Neuronal migration is a complex process that is affected in a variety of human disorders such as periventricular heterotopias and different types of lissencephalies (1–3). The Disabled-1 (Dab1) gene belongs to the Reelin signaling pathway that plays a key role during brain development in mouse and human (4–6). Inactivation of Dab1, either by homologous recombination (7) or by spontaneous mutations in scrambler or yotari mutant mice (8, 9), generates a phenotype similar to that of Reelin-deficient mice. This phenotype is characterized by a poor organization of architectonic patterns at the end of radial neuronal migration (reviewed in Ref. 10). The neurons that are the most affected include those of the cortical plate in the cortex and hippocampus, Purkinje cells, and inferior olivary neurons. In humans, mutations in Reelin result in a specific lissencephaly with mental retardation and severe abnormalities of the cerebellum, hippocampus, and brain stem (Norman-Roberts type, OMIM257320) (11), a phenotype that shows similarity to its mouse counterpart. Cognitive development is delayed, with little or no language acquisition and no ability to sit or stand unsupported. Thus far, no human disease associated with mutations in DAB1 or other genes in the Reelin pathway has been identified.

Reelin is an extracellular protein secreted by some neurons such as Cajal-Retzius cells in the marginal zone of the embryonic cerebral cortex and hippocampus, external granule cells in the cerebellum, olfactory mitral cells, and ganglion and amacrine cells in the mouse retina (10, 12) and in the spinal cord (13, 14). The response of target neurons to Reelin requires the expression of at least one of two surface receptors that belong to the lipoprotein receptor family, viz. the very low density lipoprotein receptor and apoE receptor type 2, as well as the presence of the intracellular adaptor DAB1. The DAB1 protein contains a 180-amino acid N-terminal protein interaction/phosphotyrosine-binding (PTB) domain that docks to the short cytoplasmic tail of the very low density lipoprotein receptor or apoE receptor type 2 at the level of NPYX motifs, with a preference for unphosphorylated motifs (15–19). Potential tyrosine phosphorylation sites and a 310-amino acid C-terminal region of unknown function follow the PTB domain. The binding of Reelin to the extracellular part of both receptors induces phosphorylation of tyrosine residues of DAB1, particularly Tyr188 and Tyr220 (20, 21). Mice expressing a mutant form of DAB1 in which all the potential tyrosine phosphorylation sites are mutated have a phenotype similar to reeler mice (4), and mice expressing a truncated DAB1 protein missing the C-terminal part have an almost normal phenotype (22). This shows that the PTB domain and tyrosine phosphorylation are both necessary and sufficient to fulfill most of the DAB1 functions. Cytoplasmic tyrosine kinases of the Src family are able to phosphorylate DAB1 in vitro, but the kinase(s) involved in DAB1 phosphorylation in vivo remain to be identified (23). Similarly, the other downstream effectors of the Reelin signal are not known.

* This work was supported in part by Grant 3.4533.95 from the Fonds de la Recherche Scientifique et Médicale, Grants 186 and 248 from the Actions de Recherches Concertées, and by the Fondation Reine Elisabeth (Belgium). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Chargée de Recherche of the Fonds National de la Recherche Scientifique.
‡ Postdoctoral Fellow supported by European Union Program CONCORDE Grant QLG3-CT-2000-00158.

© 2003 by The American Society for Biochemistry and Molecular Biology, Inc.
Various isoforms of the mouse DAB1 protein have been described. The main form contains an open reading frame of 555 amino acids encoding a 58-kDa protein, the predominant form expressed in the brain. Another form, 555°, contains an additional exon inserted in-frame between codons 241 and 242. Form 217 results from alternative polyadenylation, whereas isoform 271 is similar to form 555, except that an additional exon of 270 bp has been detected with a probe covering the PTB domain, and protein isoforms of 36, 45, 60, 80, and 120 kDa have been observed on Western blots (23).

In this work, we defined the genomic organization of the human and mouse Dab1 genes. The structure is highly complex and similar in both species. The gene extends over 1 Mbp of genomic DNA due to the presence of large introns and the wide dispersion of several alternative transcription initiation sites. The presence of several alternative promoters and alternatively spliced forms points to a fine regulation of Dab1 expression and further emphasizes the key position of this gene as a switch in the Reelin signaling pathway. The complexity of the gene may explain why no human disease associated with DAB1 mutations could be identified thus far.

EXPERIMENTAL PROCEDURES

Rapid Amplification of cDNA Ends (RACE)—Information on the primers used is provided in Table I. First-strand cDNA synthesis was performed at 60 °C on 1 μg of DNase I-treated RNA (embryonic human brain and P0 mouse brain) using Dab1-specific primer in exon 2 (Table I) and Thermoscript reverse transcriptase (Invitrogen). After Rnase H digestion, the cDNA was purified with a GlassMAX spin cartridge (Invitrogen) and modified by adding a polydeoxycytidine sequence to the 5′-end using terminal deoxynucleotidyltransferase (Invitrogen). cDNAs were amplified by PCR with a primer pair (5′-GCCACGGTCGACTAGTACTGGGGGIGGGIGGGIG-3′) and antisense Dab1-specific primer 2 using Taq DNA polymerase (Biolo). Nested PCRs were performed with the abridged universal amplification primer (5′-GCCACGGTCGACTAGTACTGGGGGIGGGIGGGIG-3′) and antisence Dab1-specific primer 3. The PCR products were cloned into the pCRII vector (Invitrogen) and sequenced using the BigDye terminator cycle sequencing kit (PE Biosystems) and an ABI Prism 377 sequencer.

3′-RACE was carried out on 2 μg of total RNA from E15 mouse brain. First-strand cDNA synthesis was performed using an adaptor oligo(dT) primer (5′-GCCACGGTCGACTAGTACTGGGGGIGGGIGGGIG-3′) with Superscript II reverse transcriptase (Invitrogen). cDNA was amplified using Dab1-specific primer 4 in exon 15 and the abridged universal amplification primer (5′-GCCACGGTCGACTAGTACTGGGGGIGGGIGGGIG-3′) and antisense Dab1-specific primer 3. The PCR products were cloned into the pCRII vector (Invitrogen) and sequenced using the BigDye terminator cycle sequencing kit (PE Biosystems) and an ABI Prism 377 sequencer.

5′-RNA Ligase-mediated RACE—5′-RNA ligase-mediated RACE was performed using the GeneRacer kit (Invitrogen). 250 ng of mRNA from E15 mouse brain was dephosphorylated using calf intestinal phosphatase and decapped using tobacco acid pyrophosphatase to target full-length mRNAs. An RNA oligonucleotide was ligated to the decapped mRNA, and reverse transcription was performed at 60 °C using random primers and Thermoscript reverse transcriptase. PCR was done to amplify the resultant cDNAs using the GeneRacer 5′-primer and Dab1-specific primer 6 in exon 3 or Dab1-specific primer 7 in exon 1B1. Nested PCR was done using the GeneRacer 5′-nested primer and Dab1-specific primer 8 in exon 2, primer 9 in exon 1B1, primer 10 in exon 1A, and primer 11 in exon 1B2.
**The Dab1 Gene**

and primer 11 in exon 1D. Products were cloned into pCRII and sequenced. **Genomics**—PACs containing parts of the human and mouse Dab1 genes were obtained by PCR screening of the RZPD Deutsches Ressourcenzentrum fuer Genomforschung libraries 711 (RPCI21; constructed by Dr. Pieter de Jong, Roswell Park Cancer Institute) and 709 (RPCI6). Human PACs were isolated as follows: PACs RPC18-239D12 and RPC16-233E22 with primers 12 and 13; PAC RPC16-6SP20 with primers 14 and 15; and PAC RPC16-102010 with primers 16 and 1. Mouse PACs RPC121-97L11 and RPC121-31E11 were isolated with primers 17 and 18. Intron sizes were determined by PCR on genomic or PAC DNA. Amplification of large fragments was carried out with the Expand amplification system (Invitrogen). The sequences of the exon-intron junctions of human and mouse genomic regions were determined by direct sequencing of PAC DNA or of PCR products. Exon-intron boundaries were determined by aligning cDNA and genomic sequences. Results were confirmed by comparison with the human and mouse genome sequences (human contig, NCBI accession number NT_029223.8 (build 30); and mouse contig, NCBI accession number NW_000211.1).

**Reporter Plasmid Constructs**—Fragments were cloned into the pGL3-Basic promoterless vector (Promega). A 2.9-kb mouse fragment named ACD (containing exons 1A–1D) was amplified from genomic DNA by PCR using primers 19 and 17. The PCR product was cloned into pCRII and cloned in the forward and reverse orientations into the SacI site of pGL3-Basic. Fragment AD (821 bp) was amplified using primers 20 and 21, cloned into pCRII, and cloned (SpcI-XbaI) in the forward and reverse orientations into the NheI site of pGL3-Basic. Construct A (248 bp) was derived from construct AD by MluI digestion and ligation. To obtain fragment D (574 bp), fragment AD in pCRII was digested with KpnI and MluI and cloned into the corresponding sites of pGL3-Basic for the forward construct and digested with HindIII and MluI and cloned into the corresponding sites for the reverse construct. Fragment C (1641 bp) was amplified from genomic DNA using primers 22 and 17, cloned into pCRII, and cloned in the forward and reverse orientations into the SacI site of pGL3-Basic. Fragment B (1.3 kb) was amplified with primers 23 and 24, cloned into pCRII, and cloned into the KpnI-XhoI site of pGL3-Basic. A PstII-EcoRV digestion/ligation of this construct was carried out to remove three ATG codons present in the 5′ -UTR.

**Cell Culture and Transient Transfections**—Undifferentiated P19, HepG2, and HEK293 cells were grown in Dulbecco–UTR. Primary neuronal cultures were prepared from embryonic brain—

and transfection, one-half of the medium was replaced. For transfection, 4 μg of LipofectAMINE 2000 (Invitrogen)/well in 50 μl of Opti-MEM (Invitrogen) was mixed with 1 μg of DNA in 50 μl of Opti-MEM and incubated at room temperature prior to addition to cells in 400 μl of culture medium. Following 5 h of incubation at 37 °C, the supernatant was removed and replaced with 1 ml of medium. Assay for luciferase activity was done 24 h after transfection.

**In Situ Hybridization and Northern Blotting**—The probe for exon 1A (130 bp) was amplified with primers 25 and 28. The probe for exon 1B was amplified with primers 27 and 28 and cloned into pBluescript. This 713-bp fragment contains four exons of 5′ -UTR 1B (see Fig. 1): 1B1, 1B2, 1B4, and 1B8. The probe for exon 1D (343 bp) was amplified with primers 29 and 30. The probe for alternative exon 555* was a 189-bp fragment amplified with primers 31 and 32 and cloned into pCRII. It contains both exons 555* plus 27 bp of exon 9 and 47 bp of exon 10. A 480-bp probe covering most of the PTE region was amplified with primers 16 and 33 and cloned into pCRII. In situ hybridization was carried out as explained in the legend to Fig. 5. The probe for the mouse Dabi 3′ -UTR was a 2-kb fragment amplified with primers 34 and 35.

**RESULTS**

The published physical map of the mouse Dab1 gene contains inconsistencies (8, 9). We therefore reconsidered the physical map and genomic organizations of the human and mouse Dab1 genes as summarized in Fig. 1 and Table II.

**Physical Maps**—The mouse Dab1 gene maps to chromosome 4 at 52.7 centimorgans, whereas human DAB1 maps to chromosome 1p32-p31 (24). In both species, the 5′-end is located on the centromeric side. The gene is flanked by the mouse AK008820 locus (identical to human XM_055482 or LOC115209) on the centromeric side and by the complement factor 8B (CSB) gene on the telomeric side. As shown in Fig. 1A, Dab1 is located between microsatellites D4Mit118 and D4Mit75, but is at least 3 Mbp away from marker D4Mit117, confirming the data published by Ware et al. (9). D4Mit331 is located between Dab1 exons 1B1 and 1B2 (described below), 565 kb upstream of the Dab1 ATG codon and 700 kb centromeric to D4Mit29, in agreement with the reported genetic distance of 0.6 centimorgans between D4Mit29 and D4Mit331 (9). D4Mit29 is located in intron 4, 1.5 kb from exon 4 of Dab1 (Fig. 1B). D4Mit75 maps 440 kb distal to the ATG codon and 300 kb telomeric to D4Mit29.

In *scrambler* mutant mice, a portion of an intracisternal A particle (IAP) sequence is inserted in the antisense orientation in the Dab1 mRNA by aberrant splicing. The mutation results in production of an enlarged transcript of ~7 kb, with the introduction of multiple stop codons. The defect results from the use of a cryptic splice acceptor site in intron 4 coupled with a cryptic donor site in the IAP element. In the *scrambler* mRNA, 28 bases unrelated to Dab1 or to the IAP are inserted between Dab1 exon 4 and the IAP sequence. BLAST alignment against the mouse genome localized these 28 bases in intron 4, 11 kb distal to exon 4. No IAP sequence is present in this region in the C57BL/6 DNA. Although we cannot exclude that the IAP element was present in the DCHe strain in which the *scrambler* mutation arose, it appears more likely that an IAP insertion caused the mutation. In *notari* mutant mice, 357 nucleotides corresponding to exons 5–8 are missing from the Dab1 mRNA, and the open reading frame is maintained. At the genomic level, this deletion in the mRNA is due to the insertion of a 962-bp L1 element. This insertion starts at the junction of exon 5-intron 5 and ends in the middle of exon 8 (25).

**Genomic Organization**—The mouse genomic organization was derived from direct sequencing on PACs and sequencing of PCR products derived from YACs and genomic DNA. Data from the NCBI mouse genome sequence (accession number 5804
The Dab1 Gene

The entire human DAB1 ORF (Fig. 1B) is contained in four PAC clones covering ∼300 kb of genomic DNA. Clone RPCI6-102010 (132 kb; accession number AL390243) contains 120 kb of intron 1, exon 2, and 15 kb of intron 2; clone RPCI6-65FP20 (107 kb; accession number AL138779) contains exons 3–6; clone RPCI6-239D12 (175 kb; accession number AL161740) contains exons 10–15 and the C8B gene; and clone RPCI6-225E22 (not sequenced) contains exon 5 up to at least exon 15. The human PACs have been characterized by the Sanger Center using fluorescent in situ hybridization and sequence determination and are not chimeric.

The Dab1 coding regions (from exon 2 containing the ATG codon to exon 14 containing the stop codon) are spread over 254 kb of genomic DNA for the mouse gene and over 294 kb for the human gene. The size of the major DAB1 protein is 555 amino acids, which corresponds to a coding capacity of 0.2% compared with a mean genomic coding density of ∼10%. The organization of the mouse and human genes is conserved, and all exon-intron splice junctions conform to the GT/AG rule (Fig. 1C).

With the exception of exons 12 and 15 (549 and >3300 bp, respectively), exons are relatively small, ranging in size from 39 to 140 bp. Introns in the ORF region range in size from 89 bp to 146 kb. The PTB domain is encoded in exons 3–6. Important tyrosine residues are encoded in exons 6 (Tyr185), 7 (Tyr198), 8 (Tyr200 and Tyr220), and 9 (Tyr232) (4, 21).

The 3′-end of the mouse Dab1 mRNA was determined using 3′-RACE. Using a primer in exon 15, we found a Dab1 3′-untranslated segment that extends 1252 bp downstream from the stop codon. This sequence contains several putative polyadenylation signals and aligns with several ESTs. Another set of ESTs align with genomic sequences 1 kb farther downstream (Fig. 2A). RT-PCR with primer 34 defined in this downstream EST set and other primers in exons 14 and 15 showed that the
The upper part concerns the 5'-UTRs for the mouse (Mm) and human (Hs) genes. The methods of validation were RACE, RT-PCR, and/or GenBank™/EBI Data Bank entries as indicated. The ATG column indicates the number of ATG codons present in each 5'-UTR exon. The lower part indicates the exon-intron organization of the conserved part, with human and mouse in the top and bottom lines, respectively. RLM, RNA ligase-mediated.

| Exon number | Size(bp) | 5' junction | Intron size(bp) | 3' junction | Verification procedure | ATG |
|-------------|---------|-------------|----------------|-------------|------------------------|-----|
| 2 (Mm)      | 202     | AAG AAA G gtaccgctg | 227000 | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 3 (Mm)      | 160     | CAG GCC gtaagcatg | 228000 | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 5 (Mm)      | 199     | AAG AAA G gtggctgc | 125902 | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 6 (Mm)      | 120     | TAC AGC gtaaacctg | 65763  | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 9 (Mm)      | 149     | TAG CGC gtaaccc | 461    | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 217* 267    |       | Not found in Human | 6415   | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 8 (Mm)      | 66     | TAT GTG gtaacatg | 6722   | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 10 (Mm)     | 103     | CCG CCC gtaaagtg | 2340   | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 11 (Mm)     | 109     | CCG CCC gtaaagtg | 2352   | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 12 (Mm)     | 44      | ACG AAC gtaaccc | 7346   | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 13 (Mm)     | 118     | GAA GCT gtaaccc | 355    | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 14 (Mm)     | 111     | GTC TGG gtaaagtg | 12552  | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
| 15          | 130     |                | 9344   | tggctttctccc gag act | RACE, RT-PCR, AB075525, AB055513 | 9   |
3'-UTR extends at least until another polyadenylation signal located 3325 bp from the stop codon. To verify that the Dab1 3'-UTR is >3 kb long, we performed Northern blot analysis of mouse brain mRNA using a 2 kb probe corresponding to this 3'-sequence (Fig. 2B). This probe revealed a single band of 5.5 kb, whereas probes corresponding to the Dab1 PTB coding region revealed a band of similar size plus two additional bands of lower size (see Fig. 4). In human, four EST clones (accession numbers AA5411650, AI799728, R52905, and R67274) contain a polyadenylation signal (followed by a poly(A) tail) localized 3344 bp downstream from the DAB1 stop codon.

Alternative First Exons and Organization of the 5'-Region—Comparison of the different Dab1 sequences revealed extensive variation in the 5'-regions. In mouse, the Dab1 cDNA sequence initially described (accession number Y083379) (23) contains 263 nucleotides of 5'-UTR encoded by exon 1A and 136 bp of 5'-UTR encoded by exon 2, which contains the ATG codon. The macaque AB055528 and human AK095513 DAB1 cDNAs contain another 5'-UTR (1B) of 497 and 532 bp, respectively. The human AF263547 DAB1 cDNA contains a different 5'-UTR region of 629 bp, and the human XM_010707 cDNA sequence contains yet another different first exon of 508 bp.

As these data suggest the presence of alternative first exons, we performed 5'-RACE and RNA ligase-mediated RACE on embryonic human and mouse brain RNAs. Using embryonic mouse brain RNA, four different products named 1A–1D were obtained (Fig. 1C and Table II). Mouse fragment 1A corresponds to Dab1 exon 1 in sequence Y083379 (23) and is found in 10 mouse and two rat ESTs. RACE product 1B is similar to human DAB1 cDNA sequences AF263547, AB055528, and AK095513 and is present in three mouse ESTs and one human EST. In mouse, RT-PCR and RACE experiments with primers specific for this product 1B revealed that it does not correspond to a single exon, but is composed of combinations of 10 different exons, 1B1–1B10 (Fig. 1C). Exon 1C does not correspond to any published sequence, but could be amplified by RT-PCR, whereas exon 1D is novel and is present in one mouse and one rat EST sequence; it is conserved in human and mouse and was amplified from human brain RNA by RT-PCR. Attempts to map the transcription initiation sites by primer extension were unsuccessful, possibly because of high GC content and secondary structures of the alternative first exons. Using repeated RNA ligase-mediated RACE reactions on poly(A) RNA from E17 mouse brain, we were unable to extend the Dab1 UTR sequences further and therefore considered them close to full-length. All the sequences obtained by RNA ligase-mediated RACE were also obtained using classical RACE reactions, and all RACE products were shown to be connected to Dab1 by RT-PCR.

RACE reactions on human brain RNA yielded four different products. One is similar to mouse exon 1A, with no match in human EST data bases. A second RACE product is similar to the highly complex mouse fragment 1B and is present in human cDNA clones AF263547, AB05528, and AK095513. RACE and RT-PCR experiments revealed that 5'-UTR 1B is composed of combinations of at least seven different exons. We were not able to clone the 5'-end of the reconstituted RNA sequence XM_010707 using RACE or to connect it to DAB1 exon 2 using RT-PCR on embryonic human brain RNA. Exon E is not present in EST data bases, but can be connected to DAB1 exon 2 by RT-PCR. Exon F is similar to bases 608–701 of sequence XM_060465, the rest of which is unrelated to DAB1.

Among the novel 5'-UTRs, three are conserved between human and mouse (Fig. 1C), viz. exons 1A (90% identity), 1B (mouse 1B1/human 1B1, 65% identity; mouse 1B2/human 1B4, 67% identity; and mouse 1B4/human 1B7, 79% identity), and 1D (55% identity). The sequences are present in EST data bases and have been isolated as cDNA by others, thus confirming their expression.

Genomic Organization of the 5’-Region—The Dab1 5’-UTR spreads over 850 kb in mouse and 961 kb in human (Fig. 1, B and C). Whereas exons 1A, 1C, and 1D in mouse and exons 1A and 1D–1F in human are clustered in a 1.5-kb fragment, the complex 5’-UTR 1B has a highly unusual structure. It is composed of 10 exons (1B1–1B10) in mouse and seven exons (1B1–1B7) in human, with sizes ranging from 63 to >675 nucleotides, separated by introns with sizes between a few hundred nucleotides and >300 kb. The sequence of UTR 1B is dispersed over >800 kb of genomic DNA, which is consistent with physical mapping data. YAC 175A2, which is 1500 kb in length, contains the region between exon 1B8 and coding exon 9, but does not contain exons 1B1 and 1B2 (Fig. 1C). Exons 1B1–1B10 are flanked with consensus splice sites and obey the GT/AG rule. Both in mouse and man, the alternative exons that compose UTR 1B contain numerous ATG codons and upstream ORFs (uORFs) (Table II). For example, exon 1B1 contains one uORF of 87 codons in human and of 101 codons in mouse, and the first 55 encoded amino acids are highly conserved. uORFs are common in certain genes that are involved in the control of cellular growth and differentiation. This may have implications for the control of DAB1 mRNA translation, as many examples have been described in which ORFs present in the 5'-UTR influence expression levels (26, 27).
Expression of Alternative First Exons—To assess whether the different 5'-exons have different expression patterns, PCRs were carried out on mouse brain cDNA at different stages from E11 to E18 and at postnatal stages from P0 to adult. We used forward primers in the alternative first exons and reverse primers in exons 2 and 5 (Fig. 3A). As shown in Fig. 3B, exons 1A and 1D were expressed in all stages tested. Amplification of exon 1C was weak, indicating low level expression in the tissues examined (data not shown). The complex UTR 1B was barely detectable at E11 and E12, whereas two main bands were amplified in RNA isolated from brain at E15 and later, including adult (Fig. 3B). We tested the expression of the alternative first exons in P19 cells, which differentiate into neurons in the presence of RA (Fig. 3C). Exons 1A and 1D were detected in undifferentiated and differentiated P19 cells. In contrast, the expression of UTR 1B was complex. Multiple bands were amplified in undifferentiated cells and up to 4 days after RA induction. When neural induction was complete, only two bands were visible. This developmental regulation was confirmed in vivo. As shown in Fig. 3 (C and D), a pattern of multiple bands amplified from E11 mouse RNA becomes restricted to two main amplicons at P0 and adult. Sequencing of the two main amplicons showed that they are formed of fragments 1B1, 1B2, and 1B4, with alternative inclusion of fragment 1B8. Interestingly, the sequences of fragments 1B1, 1B2, and 1B4 are conserved in mouse, man, and macaque.

FIG. 3. Expression of alternative 5’-UTRs. A, schematic representation of 5’-UTR exons (boxes) and primers used in RT-PCR, with orientation indicated by arrows. All reactions were carried out using 30 cycles of PCR and 25 ng of cDNA. UTRs 1A, 1B, 1C, and 1D are shaded differently. Exons 2–5 are coding exons. Note the complexity of UTR 1B, also shown in C–E. B, RT-PCR analysis of UTR 1A, 1B, and 1D expression during mouse brain development from E11 to adult. The following primer combinations were used: for exon 1A, primers 33 and 22 (amplicon of 666 bp); for UTR 1B, primers 33 and 37 (amplicons of 974 and 1085 bp corresponding to exons 1B1, 1B2, and 1B4 and exons 1B1, 1B2, 1B4, and 1B8, respectively); and for exon 1D, primers 33 and 38 (amplicon of 663 bp).

Glyceraldehyde-3-phosphate dehydrogenase (G3PDH) was used as a reference gene. C, RT-PCR analysis of exons 1A, 1B, and 1D during P19 cell differentiation induced by RA from days 0 to 9 and in control P0 brain. The following primer combinations were used: for exon 1A, primers 2 and 22 (375-bp product); for exon 1D, primers 38 and 2 (351 bp); and for UTR 1B, primers 39 and 2 (450 and 560 bp, respectively, and several larger products). Hypoxanthine-guanine phosphoribosyltransferase (HPRT) was used as a reference gene. D, illustration of the complexity of UTR 1B using RT-PCR (primers 39 and 2). Shown is a comparison of undifferentiated (−RA) versus differentiated (+RA) P19 cells and of developing (E11) versus mature (adult Ad) mouse brain. Note the simplification of the UTR 1B amplification pattern in parallel to neural maturation. E, organization of UTR 1B. PCR products in C and D were cloned and sequenced (exon terminology as described for Fig. 1C). RT-PCR on non-neural tissues showed the inclusion of numerous alternative exons of UTR 1B.

FIG. 4. Expression of alternative 5’-UTR 1A. Poly(A)+ RNA (2 µg) from P0 mouse brain was analyzed by Northern blotting using 32P-labeled probes. Transcripts of three different sizes (~5.5, 4, and 1.3 kb, indicated by arrowheads) were detected with a probe for the PTB domain (left). Three bands of similar sizes plus a band of 1.8 kb were detected with a probe for exon 1A (right).

Expression of Alternative First Exons—To assess whether the different 5’-exons have different expression patterns, PCRs were carried out on mouse brain cDNA at different stages from E11 to E18 and at postnatal stages from P0 to adult. We used forward primers in the alternative first exons and reverse primers in exons 2 and 5 (Fig. 3A). As shown in Fig. 3B, exons 1A and 1D were expressed in all stages tested. Amplification of exon 1C was weak, indicating low level expression in the tissues examined (data not shown). The complex UTR 1B was barely detectable at E11 and E12, whereas two main bands were amplified in RNA isolated from brain at E15 and later, including adult (Fig. 3B). We tested the expression of the alternative first exons in P19 cells, which differentiate into neurons in the presence of RA (Fig. 3C). Exons 1A and 1D were detected in undifferentiated and differentiated P19 cells. In contrast, the expression of UTR 1B was complex. Multiple bands were amplified in undifferentiated cells and up to 4 days after RA induction. When neural induction was complete, only two bands were visible. This developmental regulation was confirmed in vivo. As shown in Fig. 3 (C and D), a pattern of multiple bands amplified from E11 mouse RNA becomes restricted to two main amplicons at P0 and adult. Sequencing of the two main amplicons showed that they are formed of fragments 1B1, 1B2, and 1B4, with alternative inclusion of fragment 1B8. Interestingly, the sequences of fragments 1B1, 1B2, and 1B4 are conserved in mouse, man, and macaque.
Fig. 5. Expression of Dab1 isoforms: in situ hybridization. 32P-Labeled riboprobes were hybridized to cryostat sagittal sections of E14 (A, B, D, E, and G) and P0 (C, F, and I) mouse brain, and the signal was revealed by dipping in LMI emulsion (Amersham Biosciences) as described (40, 41). The three following probes were used. The PTB probe (A–C) covers most of the Dab1 PTB coding region and is expected to reveal all Dab1 isoforms. Probe 1B (D–F) corresponds to exons 1B1, 1B2, 1B4, and 1B8 (see Fig. 1) and revealed expression of alternative 5′-exon 1B. Probe 555′ (G–I) corresponds to exons 555′ plus some flanking cDNA sequences. The PTB probe (A–C) confirmed very strong expression in the cortical plate at E14 and P0. At P0, the signal was stronger in the outer than in the inner tiers of the cortical plate (CP). There was also moderate expression in the ventricular zone (VZ), which was more evident at E14 than at P0. The probe covering exons 1B (D–F) revealed an expression similar to that obtained with the PTB probe, with maximal expression in the cortical plate and weak expression in the ventricular zone. Probe 555′ (G–I) revealed a pattern different from that obtained with the PTB probe. The signal was the strongest in the ventricular zone, particularly at early stages (E14), and decreased at P0. This probe also revealed a signal in the cortical plate, but part of this signal may be related to the parts of the probe that correspond to Dab1 cDNA sequences adjacent to exons 555′. V, ventricle. Bars = 100 μm.

The presence of multiple 5′-exons suggests that the transcription of Dab1 is regulated by different promoters. Promoter activity was assessed by transient transfection of HEK293 and HepG2 cells, which do not express Dab1, as well as undiffer-
entiated P19 embryonic carcinoma cells and embryonic mouse primary neuronal cultures, which express Dab1 (Fig. 6) (23). A 150-bp fragment upstream of exon 1A (construct A+ in Fig. 6A) was active in all cells tested (25-fold in neurons, 8-fold in P19, 12-fold in HEK293, and 7-fold in HepG2). A comparable or higher activity was observed when this sequence was cloned in the reverse orientation (construct A in Fig. 6B), suggesting that this region may function as a bidirectional promoter, at least in vitro. This may be related to its high GC content (75% in mouse with 32 CpG dinucleotides and 80% in human with 38 CpG dinucleotides), with three SP1-binding sites conserved between mouse and human and coupled with the absence of TATA and CAAT sequences. Construct C+, containing the region upstream of exon 1C, showed weak promoter activity in all cell lines tested. This segment has a lower GC content of ~56%. The promoter prediction programs Promoter Inspector, TSSW, and TSSG detected a promoter, a degenerate TATA box, and a transcription initiation site in this region. Construct D+, corresponding to the 500-bp region upstream of exon 1D, was 6-fold more active than the promoterless vector in HEK293 cells and neurons and 3-fold more active in HepG2 and P19 cells. This region had no promoter activity when cloned in the reverse orientation. The programs also predicted a promoter and a degenerate TATA box in this segment. Construct AD+, which contains both regions upstream of exons 1A and 1D, showed promoter activity comparable to that of fragment A+. Construct ACD+, which contains exons 1A, 1C, and 1D, showed promoter activity comparable to that of fragment C+. There was no activity of this segment when cloned in the reverse orientation. Two constructs were used to assay the promoter activity of regions upstream of exon 1B1 (data not shown). A 1.3-kb construct that includes 350 nucleotides of exon 1B1 and three ATG codons was inactive in primary cortical neurons. To avoid possible interference of the ATG codons with translation of the luciferase reporter, another construct was derived by deleting these ATG triplets. However, no promoter activity was detected in primary cortical neurons, indicating that the promoter of form 1B may be located farther upstream in the 260-kb genomic interval between exon 1B1 and the AK008020 gene. Internal Alternative Splicing Events—Using PCR on human brain cDNA and alignment of genomic and EST sequences, we were unable to identify exons corresponding to mouse Dab1 forms 217 and 271 in human. Using RT-PCR, the presence of fragment 555* was confirmed in mouse and man. In both species, this sequence corresponds to two exons of 51 and 48 bp separated by an intron of 91 bp in mouse and of 89 bp in human (Fig. 7A). Both exons were consistently co-amplified. Interestingly, an alternatively spliced product of 57 bp was detected in the corresponding location in the Dab1 cDNA in lizard and chick (Fig. 7B). As shown in Fig. 7C, the peptide sequences encoded by the two small exons that form fragment 555* in mouse and man and by the single 57-nucleotide exon in lizard and chick are conserved, suggesting a duplication event during evolution. Upon Northern blotting using a probe that includes exons 555* and some adjacent sequences, a major RNA species of ~5.5 kb, presumably corresponding to the longest form of the Dab1 mRNA, was detected in poly(A) RNA from E17 mouse brain (data not shown). In undifferentiated P19 cells, the Dab1 cDNA did include fragment 555*. However, when differentiation of P19 cells was induced with RA, a proportion of Dab1 cDNA without fragment 555* appeared at day 2 and increased progressively to become the major form at day 9 (Fig. 7D). In early embryonic mouse brain (E11 and E12), the Dab1 isoform with fragment 555* was predominant, but RNAs from later developmental stages (E12 and later) and from primary neuronal cultures did not contain this fragment (Fig. 7B). In non-neural tissues such as liver and kidney, the Dab1 mRNA contained fragments 555* (data not shown). A similar pattern was
found in chick, with inclusion of the small 57-nucleotide exon in RNA from E6 or adult eye, but exclusion of that exon from brain RNA at E20 (Fig. 7B).

Using in situ hybridization with a cDNA probe covering fragments 555* and adjacent segments (Figs. 5 (G–I) and 7A), a strong signal was detected in ventricular zones of precursor proliferation; the moderate labeling of post-migratory fields indicated that a strong signal was detected in ventricular zones of precursor proliferation; the moderate labeling of post-migratory fields

**DISCUSSION**

Both in man and mouse, the Dab1 gene reveals an unusual complexity that leaves ample room for subtle regulation of its expression and function. Examples of such highly complex genomic organization are few and include the metabotropic glutamate receptor GRM8 gene, which spans >800 kb of genomic DNA for a coding length of 2.3 kb (28), and the human neurotrophin receptor genes NTRK2 and NTRK3 (29), which extend over >350 and 380 kb for coding lengths of 3.7 and 2.8 kb, respectively. Intriguingly, the Dab1 paralogous gene Dab2 (named DAB2 or DOC2 in human) is much simpler than Dab1, with an ORF extending over <50 kb of genomic sequence compared with 300 kb for Dab1 (30, 31). Apparently, this situation is not unusual. For example, the NTRK1 gene, closely related to NTRK2 and NTRK3, spreads over only 20 kb (29). A similar feature is found in the two mouse paralogous phospholipase D genes Pld1 and Pld2. Whereas Pld1 contains 28 exons and spans ~147 kb, the whole Pld2 gene is contained in 17 kb of genomic DNA (32).

From an evolutionary standpoint, it would be interesting to know whether such large differences in the genomic complexity of paralogous genes result from extension or contraction of the set of introns in one of the genes after duplication. Like most *Drosophila* genes, the fly *Disabled* gene has small introns and extends over 12 kb of genomic DNA (33), suggesting that the large size of Dab1 might result from intron extension.

Our results also reveal a remarkable diversity in the 5′-UTR of both the human and mouse Dab1 genes. We have identified six alternative 5′-UTRs in human and four in mouse, three of which are conserved. Fragments with promoter activity were defined for two of them, but we were unable to clone the promoter for 5′-UTR 1B. This 5′-UTR is unusually complex and spreads over 1 Mb of genomic DNA. It is composed of seven exons in human and 10 exons in mouse, with three exons conserved between both species and always included together in the mRNA. This results in a 5′-UTR of 1 kb or more (with the inclusion of alternative exons), which is much larger than the average size of 210 bp (34, 35). In situ hybridization using a probe specific for UTR 1B and RT-PCRs clearly showed that it is part of the Dab1 mRNA. The long 5′-UTR of Dab1 contains small uORFs and numerous upstream ATG codons that precede the major translation initiation site, some of which are conserved between both species and always included together in the mRNA. This results in a 5′-UTR of 1 kb or more (with the inclusion of alternative exons), which is much larger than the average size of 210 bp (34, 35). In situ hybridization using a probe specific for UTR 1B and RT-PCRs clearly showed that it is part of the Dab1 mRNA. The long 5′-UTR of Dab1 contains small uORFs and numerous upstream ATG codons that precede the major translation initiation site, some of which are conserved between both species and always included together in the mRNA. This results in a 5′-UTR of 1 kb or more (with the inclusion of alternative exons), which is much larger than the average size of 210 bp (34, 35).
altered regulation of gene expression in vivo.

Large gene size and complexity may be important for the production and processing of the transcripts. Based on a transcription rate of $\approx 1.4$ kb/min (37) and data on the dystrophin gene (38), transcription of Dab1 would require at least 13 h. This is close to or larger than the estimated division time of neuronal precursors (39), suggesting that the promoter associated with form 1B could not be utilized in proliferating cells. In summary, our data show that Dab1 is far more complex than expected and that further work is needed to understand better the control of Dab1 expression and the molecular machinery by which it exerts its powerful activity. The detailed genomic structure reported here should facilitate the study of human DAB1 mutations, which are predicted to yield abnormal brain phenotypes similar to those related to Reelin deficiency.

REFERENCES
1. Gupta, A., Tsai, L. H., and Wynshaw-Boris, A. (2002) Nat. Rev. Genet. 3, 342–355
2. Monuki, E. S., and Walsh, C. A. (2001) Nat. Neurosci. 4, suppl. 1199–1206
3. Lambert de Rouvroit, C., and Goffinet, A. M. (2001) Mech. Dev. 105, 47–56
4. Howell, B. W., Herrick, T. M., Hildebrand, J. D., Zhang, Y., and Cooper, J. A. (2000) Curr. Biol. 10, 877–885
5. Rice, D. S., and Curran, T. (2001) Annu. Rev. Neurosci. 24, 1005–1039
6. Feng, Y., and Walsh, C. A. (2001) Nat. Rev. Neurosci. 2, 408–416
7. Howell, B. W., Hawkes, R., Serianno, P., and Cooper, J. A. (1997) Nature 389, 733–737
8. Shieldon, M., Rice, D. S., D’Arcangelo, G., Yonemisha, H., Nakajima, K., Mikoshiba, K., Howell, B. W., Cooper, J. A., Goldowitz, D., and Curran, T. (1997) Nature 389, 730–733
9. Ware, M. L., Fox, J. W., Gonzalez, J. L., Lambert de Rouvroit, C., Russo, C. J., Chua, S. C., Jr., Goffinet, A. M., and Walsh, C. A. (1997) Neuron 19, 239–249
10. Lambert de Rouvroit, C., and Goffinet, A. M. (1998) Adv. Anat. Embryol. Cell Biol. 150, 1–106
11. Hong, S. E., Shugart, Y. Y., Huang, D. T., Shahwan, S. A., Grant, P. E., Hourihane, J. O., Martin, N. D., and Walsh, C. A. (2000) Nat. Genet. 26, 93–96
12. Rice, D. S., and Curran, T. (2000) J. Comp. Neurol. 424, 327–338
13. Carroll, P., Gayet, O., Feuillet, C., Kallenbach, S., de Bovis, B., Dudley, K., and Alonso, S. (2001) Mol. Cell. Neurosci. 17, 611–623
14. Yip, J. W., Yip, Y. P., Nakajima, K., and Caporossi, C. (2000) Proc. Natl. Acad. Sci. U. S. A. 97, 8612–8616
15. D’Arcangelo, G., Homayouni, R., Keshvara, L., Rice, D. S., Sheldon, M., and Curran, T. (1999) Neuron 24, 471–479
16. Trommsdorff, M., Gotthardt, M., Hiesberger, T., Shelton, J., Stockinger, W., Nimpf, J., Hammer, R. E., Richardson, J. A., and Herz, J. (1999) Cell 97, 669–701
17. Hiesberger, T., Trommsdorff, M., Howell, B. W., Goffinet, A., Mumbry, M. C., Cooper, J. A., and Herz, J. (1999) Neuron 24, 481–489
18. Gotthardt, M., Trommsdorff, M., Nevitt, M. P., Shelton, J., Richardson, J. A., Stockinger, W., Nimpf, J., and Herz, J. (2000) J. Biol. Chem. 275, 25616–25624
19. Howell, B. W., Lanier, L. M., Frank, R., Gertler, F. B., and Cooper, J. A. (1999) Mol. Cell. Biol. 19, 5179–5188
20. Howell, B. W., Herrick, T. M., and Cooper, J. A. (1999) Genes Dev. 13, 643–648
21. Keshvara, L., Benhayon, D., Magdaleno, S., and Curran, T. (2001) J. Biol. Chem. 276, 16008–16014
22. Herrick, T. M., and Cooper, J. A. (2002) Development 129, 787–796
23. Howell, B. W., Gertler, F. B., and Cooper, J. A. (1997) EMBO J. 16, 121–132
24. Lambert de Rouvroit, C., and Goffinet, A. M. (1998) Genomics 53, 246–247
25. Kojima, T., Nakajima, K., and Mikoshiba, K. (2000) Brain Res. Mol. Brain Res. 75, 121–127
26. van der Velden, A. W., and Thomas, A. A. (1999) Int. J. Biochem. Cell Biol. 31, 87–106
27. Morris, D. R., and Geballe, A. P. (2000) Mol. Cell. Biol. 20, 8635–8642
28. Scherer, S. W., Soder, S., Duvoisin, R. M., Huijzenga, J. J., and Tsui, L. C. (1997) Genomics 44, 232–236
29. Valee, A., Danilet, G., and Berneheim, A. (1997) Eur. J. Hum. Genet. 5, 102–104
30. Sheng, Z., He, J., Tuppen, J. A., Sun, W., Faull, Z., Smith, E. R., Dong, F. B., and Xu, X. X. (2000) Genomics 70, 381–386
31. Sheng, Z., He, J., Tuppen, J. A., Martin, W. D., Dong, F. B., and Xu, X. X. (2001) Gene 288, 31–39
32. Redina, O. E., and Frohman, M. A. (1998) Gene 222, 53–60
33. Gertler, F. B., Hill, K. K., Clark, M. J., and Hoffmann, F. M. (1993) Genes Dev. 7, 441–453
34. Pessole, G., Liuni, S., Grillo, G., Licciulli, F., Mignone, F., Gissi, C., and Sacccone, C. (2002) Nucleic Acids Res. 30, 335–340
35. Mignone, F., Gissi, C., Liuni, S., and Pesole, G. (2002) Genome Biol. 3, S0004
36. Phelps, D. E., Hsiao, K. M., Hu, N., Franklin, D. S., Westphal, E., Lee, E. Y., and Xiong, Y. (1998) Mol. Cell. Biol. 18, 2334–2343
37. Sheroen, A. W., and O’Farrell, P. H. (1991) Cell 67, 303–310
38. Tennyson, C. N., Klamut, H. J., and Worton, R. G. (1995) Nat. Genet. 9, 184–190
39. Takahashi, T., Nowakowski, R. S., and Caviness, V. S., Jr. (1995) J. Neurosci. 15, 6046–6057
40. Bernier, B., Bar, I., Paeau, C., Lambert de Rouvroit, C., and Goffinet, A. M. (1999) J. Comp. Neurol. 413, 463–479
41. Simmons, D., Arriza, J. L., and Swanson, L. W. (1989) J. Histotechnol. 12, 169–181
The Gene Encoding Disabled-1 (DAB1), the Intracellular Adaptor of the Reelin Pathway, Reveals Unusual Complexity in Human and Mouse
Isabelle Bar, Fadel Tissir, Catherine Lambert de Rouvroit, Olivier De Backer and André M. Goffinet

J. Biol. Chem. 2003, 278:5802-5812.
doi: 10.1074/jbc.M207178200 originally published online November 21, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M207178200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 41 references, 11 of which can be accessed free at http://www.jbc.org/content/278/8/5802.full.html#ref-list-1