The mouse ADP-ribosylation factor-like 4 gene: two separate promoters direct specific transcription in tissues and testicular germ cell

Stephan JACOBS*, Annette SCHÜRMANN*, Walter BECKER*, Tobias M. BÖCKERS†, Neal G. COPELAND‡, Nancy A. JENKINS‡ and Hans-Georg JOOST*[1]

*Institut für Pharmakologie und Toxikologie, Medizinische Fakultät der Rheinisch-Westfälischen Technischen Hochschule Aachen, Wendlingweg 2, D-52057 Aachen, Germany, †Institut für Anatomie, Westfälische Wilhelms-Universität Münster, Vesaliusweg 2-4, D-48149 Münster, Germany, and Leibniz-Institut für Neurobiologie, Brenneckestrasse 6, D-39118 Magdeburg, Germany, and ‡Mammalian Genetics Laboratory, ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, Frederick, MD 21702-1201, U.S.A.

ADP-ribosylation factor-like protein 4 (ARL4) is a Ras-related GTPase that has been cloned from the 3T3-L1 preadipocyte cell line as an adipocyte-specific cDNA [Schürmann, Breiner, Becker, Huppertz, Kainulainen, Kentrup and Joost (1994) J. Biol. Chem. **269**, 15683–15688]. The *Arl4* gene maps to the proximal region of mouse chromosome 12 linked to *Lamb1-1*, *Hfhbf1* and *Sos2*. Compared with all other known genes of Ras-related GTPases, the genomic organization of *Arl4* is unusual in that its entire coding region, the 3′ untranslated region (UTR) and most of the 5′ UTR are located on a single exon. This structure suggests that *Arl4* has evolved by retroposition of an *Arf* (ADP-ribosylation factor) or *Arf*-like gene. Isolation of the 5′ UTR by rapid amplification of cDNA ends (RACE)-PCR revealed heterogeneous transcription initiation sites in alternative exons 1. Both 5′-flanking regions exhibited promoter activity when expressed in COS-7 cells, indicating that the expression of *Arl4* is directed by two separate promoters. mRNA transcribed under the control of the downstream promoter was isolated by RACE-PCR from all investigated tissues. In contrast, the upstream promoter seems to drive specifically the expression of *Arl4* in adult testis. Hybridization of rat testis *in situ* indicated that *Arl4* is expressed in germ cells of puberal and adult testis, but not in prepuberal testis, suggesting that *Arl4* is involved in sperm production.

INTRODUCTION

ADP-ribosylation factors (ARFs) represent a subfamily of the Ras-related GTPases comprising six homologous genes (ARF1–6 [1,2]) that are presumably involved in vesicle trafficking [3–5] and signal transduction, e.g. phospholipase D activation [6–8]. In addition to the six ARF isoforms, several ARF-like cDNA species (*Arl1–5*, *Arp*) have been cloned and assigned to the extended ARF family [9–14]. *Arl4* was found in a PCR-cloning approach designed to identify genes that are differentially expressed in the preadipocyte cell line 3T3-L1 [12]; it was abundant in differentiated 3T3-L1 cells but was undetectable in undifferentiated cells. Thus the function of the GTPase seemed to be related to adipocyte metabolism and, because *Arl4* mRNA was also found in muscle and brain, to other cellular pathways present in the differentiated cell.

The expression of most ARF isoforms (ARF1–3) seems to be constitutive [1,2,15]. Striking exceptions are the developmentally regulated expression of *Arf2* [16] and the differentiation-dependent expression of *Arl4* [12]. The basis of this unusual regulation of expression is of particular interest, and might help to identify the physiological function of ARL4. Here we describe the unique organization of the *Arl4* gene, which contains two separate promoter regions directing its tissue-specific expression. The downstream promoter seems to be active in most tissues, whereas the upstream promoter specifically drives a large expression of *Arl4* in testicular germ cells, suggesting that the GTPase is involved in spermatid maturation.

MATERIALS AND METHODS

Interspecific mouse backcross mapping

Interspecific backcross progeny were generated by mating (C57BL/6J × *Mus spretus*) F_1 females and C57BL/6J males as described [17]. A total of 205 N_2 mice were used to map the *Arl4* locus. DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, Southern blot transfer and hybridization were performed essentially as described [18]. All blots were prepared with Hybond-N^+ nylon membrane (Amersham). The probe, a mouse genomic-DNA fragment of the 5′ flanking region (nt −244 to −780; see Figure 2), was labelled with [α-^{32}P]dCTP by using a nick translation labelling kit (Boehringer Mannheim); washing was done to a final stringency of 1.0 × SSCP/0.1 % SDS at 65 °C. A fragment of 7.6 kb was detected in *Sac*I-digested C57BL/6J DNA and a fragment of 6.2 kb was detected in *Sac*I-digested *M. spretus* DNA. The presence or absence of the 6.2 kb *Sac*I *M. spretus*-specific fragment was followed in backcross mice.

A description of the probes and restriction fragment length polymorphisms for the loci linked to *Arl4*, including *Lamb1-1*, *Hfhbf1* and *Sos2*, has been reported previously [19]. Recombination distances were calculated using Map Manager, version 2.6.5. Gene order was determined by minimizing the number of recombination events required to explain the allele distribution patterns.

Library screening and DNA sequencing

Genomic clones of *Arl4* were isolated by screening a 129 SvJ mouse genomic library (approx. 3×10^5 plaques) with a PCR-

Abbreviations used: ARF, ADP-ribosylation factor; ARL4, ADP-ribosylation factor-like protein 4; CRE, cAMP response element; EST, expressed sequence tag; RACE, rapid amplification of cDNA ends; UTR, untranslated region.

[1] To whom correspondence should be addressed (e-mail joost@rwth-aachen.de).

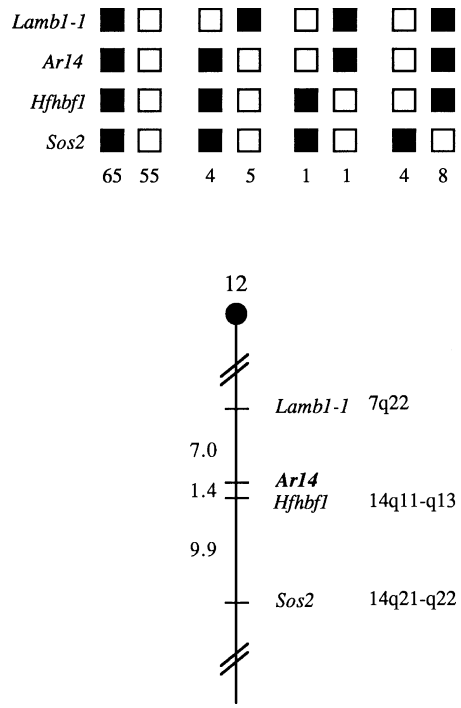

Figure 1 *Arl4* maps in the proximal region of mouse chromosome 12

Arl4 was placed on mouse chromosome 12 by interspecific backcross analysis. The segregation patterns of *Arl4* and flanking genes in 143 backcross animals that were typed for all loci are shown at the top of the figure. For individual pairs of loci, more than 143 animals were typed (see the text). Each column represents the chromosome identified in the backcross progeny that was inherited from the (C57BL/6J × *M. spretus*) F_1 parent. Filled boxes represent the presence of a C57BL/6J allele, open boxes represent the presence of an *M. spretus* allele. The number of offspring inheriting each type of chromosome is listed at the bottom of each column. A partial chromosome 12 linkage map showing the location of *Arl4* in relation to linked genes is shown at the bottom of the figure. Recombination distances (centimorgans) between loci are given at the left, and the positions of loci in human chromosomes, where known, are shown at the right. References for the human map positions of loci cited here can be obtained from the Genome Data Base (GDB), a computerized database of human linkage information maintained by The William H. Welch Medical Library of The Johns Hopkins University (Baltimore, MD, U.S.A.).

generated DNA fragment of mouse *Arl4* [12]. The clone with the longest 5′-flanking sequence was subcloned into pBluescript (Stratagene, La Jolla, CA, U.S.A.). Fragments were prepared by sonication or digestion with restriction enzyme, subcloned into pUC18 and sequenced in both directions by the method of Sanger (Thermo-sequenase fluorescent labelled primer cycle sequencing kit; Amersham Life Science, Little Chalfont, Bucks., U.K.) with the aid of an automated sequencer (LI-COR, Lincoln, NE, U.S.A.).

RNA preparation and rapid amplification of cDNA ends (RACE)-PCR

Mouse or rat tissues were homogenized with a Polytron homogenizer in 4 M guanidine thiocyanate. Murine 3T3-L1 fibroblasts (American Type Culture Collection, Rockville, MD, U.S.A.) were differentiated as described [20], frozen in liquid nitrogen and lysed with 4 M guanidine thiocyanate supplemented with 7% 2-mercaptoethanol. The lysates were layered on a CsCl cushion (5.88 M) and centrifuged at 138000 ***g*** (rotor SW40) for 22 h at 20 °C. Pelleted RNA was dissolved in 300 μl of sodium acetate/Tris buffer (0.1 M, pH 9), and was neutralized by the addition of 50 μl of 2 M potassium acetate, pH 5.5.

5′ RACE was performed with a kit from Gibco BRL (Eggenstein, Germany) in accordance with the instructions of the manufacturer. Primers for cDNA synthesis and amplification were derived from the *Arl4* cDNA: cDNA synthesis, 5′-CCA TTT CCA GCT GCA AAT G-3′; reverse primer for first amplification, 5′-TCT CAA ACT GAA CGC GT; reverse primer for second amplification, 5′-CGC GTT CTT CTT ATC TGC. PCR conditions: 35 cycles at 1.5 mM $MgCl_2$ and 0.2 mM dNTP; hot start, 1 min at 94 °C, 1 min at 55 °C, 1 min at 72 °C, 7 min at 72 °C in the last cycle. DNA fragments were isolated and subcloned into pUC18 with the Sureclone® kit (Pharmacia, Freiburg, Germany).

Transfection of COS-7 cells and luciferase assay

DNA fragments derived from the 5′-flanking regions of *Arl4* were fused to luciferase cDNA in the vector pGL3 (Promega, Madison, WI, U.S.A.) and the constructs were transfected into COS-7 cells with the aid of the cholesterol derivative DAC-30 (Eurogentec, Seraing, Belgium). Cells were lysed 30 h after transfection; luciferase activity was assayed with a kit from Promega in accordance with the instruction of the manufacturer. Transfection efficiency was determined by co-transfection of a plasmid harbouring *lacZ*, and data were normalized to β-galactosidase activity.

Northern blot analysis

Samples (10 μg) of total RNA were separated by electrophoresis on 1% (w/v) agarose gels containing 3% (v/v) formaldehyde and were transferred to nylon membranes (Hybond; Amersham-Buchler, Braunschweig, Germany). Before transfer, gels were stained with ethidium bromide to ascertain that equal amounts of total RNA had been separated. Probes were generated with the Klenow fragment of DNA polymerase I and [α-^{32}P]dCTP by random oligonucleotide priming [21]. The nylon membranes were hybridized at 42 °C, and blots were washed twice with 0.8 × SSC containing 0.1% SDS.

Hybridization *in situ*

Rat tissues were removed and frozen immediately in isopentane on solid CO_2 at −40 °C. Tissues were cut in 18 μm sections on a cryostat, mounted on Superfrost plus slides (Fisher Scientific, Düsseldorf, Germany) and stored at −70 °C. Oligonucleotides corresponding to bp 2644–2677 of the *Arl4* sequence (see Figure 1) (anti-sense oligonucleotide: 5′-CCA AGG AGT CGA TGA GCT CAG TTC ACC CAT TGC-3′) were synthesized by MWG Biotech (Ebersberg, Germany). The oligonucleotides (MWG Biotech) were 3′ end-labelled with [α-^{35}S]dATP (Amersham Buchler). Frozen sections were air-dried at room temperature and were hybridized with the oligonucleotide (5×10^5 c.p.m. per slide) at 42 °C for 16 h in a buffer [formamide/20 × SSC/0.2 M phosphate buffer (pH 7.6)/dextran sulfate/5% sarcosyl (5:2:1:1:2, by vol.)] containing 500 μg of sheared salmon sperm DNA, 250 μg/ml yeast tRNA and 100 mM dithiothreitol. The hybridized sections were washed twice with 2 × SSC for 10 min at room temperature, six times with 1 × SSC/10 mM 2-mercaptoethanol for 15 min at 55 °C and 1 × SSC for 15 min at room temperature. After being washed, the sections were dehydrated, air-dried and exposed to β-max Hyperfilm (Amersham Buchler) for 1–3 weeks. Alternatively, sections were dipped in NTB3 nuclear track emulsion (Kodak) diluted 1:1 with water, stored for 4–8 weeks at 4 °C, developed and stained with Cresyl Violet.

RESULTS

Genomic localization of the mouse *Arl4* gene

The mouse chromosomal location of *Arl4* was determined by interspecific backcross analysis with progeny derived from matings of [(C57BL/6J × *Mus spretus*)F_1 × C57BL/6J] mice. The 6.2 kb *Sac*I *M. spretus* restriction fragment length polymorphism (see the Materials and methods section) was used to follow the segregation of the *Arl4* locus in backcrossed mice. The mapping results indicated that *Arl4* is located in the proximal region of mouse chromosome 12 linked to *Lamb1-1*, *Hfhbf1* and *Sos2* (Figure 1). The ratios of the total numbers of mice exhibiting recombinant chromosomes to the total numbers of mice analysed for each pair of loci and the most likely gene order are: centromere — *Lamb1-1* — 12/172 — *Arl4* — 2/146–*Hfhbf1* — 16/161 — *Sos2*. The recombination frequencies [expressed as genetic distances in centimorgans ± S.E.M.] are *Lamb1-1* — 7.0 ± 1.9 — *Arl4* — 1.4 ± 1.0 — *Hfhbf1* — 9.9 ± 2.4 — *Sos2*. The proximal region of mouse chromosome 12 shares regions of homology with human chromosomes 7q and 14q (summarized in Figure 1), suggesting that the human homologue of *Arl4* also maps to 7q or 14q.

Sequence and organization of the mouse *Arl4* gene

A total of five genomic clones of *Arl4* were isolated and were characterized with restriction enzymes and probes derived from the 5′ or 3′ untranslated region (UTR) of the rat *Arl4* cDNA. The

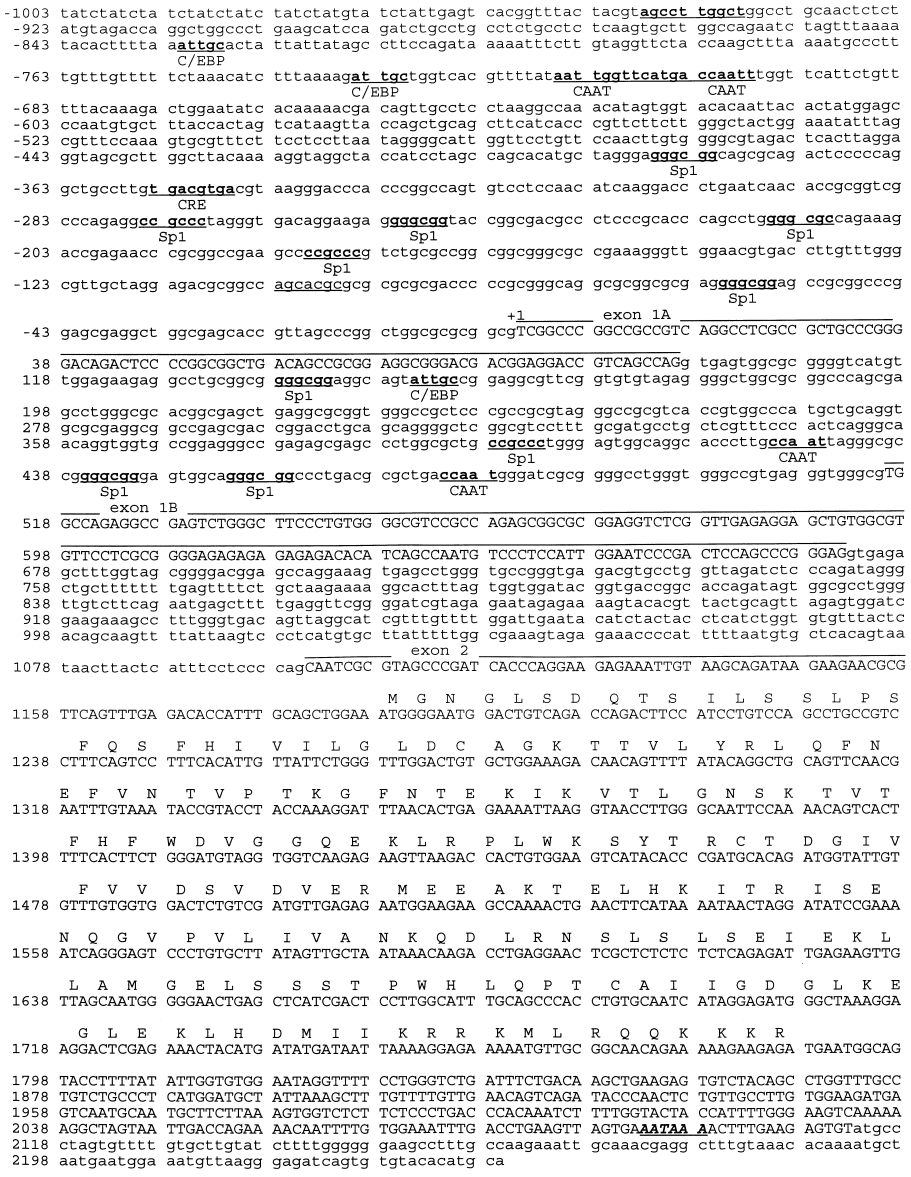

Figure 2 Nucleotide sequence of the mouse *Arl4* gene

Capital letters denote sequences found in cDNA clones; lower-case letters denote intron sequences. The deduced amino acid sequence of the open reading frame is given in single-letter code above the nucleotide sequence. The starts of exon 1A and exon 1B are deduced from the longest cDNA clones isolated with the RACE procedure (see Figure 3). The polyadenylation signal is highlighted by underlined bold italics; CAAT boxes and Sp1 consensus binding sites (GGGCCGG or CCCGCC) are shown in capital bold characters and are underlined. The EMBL, GenBank accession number for *Arl4* is Y12577.

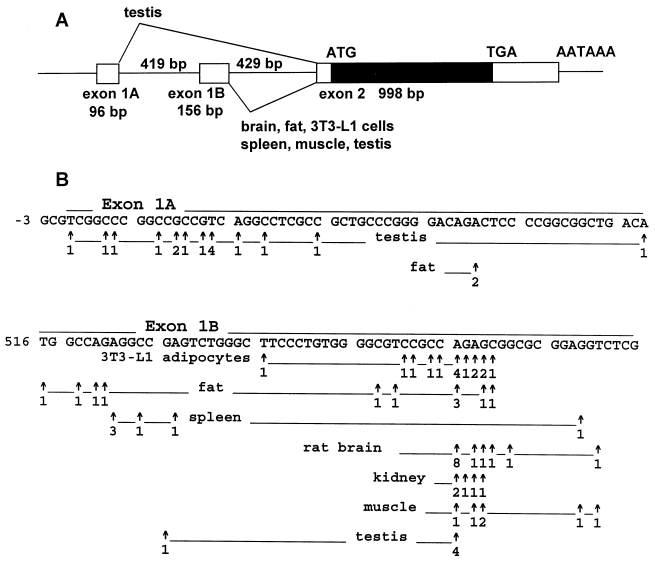

Figure 3 Alternative splicing and transcription initiation sites of *Arl4* mRNA

(**A**) Schematic representation of the exon–intron organization of *Arl4* and of its alternative splicing. Boxes represent sequences found in cDNA; the coding region is drawn as a filled box. Alternative splicing was demonstrated by the isolation of 79 cDNA clones that contained either exon 1A (testis, fat) or 1B (3T3-L1 cells, fat, spleen, brain, kidney, muscle, testis). (**B**) Determination of the transcription initiation sites in different tissues by RACE cloning. The numbers of clones isolated are given below the positions (arrows) of the 5′ end of each RACE clone.

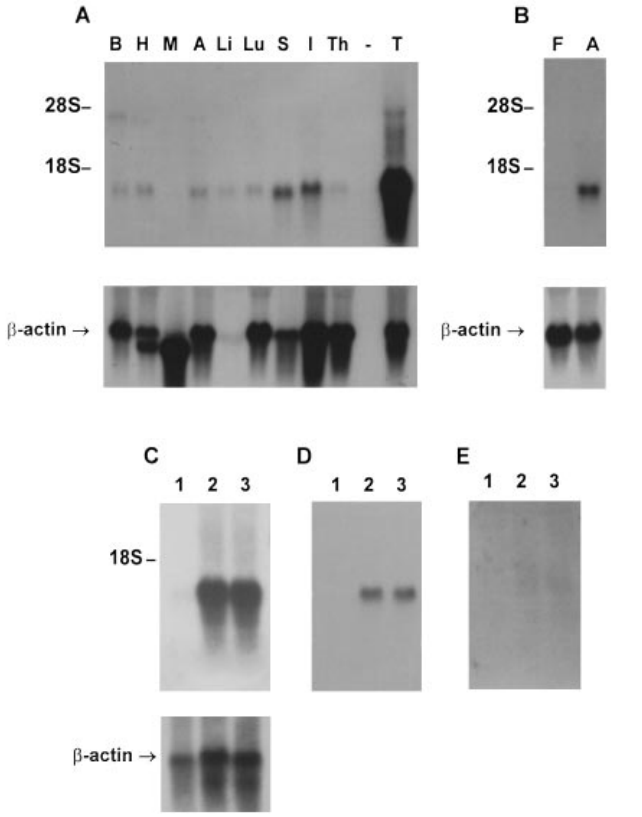

Figure 4 Northern blot analysis of the expression of *Arl4*

Samples (10 μg) of RNA from the indicated tissues or cells were separated on an agarose gel and transferred; the blot was hybridized with rat *Arl4* cDNA. In all blots, equal amounts of total RNA were loaded as assessed by ethidium bromide staining. The blots were re-probed with a human β-actin cDNA as indicated. (**A**) Expression of *Arl4* mRNA in different rat tissues. Abbreviations: B, brain; H, heart; M, skeletal muscle; A, adipose tissue; Li, liver; Lu, lung; S, spleen; I, intestine; T, testis; Th, thymus. (**B**) Differential expression of *Arl4* mRNA in 3T3-L1 fibroblasts (F) and adipocytes (A). (**C**, **D**, **E**) Expression of *Arl4* mRNA in testis from rats at 2 (lane 1), 8 (lane 2) and 16 (lane 3) weeks of age (prepuberal, puberal and adult). Blots were hybridized with a probe derived from the full-length rat *Arl4* cDNA (**A**, **B**, **C**), or with cDNA fragments corresponding to exon 1A (**D**, rat cDNA) or exon 1B (**E**, mouse cDNA).

clone with the longest 5′-flanking sequence was subcloned and sequenced (Figure 2). A comparison of the genomic sequence with the cDNA sequences of *Arl4* (rat and human) indicated that the complete coding region, the 3′ UTR including a polyadenylation signal (nt 2093–2098) and a considerable portion of the 5′ UTR are located on a single exon. This genomic organization is in striking contrast with that of related genes such as *Arf1*, *Ki-Ras* and *Rab1* (see Figure 6). The sequence of the coding region (nt 1188–1787) is 97.5% identical with that of the rat (15 nucleotide exchanges) and 100% identical with a partial mouse sequence (connected from the expressed sequence tags MMA03086, MM123891, MMAA9306). It contains four mismatches (99.3% identity) with a database entry of mouse *Arl4* cDNA (accession no. U76546).

The genomic sequence of *Arl4* predicts that Southern blots of DNA digested with the restriction enzymes *Bam*H1, *Eco*R1 and *Nco*1 should exhibit single bands hybridizing with the *Arl4*-specific probe. Indeed, these enzymes generated single bands (5.5, 10 and 12 kb respectively) under low-stringency hybridization and washing conditions (results not shown). This finding is consistent with the conclusion that no highly similar isoforms or pseudogenes of *Arl4* exist.

Tissue-specific alternative splicing of the 5′ UTR (exon 1) and alternative promoter use in the *Arl4* gene

To determine the transcription initiation site of *Arl4*, RACE-PCR was performed with cDNA from 3T3-L1 cells and various mouse and rat tissues. As anticipated, the sequences of the RACE products corresponded to portions of the genomic sequence (Figure 2). Surprisingly, however, they indicated the presence of two alternatively spliced exons (1A and 1B) encoding the 5′ end of the mRNA (Figure 3). All 79 RACE clones analysed contained either exon 1A or 1B; no clone containing both exons was isolated. RACE clones obtained with testis mRNA predominantly (16 out of 21) started with exon 1A. Of 13 clones isolated from fat tissue, only two started with exon 1A. All

Table 1 Promoter activity of the 5′-flanking regions of exon 1A and exon 1B

DNA fragments corresponding to the 5′-flanking regions of exon 1A (A, bp 54–1104) or exon 1B (B, bp 1104–1644; A + B, bp 54–1644) were fused with luciferase cDNA in the vector pGL3. Constructs were transiently expressed in COS-7 cells; luciferase activity was assayed. Results from four independent transfections were corrected for transfection efficiency and are presented as relative luciferase units (RLU) ± S.E.M.

DNA fragment	Luciferase activity (RLU)
Control plasmid	261 ± 148
Promoter A + B	6937 ± 1550
Promoter A	1954 ± 673
Promoter B	8766 ± 2450

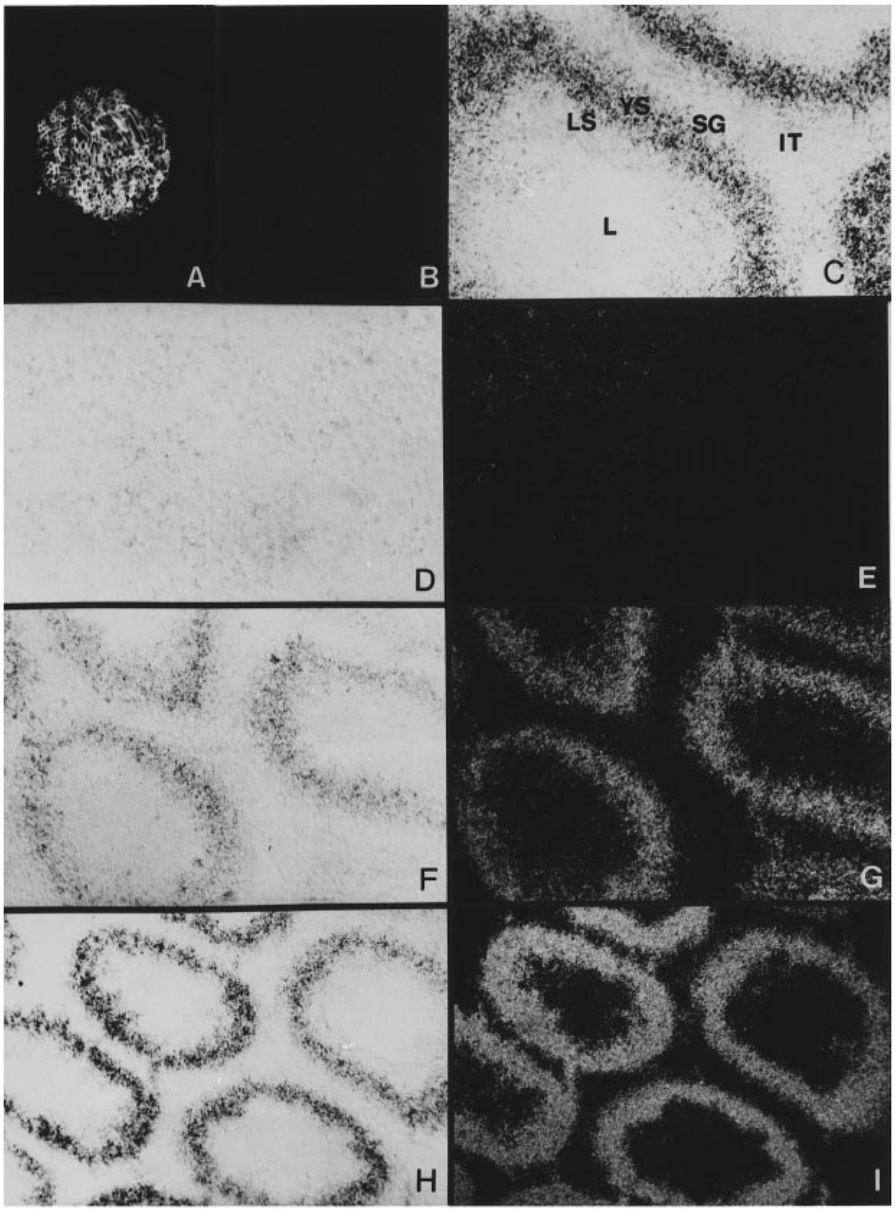

Figure 5 Localization of *Arl4* mRNA in seminiferous tubules of rat testis as detected by hybridization *in situ* with a labelled oligonucleotide

(**A**) Testis from adult rat hybridized with anti-sense oligonucleotide; dark-field photomicrograph at 2-fold magnification. (**B**) Control hybridization with the sense oligonucleotide. (**C**) Localization of *Arl4* mRNA in seminiferous tubules as detected by anti-sense oligonucleotide hybridization at a magnification of ×160. Abbreviations: L, lumen; LS, late spermatids; YS, young spermatids; SG, spermatogonia; IT, interstitial tissue. (**D–I**) Testis from different developmental stages (**D**, **E**, prepuberal; **F**, **G**, puberal; **H**, **I**, adult rat) hybridized with anti-sense oligonucleotide; magnification ×100. (**E**), (**G**) and (**I**) are darkfield photomicrographs.

RACE clones obtained from 3T3-L1 cells, brain, spleen, kidney and muscle started with exon 1B (Figure 3B). The transcription initiation sites in exon 1B were variable and seemed to cluster around two sites.

To add support to the conclusion that transcripts comprising exon 1A are predominant in testis, Northern blots of testis mRNA were hybridized with probes derived from either exon 1A or 1B. As illustrated in Figures 4(D) and 4(E), a 1.2 kb transcript was detected in puberal and adult rat testis with the exon 1A-specific probe, but not with an exon 1B-specific probe. It should be noted that both probes seemed to hybridize poorly with the *Arl4* mRNA because a probe corresponding to the full-length cDNA (Figure 4C) gave a considerably stronger signal. Figure 4(A) also illustrates the tissue distribution of the *Arl4* mRNA as assessed with the full-length cDNA probe. mRNA was detected predominantly in testis, and also in spleen and intestine; smaller amounts of *Arl4* mRNA were found in brain, heart, total fat, liver, lung and thymus. As we have shown previously [12], *Arl4* was expressed in differentiated 3T3-L1 adipocytes but was not detectable in undifferentiated fibroblasts (Figure 4B).

A search of the expressed sequence tag (EST) databases identified two mouse and three human EST entries of *Arl4* cDNA with a 5′ UTR. Both mouse ESTs started with exon 1B. However, of the three human EST entries, one entry (HS84754)

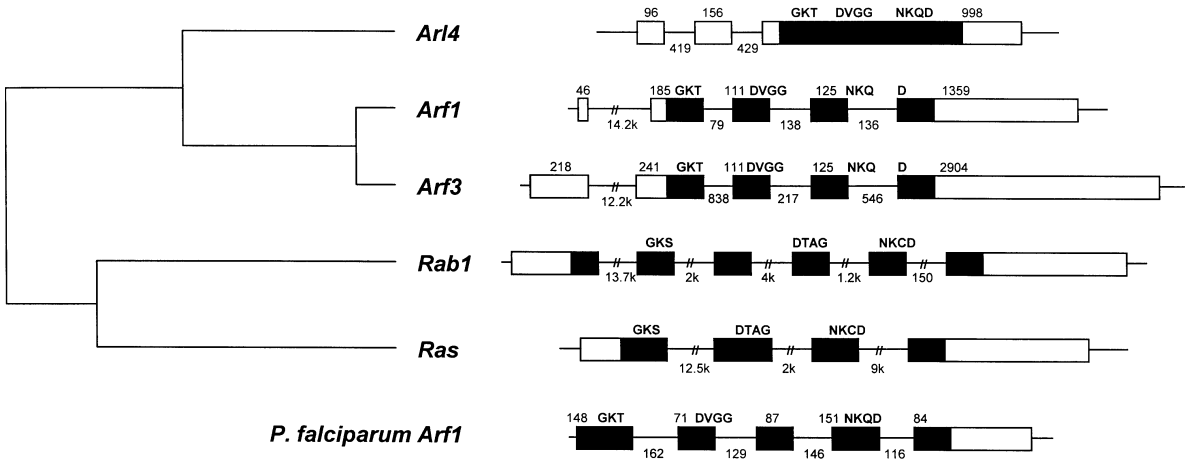

Figure 6 Comparison of the genomic organization of *Arl4* with that of other Ras-related GTPases

The dendrogram of a multiple alignment of amino acid sequences (left panel) was performed with the CLUSTAL program. Genomic organization of the GTPases (right panel) was obtained from references [24] (*Arf3*), [15] (*Arf1*), [37] (*Rab1*), [38] (*Ki-Ras*) and [25] (*Plasmodium falciparum Arf*). Boxes represent exons, filled boxes denote coding regions. The sizes of exons and introns are given above and below the structures respectively. The main GTP-binding motifs PM1 (GKT or GKS), PM3 (DVGG or DTAG) and G2 (NKQD or NKCD) are indicated above the structures at their approximate positions within the coding region.

started with a different 5′ UTR. Thus we tentatively assume that the transcription of the human *ARL4* gene also generates two isoforms with alternative exons 1.

Promoter activity of the 5′-flanking regions of exons 1A and 1B of *Arl4*

The alternative splicing of the 5′ UTR suggested that the expression of *Arl4* is controlled by two separate promoters. Indeed, both 5′-flanking regions of the exons 1A and 1B exhibited promoter activity when expressed in COS-7 cells as luciferase fusion constructs (Table 1). We had expected that COS-7 cells lacked the components required to produce a large activation of the testis-specific promoter A. Accordingly, the upstream promoter (A) seemed less active than the downstream promoter (B). Furthermore Table 1 demonstrates that the activity of promoter B is not dependent on the presence of promoter A, because B exhibited the same activity as a construct comprising both promoters.

Both promoters contain G/C-rich regions (73 % in promoter A, nt −403 to 1; 74 % in promoter B, nt 97 to 515) and elements characteristic of promoter regions (Figure 2). Two CAAT boxes and four consensus sequences for the binding of Sp1 are present in promoter B. Promoter A lacks a CAAT box in the proximity of the transcription start but contains six Sp1 consensus binding sites, two motifs matching presumed testis-specific elements of the proacrosin gene [22], a cAMP response element (CRE)-like motif [23], and a palindromic motif flanked by CAAT boxes (Figure 2). These sequence characteristics are consistent with the conclusion that the 5′-flanking regions of both exons 1 might function independently as promoters.

Localization of *Arl4* mRNA in rat testis

Northern blot analysis of mRNA from rat tissues indicated the highest levels of *Arl4* mRNA to be in testis (Figure 4A). Thus, to test a possible involvement of ARL4 in spermatogenesis, we compared mRNA levels of *Arl4* in testis from rats of different ages. As illustrated in Figure 4(C), a large expression of *Arl4* was detected in testis from adult rats (8 and 16 weeks of age), but not in testis from prepuberal rats (2 weeks of age). Because this pattern seemed to parallel the time course of spermatogenesis, hybridization *in situ* of tissue slices from rat testis was performed with a 33-mer oligonucleotide corresponding to nt 2644–2677 of the *Arl4* sequence shown in Figure 1. As is illustrated in Figure 5, the hybridization *in situ* revealed an intense labelling of all seminiferous tubules per section. The highest density of silver grains seemed to be associated with spermatogonia and/or young spermatids (Figure 5C); no *Arl4* mRNA was detected in interstitial tissue. The lack of a specific signal in testis from prepuberal rats indicates that *Arl4* expression depends on sperm production. Thus the location of the tracer and the time course of expression in testis are consistent with a specific expression of *Arl4* mRNA in germ cells.

DISCUSSION

The organization of the *Arl4* gene is unique among the known Ras-related GTPases in that its coding region, its 3′ UTR and a considerable portion of the 5′ UTR are located on a single exon (Figure 6). This unusual genomic organization suggests that the gene has evolved by retroposition and that its evolution is a relatively recent event. The dendrogram of the multiple sequence alignment of ARL4 with ARF, Ras and Rab1 suggests that the precursor of ARL4 is ARF or an ARF-like isoform. The coding regions of all known genes of Ras-related GTPases, including three mammalian ARF isoforms [15,16,24] and the *Plasmodium falciparum* ARF [25], are interrupted by several introns (Figure 6). Thus it is unlikely that the *Arl4* gene has originated from these genes by duplication of genomic DNA, unless one assumes that the intronless structure was conserved over a very long period. Furthermore a database search failed to identify a yeast homologue of ARL4 (the closest relative in *Saccharomyces cerevisiae* is the ARL1 homologue SCZ36033 with 43.7 % identical amino acids), which is consistent with the conclusion that the evolution of ARL4 is a relatively recent event.

One of the few known retroposons encoding a functional protein is the testis-specific isoform of glucose-6-phosphate

dehydrogenase-2 [26]. It has probably evolved, because the glucose-6-phosphate dehydrogenase-1 gene is X-chromosomal and cannot be transcribed in post-meiotic male germ cells. Thus we speculated that *Arl4* originated in a similar way. However, the Southern blot analysis of mouse DNA gave no indication that a highly similar *Arl4* isoform exists.

The heterogeneous transcription start with alternative exons 1 indicates that two separate promoters induce the expression of the *Arl4* gene. Many other genes are directed by separate promoters that control an independent, tissue-specific (e.g. of glucocokinase [27]) or a developmentally regulated expression (e.g. of Pax6 [28]). The *Arl4* gene is unique in that its two promoters seem to be separated by only approx. 100 bp, whereas in other genes the distance between dual promoters is usually much larger. On the basis of the preliminary analysis of the 5′-flanking regions with luciferase fusion constructs, we assume that the upstream promoter resides in the 5′-flanking region of exon 1A and that the second promoter is located in the 419 bp intron separating exon 1A and 1B. Both presumed promoter regions of the mouse *Arl4* gene are TATA-less but contain elements characteristic of promoter regions. The presumed downstream promoter is G/C-rich, contains two CAAT boxes and three Sp1 consensus binding sites. It seems to be constitutively active in differentiated tissues, e.g. fat and heart, and contains a partial C/EBP consensus motif (Figure 2) that is known to control the expression of fat-specific genes [29]. Thus the specific expression of *Arl4* in the differentiated 3T3-L1 cells might be induced by the transcription factor C/EBP. However, the possibility cannot be excluded that the differential expression of *Arl4* mRNA reflects a repression of the gene in the 3T3-L1 fibroblasts rather than its induction in the adipocytes.

The upstream promoter contains a G/C-rich region with six Sp1 consensus binding sites and one CRE-like motif. The expression of many testis-specific transcripts is regulated by CRE-like motifs, namely angiotensin-converting enzyme [23,30], phosphoenolpyruvate carboxykinase [30], calspermin [31], pro-protein convertase 4 [32], lactate dehydrogenase [33] and NO synthase [34]. The transcription factor responsible for the large expression of these genes in testis seems to be CREM-tau [23], a testis-specific isoform of the CRE-binding protein CRE modulator (CREM). Expression of this transcription factor is repressed in premeiotic germ cells and is induced by the follicle-stimulating hormone in postmeiotic spermatids [35,36]. Thus the lack of expression of *Arl4* in prepuberal testis and its large expression in germ cells of adult rat testis is consistent with the presence of the CRE-like element in the presumed upstream promoter of *Arl4*. In addition to the CRE motif, two other elements considered to be responsible for the testis-specific expression of the proacrosin gene [22] are present in the upstream promoter of *Arl4*. Thus the sequence characteristics of the upstream promoter of *Arl4* are consistent with our finding that it is specifically activated in testicular germ cells.

We have compared our interspecific map of chromosome 12 with a composite mouse linkage map that reports the map location of many uncloned mouse mutations (provided by Mouse Genome Database, a computerized database maintained at The Jackson Laboratory, Bar Harbor, ME, U.S.A.). *Arl4* mapped in a region of the composite map that lacks mouse mutations with a phenotype that might be expected for an alteration in this locus. On the basis of the present data we can only speculate on the cellular function of ARL4. However, because of its specific expression in testicular germ cells, it seems reasonable to assume that the GTPase is involved in sperm production.

Received 19 March 1998/29 June 1998; accepted 28 July 1998

We thank Petra Mühl-Zürbes and Debra J. Gilbert for excellent technical assistance. This research was supported by a grant of the Deutsche Forschungsgemeinschaft (Jo 117/10-2) and by the National Cancer Institute, DHHS, under contract with ABL.

REFERENCES

1 Kahn, R. A., Kern, F. G., Clark, J., Gelmann, E. P. and Rulka, C. (1991) J. Biol. Chem. **266**, 2606–2614
2 Tsuchiya, M., Price, S. R., Tsai, S. C., Moss, J. and Vaughan, M. (1991) J. Biol. Chem. **266**, 2772–2777
3 Stearns, T., Willingham, M. C., Botstein, D. and Kahn, R. A. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 1238–1242
4 Helms, J. B. and Rothman, J. E. (1992) Nature (London) **360**, 352–354
5 Kahn, R. A., Randazzo, P., Serafini, T., Weiss, O., Rulka, C., Clark, J., Amherdt, M., Roller, P., Orci, L. and Rothman, J. E. (1992) J. Biol. Chem. **267**, 13039–13046
6 Brown, H. A., Gutowski, S., Moomaw, C. R., Slaughter, C. and Sternweis, P. C. (1993) Cell **75**, 1137–1144
7 Kahn, R. A., Yucel, J. K. and Malhotra, V. (1993) Cell **75**, 1045–1048
8 Rümenapp, U., Geiszt, M., Wahn, F., Schmidt, M. and Jakobs, K. H. (1995) Eur. J. Biochem. **234**, 240–244
9 Tamkun, J. W., Kahn, R. A., Kissinger, M., Brizuela, B. J., Rulka, C., Scott, M. P. and Kennison, J. A. (1991) Proc. Natl. Acad. Sci. U.S.A. **88**, 3120–3124
10 Clark, J., Moore, L., Krasinskas, A., Way, J., Battey, J., Tamkun, J. and Kahn, R. A. (1993) Proc. Natl. Acad. Sci. U.S.A. **90**, 8952–8956
11 Cavenagh, M. M., Breiner, M., Schürmann, A., Rosenwald, A. G., Terui, T., Zhang, C., Randazzo, P. A., Adams, M., Joost, H. G. and Kahn, R. A. (1994) J. Biol. Chem. **269**, 18937–18942
12 Schürmann, A., Breiner, M., Becker, W., Huppertz, C., Kainulainen, H., Kentrup, H. and Joost, H. G. (1994) J. Biol. Chem. **269**, 15683–15688
13 Schürmann, A., Massmann, S. and Joost, H. G. (1995) J. Biol. Chem. **270**, 30657–30663
14 Breiner, M., Schürmann, A., Becker, W. and Joost, H. G. (1996) Biochim. Biophys. Acta **1308**, 1–6
15 Lee, C. M., Haun, R. S., Tsai, S. C., Moss, J. and Vaughan, M. (1992) J. Biol. Chem. **267**, 9028–9034
16 Serventi, I. M., Cavanaugh, E., Moss, J. and Vaughan, M. (1993) J. Biol. Chem. **268**, 4863–4872
17 Copeland, M. G. and Jenkins, M. A. (1991) Trends Genet. **7**, 113–118
18 Jenkins, N. A., Copeland, N. G., Taylor, B. A. and Lee, B. K. (1982) J. Virol. **43**, 26–36
19 Fletcher, C. F., Okano, H. J., Gilbert, D. J., Yang, C., Copeland, N. G., Jenkins, N. A. and Darnell, R. B. (1997) Genomics **45**, 313–319
20 Weiland, M., Schürmann, A., Schmidt, W. E. and Joost, H. G. (1990) Biochem. J. **270**, 331–336
21 Feinberg, A. P. and Vogelstein, B. (1983) Anal. Biochem. **132**, 6–13
22 Keime, S., Adham, I. M. and Engel, W. (1990) Eur. J. Biochem. **190**, 195–200
23 Zhou, Y., Sun, Z., Means, A. R., Sassone-Corsi, P. and Bernstein, K. E. (1996) Proc. Natl. Acad. Sci. U.S.A. **93**, 12262–12266
24 Tsai, S. C., Haun, R. S., Tsuchiya, M., Moss, J. and Vaughan, M. (1991) J. Biol. Chem. **266**, 23053–23059
25 Stafford, W. H. L., Stockley, R. W., Ludbrock, S. B. and Holder, A. A. (1996) Eur. J. Biochem. **242**, 104–113
26 Hendriksen, P. J., Hoogerbrugge, J. W., Baarends, W. M., de-Boer, P., Vreeburg, J. T., Vos, E. A., van der Lende, T. and Grootegoed, J. A. (1997) Genomics **41**, 350–359
27 Magnuson, M. A. and Shelton, K. D. (1989) J. Biol. Chem. **264**, 15936–15942
28 Plaza, S., Dozier, C., Turque, N. and Saule, S. (1995) Mol. Cell. Biol. **15**, 3344–3353
29 MacDougald, O. A. and Lane, M. D. (1995) Annu. Rev. Biochem. **64**, 345–373
30 Goraya, T. Y., Kessler, S. P., Stanton, P., Hanson, R. W. and Sen, G. C. (1995) J. Biol. Chem. **270**, 19078–19085
31 Sun, Z., Sassone-Corsi, P. and Means, A. R. (1995) Mol. Cell. Biol. **15**, 561–571
32 Seidah, N. G., Day, R., Hamelin, J., Gaspar, A., Collard, M. W. and Chretien, M. (1993) Mol. Endocrinol. **6**, 1559–1570
33 Zhou, W. and Goldberg, W. (1996) Biol. Reprod. **54**, 84–90
34 Wang, Y., Goligorsky, M. S., Lin, M., Wilcox, J. N. and Maseden, P. A. (1997) J. Biol. Chem. **272**, 11392–11401
35 Foulkes, N. S., Mellström, B., Benusiglio, E. and Sassone-Corsi, P. (1992) Nature (London) **355**, 80–84
36 Foulkes, N. S., Schlotter, F., Pevet, P. and Sassone-Corsi, P. (1993) Nature (London) **362**, 264–267
37 Wichmann, H., Disela, C., Haubruck, H. and Gallwitz, D. (1989) Nucleic Acids Res. **17**, 6737–6738
38 McGrath, J. P., Capon, D. J., Smith, D. H., Chen, E. Y., Seeburg, P. H., Goeddel, D. V. and Levinson, A. D. (1983) Nature (London) **304**, 501–506