Genomic organization and transcription of the human retinol dehydrogenase 10 (RDH10) gene

P. Picozzi[a], A. Marozzi[a], D. Fornasari[b], R. Benfante[b], D. Barisani[c], R. Meneveri[c], E. Ginelli[a,*]

[a] *Department of Biology and Genetics for Medical Sciences, University of Milan, Via Viotti 3/5, 20133 Milan, Italy*
[b] *CNR Cellular and Molecular Pharmacology Center, Department of Medical Pharmacology, University of Milan, Via Vanvitelli 32, 20129 Milan, Italy*
[c] *Department of Experimental, Environmental Medicine and Medical Biotechnology, University of Milan-Bicocca, Via Cadore 48, 20052 Monza, Italy*

Received 24 July 2003; revised 18 September 2003; accepted 20 September 2003

First published online 3 October 2003

Edited by Horst Feldmann

Abstract **A cDNA clone up-regulated in hydraulic lung edema in rabbit showed high similarity with human RDH10 mRNA, which encodes a protein involved in retinoic acid metabolism. We defined the organization of the human gene, which includes a unique transcriptional start site, a coding region with six translated exons and a 3′ untranslated region containing at least two used polyadenylation sites. The two poly(A) signals are responsible for the production of the 3 and 4 kb RDH10 mRNA isoforms detected in several human tissues and cell lines.**

Key words: Short chain dehydrogenase reductase; 3′ Untranslated region; Polyadenylation signal

1. Introduction

Retinoic acid (RA), an active form of vitamin A (retinol), regulates vertebrate development, vision, epithelial maintenance, cellular differentiation and proliferation [1,2], through its binding to retinoic acid receptors, members of the superfamily of nuclear ligand-activated transcriptional regulators [3]. In non-ocular tissues, RA arises from two subsequent oxidations in which retinol is first oxidized to retinal and then to RA [4]. Several enzymes involved in the metabolism of retinal or retinol have been identified and classified into four distinct families: alcohol dehydrogenase, short chain dehydrogenase/reductase (SDR), aldehyde dehydrogenase and cytochrome P450. The reversible oxidation of retinol to retinal, suggested to be the rate-limiting step in the conversion of retinol to RA [5], is catalyzed by several members of the SDR family. The SDR superfamily consists of multiple prokaryotic and eukaryotic enzymes, which share an oxidoreductase activity, and preferentially use NAD(P)H as cofactor [6]. These enzymes are characterized by a low sequence similarity and widely variable functions [7]. The determination of the cDNA and amino acid sequences, as well as the crystallographic three-dimensional structures of members of this family [8], made it possible to identify two typical motifs required for the enzymatic activity [9], a 'Rossman fold' structure with a conserved Gly-XXX-Gly-X-Gly pattern (residues 13–19 in the 13β/17β-HSD sequence) [10], and a conserved Tyr-XXX-Lys segment (residues 151–155 in the 13β/17β-HSD sequence) [11].

The cDNA encoding a novel SDR, retinol dehydrogenase 10 (RDH10), has recently been cloned in man, cow and mouse [12]. The identified human RDH10 transcript (accession number AF456765) includes 260 bp of the 5′ untranslated region (5′UTR), a coding region of 1023 bp and 162 bp of the 3′UTR, with a total length of 1445 bp. However, in this paper [12], neither an analysis of the expression in human tissues nor the exact size of the human RDH10 mRNA was reported.

In the present study, we demonstrate that the human RDH10 gene transcribes at least two mRNAs of 3 and 4 kb, and shows a non-ubiquitous tissue distribution. Neither alternative splicing nor different promoter usage accounts for the observed variability of the mRNA size. An analysis of the 3′UTR revealed the presence of two canonical and several non-canonical polyadenylation signals. Among these, a non-canonical and a canonical signal were preferentially used, a choice which seems to be correlated with the mRNA stability.

2. Materials and methods

2.1. Lung tissue sample preparation

Rabbit lung tissue samples were obtained as previously described [13].

2.2. Cell cultures

A549 lung adenocarcinoma, H460 non-small-cell lung cancer, SKMES squamous lung cancer, and SCLC small-cell lung carcinoma cell lines were grown in RPMI 1640 medium supplemented with 10% fetal bovine serum, 1% L-glutamine, and 1% penicillin/streptomycin (Celbio).

2.3. Total and poly(A)$^+$ RNA isolation

Total RNA was extracted using the Rna Witz™ reagent (Ambion) and poly(A)$^+$ RNA isolated by means of the Poly(A)$^+$ Pure mRNA Isolation Kit (Ambion), according to the manufacturer's instructions.

2.4. Differential display analysis and cloning

Differential display (DD) was performed as previously described [14] utilizing the RNA populations obtained from control and treated rabbit lungs. One clone (ED6, 592 bp), identified as up-regulated in lung edema, was cloned into pGEM T-Easy Vector (Promega).

2.5. Northern blot analysis

Two micrograms of poly(A)$^+$ RNAs from rabbit lung, A549, H460, SKMES and SCLC lung tumor cell lines were size-fractionated on a 1.2% formaldehyde agarose gel and capillary transferred to Nytran SuPerCharge membrane (Schleicher and Schuell). The blot obtained

*Corresponding author. Fax: (39)-02-50315864.
E-mail address: enrico.ginelli@unimi.it (E. Ginelli).

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)01089-5

as well as that containing poly(A)$^+$ from adult human tissues (MTN blot, Clontech) were hybridized with α-^{32}P-labeled probes (ED6 clone and Exon 1–2) in ULTRAhyb solution (Amersham) according to the manufacturer's instructions. The membranes were stripped using boiling 0.5% sodium dodecyl sulfate for 30 min and re-probed with a β-actin probe, corresponding to a partial coding region of the β-actin gene.

2.6. RT-PCR analysis

Reverse transcription (RT) was performed with 1 μg of total RNA as template using an oligo(dT) primer and Superscript II reverse transcriptase (Life Technologies). Polymerase chain reactions (PCRs) were performed using BioTaq DNA polymerase (Bioline). The sequences of the PCR primers are provided in Table 1.

2.7. Primer extension analysis

Primer extension analysis was performed using a 30-base oligonucleotide primer (indicated as Extension 1 in Table 1) end-labeled with γ-[^{32}P]ATP and 20 μg of total RNA from cell line A549 or 20 μg tRNA from yeast (negative control), as previously described [15]. The sequencing reaction product of pGEM T-Easy vector was run in parallel.

2.8. 3′ End analysis of RDH10

One microgram of total RNA obtained from the A549 cell line was used to synthesize the 3′ end of the first-strand cDNA according to the instruction manual supplied with the First Choice™ RLM rapid amplification of cDNA ends (RACE) kit (Ambion). Two microliters of cDNA were subjected to PCR amplification by the Expand 20Kb^{plus} PCR system (Roche), using a forward gene-specific primer (4Fw, Table 1) and a 3′RACE outer primer as reverse (5′-GCGAGCACAGAATTAATACGACT-3′). PCR products were loaded and separated on 0.8% agarose gel, and then transferred onto Hybond N$^+$ nylon membrane according to the manufacturer's protocol (Amersham). The blot was hybridized with two different probes obtained by RT-PCR on total RNA from A549 cells using primer pairs 3′RNA R1-3′RNA R2 (probe 1) and 6Fw-TERM1 (probe 2), respectively (Table 1). Hybridization was carried out at 60°C overnight in 5×Denhardt's, 0.5% SDS and 5×SSPE. The membrane was washed twice in 2×SSC, 0.1% SDS at room temperature and twice in 1×SSC, 0.1% SDS at 60°C before being exposed to phosphoimager.

The 3′RACE PCR products were cloned into pGEM T-Easy Vector (Promega) and subsequently sequenced.

2.9. Fluorescence in situ hybridization (FISH) of human chromosome spreads

Metaphase spreads were obtained by standard methods from peripheral blood lymphocytes from normal human donors. The BAC RP11-434I12 (RZPD) probe was nick-translated with digoxigenin-11-dUTP according to the protocol of the vendor (Boehringer Mannheim) and hybridized to the chromosomal preparations for 2 days as previously described [16]. Digital images were captured by using a Leitz microscope equipped with a charge-coupled device camera and analyzed by means of Chromowin software (Casti Imaging).

Table 1
Sequence primers

Primer name	Location	Primer sequence (5′ to 3′)
P1	5′UTR	CGGCCCGGAGCGCTCTGACTTG
Extension 1	Exon 1	AAAGTGACCACGAAGAACTCCACCACGATG
Oligo ATG	Exon 1	TCGCGATGAACATCGTGGTGGA
1Fw	Exon 1	TTTCAAAGTGCTCTGGGCGTTCGT
2Fw	Exon 2	GTGGGGAAGAGGGAGAACG
2Rev	Exon 2	GCCCAGAGACCACACCAGCATTAT
3Fw	Exon 3	ATCATGGTCATATTGTGACAGTTGCA
3Rev	Exon 3	AATATGACCATGATTAATCTCCAGC
4Fw	Exon 4	GAATCCCTGAGCCATGAACTAAAG
4Rev	Exon 4	TCTGAACATGCCAGTGTCTACAAGGT
5Fw	Exon 5	CTCTGAAGCCTGATTACTGTGTGA
5Rev	Exon 5	GAGTGCAGATCATGGGCTGGTCAG
6Fw	Exon 6	GCAGTTGTGTGCATGTATCGGTTC
6Rev	Exon 6	GCTTCATTATTGTTTGTGGCTTGC
TERM1	3′UTR	CCAAGTACTCTAGTCGGTTA
3′RNA R1	3′UTR	TCTGGTGTACTCCAGGCTCGCTGG
3′RNA R2	3′UTR	ATAGCCGATGTTACTGTCCCTAGA
BA1		GCGGGAAATCGTGCGTGACATT
BA2		CTAGAAGCATTTGCGGTGGA

3. Results

3.1. Identification of RDH10 by DD

DD analysis of hydraulic lung edema in rabbit [13,14] detected at least three mRNAs that were strongly up-regulated. One of these PCR products (594 bp) was cloned (clone ED6) and further investigated by Northern blot hybridization and sequencing. In accordance with the result obtained with DD, the expression of the ED6 clone was highly increased in rabbit pulmonary edema (Fig. 1A), where one transcript of approximately 3.0 kb was identified. Searches of the GenBank database (NCBI) using the ED6 sequence revealed 94% and 91% similarity with, respectively, human (AF456765) and bovine (AF456766) RDH10 mRNAs (Fig. 1B). The ED6 sequence, corresponding to nt 820–1410 of the human RDH10 mRNA, contains an open reading frame (ORF) of 462 nt. The similarity of the amino acid sequence of this region (154 aa) with that of the human RDH10 protein was 100%.

3.2. Predicted organization of the RDH10 gene

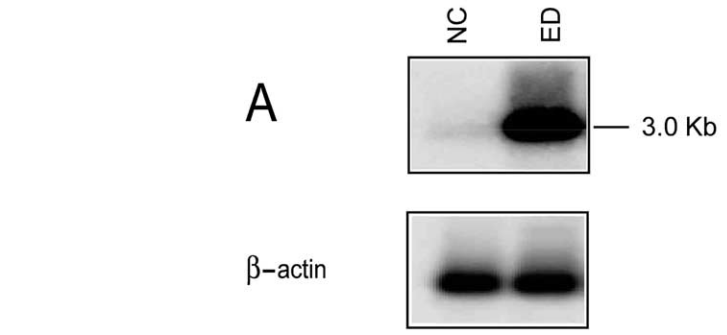
Similarity search of DNA database with the human RDH10 mRNA made it possible to identify the BAC clone RP11-434I12 (accession number AC111149). The in silico analysis of the genomic sequence revealed that the putative RDH10 gene covers approximately 30 kb and contains six exons and five introns (Fig. 2A), as predicted by Ensembl automatic analysis (www.ensembl.org). Exon 1 contains one putative start codon ATG, and exon 6 the translation termination codon TAA. Two canonical (AATAAA) polyadenylation signals were found, at 139 bp and 2233 bp, respectively, downstream of the stop codon (http://genes.mit.edu). Two transcription start sites were also predicted by the Promoter prediction software (www.fruitfly.org/seq_tools/promoter.html), located at 688 and 216 nt upstream of the ATG start codon. The ORF encodes a putative protein of 341 amino acids.

The BAC RP11-434I12 was mapped by FISH only to chromosome 8q21.11 (Fig. 2B).

3.3. Expression of the RDH10 gene in human tissues

Fig. 3A shows the Northern blot hybridization of a panel of human tissues using the ED6 probe. Two transcripts, of approximately 3 and 4 kb, were detected in kidney, liver, small intestine, placenta, lung, heart, and skeletal muscle. Placenta, kidney and liver exhibited the strongest expression. Similar to what was previously observed in rabbit, a weak but detectable signal was present in normal lung. In all the tissues expressing the RDH10 mRNAs, the 3 kb isoform was the most abundant one.

RDH10 mRNA expression was also evaluated in various human lung cancer cell lines. A detectable expression of both mRNAs was observed in cell line A549 (lung adenocarcinoma), but not in H460 (non-small-cell lung cancer) (Fig. 3B). A very low level of expression was also detected in cell lines SKMES (squamous lung cancer) and SCLC (small-cell lung carcinoma) (not shown). Expression patterns superimposable on those shown in Fig. 3A,B were also obtained by hybridizing the two filters with a probe corresponding to the first 525



B

```
                                                                                                  889
Human   ttaatcatgg tcatattgtg acagttgcaa gttccttggg attgttcagt actgccggag ttgaggatta
Bovine  .......... .......... .......... .......... .........c .....t.... .......c..
Rabbit  .......... .......... .......... .......... ...a...... .......... ..........
                                                                                                  959
Human   ctgtgccagt aaatttggag ttgtgggttt tcatgaatcc ctgagccatg aactaaaggc tgctgaaaag
Bovine  ...c...... .......... .c........ .......... .......... .gt....... c..a..g...
Rabbit  t......... .......... .......... .......... ........c. .......... ..........
                                                                                                 1029
Human   gatggaatta aaacaacctt ggtttgccct tatcttgtag acactggcat gttcagaggc tgccgaatca
Bovine  .......... .......g.. ...c...... ..c..g..g. .......... ......g... ..........
Rabbit  .......... ....c....a .......... ..c....... .......... .......... a.........
                                                                                                 1099
Human   ggaaagaaat tgagcctttt ctgccacctc tgaagcctga ttactgtgtg aagcaggcca tgaaggccat
Bovine  .......... c........c .....c.... .......c.. c.....c... .......... ..........
Rabbit  .......... a......... .....c.... .......c.. c......... .......... ..........
                                                                                                 1169
Human   cctcactgac cagcccatga tctgcactcc ccgcctcatg tacatcgtga ccttcatgaa gagcatccta
Bovine  ......c... .......... .......g.. .......... .......... .......... .........c
Rabbit  t.....a... .......... .......c.. .......... .......... .......... .........t
                                                                                                 1239
Human   ccatttgaag cagttgtgtg catgtatcgg ttcctaggag cggacaagtg tatgtacccc tttattgctc
Bovine  ..t....... .......... .......... .......... .a........ .......... .....c....
Rabbit  ..t....... .......... .......... .......... .......... .........t ..........
                                                                                                 1309
Human   aaagaaagca agccacaaac aataatgaag caaaaaatgg aatctaagaa tc-tttttgt atggaatatt
Bovine  ....g..a.. .......... ..g..c.... .g.....c.. .....g..gg ..t....... .........c
Rabbit  .g.....a.. .......... .......... .......... .......... ..-....... ........ac
                                                                                                 1379
Human   acttctatca gaagatgatc aagatgtttc agtccagtgc acatcagcat tg-ctgacat tttatggatt
Bovine  .t.....c.. .....c.... .ga..a.c.t .....gacc. ........-. ..a....t.c .c.c......
Rabbit  .t......t. .....c.... ......c... .......... .......... .-a....... .c.c......
                                                                1423
Human   c-ta-a-ac- t-t-g-tgtt -gtttctttt tt-aaatcaa cttt
Bovine  .-..c.tc.a .g.ttt.t.. g....t.... ..c......g ....
Rabbit  .a..-.tc.- .g-t-g.-.- ga...t.... ..-.-..... ....
```

Fig. 1. Characterization of clone ED6. A: Northern blot analysis of ED6 expression in rabbit lung during hydraulic edema (ED) and in physiologic condition (NC). Sizes in kilobases are indicated on the right. The same blot was probed with the β-actin cDNA fragment as internal control for RNA loading. B: Partial comparison of the human RDH10 mRNA (nt 820–1423; AF456765) to bovine (AF456766) and rabbit (ED6) RDH10 transcripts. The dots indicate agreement between the aligned sequences, the dashes indicate gaps introduced to maximize alignment.

nt of the human RDH10 cDNA coding region (AF456765) (data not shown). The cDNA probe was derived by RT-PCR on total RNA from the A549 cell line using the PCR primers 1Fw and 2Rev (Table 1).

3.4. Characterization of the 3 and 4 kb RDH10 mRNAs

To determine the transcription start site of the 3 and 4 kb mRNAs, we first performed a RT-PCR reaction on total RNA from cell line A549 with a pair of primers located, respectively, at the distal predicted transcription start site (P1, Table 1) and within exon 2 (2Rev, Table 1). A single band of 1138 bp was detected (Fig. 4A, lane 1), whereas the amplification of genomic DNA with the same pair of primers produced a larger band (Fig. 4A, lane 2). The sequence of the RT-PCR product was totally in agreement with the genomic sequence of the RDH10 gene shown in Fig. 4C. This result indicated the utilization of the distal transcription start site. To also verify the functionality of the proximal promoter, a primer extension assay was performed on the A549 total RNA with an oligonucleotide complementary to a region within the first predicted exon of the RDH10 mRNA (Extension 1, Fig. 4C). This experiment showed the occurrence of a single band of 723 bp in length (Fig. 4B), deriving from the distal transcription start site, thus suggesting that the proximal promoter was not used.

Fig. 4C depicts the genomic organization of the 5′ end of the RDH10 gene, including a part of exon 1, and 795 bp of the flanking untranscribed region. The latter region included a TATA box at −24 nt and a GC box at −95 nt. Using the MatInspector program (www.genomatix.de), regulatory sequences with a core similarity of 100% and a matrix similarity of >0.90 were predicted. Multiple putative binding sites, in-

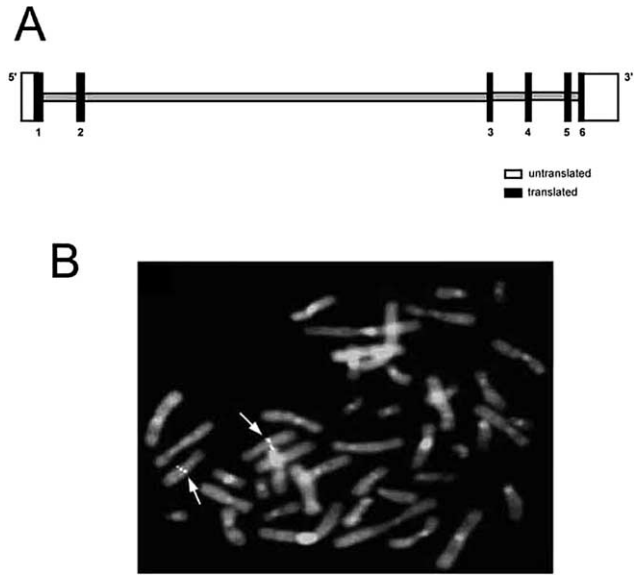


Fig. 2. Organization and chromosomal localization of the RDH10 gene. A: Schematic representation of the RDH10 gene: shaded and open boxes represent exons and UTR, respectively; introns are represented by the connecting lines. B: Cytogenetic localization of the human RDH10 gene by FISH analysis, using BAC RP11-434I12 as probe on normal human metaphase chromosomes.

cluding ADR1, Sp1, HSF and cap, were found. Other less frequent regulatory sequences, including v-Myb, SRY, Ttk-69, AML-1a, AP-1 and 2, GATA1 and 2, were also detected.

To verify whether a mechanism of alternative splicing could be responsible for the generation of the 3 and 4 kb mRNAs, four pairs of primers located within the six predicted exons were used in a RT-PCR assay on total RNA from A549 cells (Fig. 5A). For each primer pair, a single band was detected; in particular, the length and the sequence of the PCR products (346 bp, B1; 305 bp, B2; 401 bp, B3; 543 bp, B4) perfectly matched the predicted exon/intron organization of the RDH10 gene, and ruled out the occurrence of mRNA variants within the coding region. In agreement with these data, the RT-PCR reaction carried out with a primer pair defining the almost complete coding region produced a single band of 1014 bp (Fig. 5B, lane B5).

Finally, to ascertain whether the two mRNAs differed in their 3′UTR, we performed a 3′RACE reaction using a primer (4Fw, Table 1) located within exon 4 (Fig. 6A). The RACE products were analyzed by Southern blot and hybridized with two probes, corresponding to different regions of the 3′UTR (probes 1 and 2 in Fig. 6A). The probes were derived by RT-PCR on total RNA from A549 cells, using primer pairs 3′RNA R1-3′RNA R2 (probe 1) and 6Fw-TERM1 (probe 2), respectively (Table 1). The hybridization experiments identified one and three bands with, respectively, the distal and proximal probe (Fig. 6B). The longer band (2600 bp), detected by probe 1 (Fig. 6B, lane 1), was subcloned and sequenced; the sequence ended at the distal predicted polyadenylation signal (polyA 1 of Fig. 6A). The same characterization was performed for the bands with a length of 1900 and 490 bp detected by probe 2 (Fig. 6B, lane 2). The 490 bp band corresponded to a transcript ending at the proximal predicted polyadenylation signal (polyA 3 in Fig. 6A), whereas the 1900 bp band identified a non-canonical poly-A signal (ATTAAA) (polyA 2 in Fig. 6A). Fig. 6D reports the genomic organization of the 3′UTR of the RDH10 gene, including part of exon 6; the experimentally identified polyadenylation signals are bold and underlined.

To derive the almost complete sequence organization of the

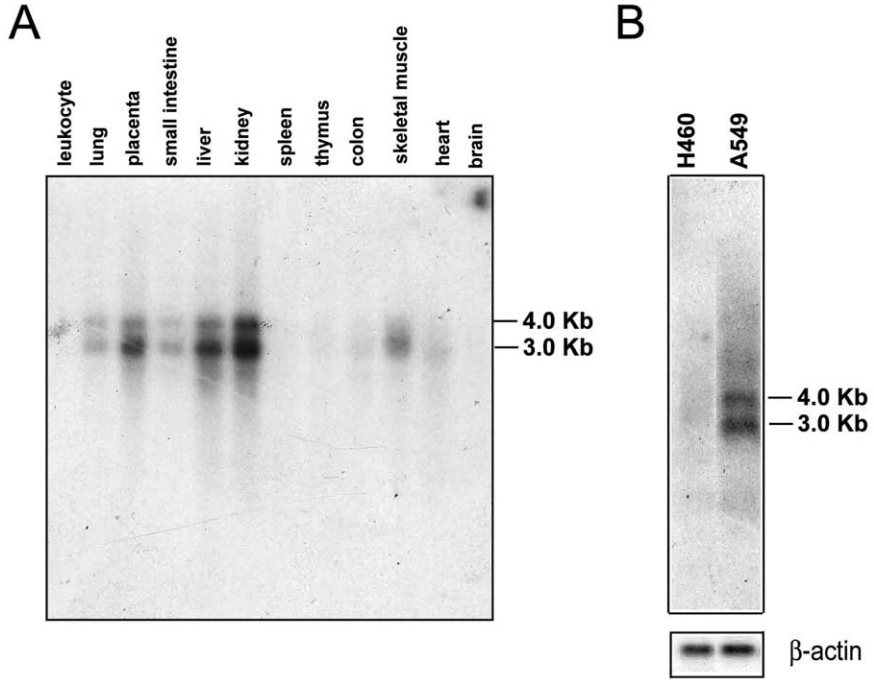
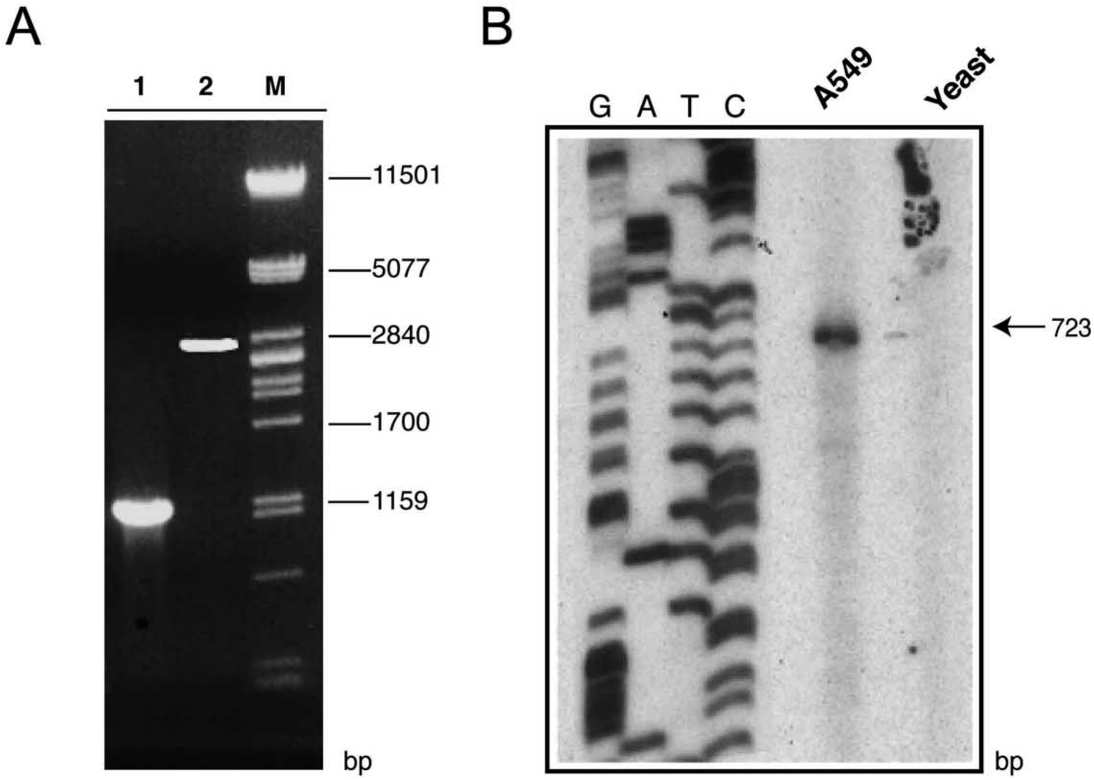


Fig. 3. Analysis of RDH10 gene expression in human tissues. A: Northern blot analysis of poly(A)$^+$ RNA from human tissues hybridized with probe ED6. Each lane contains 2 μg of poly(A)$^+$ RNA. B: Northern blot analysis of poly(A)$^+$ RNA from H460 (non-small-cell lung cancer) and A549 (lung carcinoma) cell lines, hybridized with probe ED6; to evaluate poly(A)$^+$ RNA loading on each lane, the hybridization with human β-actin cDNA is shown below. Each lane contains 2 μg of poly(A)$^+$RNA. Sizes in kilobases are indicated on the right.



C

```
                                                                        HSF
                                                                       gaaaa  - 795
        SRY                                                             Bcd
tatatgcatttgattagacttttaatcaattcaagagcaaagagtgctgacaccttggcaactccgggtcct aatcc  - 790
                            HSF  ADR1                        AD1
ccaggctgcgctgacaggattaggctccgttcctccccataatgttcccaggacgagcctcatggggacgaactacaaa  - 711
          cap                     v-Myb      Ttk-69 AML-1a
tcccagcatgcaccagtcttcgcccgcccggcgggagggcaacggctgaccaggaccgcaggcaagcaccgcggcgacg  - 632
 HSF      cap                     HSF                   AP-1    cap
gttccagccaggaaaatgagagcctcttgggccacgttccaaacggctagctgcaattcgaagtcacttttggccgccc  - 553
                               ADR1  HSF
tcccggctccctgccgcggtagtcgcagactacacccctttcgccgctccgcgaggtccttgtcgtcgctccagcccac  - 474
                              ADR1      ADR1                                MZF1
agggctgtcttccgcagcgcctgatcgcaggcggggcggtctggagtgctctagcttacgaaaagtggctagccagacc  - 395
                    GATA-2   ADR1                                           ADR1
agtggggaccggacggccggaggaggatggccggggctctgctgactgccgagctcggcaggcggaggccaggacaccc  - 316
                                      cap                     ADR1
cacgggagggggcgtgagctgcgaagggaggtgcgaggagggctgaggttgcggccacggccgtgtgagaggggctacc  - 236
                                          Sp1          GC    box
cggctcggctccgcggacggctggcgggctgggggctgggccgggggcggggtgcgcccgggggcggggaggccgggc  - 158
   GATA-1                                             TATA BOX
ggagggagggccgatgccgaggctggcgcgggctcttcccggtcgcggggttatatagcgcggagcgtggagcccgctc  -  79

AGAGCCGGCCCGGAGCGCTCTGACTTGCAAGCGGGCTGCGCTGCGGAGCCCAGTGCCCGAGTGACACCCGCGGAGAGTG  +  79
CAGGGCCGGGGAACGCGAGCCCTCGGGGGCAGCTGCAAGGCGTTGGGCAGCGCTTGCCTGCGCCGAGCGAGTCTCCCCT  + 158
TCCCGGCGCTCCGCCGCCCCGCACCCCACTCTCCCACCCTCTCGCAACTTGGGTCGAGTTGACAACTCCCGCGGCAGCC  + 237
CGCTGGCCCGTGCCGCCTCCGCTGCGCACCCCTCCCCCGGGGTGAGAGGGAGCCGGCGCGCCGGTTCCGGGGACGCTCG  + 316
GGCGGCAGCAGCTTGGCCATGAGGGCAGTTCGAGTAGTCTAACTCGCGGCTGTCACCGCCACTGCAGCGGAGCCGGCCG  + 395
GCCGGGCGCTGCGGGACGGGCGGGCGGCTGCCGGCAGGAGGCGCCGAGCCGGGTGACTGCCGCGGCGGGCACAGTCCGG  + 474
GGCCACAAGCGCCGAGCCCGGGCGGGAGTGGCCCCGCGCAGGCAGGGAGCGGCGCCGCGCACTCCAACCCGGCGGGCAC  + 553
CTCGGGGGCGGGCGCGGGGCGCAGCCTTCTCGTCCCGGCCTCTGTGACAAGCGCCCCGGAGCCGGGAGCCCGATTGCCG  + 632
GGCTCGGGGTGGGCGCGGACGCAGGCACTGGGCTCGTGCGGGGCCCCGGGCGTCGCG  atg  aac atc gtg gtg  + 704
gag ttc ttc gtg gtc act ttc aaa gtg ctc tgg gcg ttc gtg ctg gcc gcg gcg cgc tgg  + 764
     Extension 1
```

Fig. 4. Analysis of the RDH10 5′UTR. A: RT-PCR analysis of total RNA extracted from cell line A549 (lane 1) and PCR on genomic DNA as a control (lane 2) using a pair of primers located in the 5′ end (P1, 5–27 nt) and exon 2 (2Rev, 1119–1142 nt). B: Transcription start site of the RDH10 mRNA. Twenty microgram of total RNA from cell line A549 and yeast tRNA, as negative control, and a 30 base oligonucleotide primer Extension 1 were used for the assay. A single band of 723 bp was detected in cell line A549. The sequence reaction product of pGEM T-Easy vector was run in parallel. C: The RDH10 5′-flanking sequence. The 5′UTR is reported in upper case, while the untranscribed region is in lower case. Numbers on the right indicate the nucleotide position relative to the transcription initiation site +1. The transcription starting nucleotide (adenine) and the translation initiation codon ATG are in bold. The potential transcription factor binding sites predicted with Mat-Inspector program are indicated. The oligonucleotide used in the primer extension assay is underlined (Extension 1).

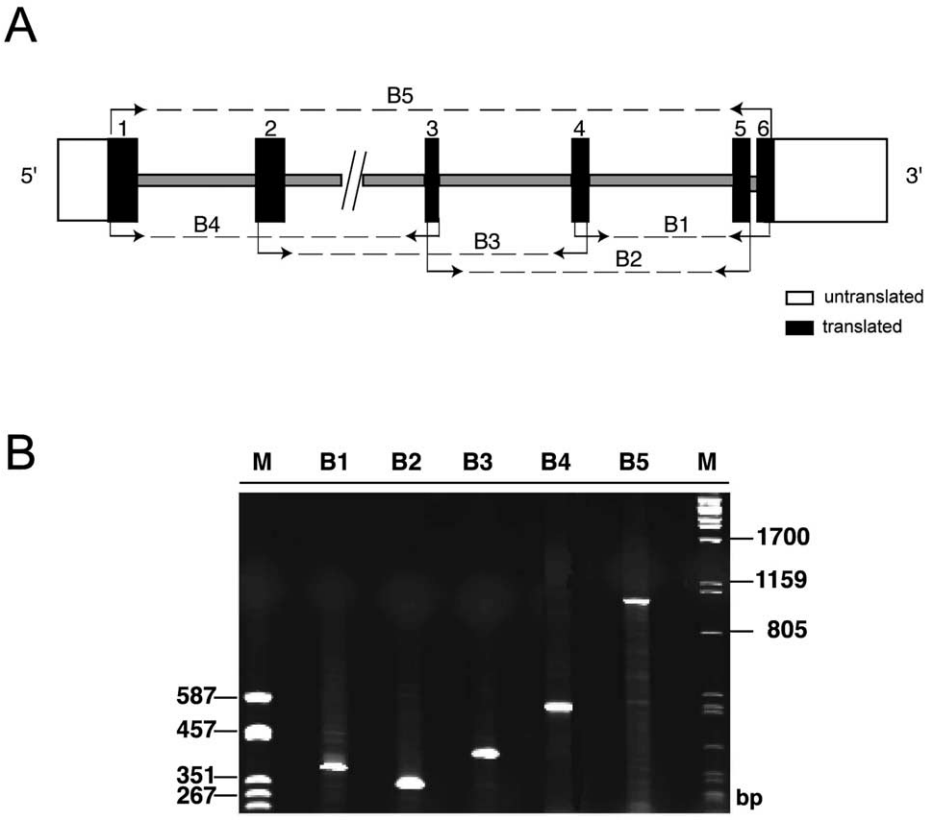


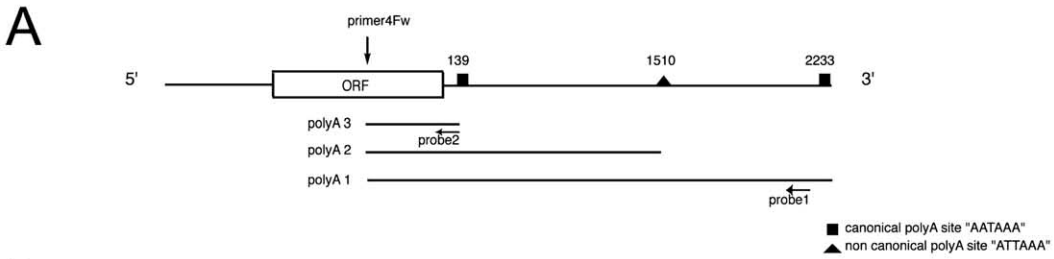
Fig. 5. Analysis of the RDH10 ORF by RT-PCR assay on total RNA from A549 cells. A: Schematic representation of intron/exon structure of the RDH10 gene, with the orientation and position of the exon-specific RT-PCR primers. The RT-PCR products are shown in B. PCR 1Fw-3Rev (lane B4, 543 bp), PCR 2Fw-4Rev (lane B3, 401 bp), PCR 3Fw-5Rev (lane B2, 305 bp), PCR 4Fw-6Rev (lane B1, 346 bp). RT-PCR reaction was carried out also with a primer pair defining the almost complete coding region ATG-6Rev (lane B5, 1014 bp).

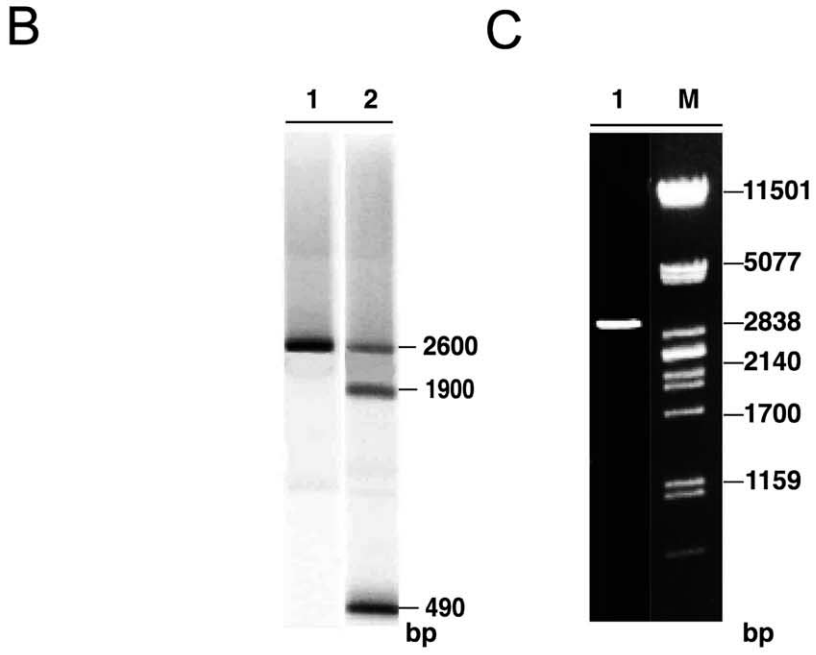
4000 bp mRNA, a pair of primers spanning from the ATG to the distal polyadenylation site was used to amplify, by RT-PCR, total RNA from A549 cells (Fig. 6C). A single band of approximately 3090 bp was detected; the sequence corresponded to the predicted genome organization of the RDH10 gene, including the 1023 bp coding region and 2067 bp of the 3′UTR (not shown).

4. Discussion

This paper demonstrates that the human RDH10 gene transcribes at least two mRNAs, of approximately 3 and 4 kb, and that the two variants are expressed at a level detectable by Northern blot in several human tissues. Human, bovine and mouse transcripts from the RDH10 gene have recently been cloned [12], providing partial cDNA sequences of approximately 1.4 kb, while the transcript size derived by Northern blot for the bovine was approximately 3 kb [12]. The occurrence of a single 3 kb mRNA also in rabbit lung seems to confirm this finding. On the other hand, in man the transcription of the RDH10 gene, in addition to the 3 kb message, yields a longer mRNA (4 kb). Among the evaluated mammalian species, only human tissues show the occurrence of two RDH10 mRNA isoforms. FISH analysis performed with the BAC clone carrying the RDH10 gene excluded that the observed additional transcript could derive from gene duplication. Thus, RDH10 mRNA variants are the products of other mechanisms involved in the synthesis (alternative promoter usage) and/or the maturation of the pre-mRNA.

To derive the organization of the human RDH10 mRNAs we carried out 5′ extension, 3′RACE and a series of RT-PCRs using primers located within the ORF. These experiments indicated that both transcripts are characterized by the same 5′UTR of 688 bp, demonstrating that the transcription of the RDH10 gene utilizes only the distal one of the two predicted promoters. No alternative splicing occurs within the coding region, and in both transcripts the ORF is 1023 bp. Regarding the 3′UTR, the results obtained indicate that the stop of transcription occurs at two different polyadenylation sites, yielding mRNAs with 3′UTRs of 1510 and 2233 bp, respectively. The polyadenylation sites belonged to the canonical (AATAAA) and non-canonical (ATTAAA) classes of signals [17]. Furthermore, an additional highly conserved AATAAA signal, located 139 bp downstream of the stop codon, was identified by the analysis of the RDH10 gene organization. In fact, a mRNA species with a transcription termination at this signal was identified only by 3′RACE, but not by Northern blot. Supposing that the transcription for all the RDH10 messages starts at the distal promoter and the absence of alternative splicing in the coding region, this finding indicates the possible occurrence of a RDH10 mRNA of 1.853 kb. However, in the analyzed tissues, the abundance of the transcript was under the detection threshold of the Northern blot, thus suggesting a low level of usage of this polyadenylation signal or, alternatively, the instability of the mRNA. A similar situation may occur in the bovine, since the RDH10 mRNA sequence shows a polyadenylation signal located at 139 nt downstream of the stop codon, but Northern blot hy-





D

```
AAG CAA GCC ACA AAC AAT AAT GAA GCA AAA AAT GGA ATC TAA gaatctttttgtatggaatattacttctatcagaagatg +1753
gatcaagatgtttcagtccagtgcacatcagcattgctgacattttatggattctaaacttgtgttgtttctttttttaaatcaacttttaaaaa +1850
AATAAAgtgtaaattaaccgactagagtacttggaaaatgtgatcagtacaagtgaacttaggttgttgccaacagggtccttttaggcagaacc +1945
cagaaaccagtcaaatctgtagagaagcagtgtgacatcttcaggttaccattatttttttaatgagcaggaagtctagaaatgataactagactg +2040
tatgtttcatgtgtgtgatttttcagaattcccagagtttactcattcttgttattaaactctagccagttgacatcttcgcaatttcaaggact +2135
gatagtgctgtattttctcacgttttctaagtttccgttttgcaaggcctaggtgacttttttcatggtgtttgtatgtttagctcttttgaaaag +2230
gaattttgaaatctccatcaactgaagtaaatgatgtctgagtgttacagtaaaggtgaccaagtctctttcttaaagtcacaatgactaaagta +2325
ttagttgaatttttttttttttttttttgatggagtctcgctctgtcaccaggctggagtgcagtagcacaatcacggctcactgcaatctctg +2420
cctcccggtttcaagtgattctgctgtctcagcctcccaagtagctgggactacaggcatgcgccaccacgcccagctaatttttgtatttttag +2515
tagagacggggtttcaccatgttggtcaggatggtctccatctcttgacattgtgatccacctgcctcggcctcccaaagtgctgggattacagg +2610
catgagccactgcacccagccttgaatttttaattttatctctgatatacttcattaagtgtctggagacctaattatcctaaaagatcatacat +2705
tttctacctatgaattttgctgcatacagaaagtgccctttcctcaggaagttgctgtgtttcatttctttggatggactcttatctagaataca +2800
tagcagctctgcaaagaaacagtttttaaaaatgggaacttctacattgaaaagtcccctatttttgtgccaactatgattagtgagaggaagaaa +2895
tcttattctatggcatatgtatggaagggtgtaaagattcttttgaaaggtttattcacattgtagaacagcaaatgacatttttacagtatttt +2990
tttgtaaagcaaactatttttgtgccttgaatttggtatatgtgtattagtgaaacattgtaaaggtgaacttctacctctgtatctaaatgtata +3085
ccatccacttgtaaatgactatataactattatgtgattgctttttttttagaatgtcttgtttaaatagtggccaatgtttaaggctgttaaaa +3180
taagccaacttttactaattggggagttttataaatgactgATTAAAttaaagaattaacttacatgcaattgtgtgattattagttatcagca +3275
gtgttgtaaggaaaattattgtgttttttctatgatcattatcccactttaggtaaagaaaaatattggaatggaatagtgttgggaaacagac +3370
attaacaacctagggtgcctgcactcaaatagccgatgttactgtccctagattagagacttgattaagggcttgtttgtaccaaaagtggggaa +3465
acaatgccatgacctgtgttttagtttggctgcaccacagatcaaatctgcactgtgtctacatataggaaaggtcctggtgtgtgctaatgttc +3560
ccaatgcaggacttgaggaagagctctgttatatgtttccatttctctttatcaaagataaccaaaccttatggcccttataacaatggaggcac +3655
tggctgcctcttaattttcaatcatggacctaaagaagtactctgaagggtctcaacaatgccaggtggggacagatatactcagagattatcca +3750
ggtctgcctcccagcgagcctggagtacaccagaccctcctagagaaatctgttataatttaacaacccacttatccaccttaaaactgaggaaa +3845
gtctctttacatctaatttttattcttgtgtgttataacttaaacctatttctatttttgtttgttattgcccttataagggtgtccatctccaag +3940
ttcAATAAA +3949
```

Fig. 6. Analysis of the RDH10 3′UTR. A: Schematic illustration of the RDH10 cDNA structure. The rectangle represents the protein coding region (ORF). The horizontal lines represent the 5′ and 3′UTR, respectively, with the canonical and non-canonical polyadenylation signals. PolyA1, PolyA2, PolyA3 represent the 3′RACE products obtained using the primer 4Fw as gene-specific primer. The relative position of the fragments used as probes in Southern blot analysis are indicated by arrows. B: Analysis of the 3′RACE products by hybridization with radioactively labeled probe 1 (lane 1) and probe 2 (lane 2). C: RT-PCR analysis of the RDH10 mRNAs using primers located at the first translation codon and at the most downstream polyadenylation signal (see Table 1). D: Genomic organization of the 3′UTR of the RDH10 gene, including a part of exon 6. The three experimentally identified polyadenylation signals are in bold and boxed. The 12 iron-responsive elements (IRE) are in bold. The numbers on the right refer to the nt position relative to the transcription start site (+1).

bridization evidenced a single transcript of approximately 3 kb [12]. Thus the human RDH10 gene transcription generates at least two mRNAs which share the same 5′UTR and the same ORF, but differ in the length of their 3′UTR. The use of two different polyadenylation sites yields the contemporary presence of two RDH10 variants in the same human cell. Three main hypotheses could explain this phenomenon: (i) the stability of the transcripts, (ii) a different subcellular localization [18], (iii) a phenotypic protective effect [19]. Regarding the mRNA stabilization, the sequence analysis of the full length RDH10 3′UTR revealed the presence of at least 12 iron-responsive elements (IRE), preferentially concentrated in 1 kb downstream of the first polyadenylation signal. IREs play a critical role in mRNA stability by binding iron-regulatory proteins, which mask potential endonucleolytic cleavage sites [18], thus possibly explaining the greater stability of the 3 and 4 kb mRNAs as compared to the 1.853 kb message.

Thus human cells, compared to the other analyzed mammalian species, show the occurrence of more than one RDH10 transcript, differing in the choice of the polyadenylation site. This feature strongly suggests the possibility of evolutionary changes of the 3′UTR, which is characterized by more degrees of freedom than the 5′UTR and the ORF [18].

Interestingly, preliminary indications of the modulation of RDH10 gene expression were derived from the analysis of injured rabbit lung (edema) and of different human lung cancer cells. Although these observations require further investigations, a possible explanation of this modulation can be derived by considering the function of the RDH10 gene product. The RDH10 protein, belonging to the SDR family, is responsible for the oxidation of all-*trans*-retinol to all-*trans*-retinal, which is a precursor of RA [12]. In man, retinol administration to premature infants reduces the incidence of bronchopulmonary dysplasia, suggesting that it has a protective or reparative role in structurally immature lungs [20]. In rats, retinol and RA have been shown to affect normal lung development [21]. Furthermore, mice bearing a deletion of RA receptor γ have decreased lung elastin, decreased alveolar number and increased alveolar size suggesting a link between retinoids, alveolarization and elastin [22]. Thus, the up-regulation of RDH10 in lung edema may reflect its role in RA metabolism [12], which in turn could play a reparative role in the damaged tissue.

References

[1] Cheryl, M.S., Di, Y.-P., Hayden, L., Zhao, Y.H. and Satre, M.A. (2001) J. Biol. Chem. 276, 24194–24202.
[2] Kurlandsky, S.B., Xiao, J.-H., Duell, E.H., Voorhees, J.J. and Fisher, G.J. (1994) J. Biol. Chem. 269, 32821–32827.
[3] Chambon, P. (1996) FASEB J. 10, 940–954.
[4] Romert, A., Tuvendal, P., Tryggvason, K., Dencker, L. and Eriksson, U. (2000) Exp. Cell Res. 256, 338–345.
[5] Napoli, J.L. (1986) J. Biol. Chem. 261, 13592–13597.
[6] Terada, T., Sugihara, Y., Nakamura, K., Sato, R., Inazu, N. and Maeda, M. (2000) Eur. J. Biochem. 267, 6849–6857.
[7] Su, J., Chai, X., Kahn, B. and Napoli, J.L. (1998) J. Biol. Chem. 273, 17910–17916.
[8] Nakanishi, M., Matsuura, K., Kaibe, H., Tanaka, N., Nonaka, T., Mitsui, Y. and Hara, A. (1997) J. Biol. Chem. 272, 2218–2222.
[9] Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzalez-Duarte, R., Jeffery, J. and Ghosh, D. (1995) Biochemistry 34, 6003–6013.
[10] Opperman, U., Filling, C., Berndt, K.D., Persson, B., Benach, J., Ladenstein, R. and Jornvall, H. (1997) Biochemistry 36, 34–40.
[11] Ensor, M.E. and Tai, H.H. (1994) Biochim. Biophys. Acta 21, 151–156.
[12] Wu, B.X., Chen, Y., Chen, Y., Fan, J., Rohrer, B., Crouch, R.K. and Ma, J.-X. (2002) Invest. Ophthalmol. Vis. Sci. 43, 3365–3372.
[13] Sabbadini, M., Barisani, D., Conforti, E., Marozzi, A., Ginelli, E., Miserocchi, G. and Meneveri, R. (2003) Biochim. Biophys. Acta 1638, 149–156.
[14] Barisani, D., Meneveri, R., Ginelli, E., Cassani, C. and Conte, D. (2000) FEBS Lett. 469, 208–212.
[15] Flora, A., Lucchetti, H., Benfante, R., Goridis, C., Clementi, F. and Fornasari, D. (2001) J. Neurosci. 21, 7037–7045.
[16] Ballarati, L., Piccini, I., Carbone, L., Archidiacono, N., Rollier, A., Marozzi, A., Meneveri, R. and Ginelli, E. (2002) Gene 296, 21–27.
[17] Beaudoing, E., Freier, S., Wyatt, J.R., Claverie, J.-M. and Gautheret, D. (2000) Genome Res. 10, 1001–1010.
[18] Grzybowska, E.A., Wilczynska, A. and Siedlecki, J.A. (2001) Biochem. Biophys. Res. Commun. 288, 291–295.
[19] Qu, X., Qi, Y. and Qi, B. (2002) Arch. Biochem. Biophys. 400, 233–244.
[20] Venness-Meehan, K.A., Pierce, R.A., Moats-Staats, B.M. and Stiles, A.D. (2002) Am. J. Physiol. Lung Cell Mol. Physiol. 283, L971–L980.
[21] Cardoso, W.V., Mitsialis, S.A., Brody, J.S. and Williams, M.C. (1996) Dev. Dyn. 207, 47–59.
[22] McGowan, S., Jackson, S.K., Jenkins-Moore, M., Dai, H.-H., Chambon, P. and Snyder, M.J. (2000) Cell. Mol. Biol. 23, 162–167.