finger domain is well characterized and present in several transcription factors [4]. It was originally discovered in TFIIIA, a 5S gene-specific transcription factor from *Xenopus*. The zinc finger domains can be classified into different classes, the $Cys_2$-$His_2$ class being the most common [5].

The auditory epithelium, the organ of Corti, is an example of differentiation and biological adaptation. Due to its complex function, it possesses a highly organized morphological structure, with sensory hair cells changing in height, number, width and distribution of stereocilia depending on their location in the cochlea [6]. Little is known about the molecular process that leads to the development of the organ of Corti and to the maintenance of such an elaborate phenotype. Therefore, understanding the mechanisms of gene regulation in this system becomes an important goal in neurosensory research. As a primary approach, we have pursued the isolation of transcription factor encoding genes expressed in the organ of Corti which may be involved in cellular differentiation and regulation.

Abbreviations: DAPI, 4', 6-diamidino-2-phenylindole.

* Corresponding author. Present address: Department of Physiology, School of Medicine, University of Bristol, University Walk, BS8 1TD, Bristol, UK. Fax: +44 117 9288923; e-mail: m.n.rivolta@bristol.ac.uk.

[1] The sequences reported in this paper have been deposited in the GenBank data base (accession no. L26335 and L41669).


*PII* S0167-4781(96)00022-X

The present report describes the molecular cloning from the organ of Corti of a cDNA coding for a new zinc finger protein which consists only of zinc finger domains. Since this gene is the first transcriptional regulator cloned from the auditory organ, it has been designated *zfOC1* (*z*inc *f*inger *O*rgan of *C*orti *1*).

Genes involved in regulatory activities have been shown to be responsible for hereditary diseases in general and for inherited disorders involving sensorial disabilities in particular. The zinc finger gene GLI3, located at the human chromosome 7, has been characterized as responsible for the Grieg cephalopolysyndactyly syndrome (GCPS) [7]. Mutations in the sequences of two transcription factors (PAX3 and MITF) have been described in patients affected with Waandenburg syndrome type I and II [8–10]. Because of the possible implications of this gene on hereditary disorders, a human genomic fragment encoding for *ZFOC1* was also cloned and mapped to a human chromosome.

Zinc finger genes expressed in the organ of Corti were sought by means of PCR. Total RNA was prepared from microdissected guinea pig organ of Corti, following the guanidinium thiocyanate-acidic phenol purification [11]. DNAse digestion of any genomic contamination was performed before the reverse transcription by adding 70 U of DNAse (RNAse free) to the RNA pellet. cDNA was synthesized using oligo-dT and SuperScript reverse transcriptase (BRL). PCR primers were identical to those described and characterized by Pellegrino and Berg [12]. The upper primers were generated from either the Cys-$X_2$-Cys or the Cys-$X_4$-Cys regions, while the lower primers came from the His-$X_3$-His region and the His-Cys link. cDNA obtained from 7 organs of Corti was used as a template per reaction mixture. Reactions were performed as follows: 1 min at 94°C, 2 min at 52°C and 1 min at 72°C for 35 cycles. Products were separated on agarose gels, the bands excised and electroeluted. Amplification products of the expected sizes were subcloned into pBluescript SK II and *Escherichia coli* transformed with the constructs. Sequencing and data bank analysis of the recombinant clones revealed one 150 bp cDNA fragment encoding a fragment of a novel zinc finger protein. The deduced peptide had two consecutive zinc finger domains of the $C_2H_2$ subtype. This fragment was used as a probe to screen an organ of Corti cDNA library.

Approximately $1.5 \times 10^6$ colonies of a guinea pig organ of Corti cDNA library [13], constructed in pSPORT, were screened by hybridization. Bacteria were plated on LB-Amp plates, and a replica of each plate was made with nitrocellulose filters. Hybridization was carried out at 65°C in a solution containing $6 \times$ SSPE, $10 \times$ Denhardt's solution, 1% SDS, 100 $\mu$g/ml of denatured salmon sperm DNA and 50 ng of a 158 bp PCR generated zinc finger probe labeled by random priming with [$^{32}$P]dCTP at $> 5 \times 10^8$ cpm/$\mu$g. Filters were washed at 65°C in $0.4 \times$ SSPE, 0.5% SDS. After three cycles of screening, positives were picked from isolated single colonies. Plasmid DNA was prepared using DNA binding columns (Qiagen). After a first screening, fifteen cDNA clones were obtained. These clones were characterized by restriction mapping and sequence analysis. Plasmid DNAs were sequenced by using the dideoxy chain-termination method [14]. Reactions were performed initially with fluorescent dye-labeled M13-21 and reverse primers, and then with sequence-specific primers using fluorescent dye-labeled dideoxy terminators as per manufacturer recommendations (Applied Biosystems). Samples were run on a 373A automated DNA sequencer (Applied Biosystems). Sequences were analyzed using the DNASTAR software and data base searches were run against the EMBL-GenBank and the Swiss-PIR-Translated Genbank Release 86. An open reading frame was found, encoding 9 consecutive zinc finger domains. In attempting to obtain a full-length transcript, a guinea pig retina cDNA library was also screened. A total of 25 clones from the organ of Corti and 15 from the retina cDNA library were isolated. All of them were characterized by restriction digestion and partial sequencing.

The nucleotide sequence and deduced polypeptide of

Fig. 1. Nucleotide and predicted amino acid sequences of *zfOC1* (accession no. L26335). Asterisk shows the stop codon. The 9 $Cys_2$-$His_2$ zinc fingers are underlined.

the *zfOC1* cDNA is shown in Fig. 1 (sequence accession #L26335). The obtained sequence was 3703 nucleotides long and contains a canonical AATAAA polyadenylation signal located at position 3660 in the 3′ untranslated region, 17 nucleotides upstream the beginning of the poly $(A)^+$ tail. Despite the length of the transcript, a unique and relative short open reading frame of 783 bp was found. It extends from an ATG codon, located at nucleotide 687, to a TAA termination codon at position 1470. The initiation methionine is flanked by sequences conforming to the second most frequent bases described by Kozak [15]. The 5′ untranslated region, upstream the initiation ATG, possesses stop codons in all three possible reading frames. The *zfOC1* cDNA encodes a basic polypeptide (calculated pI, 8.69) of 261 amino acids, with a predicted molecular weight of 30284 Da. A remarkable feature is that 95.7% of the protein is organized into 9 tandem zinc finger motifs, sparing only 4 residues at the amino- and 7 at the carboxy-terminus. The 9 fingers fit the canonical motif of Cys-$X_2$-Cys-$X_3$-Phe-$X_5$-Leu-$X_2$-His-$X_3$-His. When *zfOC1* was compared to the sequence databases, one of the highest scores of similarity (66%) was obtained with MOK-2. This is a protein which is composed only of zinc finger domains and is preferentially expressed in transformed cell lines [16].

To determine the pattern of expression of *zfOC1*, mRNAs from a comprehensive panel of adult tissues were analyzed by Northern blot (Fig. 2). Total RNA was extracted from several different tissues by the guanidinium thiocyanate method and mRNA purified by using magnetic beads coupled to oligo-dT (Dynabeads, Dynal A.S., Oslo, Norway). 2 $\mu$g of Poly $(A)^+$ RNA from each tissue was electrophoresed in an agarose denaturing gel and blotted onto a Nytran (Schleicher and Schuell) membrane. The RNA blot was tested and normalized by probing it with a $\beta$-actin probe. The Northern blot was then stripped and probed with 120 ng of a 700 bp fragment from the 5′ end of *zfOC1*, labeled by random priming with [$^{32}$P]dCTP at $4 \times 10^9$ cpm/$\mu$g. Hybridization was carried out at 68°C with the same buffer conditions as those used for screening. Densitometric analysis was performed by using the NIH Image 1.54 software.

The *zfOC1* probe hybridized to a transcript of ~4.3 Kb in organ of Corti, stria vascularis, auditory nerve, retina and with less intensity, in cerebellum, tongue, small intestine and kidney. From all the tissues analyzed, the levels of transcription are clearly highest in retina and organ of Corti (Fig. 2).

In order to generate a human *ZFOC1* specific probe, a 550 bp cDNA fragment from the coding region was cloned from human retina cDNA by PCR. Two primers (5′ AAG TCC ACA CAG CCG AGA AG 3′ and 5′ GCC TTC CCA CAC CCA CAA CA 3′), located at position 754-773 and 1287-1306 of the guinea pig sequence, were selected. The PCR reactions used 100 pmol of primers in a 100 $\mu$l volume. Reactions were: 95°C 1 min; 50°C 2 min and

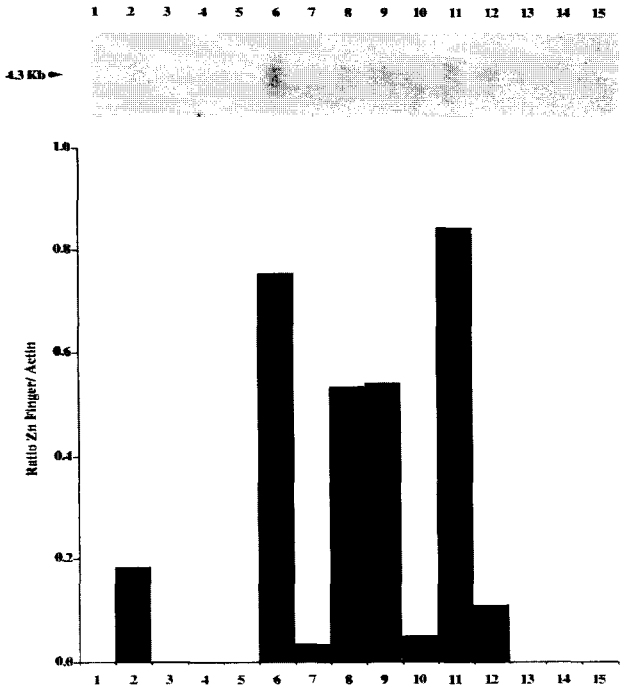


Fig. 2. Pattern of expression of *zfOC1* transcripts. A Northern blot was probed as described in materials and methods. The probe hybridizes to a ~4.3 Kb transcript expressed in organ of Corti, stria vascularis, auditory nerve, and retina. A weaker signal is detected in cerebellum, small intestine, kidney and tongue. Lanes are: 1. lung, 2. cerebellum, 3. cerebrum, 4. olfactory epithelium, 5. liver, 6. organ of Corti, 7. small intestine, 8. stria vascularis, 9. auditory nerve, 10. kidney, 11. retina, 12. tongue, 13. skeletal muscle, 14. heart, and 15. pancreas. The bar graphic below shows the relative abundance of the *zfOC1* mRNA, after being normalized using a $\beta$-actin probe. Highest levels are detected in retina (lane 11) and organ of Corti (lane 6).

72°C 2 min, for 35 cycles. The PCR product was gel purified and subcloned into pCR II (Invitrogen), using the TA cloning system. The constructs were introduced by electroporation into *E. coli* XL1-Blue (Stratagene) and recombinant clones were sequenced. The human cDNA fragment (sequence accession #L41669) contains sequences spanning the nucleotide positions 754 to 1306 of the guinea pig cDNA, showing a 92% sequence similarity, with most of the differences located at the wobble position of the codons. The deduced amino acid sequence of the human *ZFOC1* evidences a high degree of evolutionary conservation, with a similarity score of >98% when compared with the guinea pig homologue using the Lipmann-Pearson algorithm.

Initially, the chromosomal location of the *ZFOC1* gene was determined by Southern blot analysis of *Taq I*-digested DNA from human–hamster somatic hybrid cell lines (BIOS Laboratories). A human–hamster somatic cell hybrid DNA panel (Bios Laboratories) was probed with the 550 bp human *ZFOC1* cDNA fragment, radiolabeled with $^{32}$P-dCTP to a specific activity of $> 5 \times 10^8$ cpm/$\mu$g. Highly stringent conditions for hybridization were as follows: 6 × SSPE, 10 × Denhardt's solution, 1%

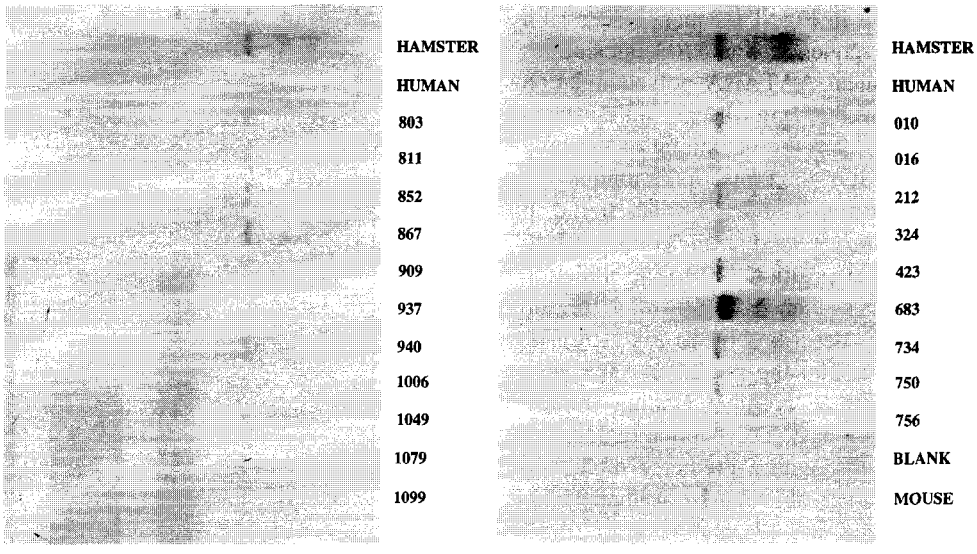


Fig. 3. Segregation of the *ZFOC1* sequences in a human × hamster somatic cell hybrid panel. The blot was hybridized with a *ZFOC1* cDNA fragment (see materials and methods). A control lane with mouse genomic DNA is included, since the cell line 016 is indeed a human × mouse hybrid. An hybridized fragment of the size corresponding to the human control lanes is only visualized in the cell line 683. The only human chromosome shared by the positives lanes, and absent from the others, is Chromosome 12.

SDS, and 100 $\mu$g/ml of denatured salmon sperm DNA, at 68°C. Washes were done at 68°C in 0.4 × SSPE, 0.5% SDS, for 90 min. Genomic DNA from hybrid cell line 683 hybridized to a fragment identical in size to those in control lanes containing total human genomic DNA (Fig. 3), indicating the presence of human *ZFOC1* sequences in this cell line. A weak signal was also seen for hybrid cell line 756, but only after an extended exposure (data not shown). The only chromosome present in the positive hybrid cell lines, and absent from the remaining of the panel is Chromosome 12. Since Chromosome 12 is present in only 5–30% of the cells in hybrid line 756, the weak

Fig. 4. Mapping of *ZFOC1* by fluorescent in situ hybridization. Highly specific signals are seen on both chromatids of both homologues of 12q24.3 (arrows). Texas Red signal is the chromosome 12 centromere specific probe (D12Z1).

hybridization signal was apparently due to the loss of this human chromosome in the majority of the cells.

A more precise mapping of the *ZFOC1* gene was performed by fluorescent in situ hybridization (FISH). In order to pursue that, a human PAC genomic library (Genome Systems) was screened with the 550 bp cDNA fragment, using the same conditions for hybridization described above. A single, positive clone was isolated and its DNA labeled with digoxigenin dUTP. Labeled probe was combined with sheared human DNA and hybridized to normal metaphase chromosomes derived from PHA stimulated peripheral blood lymphocytes in a solution containing 50% formamide, 10% dextran sulfate and 2 × SSC. Specific hybridization signals were detected by incubating the hybridized slides in fluoresceinated antidigoxigenin antibodies, followed by counterstaining with propidium iodide for one color experiments. Probe detection for two color experiments was accomplished by incubating the slides in fluoresceinated antidigoxigenin antibodies and Texas Red avidin followed by counterstaining with DAPI. The initial experiment resulted in specific labeling of the long arm of a group C chromosome. A second experiment was conducted in with a biotin-labeled probe which is specific for the centromere of chromosome 12 (D12Z1) was cohybridized with the *ZFOC1* clone DNA. This experiment resulted in the specific labeling of the centromere in red and the long arm in green of chromosome 12 (Fig. 4). Measurements of 10 specifically hybridized chromosomes 12 demonstrated that *ZFOC1* is located at a position which is 89% of the distance from the centromere to the telomere of chromosome arm 12q, an area that corresponds to band 12q24.3. A total of 80 metaphase cells were analyzed with 77 exhibiting specific labeling.

The identification and characterization of genes coding for proteins with potential regulatory properties is one of the primary approaches used to study the mechanisms involved in the control of gene expression in a particular system. In this manuscript the isolation of *zfOC1*, a cDNA which encodes a polypeptide with 9 zinc fingers of the $Cys_2$-$His_2$ subtype, is reported.

The zfOC1 protein is constituted of only zinc finger motifs. It is similar to the mouse protein MOK-2, which consists of 7 zinc finger domains and 5 additional amino acids. An interesting similarity between these two polypeptides is the degree of internal homology of the fingers. For example, the motif SQSS located between the conserved phenylalanine and leucine residues, is shared by five repeats in the MOK-2 protein and by two in zfOC1. This region of the finger constitutes what is called the 'shoulder', and some of their residues have been shown to be the base recognition positions in many zinc fingers [17]. Furthermore, by comparing some of those base recognition positions named by Jacobs [17] as s3 and m3, they are shown to be identical between zfOC1 and MOK-2 in the three carboxyl located fingers. These findings raise the possibility that both protein products may bind to similar core sequences and, based on the overall structural homology, compose a subfamily of zinc finger proteins.

The evidence presented in this paper shows that *zfOC1* is preferentially expressed in two major sensorineural structures, the retina and the organ of Corti. The role that *zfOC1* could play in these tissues remains unknown. Revealing the identity of the target genes under *zfOC1* regulation becomes an interesting next step towards elucidating its function.

Zinc finger sequences of the $Cys_2$-$His_2$ type have been proposed to be unevenly distributed through the human genome, many of them arranged in clusters [18,19]. The zinc finger genes so far characterized localize predominantly either to telomeric or to fragile regions within a few chromosomes. One of them, Chromosome 19, carries the largest fraction of mapped zinc finger loci. The possibility of a cluster of zinc finger sequences in Chromosome 12 has been described by Rousseau-Merck et al. [20]. They have shown that ZNF10 (KOX1) maps to 12q24.33 and is adjacent to ZNF26 (KOX20), within a PFGE fragment less than 300 Kb long. In this work, *ZFOC1* is assigned to Chromosome 12q24.3; reinforcing the idea that a cluster of zinc finger sequences might be present in this chromosome. Other members of the $Cys_2$-$His_2$ subfamily of zinc finger genes that locates to Chromosome 12 are the glioblastoma gene (GLI1) [21] and the transcription factor SP1 [22].

Experiments currently in progress are indicating that the zfOC1 protein indeed binds to DNA sequences. Because of its possible role as gene regulator and its enhanced expression in sensorial structures, *zfOC1* is an attractive candidate for syndromic disorders involving hearing and/or sight impairments that link to Chromosome 12q24.3.

## References

[1] Johnson, P.F. and McKnight, L. (1989) Annu. Rev. Biochem. 58, 799–839.
[2] Pabo, C.O. and Sauer, R.T. (1992) Annu. Rev. Biochem. 61, 1053–1095.
[3] Blau, H.M. (1992) Annu. Rev. Biochem. 61, 1213–1230.
[4] Berg, J.M. (1990) J. Biol. Chem. 265, 6513–6516.
[5] El-Baradi, T. and Pieler, T. (1991) Mech. Dev. 35, 155–169.
[6] Tilney, L.G. and Saunders, J.C. (1983) J. Cell Biol. 96, 807–821.
[7] Vortkamp, A., Gessler, M. and Grzeschik, K. (1991) Nature 352, 539–540.
[8] Baldwin, C.T., Hoth, C.F., Amos, J.A., da-Silva, E.O. and Milunsky, A. (1992) Nature 355, 637–638.
[9] Tassabehji, M., Read, A.P., Newton, V.E., Harris, R., Bailling, R., Gruss, P. and Stratchan, T. (1992) Nature 355, 635–636.
[10] Tassabehji, M., Newton, V.E. and Read, A.P. (1994) Nature Genetics 8, 251–255.
[11] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[12] Pellegrino, G.R. and Berg, J.M. (1991) Proc. Natl. Acad. Sci. USA 88, 671– 675.
[13] Wilcox, E.R. and Fex, J. (1992) Hearing Res. 62, 124–126.
[14] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5465.

[15] Kozak, M. (1987) Nucleic Acids Res. 15, 8125–8148.
[16] Ernoult-Lange, M., Kress, M. and Hamer, D. (1990) Mol. Cell. Biol. 10, 418–421.
[17] Jacobs, G.H. (1992) EMBO J. 11, 4507–4517.
[18] Litcher, P., Bray, P., Ried, T., Dawid, I. B., Ward, D.C. (1992) Genomics 13, 999–1007.
[19] Hoovers, J.M.N., Mannens, M., John, R., Bliek, J., Van Heyningen, V., Porteous, D.J., Leschot, N.J., Westerveld, A. and Little, P.F.R. (1992) Genomics 12, 254–263.
[20] Rosseau-Merck, M.F., Hillion, J., Jonveaux, P., Couillin, P., Seité, P.,Thiesen, H.J. and Berger, R. (1993) Hum Genet. 92, 583–587.
[21] Kucherlapati, R., Craig, I. and Marynen, P. (1994) Cytogenet. Cell Genet. 67, 246–276.
[22] Szpirer, J., Szpirer, C., Riviere, M., Levan, G., Marynen, P., Cassiman, J.J., Wiese, R. and DeLuca, H.F. (1991) Genomics 11, 168–173.